

UNIVERSIDAD PONTIFICIA COMILLAS DE MADRID

ESCUELA TÉCNICA SUPERIOR DE INGENIERIA
(Instituto de Investigación Tecnológica)

DISEÑO REGULATORIO DE LA ACTIVIDAD DE TRANSPORTE EN MERCADOS DE ELECTRICIDAD REGIONALES

Tesis para la obtención del grado de Doctor
Director: Prof. Dr. D. José Ignacio Pérez Arriaga
Autor: Ing. D. Luis Olmos Camacho



Madrid 2006

**DISEÑO REGULATORIO DE LA
ACTIVIDAD DE TRANSPORTE EN
MERCADOS DE ELECTRICIDAD
REGIONALES**

**CONSTANCIA REGISTRAL DEL TRIBUNAL DEL
ACTO DE LA DEFENSA DE TESIS DOCTORAL**

TÍTULO:

AUTOR:

DIRECTOR:

TUTOR-PONENTE:

DEPARTAMENTO:

FACULTAD O ESCUELA:

Miembros del Tribunal Calificador:

PRESIDENTE:

Firma:

VOCAL:

Firma:

VOCAL:

Firma:

VOCAL:

Firma:

SECRETARIO:

Firma:

Fecha de lectura:

Calificación:

Agradecimientos

Son muchos y muy merecidos los agradecimientos que he de dar a personas e instituciones que de un modo u otro han tenido que ver con la elaboración de esta tesis.

Comenzar por mi familia, sin cuyo apoyo, comprensión y sacrificios esto no hubiese sido posible. Tampoco puedo olvidarme de Cristina, que me ha animado día a día y me ha apoyado cuando todo parecía complicarse.

Ignacio, mi director de tesis y supervisor, ha sabido guiarme en toda esta travesía con una mente preclara y su gran conocimiento. Sin su ayuda nunca podría haber, siquiera abordado, el complejo problema al que me he enfrentado en la tesis.

El director del IIT, Michel, y todo el personal del centro, con su gran labor, han contribuido a crear un marco incomparable para el desarrollo de mi trabajo cotidiano. Los miembros del IIT forman un grupo humano impresionante. Nunca han dudado en prestarme su apoyo cuando lo he necesitado. Debo dar las gracias, en especial, a Carlos Vázquez por su generosidad y lo mucho que me ha enseñado. Él ha sido para mí un segundo director de tesis. Juan Luis y Ramón me brindaron la posibilidad de entrar en el IIT y aprendí mucho de ellos en el año en que estuvimos trabajando juntos, tanto en lo profesional como en lo humano.

También es de recibo agradecer a Red Eléctrica de España la financiación que he recibido durante estos dos últimos años, y a José Luis Fernández y Luis Villafruela, su trabajo y colaboración en aquellos proyectos en que hemos coincidido. En la financiación de la tesis también han contribuido la Comisión Europea, a través de la Dirección general de Transporte y Energía y el Fondo Social Europeo, y la Comunidad de Madrid, mediante la concesión de una beca FPI.

Por último dar las gracias a todos aquellos que me he dejado por el camino. No penséis que me olvido de vosotros, simplemente sería imposible enumeraros a todos.

Table of contents

1	Introduction.....	1
1.1	Presentation	2
1.2	Transmission of electricity	2
1.2.1	Investment in new transmission facilities	4
1.2.2	Access to the transmission grid.....	5
1.2.3	Computation of transmission tariffs.....	6
1.3	The regulation of the transmission activity in the regional electricity markets	8
1.3.1	The Internal Electricity Market of the European Union.....	10
1.3.1.1	The Florence Regulatory Forum.....	13
1.3.2	The Regional Electricity Market of Central America	13
1.3.3	The Regional Electricity Markets in the US.....	18
1.3.4	The Market in South America.....	21
1.4	Objectives of the thesis	23
1.5	Layout of this document.....	26
	Bibliography	28
	Annex 1. Brief introduction to the concept and properties of nodal prices.....	29
	Mathematical Proof.....	29
	The case of generators	29
	The case of consumers.....	33
	Comments on the proof	36
	Other properties of nodal prices.....	37
	Decomposition of spot prices	37

Recovery of the cost of constructing, operating and maintaining the grid	39
2 Investments.....	41
2.1 Introduction.....	42
2.2 Regulatory paradigms in transmission investment.....	44
2.2.1 Centralized expansion of the transmission grid.....	44
2.2.2 The expansion of the grid is left to a private for-profit company.....	47
2.2.3 Network users form coalitions to promote the construction of new lines	49
2.2.4 Merchant lines.....	52
2.3 Selection of the optimal scheme for the regulation of the grid expansion.....	54
2.3.1.1 General considerations	54
2.3.1.2 Recommendations.....	55
2.3.2 The European Electricity Market.....	58
2.3.2.1 The situation in the region	58
2.3.2.2 The proposed regulatory mechanism for the development of the grid	58
2.3.3 The regional markets in the US	62
2.3.3.1 The situation in the region	62
2.3.3.2 Regulatory mechanism proposed for the development of the grid	64
2.3.4 The Central-American regional market.....	68
2.3.4.1 The situation in the region	68
2.3.4.2 Regulatory mechanism proposed for the development of the grid	70
2.3.5 Possible mechanisms to deal with conflicts among countries or states that are related to the construction of lines that belong to the regional grid.....	74

2.4	The regulatory test	75
2.4.1	Description of the regulatory test to be applied in different situations.....	78
2.4.1.1	Proposal of a (set of) reinforcement(-s) by a SO	78
2.4.1.2	Assessment by the regulator of a proposal by another party	79
2.4.2	Cost factors to be taken into account when assessing potential grid reinforcements.....	80
2.4.2.1	Criteria not directly related to the investment project.....	81
2.4.2.2	Criteria directly related to the investment project	83
2.4.3	Discussion on the treatment to be given to the negative benefits brought about by the installation of a new line.....	83
2.4.3.1	Treatment to be given to the negative benefits associated with a regulated grid investment.....	84
2.4.3.2	Treatment to be given to the negative benefits associated with grid investments by private parties	86
2.4.4	Some final remarks on the regulatory test.....	87
2.4.5	Classification of the existing algorithms to compute an expansion plan for the grid	87
2.4.5.1	Solution method used	88
2.4.5.2	Timing of the investments proposed	89
2.4.5.3	Ability to adapt to the situation emerging from the restructuring of the electricity sector.....	90
2.5	Compatibility of investment signals in distribution, transmission and generation.....	91
2.5.1	Compatibility of investment signals in transmission and generation	92
2.5.1.1	The difficulties	92
2.5.1.2	Signals derived from transmission pricing.....	93
2.5.1.3	Recommendations for practical implementation	95

2.5.1.4 A proposal to compute incremental transmission charges for new network users	98
2.5.1.5 Other locational signals for generators and network planners	103
2.5.2 Compatibility of investment signals in transmission and distribution.....	105
2.5.3 Case examples.....	109
2.5.3.1 Numerical results resulting from the application of the method to compute incremental transmission charges for new network users	110
2.5.3.2 Comparison of transmission charges to other locational signals to market agents in the IEM.....	114
2.5.3.3 Locational Signals affecting the decision by a CCGT on where to install	117
2.6 Conclusions	121
2.6.1 Interaction between the regulation of the expansion of the grid and that of other aspects of transmission regulation.....	124
Bibliography	126
Annex 1: Mathematical formulation of the problem of centrally planning the expansion of the grid.....	134
Centralized planning of the expansion of the grid in a traditional regulatory setting.....	134
Differences and similarities in centralized transmission planning under the traditional and competitive regulatory frameworks.....	136
3 Transmission tariffs	139
3.1 Introduction.....	140
3.2 Conceptual analysis of the computation of transmission tariffs.....	141
3.2.1 Discussion of several issues related to the implementation of a method to allocate the grid costs	146
3.2.1.1 Treatment to be given to the unused fraction of the capacity of a line.....	146

3.2.1.2 Treatment of the generation and load that are located within the same node.....	146
3.3 Review of network cost allocation methods	147
3.3.1 Connection charges.....	148
3.3.2 Network use of the system charges.....	148
3.3.2.1 Conceptually sound methods	149
3.3.2.2 Other methods.....	155
3.3.2.3 Methods recently proposed in the technical literature.....	158
3.3.3 Application to regional markets.....	160
3.3.3.1 Regional markets in the US.....	160
3.3.3.2 The IEM in the European Union	161
3.3.3.3 The market in Australia	163
3.3.3.4 The MER in Central America	163
3.3.3.5 The MERCOSUR market in South America.....	166
3.3.4 Different aspects of the implementation of any transmission tariff. Structure of the tariff and others.....	166
3.3.4.1 Structure of transmission charges.....	166
3.3.4.2 Recovery of the cost of the unused fraction of the capacity of a line.....	168
3.3.4.3 Number of snapshots considered when computing transmission tariffs	169
3.3.4.4 Global split of transmission charges between generation an demand	169
3.4 Proposed solution to the problem of allocating the grid costs.....	169
3.4.1 Main usage-based methods.....	170
3.4.1.1 Average Participations	171
3.4.1.2 Marginal Participations (MP)	173
3.4.2 Assessment of the main usage-based methods.....	176

3.4.2.1 Soundness of the method from a technical point of view	177
3.4.2.2 Compatibility with the remaining aspects of the regulation of the transmission service	181
3.4.2.3 Compliance with the Single System Paradigm principle	182
3.4.2.4 Soundness of the method from an economic point of view	183
3.4.2.5 Data and algorithmic requirements of the different methods.....	184
3.4.2.6 Complexity of the method, dependence of the method on commonly agreed definitions and experience with its utilization	185
3.4.2.7 Assessment of the results of application of the method to a representative regional system	185
3.4.2.8 Conclusion	193
3.4.3 Foundation of the method of Average Participations.....	194
3.4.3.1 Rationale of the method from the point of view of economics theory	194
3.4.3.2 Rationale of the method from the point of view of the physics of power systems	199
3.4.4 Structure of the tariffs.....	201
3.4.4.1 Connection charges.....	201
3.4.4.2 Access charges	201
3.4.5 Numerical results.....	202
3.4.5.1 Results for the AP method	203
3.5 Extension of the computation of transmission tariffs to the case of regional systems	212
3.5.1 The European system	212
3.5.1.1 The market context	212
3.5.1.2 Methods considered for the computation of inter-TSO compensations.....	214
3.5.1.3 Further considerations related to the implementation of any method	227

3.5.1.4	Coherence of the selected method with other aspects of transmission regulation	231
3.5.1.5	Results for the European system.....	237
3.5.1.6	Comparative analysis of the preferred methods	263
3.5.1.7	Final recommendation of a procedure.....	274
3.5.2	The regional electricity market in Central America.....	275
3.5.2.1	Description of the adopted regulation	275
3.5.2.2	Description of the proposed solution	277
3.5.2.3	Numerical results	281
3.5.3	Other regional markets.....	284
3.5.3.1	The MERCOSUR regional electricity market.....	285
3.5.3.2	The Standard Market design in the US	287
3.6	Conclusions	290
	Bibliography	294
	Annex 1: Mathematical properties of the participations of agents in the use of lines according to MP.....	303
4	Access to the grid.....	305
4.1	Introduction.....	306
4.2	Treatment of local grid restrictions.....	307
4.2.1	Regulation of the connection of conventional, non-renewable, generation capacity.....	307
4.2.2	The connection of renewable energy producers to the grid.....	310
4.3	Congestion management mechanisms in the short term and the long term	311
4.3.1	Discussion of different algorithms for the short term energy dispatch	313
4.3.1.1	Terminology and approach classification	315
4.3.1.2	Formulation of the implicit auction scheme	323

4.3.1.3 Coordinated explicit auction of transmission capacity followed by an energy auction where the grid is not considered	337
4.3.1.4 Dealing with Market Power in the short term. Implications of the use of different types of short-term auctions for energy and capacity in the exercise of market power.....	341
4.3.1.5 Managing congestion in the short term, the case of three different regional markets	343
4.3.2 Definition and subsequent trading of capacity contracts previous to the short term energy dispatch.....	351
4.3.2.1 Introduction to the use of transmission rights.....	352
4.3.2.2 Different types of transmission rights	353
4.3.2.3 Sequence of capacity markets leading to the short term energy dispatch.....	357
4.3.2.4 Effect that the existence and allocation of congestion rights may have on the exercise of market power	359
4.3.2.5 Formulation of a long term capacity auction including both options and obligations.....	362
4.3.2.6 Regulatory recommendations on the implementation of a system of long-term transmission contracts in a regional market	364
4.4 Treatment of some specific issues related to the management of congestion.....	367
4.4.1 Allocation of congestion rents to the lines in the system. Marginal value of a line in the dispatch.....	367
4.4.1.1 Extension of the analysis to the case where contingencies are considered in the dispatch. Application to an auction for transmission rights that includes options	370
4.4.1.2 Possible applications of this formula in the regulation of transmission systems	373
4.4.2 Treatment of special restrictions regarding the allocation of the corresponding revenues.....	374
4.4.3 Definition of Single Price Areas.....	377
4.4.3.1 Conceptual analysis of the problem.....	378

4.4.3.2 Identification of Single Price Areas in the Iberian Peninsula and the system Spain-France-Portugal	387
4.4.3.3 Conclusions of the work on the identification of SPAs	408
4.4.4 Identification of a balancing point for electricity transmission contracts.....	409
4.4.4.1 Methodology for the computation of price sensitivities	412
4.4.4.2 Identification of a balancing point in the European system	417
4.4.4.3 Some final remarks on the definition of a balancing point	426
4.5 Conclusions	427
4.5.1 Interaction between the regulation of the access to the grid by market agents and other aspects of transmission regulation.....	429
Bibliography	431
Annex 1: Equivalence, under ideal conditions, between regional explicit auctions followed by local implicit ones and regional implicit auctions	439
Annex 2: Equivalence between the net revenue from the application of nodal prices and that associated with the aggregate marginal cost of restrictions limiting the flow on lines	442
5 Conclusions.....	445
5.1 Summary and conclusions.....	446
5.1.1 Reference regulation of the transmission activity	447
5.1.1.1 Investments in new transmission facilities	447
5.1.1.2 Computation of transmission tariffs	449
5.1.1.3 Access to the grid	452
5.1.2 Regulation of transmission in regional markets.....	454
5.1.2.1 The Internal Electricity Market of the European Union	456
5.1.2.2 The regional market in Central America (MER).....	458
5.1.2.3 Development of regional markets in the US	460

5.2	Original contributions.....	464
5.2.1	Contributions related to the design of transmission tariffs.....	464
5.2.2	Contributions related to the management of congestion.....	465
5.2.3	Contributions related to the regulation of the expansion of the grid	466
5.3	The way forward.....	467
5.3.1	Allocation of the cost of construction of the grid	467
5.3.2	Management of congestion in the grid.....	467
5.3.3	Investments in new transmission facilities.....	468
	Bibliography	469

List of Figures

Figure 2.5-1: Process of computation of locational signals	101
Figure 2.5-2: Geographical distribution of locational signals related to the cost of transmission lines when computed using the ALF method. Values obtained for the year 2008 in the Spanish power system.....	114
Figure 2.5-3: Comparison between the transmission tariff and the net inter-TSO payment for 17 European countries.	116
Figure 2.5-4: Evolution of average quarterly energy prices of several Power Exchanges belonging to Europex from January 2000 to November 2004.....	116
Figure 3.4-1: Proportionality principle in average participations.....	173
Figure 3.4-2: Participation of generation and demand nodes in the lines of the UCTE system according to AP	179
Figure 3.4-3: Two areas example.....	181
Figure 3.4-4: Stylized system used to show the properties of the results produced by the AP and MP methods.....	187
Figure 3.4-5: Path followed by the flow injected by the generator in node 2 when the slack node is in node 8.....	191
Figure 3.4-6: Path followed by the flow injected by the generator in node 2.	193
Figure 3.4-7: allocation of an inflow ϕ to a node among the outflows.....	196
Figure 3.4-8: Original and new L tariffs within Spain for the IEM-13 system.....	205
Figure 3.4-9: Original and new G tariffs within Spain for the IEM-13 system	206
Figure 3.4-10: Geographical distribution over the Spanish system of the nodal L transmission tariffs computed using AP for the IEM-13 system	206
Figure 3.4-11: Geographical distribution over the Spanish system of the nodal G transmission tariffs computed using AP for the IEM-13 system	207
Figure 3.4-12: Geographical distribution over the Spanish system of the nodal L transmission tariffs computed for the IEM-13 system using AP and applying the smoothing process afterwards.....	209

Figure 3.4-13: Geographical distribution over the Spanish system of the nodal G transmission tariffs computed for the IEM-13 system using AP and applying the smoothing process afterwards.....	210
Figure 3.4-14: Geographical distribution of nodal L tariffs obtained using AP to allocate the cost of the used fraction of each line. The remaining part is socialized.....	211
Figure 3.4-15: Geographical distribution of nodal G tariffs obtained using AP to allocate the cost of the used fraction of each line. The remaining part is socialized.....	211
Figure 3.5-1: Definition of the flow pattern of the transit at the borders.....	221
Figure 3.5-2: Transit according to ETSO for two different network topologies	221
Figure 3.5-3: The influence of political borders on the definition of transits.....	222
Figure 3.5-4: Selection of the scenarios to be taken into account in the computation of ITCs.....	231
Figure 3.5-5: Average nodal L and G tariffs for the 13 countries in the IEM network.....	252
Figure 3.5-6: Nodal tariffs for all the consumers in the IEM-13 network.....	252
Figure 3.5-7: Nodal tariffs for all the generators in the IEM-13 network.....	253
Figure 3.5-8: Vectors corresponding to the representative and all the samples within each of the 12 clusters of net cross-border flows.....	256
Figure 3.5-9: Representation of the samples of 12 clusters of net cross-border flows among countries according to the time of the year when they occur	258
Figure 3.5-10: Geographical distribution of the net inter-TSO payments in 2002 for AP.....	259
Figure 3.5-11: Geographical distribution of the net inter-TSO payments in 2002 for WWT.....	260
Figure 3.5-12: Geographical distribution of the net inter-TSO payments in 2002 for APT	260
Figure 3.5-13: Geographical distribution of compensations to countries in 2002 for the AP method.....	261

Figure 3.5-14: Geographical distribution of compensations to countries in 2002 for the WWT method	261
Figure 3.5-15: Geographical distribution of compensations to countries in 2002 for the APT method.....	262
Figure 3.5-16: Geographical distribution of charges to be levied on countries in 2002 for the AP method.....	262
Figure 3.5-17: Geographical distribution of charges to be levied on countries in 2002 for the WWT method	263
Figure 3.5-18: Geographical distribution of charges to be levied on countries in 2002 for the APT method.....	263
Figure 3.5-19: stylized representation of the IEM used in the case example.....	272
Figure 3.5-20: Net power balance for each country in Central America over the year 2002.....	282
Figure 4.3-1: 10 node system comprising 3 different control areas: A1, A2 and A3....	316
Figure 4.4-1: multivariable Kohonen map and maps corresponding to each of the classification variables for the division into areas of the Iberian peninsula when 12 cells are used.....	383
Figure 4.4-2: void map showing the identity of the samples belonging to each of the cells in the previous <i>kohonen</i> map. Two clusters (SPAs) have been identified. We show in bold the identity of the samples within the smaller area.....	383
Figure 4.4-3: Maximum errors committed for the different classifications obtained corresponding to the division of the Iberian peninsula into two areas....	389
Figure 4.4-4: Average errors committed for the different classifications obtained corresponding to the division of the Iberian peninsula into two areas....	389
Figure 4.4-5: Maximum error committed within each SPA for all the classifications computed corresponding to the division of the Iberian peninsula into two SPAs.....	390
Figure 4.4-6: Average error committed over all the intra-area transactions within each SPA for all the classifications computed corresponding to the division of the Iberian peninsula into two SPAs.....	390

Figure 4.4-7: Comparison of the maximum and average errors for all the classifications considered corresponding to the division of the Iberian peninsula into two SPAs.....	391
Figure 4.4-8: Comparison of the maximum transaction errors committed for the division of the Iberian system into 2, 3 and 4 areas	393
Figure 4.4-9: Comparison of the average transaction errors committed considering 2, 3 and 4 SPAs in the Iberian peninsula	393
Figure 4.4-10: Map of the French system where the geographical areas considered have been represented.....	395
Figure 4.4-11: Maximum transaction errors committed for the different classifications corresponding to the division of the system comprised of France, Spain and Portugal into three areas	395
Figure 4.4-12: Average transaction errors committed for the different classifications corresponding to the division of the system comprised of France, Spain and Portugal into three areas	396
Figure 4.4-13: Maximum transaction error committed for each area and classification corresponding to the division into three areas of the system comprised of France, Spain and Portugal.....	396
Figure 4.4-14: Average transaction error committed for each area and classification corresponding to the division into three areas of the system comprised of France, Spain and Portugal.....	397
Figure 4.4-15: Comparison of the maximum and average transaction errors in each of the classifications obtained for the division into three areas of the system comprised of France, Spain and Portugal.....	397
Figure 4.4-16: Division into three SPAs adopted for the system comprised of Spain, France and Portugal	398
Figure 4.4-17: Maximum transaction error committed for the different classifications obtained corresponding to the division of the system comprised of France, Spain and Portugal into four areas.....	399
Figure 4.4-18: Average transaction errors committed for the different classifications computed corresponding to the division of the system comprised of France, Spain and Portugal into four areas.....	399
Figure 4.4-19: Maximum transaction error committed within each area and classification for the division into four areas of the system comprised of France, Spain and Portugal.....	400

Figure 4.4-20: Average transaction error committed within each area and classification for the division into four areas of the system comprised of France, Spain and Portugal.....	400
Figure 4.4-21: Maximum and average transaction errors for each of the classifications corresponding to the division into four areas of the system comprised of France, Spain and Portugal.....	401
Figure 4.4-22: Division into four SPAs adopted for the system comprising Spain, France and Portugal	401
Figure 4.4-23: Geographical distribution of the 5 SPAs that have been defined in the system that is comprised of the French, Spanish and Portuguese ones.....	403
Figure 4.4-24: Maximum transaction errors committed when dividing the system comprised of France, Spain and Portugal into 3, 4 and 5 areas.....	404
Figure 4.4-25: Average transaction errors committed when dividing the system comprised of France, Spain and Portugal into 3, 4 and 5 areas.....	404
Figure 4.4-26: Number of nodal transactions in the Iberian system for different threshold levels.....	406
Figure 4.4-27: Average intra-area transaction error in the Iberian system for different threshold levels.....	406
Figure 4.4-28: Number of nodal transactions in the system comprised of France, Spain and Portugal for different threshold levels.....	407
Figure 4.4-29: Average intra-area transaction error in the system comprised of France, Spain and Portugal for different threshold levels	407
Figure 4.4-30: Geographical location of 17 UCTE countries considered in the study	418
Figure 4.4-31: Impact on energy prices of a marginal decrease in the power output of an agent located in Germany (node i).....	419
Figure 4.4-32: Example three-node network creating negative cross-price sensitivities	420
Figure 4.4-33: Relative cross-price sensitivities sorted by magnitude, with the generation node located in KKPhili (Germany).....	421

Figure 4.4-34: Effect of the number of generation nodes considered on the shape of the curve representing the maximum relative cross-price sensitivity for each node in the system	422
Figure 4.4-35: Curves representing the 95 percentile of the probability distribution of relative cross-price sensitivities for each node in the system. Each curve has been obtained for a different number of generation nodes.....	423
Figure 4.4-36: Election of the balancing point for the European system using both the maximum value of the distribution and the 95 percentile.....	424
Figure 4.4-37: Geographical distribution of the 20 best candidates to become the European balancing point	425

List of Tables

Table 2.5-1: locational signals for new generators in the Spanish system coming from grid charges and transmission losses. Values estimated for the year 2008.....	112
Table 2.5-2: Impact of different factors on the total generation capacity needed to supply a 384 MW load, located close to a main consumption centre, from two different locations, one close to the load centre and the other one close to an entry point for LNG	120
Table 2.5-3: Comparison of the cost savings involved in supplying a 384 MW load located close to a main load center.....	121
Table 3.4-1: Participations by agents in the usage made of lines, resulting from the MP method, when we take node 8 as the slack bus.....	188
Table 3.4-2: Participations of countries in the use made of lines, resulting from the MP method, when we take node 8 as the slack bus.....	188
Table 3.4-3: Compensations among countries produced by the MP method corresponding to the usage that external agents make of the grid in each country. We take node 8 as the slack bus.....	188
Table 3.4-4: Participations by agents in the usage made of electrical lines, resulting from the MP method, when we take node 2 as the slack bus....	189
Table 3.4-5: Participations of countries in the use made of lines, resulting from the MP method, when we take node 2 as the slack bus.....	189
Table 3.4-6: Compensations among countries produced by the MP method corresponding to the usage that external agents make of the grid in each country. We take node 2 as the slack bus.....	190
Table 3.4-7: Participations by agents in the usage made of electrical lines resulting from the AP method.....	191
Table 3.4-8: Participations of countries in the use made of lines resulting from the AP method.....	192
Table 3.4-9: Compensations among countries produced by the AP method corresponding to the usage that external agents make of the grid in each country	192
Table 3.5-1: Aggregated data for the considered UCTE scenario.....	238

Table 3.5-2: Compensations due to countries because of network usage for the year 2001	240
Table 3.5-3: Charges to be paid by countries because of network usage for the year 2001	240
Table 3.5-4: Inter-TSO payments because of network usage for the year 2001	240
Table 3.5-5: Total economic compensations (M€) related to network infrastructure costs for the year 2001	241
Table 3.5-6: Total economic charges (M€) related to network infrastructure costs for the year 2001	241
Table 3.5-7: Total volume of inter-TSO payments (M€) related to network infrastructure costs for the year 2001	242
Table 3.5-8: Aggregate data for the average scenario of 2002 in the 13-IEM system....	244
Table 3.5-9: Annual compensations, charges and net payments because of network usage produced by the AP method for the year 2002	245
Table 3.5-10: Compensations, charges and net payments for the first scenario of the year 2002 obtained using the basic AP method	246
Table 3.5-11: Compensations, charges and net payments yielded by AP for the first scenario of the year 2002 when only the cost of the used fraction of each line is considered to compute ITCs.....	246
Table 3.5-12: Annual compensations, charges and net payments because of network usage yielded by the WWT method for the year 2002	247
Table 3.5-13: Annual compensations, charges and net payments because of network usage yielded by the APT method for the year 2002.....	248
Table 3.5-14: Annual compensations, charges and net payments because of network usage yielded by the PM method for the year 2002	248
Table 3.5-15: Summary of the results of inter-TSO payments among countries for the AP, WWT, APT and PM methods.....	250
Table 3.5-16: A plausible selection of 12 scenarios representative of the system operation over a whole year.....	257
Table 3.5-17: inter-TSO compensations produced by the WWT method when applied to the stylized representation of the IEM system.....	272

Table 3.5-18: inter-TSO compensations produced by the WWT method when applied to the stylized representation of the IEM system and P is split into two new systems P1 and P2.....	273
Table 3.5-19: Compensations among countries produced by the AP method corresponding to the usage that external agents make of the grid in each country	273
Table 3.5-20: Compensations among countries produced by the MP method corresponding to the usage that external agents make of the grid in each country. We take node 8 as the slack bus.....	273
Table 3.5-21: comparison of the net compensation to be received by countries in the stylized example according to MP, considering two different choices of the slack bus, AP and WWT.....	274
Table 3.5-22: results provided by the MFD method for Central America in 2002.....	282
Table 3.5-23: results provided by the AP method for Central America in 2002	283
Table 4.4-1: Comparison of the maximum and average errors committed in each of the two classifications selected for dividing the Iberian peninsula into 4 areas. Values are expressed in %.....	392
Table 4.4-2: Maximum and average error for each of the 5 areas identified in the system comprised of France, Spain and Portugal.....	403

1 Introduction

1.1 Presentation

This thesis is about the regulation of the activity of transmission in regional electricity markets. Transmission regulation has frequently been a contentious topic when trying to introduce competition in national markets and, without any doubt, is the major stumbling block in the creation of regional markets. This thesis draws from multiple experiences in the development of supra national markets, and the best conceptual contribution from regulatory economics applied to transmission networks, in order to propose sound mechanisms for the organization of the transmission activity in regional electricity markets. Besides presenting a comprehensive regulatory package for transmission, the thesis also examines how this reference regulation could be applied in several existing representative regional markets, such as the Internal Electricity Market of the European Union (IEM), the regional electricity market in Central America (MER) or the regional markets under development in the USA. Besides, several specific topics in transmission regulation are explored in more depth. These include the choice of an adequate method to compute compensations among countries for the external use that is made of their networks, the design of the most adequate locational signals to be embedded in transmission charges to be sent to market agents under different circumstances, the identification of Single Price Areas (SPAs) for the successful application of coordination schemes in congestion management and the identification of a balancing point for the definition of transmission rights.

1.2 Transmission of electricity

The electricity transmission grid is the meeting point for generators and consumers. This grid allows power producers to supply loads with the energy they need to carry out their economic activities. Similarly to what happens in other network industries (the gas industry or the transport industry for example), developing alternative transmission networks that compete against each other in order to provide the electricity transmission service makes no sense economically, socially and environmentally. Therefore, the electricity transmission grid in any system must be unique.

Electricity cannot be stored. Consequently, generation and load in any system must be balanced at any time. The power consumed by loads must be produced at the same time by generators. Mainly because of this reason, the transmission grid is a critical element of the power system. The correct functioning of the transmission grid is essential to preserve the system integrity. Concepts like reliability or safety in the system operation are intimately linked to the performance of the transmission grid. The failure of any component of the grid may affect the whole system if protection devices do not isolate it quickly from other connected components.

In the traditional environment, generation, transmission, distribution and retailing were vertically integrated in a single company. In this context, electricity transmission went unnoticed. Additionally, generation companies did not compete to get access to the grid since, either there was only one producer (the incumbent company), or power producers were dispatched based on their audited variable costs and their remuneration was set to cover their costs and allow for a reasonable rate of return on their investments. Besides, transmission typically has had a modest impact on the cost of electricity (between 5 and 10% of the bill in those systems where the imbalance between generation and demand in each area was not particularly important).

With the liberalization of the electricity sector, the transmission grid has become the place where generators meet to compete against each other to supply the load. What is more, the development of transmission grids allows relevant markets to increase in size, which has made possible competition in generation. Because of all this, the regulation of the transmission activity must become much more sophisticated than it traditionally was in order for electricity markets to achieve their goals.

The first issue to be addressed is determining how much the market agents must pay for using the transmission grid or, alternatively, how much they have to pay for benefiting from the existence of this grid. These charges may affect the competitive position of agents in the market. Another issue is designing rules that govern access to the transmission grid whenever the set of transactions requested by agents are incompatible with transmission constraints. Lastly, a mechanism must be put in place so as to guarantee that the development of the transmission grid is consistent with the system needs. According to this, the regulation of the transmission activity in the new context consists of three main aspects: investments in new transmission facilities, access to the transmission grid and the design of transmission tariffs. All these aspects cannot be considered in isolation, since they mutually influence one another.

The main objectives of the regulation of the transmission service are:

- To promote the economic efficiency of system operation, both in the short term (the dispatch of generation and load should be as close as possible to one where the aggregated surplus of producers and consumers is maximized) and in the long term (locational signals aimed at siting and retirement decisions of generators and loads should be sound).
- To provide an adequate remuneration for the transmission activity that is consistent with the incurred costs and risks.
- To promote the efficient operation and maintenance of the transmission grid.
- To promote the optimal development of the transmission grid.

The following paragraphs briefly describe each one of the three main aspects of transmission regulation. Here, we will only introduce the main problems that the

regulator encounters when developing the set of rules and mechanisms to be applied in order to deal with each of the aforementioned aspects. Other chapters of this thesis carry out a much more detailed description of these same problems and the possible solutions to them. An annex briefly explains the concept and properties of nodal prices, which must be regarded as a reference for the pricing of transmission.

1.2.1 Investment in new transmission facilities

Regulation must guarantee that optimal grid investments are undertaken in due time and that they are operated and maintained at the minimum possible cost. Optimal investments are generally characterized as those that maximize social welfare, which may be computed as the net aggregated benefit of all the consumers and generators in the system. This optimal investment criterion is consistent with the one that was employed in the traditional regulatory context, where investments in transmission facilities were planned in order to minimize the total cost incurred when serving the load.

As mentioned above, costs corresponding to regulated grid investments that must be recovered through tariffs include those incurred when constructing, operating and maintaining these facilities as well as the administrative ones. On the other hand, the cost of transmission losses, congestion in the grid and operation of the system are not grid costs and they must be dealt with separately.

As it will be explained later in chapter 2, the remuneration of the transmission activity depends on the regulatory paradigm that is chosen for the expansion of the grid. One first option involves organizing the development of the grid as a regulated, centrally planned, activity. In this case, some specialized institution may be in charge of proposing the construction of new lines, which must be approved by the competent authority. The remuneration of the transmission service should in this case refer to the grid that actually exists. Another option would be leaving the expansion of the grid in the hands of a private for-profit company whose decisions would not be formally subject to a regulatory approval. In this case, the remuneration of this company should refer to an efficient or 'perfectly adapted' grid. In a third option, the users of the transmission grid (generators and consumers) themselves are in charge of promoting the construction of new lines. If these lines get the regulatory approval, their cost may be financed either by the system as a whole through the regulated transmission tariffs or by the (coalitions of) network users that proposed their construction. Finally, the construction of new lines could be regulated as a competitive business whereby private entrepreneurs promote the construction of new lines and pay for them. Afterwards, these promoters would sell the capacity of the lines they own either through long-term contracts or in the spot market. All these alternative schemes will be discussed in detail in chapter 2 of this document, together with many other issues.

1.2.2 Access to the transmission grid

Any sound regulatory scheme should grant non discriminatory access to all the users of the grid. Rules applied to authorize the connection of market agents to the grid, as well as those used to allocate transmission capacity when it is scarce, must be efficient, transparent and objective. In principle, every agent should have access to the transmission grid under the new regulatory environment. Obviously, when flows resulting from the commercial transactions are incompatible with the available transmission capacity, regulation must determine which commercial transactions can take place and which ones not.

As far as the authorization of the connection of new generation agents to the grid is concerned, we must make a distinction between those schemes that give priority to already existing market agents over new entrants, and those where every agent is authorized to connect to the grid and, afterwards, market mechanisms are employed to solve local congestions. On the other hand, consumers that are not allowed to connect at a certain node must be given the possibility to connect at other nodes.

As for the allocation of the scarce transmission capacity among agents, several options exist. One may think of implementing a nodal pricing scheme. Nodal prices are spatially and temporally differentiated energy prices that promote the efficient operation of the system and implicitly take care of congestions and losses in a complete form. When congestion in the system grid is not significant, efficiency gains resulting from the use of nodal prices may not be worth the effort required to implement such a technically complex system. Computing different energy prices for each one of a set of zones whenever congestion occurs among them is an alternative. Zonal pricing schemes may be of any of two types: they can deal with energy and network capacity at the same time (implicit auctions) or they can manage them separately (explicit auctions). Finally, when congestion in the grid is not systematic, one may think of clearing the energy market neglecting the existence of congestion and dealing with it afterwards using ad-hoc mechanisms (redispatch). In the last case, the extra-cost incurred with respect to the outcome of the unconstrained dispatch may be socialized or, alternatively, assigned to those agents that are responsible for the congestion.

Congestion management mechanisms may be complemented using long-term capacity contracts, which allow market agents to hedge against the risk of volatility in the price that is paid to access the grid. Different types of capacity contracts exist: contracts may be physical or financial; they can either be defined between two nodes or be related to a particular flow-gate, etc. In any case, regulation must make sure that the use of congestion contracts should not aggravate the situations of market power abuse that may already exist.

1.2.3 Computation of transmission tariffs

Transmission tariffs must allow the owners of regulated transmission lines to recover the cost of constructing, operating and maintaining them. Additionally, market agents must receive both long and short term efficient economic signals related to their location in the grid. Short term signals are aimed at promoting the correct functioning of the market by taking into account both transmission losses and congestion. Long term signals should make market agents internalize in their investment decisions the cost incurred by the system because of their installation being connected to one part of the grid or another. Finally, tariffs must be non-discriminatory, easy to understand and fair.

Theoretically, transmission tariffs must be computed according to the responsibility of each agent in the development of the grid. One first attempt to design transmission related tariffs involves applying nodal energy prices. Nodal energy prices are different for each node of the grid and internalize the effect of grid congestion and losses. However, computing locationally differentiated energy prices may not be acceptable under certain regulatory regimes. Hence, in most cases, a single energy price is obtained for the whole system. Additionally, and most important, application of nodal energy prices results in a net revenue for the system which typically falls far short of the regulated revenues of the transmission activity. Therefore, one must decide how to allocate the remaining part of the cost of the grid, or its whole cost if a single energy price is used.

The transmission tariff setting process is mainly concerned with the allocation of the fraction of the grid cost that is not recovered through the application of other charges. Thus, it focuses on the computation of the so-called “complementary charge” to be paid by agents. Many different methods have been used and/or proposed to compute complementary charges. Conceptually speaking, one should allocate the grid cost in proportion to the benefits that agents obtain from the existence of each line of the grid. However, often the determination of the economic benefits that are associated with a certain line is plagued with practical difficulties. Trying to estimate the benefits that market agents will obtain from a new line may only be worth it when dealing with new and expensive transmission lines. Another method that is conceptually correct, but hard to implement, is the allocation of the cost of a line based on the responsibility of each agent in causing the investment. Consequently, alternative simpler approaches have been devised. One possibility is computing transmission tariffs based on the electrical usage that agents make of the grid. Another one is determining long-term transmission marginal prices.

In those systems where the transmission grid is well meshed, sending locational signals to market agents may be ignored. In these cases, transmission tariffs should just aim to recover the cost of the grid without interfering with any efficient short term operational signals already sent to these agents. In addition, if one wants to minimize the economic distortion caused by the application of transmission tariffs, market agents

whose operational decisions exhibit a higher level of elasticity to the amount of transmission charges they are allocated (generators and large consumers) should pay lower tariffs, whereas those whose operational decisions are less conditioned by the level of transmission charges (small consumers in developed countries) should pay higher tariffs.

1.3 The regulation of the transmission activity in the regional electricity markets

Throughout the last decade there has been a general trend towards the creation of supra-national or regional markets, i.e. markets comprising several countries or states. Examples of these markets are the Internal Electricity Market of the European Union (IEM); the market that is developing in South-East Europe, which will sooner or later join the IEM; the NORDEL market in Scandinavia; MERCOSUR and the Andean market in South America, which will eventually join as well; the Standard Market Design (SMD) in the US, in the context of creation of regional markets across this country; the Australian national market, the market in the Mekong Delta region, etc.

Regional markets are created in order for control areas or countries to take advantage of the benefits that are derived from being part of a larger market. A list of the main reasons to create a regional market follows:

- Reduction of the cost of operation of the system due to the fact that modern, more efficient generators located in some areas of the region are dispatched instead of older, less efficient ones that are located in other areas.
- Additionally, due to the increase in the size of the relevant market, larger, more efficient generators may be constructed which otherwise would not be profitable. This is typical of regions where the control areas are too small to build large generators aimed only at serving the local load (Central America, for example).
- Competition among power producers increases, which drives prices down.
- The cost of preserving the security of supply, both in the long and the short term, is reduced because of the diversification of available primary energy sources and the utilization of fewer and more efficient generation units to provide the system reserve.
- Larger power systems will be more robust than smaller ones if coordination among the different control areas takes place. This would result in an improvement in general reliability conditions.

However, the interconnection of several power systems presents significant technical challenges. For instance, power systems must cooperate and coordinate to keep the system frequency stable in the region. In addition, control areas must be able to comply with the power exchanges that have been scheduled among them. Finally, they must help others when they run into stability problems that put at risk the integrity of the system, or they face a shortage of available generation at local level. In the past, these

problems were easily kept under control because interaction among control areas, which were served by vertically integrated utilities, was limited to emergency situations and a reduced number of commercial transactions. These technical problems have increased in importance in the recently developed regional context, where transactions taking place between parties from different areas have increased in number and size and they are normally agreed bilaterally, or multilaterally, among market agents that have access to the whole regional grid.

Additionally, the integration of different local markets poses new challenges that did not exist previously. Control areas must be compensated for the use that external agents make of their grids, or the benefits that these agents obtain from them. Efficient, market based mechanisms must be put in place to allocate the scarce transmission capacity used by regional transactions. A regulatory paradigm for the expansion of the grid must be implemented which results in an optimal development of the regional grid. Finally, a leveled playing field must be created for market agents to compete against each other.

In order to provide an appropriate response to these challenges, countries must promote the integration of their energy markets. In addition, existing markets must be rearranged to match the particular characteristics of each region. The economic criteria applied must focus on general, not individual, interests. Lastly, some regulatory aspects must be harmonized at regional level including the set of rules applied for the expansion of the grid and the operation of the system, the market rules in general and some basic aspects of the tariff setting process like the split of transmission charges between generation and load.

The general recommended approach to be followed in order to achieve a truly integrated regional market involves regulating the functioning of this market as if it were a single national one. The resulting regional regulatory framework will probably have to be adapted to suit the characteristics of the region so that it is politically acceptable.

In many cases, the process of formation of a regional market consists of the following three stages:

1. Development of the required interconnection capacity among the national systems in the region. Local producers in a system would offer the power they produce at the border with neighboring systems, so that this power can be used by consumers in the latter systems.
2. Some degree of coordination among national systems is achieved, affecting the operation of the global regional system and facilitating the creation of wholesale international markets.
3. National markets totally integrate into a single regional one where every agent can freely arrange commercial transactions with any other agent in the region.

Some systems may decide to stay permanently within any of these stages.

As mentioned earlier, several regional markets presently exist. The driver or general philosophy adopted in different regional markets may be different. Thus, some regions have opted for a system whereby a regional market superimposes on the local markets that already existed instead of replacing them. In others, agents from any part of the region gain access to the entire regional grid through the payment of a local G (for generators) or L (for consumers) transmission charge, which may be different from one area (country) to another, and these agents can freely negotiate transactions. In other cases, a common market design is adopted for the whole region (PJM, MISO or ERCOT in the USA).

The following paragraphs describe three representative samples of these markets: The IEM in Europe, the Standard Market Design in the US and the MER (for Mercado de Electricidad Regional in Spanish) in Central America.

1.3.1 The Internal Electricity Market of the European Union

The IEM presently consists of 27 countries, 25 belonging to the EU plus Norway and Switzerland. The electricity directive (2003/54/CE) establishes the regulatory framework for this regional market, see (EP 2003). The regulation related to cross-border transactions aims to reduce or eliminate the existing barriers to power trade among agents from different countries, see (EP 2003).

Institutions in charge of developing the regulation relevant to the IEM are the European Parliament and the European Commission. Directives are issued that must be transposed into national laws by the national governments. Other organizations of European scope include the European association of Transmission System Operators (ETSO), the Council of European Energy Regulators (CEER), the European Regulators Group (EREG), which is an advisory group of independent national regulatory authorities to assist the European Commission in consolidating the Internal Market for Electricity and Gas, or the European association of Power Exchanges (EuroPEX). All the stakeholders involved in the process of creation of the Internal Electricity Market meet periodically in a series of fora where they discuss the next steps to take in order to best advance towards the integration of national markets. These are the Florence Regulatory fora. Section 1.3.1.1 provides a description of the Florence Regulatory Forum and the main agreements reached in it.

The guidelines issued by the European Commission lay down that the transmission grid within each country must be operated by legally separate entities where vertically integrated companies exist. In addition, agents must have regulated third party access to the transmission grid in any member state. Electricity companies must keep separate accounts for all their distribution and transmission activities in order to avoid discrimination, cross-subsidisation or distortion of competition. Guidelines also draw up an agenda for the introduction of competition in the retail market so that in the end

all consumers are free to choose their supplier. The “prime driver” behind the IEM is the ability of agents in the region to freely negotiate power transactions. There is free access to the transmission grid contingent on the payment of the local access charge.

Strictly speaking, the IEM is not a truly integrated regional market but a collection of national markets that must coordinate their operation and planning processes. The European regulation includes a set of guidelines that members must comply with. However, the functioning of the market within each country is left in the hand of the competent national authorities or institutions. Thus, decisions relevant to the implementation of the guidelines produced by the European Commission and Parliament institutions are made at the national rather than at the European level. This organizing principle is commonly known as subsidiarity. In line with this principle, there is not a unique system operator in the region, since the operation of the system within each country is in the hands of the corresponding SO. Similarly, several wholesale markets exist already involving a few neighboring countries (e.g. NORDEL) and the coordination between them has been very limited to date.

Main obstacles to the development of the IEM include the significant differences that exist among the rates of liberalization of the national markets in the region, which results in a distortion of competition among agents from different countries; the lack of harmonization of access tariffs; the lack of transparency of the tariff setting process; the high level of vertical and horizontal integration of the industry within some sub-regions (countries or groups of countries) and the lack of adequate interconnection capacity between most countries.

Now, we briefly describe the situation in the region regarding some major regulatory aspects, namely investment in new transmission infrastructures, access to the grid, transmission tariffs and security of supply.

Investment in new transmission facilities

Interconnections among national transmission systems in the IEM are weak in general. This results in systematic congestion that breaks the system into several sub-regional load pockets. Differences among wholesale prices in these load pockets are significant. There is not a single administrative process to be followed in order to decide on the construction of new cross-border (regional) lines and cooperation among national entities to build lines that are of interest for cross-border trade is insufficient. In fact, the construction of new lines usually is plagued by difficulties, due to the disparity of regulatory arrangements in different countries and to environmental concerns. The grid used by cross-border transactions (the Horizontal Network in the jargon used in the region) must be significantly reinforced in some critical areas so as to realize the potential benefits of a European market.

Access to the grid

Connection to the grid must be open to any agent that requests it. If connection at a certain node is not possible, the corresponding agent must be provided with some alternative. Regarding the management of congestion in the grid, guidelines lay down that efficient coordinated market based methods must be used to allocate the scarce transmission capacity among market agents. However, cross-border transmission capacity on the majority of borders is allocated through inefficient uncoordinated methods such as dividing this capacity pro-rata of the volume of capacity requested by each agent or assigning it on a first come, first served basis. Other, more elaborated congestion management mechanisms presently used are explicit auctions, where capacity is allocated to those agents who place a highest value on it, or market splitting, which is an implicit zonal pricing scheme where energy is dispatched jointly with the interconnection capacity between several pre-defined areas (presently applied in the NORDEL market). Recently, the European Association of Transmission System Operators and the European Association of Power Exchanges have jointly proposed a hybrid implicit-explicit auction to manage congestion in the region in a coordinated way. This mechanism is called Flow Based Market Coupling. Chapter 4 presents a comprehensive description of congestion management mechanisms.

Transmission tariffs

Instead of computing pan-European transmission tariffs, countries and European institutions have agreed to implement a system of compensations among countries whereby each country is compensated for the use that others make of their grids. Authorities in each country are responsible for computing local transmission tariffs for generators and consumers. These tariffs must be modified according to the net compensation to be paid by the corresponding country. National authorities or institutions are responsible for allocating the global net compensation to be received by each country (difference between the compensation received from other countries and the payments due to others) to their generators and consumers. Countries have agreed that some level of harmonization among national transmission tariffs is necessary in order not to distort competition in the regional market, but final rules have not been approved yet.

Security of supply

Energy policies aim to achieve a balance between security of supply, protection of the environment and economic efficiency. Countries wish to be as energetically independent from others as possible. This is due to the fact that, in case of emergency, countries give priority to the supply of the local load over the contractual commitments they have with external agents. A new EU directive has been drafted that tries to get closer to an EU-wide scheme for security of supply.

1.3.1.1 The Florence Regulatory Forum

The Florence Regulatory Forum, —which has been organized by the European Commission—, has succeeded in bringing together the different stakeholders of the IEM to discuss and to build some consensus on the issue that is probably most difficult to solve when trying to achieve a working regional market: regulation of access, investment and pricing of the transmission network of the 27 countries that participate in the IEM.

Slowly, but gradually, the Florence process has built a preliminary vision or model of how transmission issues should be approached within the IEM. This model comprises four major elements. Each element provides an answer to a basic regulatory question within the context of the IEM: cross-border tariffication, network tariff harmonization, congestion management and transmission network investment.

The pragmatic Florence approach appears to have now arrived at a juncture where key orientations have been adopted with regard to: a) the trend towards adopting the concept of the “single system paradigm” so that political borders will not finally matter in the design of the rules of functioning of the IEM; b) the need for a certain level of regulatory harmonization in the 27 IEM countries; c) the compliance with the laws of physics and the maintenance of sound criteria for security of supply.

1.3.2 The Regional Electricity Market of Central America

The regional electricity market of Central America is comprised of six Central American countries: El Salvador, Guatemala, Honduras, Nicaragua, Costa Rica and Panama. This regional market is superimposed on the six national markets that exist in the region (one per country). Countries in the region agreed not to replace their national markets with a regional one but to provide market agents with an additional, larger, market where they could freely engage in commercial transactions both with the agents of their own country and with those from others. The development of the regional market should not hinder the process of reform that many national markets are undergoing simultaneously. Some minimal regulatory harmonization within the countries is necessary because of the very different level of development that local markets across the region have achieved. Thus, in some countries, there is a national market properly speaking, where suppliers compete to serve the local load, while others mostly maintain the characteristics of a vertically integrated structure and traditional cost of service regulation.

The process of creation of the regional market in Central America (MER for *Mercado de Electricidad Regional*) has taken about 18 years by now. The first milestone in this process was the signature in 1996 of the Treaty providing the basic framework for the development of the MER. In 1998, this treaty was ratified by the national parliaments

of the six countries. The BID (for *Banco Interamericano de Desarrollo*, in Spanish)¹ and other financial institutions have committed funds for building new transmission infrastructure and achieving a higher level of technical coordination among the national systems. The general terms of the design of the regional market were developed and approved in 2000, see (PHB-HAGLER/SYNEX 2000). While the detailed regulation was being developed, a transitory version of the regional market began to operate in 2002.

Since 2002, consultants have been working on the detailed design of the regional market and its relationship with national markets in the countries of the region, the development of the regulatory framework for the transmission of electricity and the institutional arrangements related to the different regional entities, see (KEMA 2002; KEMA 2002; SNC/Statkraft 2004). This detailed regulation has been approved at the end of 2005.

There are several regional institutions related to the functioning of the regional market in one way or another. A regulatory body of regional scope, the CRIE (for *Comisión Regional de la Interconexión Eléctrica* in Spanish), is in charge of developing and supervising the implementation of the regulation that applies to the electricity commerce taking place at a regional level. This entity must design and apply any preventive or corrective rule that is necessary to guarantee that the regional market works properly. The EOR (for *Ente Operador Regional* in Spanish) is concerned with the technical coordination of the regional interconnected system as well as with the commercial administration of the regional market (MER). The regional grid is owned by a number of companies including national transmission companies, private investors and grid users and the EPR (for *Entidad Propietaria de la Red* in Spanish), which is the owner of the future SIEPAC line crossing the region from north to south. Therefore, the regional system operator is not the owner of the grid it operates. As we shall briefly discuss in Chapter 2 of this document, this organizational model has both advantages and drawbacks with respect to the TSO model (also called Transco model in the US, where both models coexist), where the system operator owns the transmission grid.

General market design

Due to the fact that the regional market does not replace but complements national markets, the interface between the two kinds of markets is central to the successful implementation of the MER. The next paragraphs describe how the coupling between both markets takes place.

¹ The Interamerican Development Bank (IDB) is a financial institution aimed at promoting the development of countries throughout America.

The final dispatch in the region is the result of two different market clearing processes taking place one after the other. First, national market and system operators compute the local dispatch within each country. Second, the EOR carries out the regional market clearing process. This entity computes the optimal system dispatch from the bids submitted by the national system operators. These bids must be assigned to a specific node of the regional grid. A different price is computed for the energy negotiated at each node of the regional transmission grid (RTR for *Red de Transporte Regional* in Spanish). Every transaction taking place within the regional market is priced using the corresponding nodal prices.

There are two main ways for an agent to trade its energy in the regional market. It may sign a bilateral contract with any other agent in the region, i.e., it may participate in the regional contract market (*Mercado de Contratos Regional* in Spanish). In order for two market agents in the region to sign a bilateral contract they must gain access to the capacity of the regional transmission grid they need to use. The unit price of this capacity is equal to the difference between the nodal energy prices, obtained in the regional dispatch, for the withdrawal and injection points of the corresponding bilateral transaction.

Alternatively, any authorized agent may participate in the regional spot market (*Mercado de Oportunidad Regional* in Spanish). In order to participate in the regional spot market, an agent would have to submit a bid to the corresponding national market. Afterwards, the SO in the country where the agent is located would market its energy in the regional spot market. As a result of the regional dispatch carried out by the EOR, modifications (increments and /or reductions) are computed with respect to the power injections and withdrawals scheduled in each one of the national markets. These incremental power injections and withdrawals are priced at their corresponding nodal energy prices and result in a net aggregated payment to be made by each national SO. Finally, the national SO must allocate this payment to those generators and loads in the country that have marketed their energy in the short-term. National SOs or authorities can exercise their own discretion in charging individual agents within their country. However, it would be desirable for these agents not to be discriminated based on the market (national or regional) where their energy has been marketed. In other words, a single energy price should be computed for each node in the country and any short-term power injection and withdrawal taking place at this node should be valued using this price. But this is not the reality at the present time. The main hurdle that the MER is currently facing for its implementation to be successful is a correct design of the interfaces between the six national markets and the regional market.

Finally, we shall briefly discuss how the major aspects of transmission regulation are dealt with in the MER. The treatment given to security of supply will be discussed as well.

Access to the regional grid

All the market agents that pay the corresponding transmission charge have access to the whole regional grid on equal terms. The MFD method (for *Método del Flujo Dominante* in Spanish) is used to compute regional transmission charges. Because the transmission grid in the region is very poorly developed, congestion is frequent and very significant, specially in the interconnections between countries. Therefore, an efficient, market based, mechanism must be used to determine which agents have access to the scarce transmission capacity.

In the short term, transmission capacity and energy are allocated in an integrated manner in the regional market, since the EOR computes the optimal economic dispatch for the whole region. Both incremental energy bids, and bids for capacity associated with short-term bilateral contracts, must participate in the regional spot market. This market results in the regional short-term dispatch. Since every power injection and withdrawal is valued using the corresponding nodal energy price, efficient short-term signals are sent.

In addition to short-term bilateral contracts, long-term firm contracts exist. Firm contracts must secure access to the transmission grid in the long term. These contracts must reserve transmission capacity on the likely-to get-congested lines or corridors that they need to use in order to access the area where the consumer is located. Thus, long-term capacity auctions are held so that the agents can buy long term transmission capacity rights. Rights bought by the generators signing firm contracts must be physical since they must have access to the physical capacity of the grid in case they are requested to actually deliver the energy contracted. This is due to the fact that generators signing a firm contract with a consumer in a country A have the right to earn the capacity payment of this country.

Agents that do not want to sign a firm contract may buy financial instead of physical rights. Auctions of long-term physical and financial contracts may, or may not, take place separately. We must bear in mind that physical and financial transmission rights are different products with a different market value. Firm contracts have priority over bids into the spot market and other types of contracts when accessing the transmission grid. Therefore, whenever some contingency occurs and there is not enough transmission capacity for all the scheduled transactions to take place, firm contracts must be the last transactions to be curtailed.

Transmission tariffs

As mentioned above, any agent automatically gains access to the regional transmission grid by paying the corresponding regional transmission charge. Regional transmission charges, together with the revenues from the application of nodal prices, must recover the regulated revenue of the activity of transmission in the regional market. The regulated revenue in the regional market that is associated to each line depends on the type of the line. Thus, the regulated revenue associated to the facilities that belong to

the primary regional transmission system (the SIEPAC line) is equal to the debt contracted by the EPR in order to finance these investments (there is a limit to the amount that can be collected annually by the EPR). The regulated revenue associated with the facilities belonging to the RTR that have been centrally planned by the EOR is equal to the total cost of the line including a reasonable rate of return on the investment. Existing transmission lines and investments at risks belonging to the RTR earn a regulated revenue related to the reliability value of the line, which represents a minor part of the cost of the line. Finally, the owners of investments at risk that are beneficial for the regional system can negotiate with authorities earning a certain regulated R . This revenue R is equal to the fraction of the cost of the line represented by the ratio of the benefit that third agents obtain from the line to the total benefits produced by the line. In return for earning this revenue, the private promoter would have to transfer the property of part of the transmission rights over the capacity of the line. If private promoters do not negotiate the collection of a regulated revenue, their revenue is equal to the market value of the lines they own, i.e. these lines earn the corresponding congestion rents in the regional dispatch.

Apart from this, regional transmission charges should ideally provide long-term locational signals and not affect the short-term dispatch. A usage-based method is applied to compute regional transmission charges. This method is a combination of two others: the Dominant Flow Method MFD (*Método del Flujo Dominante* in Spanish) and the Average Participations method. Transmission charges are not directly applied to consumers or generators. Instead, transmission charges computed are aggregated at country level to determine the total transmission charge to be paid by each country. The MFD and AP methods are explained in chapter 3 of this thesis.

Lastly, in addition to regional transmission tariffs, national tariffs may also exist. Local authorities are in charge of computing national tariffs.

Investments in the transmission grid

The transmission grid in the region is very poorly developed. This results in frequent and significant congestion. The EOR is in charge of computing an indicative expansion plan for the transmission and generation in the region. Once an investment in the regional grid is proposed by the EOR, and its construction is approved by the CRIE, national authorities cannot oppose the project.

At the same time, market agents and private promoters may propose the construction of transmission lines. These investments would also have to be approved by the CRIE. A generator willing to sign a firm contract with another party, and earn the corresponding capacity charge, may need to build new transmission capacity in order to obtain the physical transmission rights that are needed for this contract to be considered valid. Hence, some investments in new transmission capacity may be associated with firm contracts. Promoters of merchant lines build them with the aim of

appropriating the corresponding congestion rents, which result from the application of nodal prices in the regional dispatch.

In principle, investments at risk are given priority over regulated investments. That is to say, any line must be constructed as an investment at risk as long as a coalition of users forms to build it. Only those lines whose benefits are too dispersed for coalitions of users to want to build them should be constructed as regulated projects.

Security of supply

Both local authorities in each country and the regional ones must work to guarantee that enough generation and transmission capacity is available to serve the local load, both in the short and the long term. The regional system operator, and some national operators as well, produce periodical indicative expansion plans for generation and transmission. In many cases (including that of the regional grid), transmission lines deemed to be necessary are constructed as regulated lines. In addition, countries may make use of capacity payments or other regulatory instruments in order to encourage market agents to invest in new generation capacity.

1.3.3 The Regional Electricity Markets in the US

Contrary to what happens in the European and Central American cases, we cannot talk about a single regional market developing in the US. Several markets have been created in different parts of the country. Some of the most relevant regional markets that presently exist are the PJM market in the north east of the country (as well as NEPOOL in New England and NYPP in New York), the system managed by the Midwest Independent System Operator (MISO), the Californian market, the Southwest Power Pool or the ERCOT system in Texas. Regional markets in the US are not of the same size as the Internal Electricity Market of the EU. Thus, comparing both types of markets as equals would be unfair. Most regional markets in the US are approximately of the same size as a large European national market. However, one should not think of complexity in terms of market size. The complexity of designing and implementing a “tight” regional market such as the MER, which encompasses six sovereign countries, should not be underestimated, even if the regional peak load may be of the order of only 10,000 MW.

The objective to be fulfilled by regional markets in the US is the same as that of the markets in Europe and Central America. Regional markets in the US are aimed at increasing system efficiency through the introduction of competition. Hopefully, this will contribute to achieving a reduction in the prices paid by final consumers as well as an increase in the quality of supply.

The organizational arrangements made in the US are quite different from those in place in other markets. Thus, contrary to what happens in the EU and Central America, a regulatory framework of regional scope does not exist. However, an entity exists that

has authority over the development of energy trade at a regional level (between states). This is the Federal Energy Regulatory Commission (FERC). Rules regulating the functioning of any regional market in the country must be approved by FERC. However, this entity does not have the authority to impose a common market design for all these markets. The National Energy Reliability Council is another regional institution which develops technical standards that power systems across the nation must comply with. States have authority over local markets, including retailing.

According to this general scheme, the development of regional markets should only be overseen by the FERC. However, many of the companies supplying the final consumers are vertically integrated and they are deeply rooted in their service area. This provides state authorities with the opportunity to participate in the creation of regional markets. Therefore, conflicts over the design of regional markets have arisen between the FERC and state authorities.

Authorities and the industry in the US have been discussing the restructuring of electricity markets over more a decade now. Federal authorities have tried to introduce competition at wholesale level and to change the structure of the power sector, promoting the creation of fully integrated regional markets throughout the country. According to a proposal by FERC, regional markets should comply with a standard defined previously, the so-called Standard Market Design, see (FERC 2002). Opposition to this initiative from some state authorities forced the FERC to water down the scope of the reform. Thus, in the end, this federal entity only issued a set of recommendations over the future evolution of regional markets, see (FERC 2003). We provide below a general description of the Standard Market Design initially proposed by FERC.

As briefly mentioned above, each regional market in the US covers several control areas that had traditionally been served by different vertically integrated utilities. Utilities and other companies operating in a regional market remain independent from one another. However, the market in the region may be managed by a single entity, the regional market operator. When this entity is also in charge of operating the regional system we talk about regional system operators. Regional Transmission Operators (RTOs) are regional system operators that own the regional transmission grid.

The characteristics of markets in the US vary widely from one case to another. Some markets have implemented design options that are very advanced, conceptually speaking, and close to the ones proposed by FERC as part of the Standard Market Design. This is the case of the PJM market. On the other hand, market arrangements adopted in areas where vertically integrated utilities still hold a strong position are far less ambitious. Given that the policy of the FERC may be adopted as a reference for the development of the relevant regulation affecting the regional trade of energy, we now summarize the views of the FERC on the three main aspects of transmission regulation, see (FERC 2002).

Computation of Transmission Tariffs in the SMD

The FERC is responsible for the design of regional transmission tariffs. This entity makes a clear distinction between long-term and short-term economic signals. Short-term grid signals include those related to transmission losses and congestion in the grid. These signals are related to the short term value of energy and should be taken into account when computing the economic dispatch in the region. Long-term grid signals include transmission tariffs and they are meant to provide locational signals to new generators and load. However, short-term signals may also affect the decision by agents on where to install. In addition to providing long-term locational signals, transmission tariffs must contribute to the recovery of the cost of the grid.

According to the SMD, market agents should pay a single transmission tariff, which would grant them access to the entire regional grid. The method applied to compute transmission tariffs should be the same for all the market agents in the region. The FERC proposes two alternative methods: the adoption of a postage-stamp (uniform) tariff in the region or the utilization of a “license plate rate” whereby transmission charges would vary depending on the area where the agent is located. Tariffs of the second type would send rough locational signals between areas whereas tariffs of the first type would not. In any case, transmission tariffs should be exit charges, i.e. only consumers would pay charges. The SMD also provides for a system of compensations among regions for the use made of their grids by external agents. According to this system, the compensation that a region *A* owes another one *B* would be proportional to the net amount of energy that the former imports from the latter.

Access to the transmission grid in the SMD

Access to the transmission grid should be open to every market agent provided it pays the transmission tariff. Transmission capacity and energy would be jointly allocated in a regional balancing market. Additionally, authorities should promote the creation of day-ahead and real time markets. Locationally differentiated marginal prices would be computed for each node of the grid. Therefore, congestion and losses would be correctly internalized in energy nodal prices. Locational Marginal energy Prices should be complemented by a system of financial transmission rights that would allow market agents to hedge against the risk associated with the volatility of transmission prices.

Investment in new infrastructures in the SMD

Traditionally, the expansion of the transmission grid within each state was planned in order for the local generation to serve the local load. Interstate commerce was very limited and cooperation between control areas was aimed at providing support in emergency situations. Therefore, neither the grid within each state nor the interconnections between state grids are prepared to support a great deal of interstate commerce resulting from the creation of regional markets. Therefore, large investments in new transmission lines are necessary.

The FERC has expressed the view that those mechanisms that encourage the institution or party in charge of the expansion of the grid to perform its function efficiently are superior to others. Thus, ideally, transmission investments should be remunerated according to the increase in the social benefit they yield. However, the FERC acknowledges that achieving an adequate development of the grid is essential for the successful implementation of regional markets. Consequently, according to this entity, an institution (probably the RTO) should identify the most beneficial grid reinforcements from the point of view of the system, see (Abraham 2002). Afterwards, the private initiative would be given the opportunity to meet the system needs previously identified. Thus, the construction of those lines that are deemed to be beneficial could be assigned through a competitive bidding process. Private promoters building these lines would recover their investment and they would earn an adequate rate of return on it. As for the merchant investments, the FERC highlights that they have the ability to introduce competition in the activity of grid planning, which has traditionally been regarded as a natural monopoly. However, these type of investments are unlikely to provide the reinforcements to the grid that are required.

Other regulatory aspects

According to the FERC, adequate mechanisms should be put in place to encourage an adequate level of investment in new generation capacity. In addition, provisions should exist to promote the reliability in the operation of the system as well as the quality of service.

1.3.4 The Market in South America

Regional markets in South America are still in an early stage of development. The authorities in the region have not adopted yet specific, clearly defined, regulatory models and designs to address most aspects of the regulation of transmission. We now only provide a brief description of the general market design. Similarly to what happens in other regional markets, the development of a multinational market in South America is aimed at moving forward in the process of integration of the national markets in the region.

The South American market is comprised of two sub-regions. The first one includes Argentina, Brazil, Uruguay and Paraguay. It is the MERCOSUR area properly speaking. The second one comprises Colombia, Peru, Ecuador, Bolivia and Venezuela. This is the so called Andean region. Chile is an associated country to the MERCOSUR area. The market in the MERCOSUR sub-region has achieved a higher level of maturity than that in the Andean sub-region. Therefore, the agenda for implementation of a regional market must make a distinction between these two areas. Basically, this process consists of two stages:

1. First Stage: promotion of cross-border power exchanges among countries. This stage is only necessary in the Andean market but not in the MERCOSUR area,

which has already gone through it. Two types of power exchanges may occur: those corresponding to long term contracts and those that are scheduled in the short run within the daily or intra-daily markets.

2. Second stage: it includes the launching of the regional market and its subsequent monitoring. Several sub-stages may be defined here:

- All the parties involved must reach an agreement at regional level on the market design.
- Transitory phase of implementation of the market. The regulation is developed.
- Consolidation phase where national regulations must be adapted so that they conform to the regional one.

As for the institutional arrangements, no regional entities exist at the moment. Authors in (WB 2000), a report commissioned by the World Bank, envisage the creation of three regional entities: a regional system operator, a regional market operator and a regional regulatory commission. The arrangements proposed in this document resemble most those of the regional market in Central America (MER). However, the situation in the region and the level of development of the market is closer to that in Europe.

According to the proposal by the World Bank, the regional market would not replace national markets in South America. National markets and the regional one would coexist. Long term firm contracts in the envisaged market would take precedence over short term commercial transactions taking place in the spot market. Apart from that, market agents might be able to provide ancillary services outside their country and quality transfers between countries could also occur.

1.4 Objectives of the thesis

The academic community and the industry have already produced sound regulatory designs and models to address the majority of the problems related to the transmission of electricity. Schweppe proposed the use of nodal energy prices to address the dispatch of energy in a system, see (Schweppe, Caramanis et al. 1988), while Hogan proposed the use of financial point-to-point transmission rights for agents to be able to hedge from the risk associated with the price of transmission, see (Hogan 1992). As for the computation of transmission tariffs, Pérez-Arriaga and Rubio proved in (Pérez-Arriaga and Rubio 1995) that nodal energy prices are not capable of recovering the whole cost of the grid under realistic conditions. Latter, they compared the method of cost allocation to beneficiaries to the two most representative usage-based methods at the moment and explained that, while the beneficiaries method is the most efficient to allocate the cost of the grid conceptually speaking, usage-based method like AP or MP should be considered as a suitable alternative when assigning the grid costs to the beneficiaries of the grid is not possible, see (Rubio and Pérez-Arriaga 2000). Lastly, Joskow and others have extensively analyzed the different possible models for the expansion of the grid. Joskow concludes that, despite having some limitations, centralized planning must be in charge of promoting those reinforcements that are necessary for transmission systems of different control areas to integrate as well as to maintain the system reliability, see (Joskow 2005). Merchant investors and network users should be left to promote those lines that the central planner fails to build in due time. In spite of all these achievements, the regulation of transmission has traditionally been regarded in many systems as a collection of different parts rather than as a whole, see (Jong 2004). This thesis, as opposed to previous work, aims to develop a reference regulatory framework that deals with all the aspects of the transmission service in a regional context, while taking into account the existing interactions among the solutions proposed to address them. This framework should include the best practices, designs and models available at the moment.

The reference regulation, which may be suitable to address the case of a perfectly integrated market, may not be compatible with the institutional framework and market philosophy adopted in many regions. However, the solutions that have been implemented to date in most regional markets are, by no means, the best possible ones. A critical factor for this situation is the absence of a common energy policy for the whole region, see (Jong 2004). Besides, the different degree of maturity of local markets in the region, and the unwillingness of some parties to let go part of the responsibilities that have traditionally enjoyed, have prevented the adoption of solutions that are conceptually sounder than those actually implemented. As a consequence of all this, we shall try to adapt the reference regulatory framework that has been developed to the different circumstances that exist in each region. In other words, we shall develop a specific set of rules to organize the transmission of electricity in each of the regional

markets that we have considered, which include the IEM of the European Union, the Regional Electricity Market in Central America (MER) and the Regional Transmission Organizations being developed in the US. Other regional markets exist, like the MERCOSUR market in South America or the Australian market. However, we have thought that the three chosen markets are representative enough of the different conditions of any type (political, institutional, economic, etc) that may exist in a region.

In addition to analyzing the different regulatory practices that have been proposed, or actually implemented, to date, we shall focus on designing solutions to several specific problems that have not been addressed satisfactorily yet. Here, we shall propose regulatory models and mechanisms to address these issues. These are the main contributions of this thesis work, and include:

- Selection of the most suitable method to compute transmission tariffs and its application in the context of the computation of compensations among countries in the IEM or elsewhere. Several novel methods have been proposed throughout the course of this thesis.
- Development of a system of long-term locational signals, which are embedded in the transmission access tariffs, and encourage market agents to make efficient investment decisions.
- Design of a method to identify Single Energy Price Areas (SPAs). This method should be applied to define the optimal set of SPAs within each regional market, as well as the accuracy of the division into areas that is obtained (which may be used to determine whether defining SPAs is feasible). Identifying accurate enough SPAs is central to the implementation of any zonal congestion management scheme. Only zonal schemes leave the local dispatch within each area in the hand of the corresponding local authorities.
- Analysis of a method to choose a balancing point for the definition of transmission rights. Application of this method to select a balancing point in the IEM, see (Olmos and Neuhoff 2006).

Despite having focused on some specific aspects, we shall discuss, to a larger or smaller extent, all the aspects that are related to the transmission activity. Some aspects shall be analyzed in much more depth than others. Thus, a significant part of the work of this thesis shall focus on the analysis of the recovery of the sunk cost of the grid through the computation of transmission tariffs and compensations among countries. On the other hand, the analysis that shall be carried out on the topic of network investment patterns is not so ambitious. However, we shall aim to provide a comprehensive picture of the regulation of regional transmission networks so that the treatment that is given to different aspects of the problem is consistent.

We must highlight again the fact that, given the regulatory nature of the work presented here, the solutions that will be proposed for some problems may not be the

most efficient ones, either from the point of view of economic theory or from the point of view of the theory over the functioning of power systems. This is due to the fact that the regulatory designs and models proposed here must not only be efficient but also politically acceptable. Thus, after making reference to the, conceptually speaking, most efficient regulatory design, we have always focused on providing some recommendations or designs that are implementable in the context of the existing regional markets.

Lastly, because each one of three main aspects of the transmission regulation (regulation of the expansion of the grid, computation of transmission tariffs and congestion management) represents a different problem to be solved, and despite the fact that all of these problems are interrelated, we have decided to separately review the state of the art corresponding to each one of them. Thus, each one of the three main chapters of this document is devoted to discussing one of these three main regulatory problems, and each of them includes a section on the relevant state of the art.

1.5 Layout of this document

The presentation of the thesis is divided into 5 chapters.

After this introductory Chapter 1, Chapter 2 discusses the definition of a regulatory framework for the expansion of the transmission grid. Additionally, it analyses the compatibility of the investment signals that are used to drive the development of the generation, transmission and distribution systems. First, we present and analyze the different paradigms that exist for the regulation of the expansion of the grid. Based on this analysis, we propose a model for organizing the expansion of the grid that combines the best features of the ones previously presented. Afterwards, we discuss the implementation of this model in the three regional markets that are considered. The next section describes the basic features of the regulatory test to be applied in order to decide about the construction of a proposed reinforcement. The chapter ends with the study of the relationships that exist between the investment signals in generation, transmission and distribution.

Chapter 3 focuses on the computation of transmission tariffs. First we provide the economic rationale leading to the computation of efficient transmission tariffs. Then, we examine the different methods that exist to calculate transmission charges. Based on the analysis carried out in the previous two sections, we propose a method to compute transmission tariffs and theoretically justify it. Finally, we address the problem of allocating the cost of transmission in the context of the IEM, the MER, the SMD and other regional markets.

Chapter 4 analyses the regulation of the access of market agents to the transmission grid. Two main problems are discussed here: the connection of market agents to the grid and the management of congestion. Much of the chapter is devoted to the latter issue. Regarding the management of congestion, first we describe and compare the methods used to allocate the scarce transmission capacity both in the short and in the long term. Based on this analysis, we make some specific recommendations targeted at the three studied regional markets. Finally, we discuss some particular issues related to the allocation of capacity including the treatment to be given to special restrictions, the computation of the value of a transmission line in the dispatch and, quite importantly, the definition of Single Price Areas for congestion management purposes.

Finally, chapter 5 draws the most relevant conclusions from the analysis carried out in the previous chapters, highlights the main contributions of this thesis work and propose further developments, based on the research work carried out, which might be deemed interesting.

Bibliography

- Abraham, S. (2002). National Transmission Grid Study. Washington DC, US Department of Energy. Web page: <http://www.energy.gov/engine/content.do>: 108.
- EP (2003). Directive 2003/54/EC of the European Parliament and of the Council concerning common rules for the internal market in electricity and repealing Directive 96/92/EC.
- EP (2003). Regulation (EC) No 1228/2003 of the European Parliament and of the council of 26 June 2003 on conditions for access to the network for cross-border exchanges in electricity. Web page: http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_176/l_17620030715en00010010.pdf: 10.
- FERC (2002). Notice of Proposed Rulemaking, Remediating Undue Discrimination through Open Access Transmission service and Standard Electricity Market Design. Washington, Federal Energy Regulatory Commission. Web page: <http://www.ferc.gov/industries/electric/indus-act/smd/nopr.asp>.
- FERC (2003). White Paper: Wholesale Power Market Platform, Federal Energy Regulatory Commission (FERC). Web page: <http://elibrary.ferc.gov/idmws/nvcommon/NVViewer.asp?Doc=9686957:0>: 32.
- Hogan, W. W. (1992). "Contract Networks for Electric Power Transmission." Journal of Regulatory Economics **4**(3): 211-242.
- Jong, J. d. (2004). The "regional approach" in establishing the internal EU electricity market, Clingendael International Energy Programme. Netherlands Institute of International Relations. Web page: http://www.clingendael.nl/publications/2004/20041200_ciep_paper_dejong.pdf.
- Joskow, P. L. (2005). "Transmission Policy in the United States." Utilities Policy **Vol. 13**: 95-115.
- KEMA (2002). Desarrollo Institucional y Organización de la Comisión Regional de Interconexión Eléctrica - CRIE y el Ente Operador Regional - EOR, Reprt prepared by Kema Consulting and ISA Interconexión Eléctrica for the Unidad Ejecutora del proyecto SIEPAC, in Spanish: 75.
- KEMA (2002). Estudios para el Desarrollo del Reglamento de Operación Técnica y Comercial del MER, Report prepared by KEMA Consulting and ISA

Interconexión Eléctrica for the Unidad Ejecutora del Proyecto SIEPAC, in Spanish: 117.

Olmos, L. and K. Neuhoﬀ (2006). "Identifying a Balancing Point for Electricity Transmission Contracts." IEEE Transactions on Power Systems **Vol. 21**(No. 1): 91-98.

Pérez-Arriaga, J. I. and F. J. Rubio (1995). "Marginal pricing of transmission services: An analysis of cost recovery." IEEE Transactions on Power Systems **Vol. 10**(No. 1): 546-553.

PHB-HAGLER/SYNEX (2000). Diseño General del Mercado Eléctrico Regional. Buenos Aires, Informe elaborado por Mercados Energéticos S.A., PHB-HAGLER BAILLY y SYNEX Ingenieros Consultores para Unidad Ejecutora del Proyecto SIEPAC.

Rivier, M. and J. I. Pérez-Arriaga (1993). Computation and decomposition of spot prices for transmission pricing. 11th PSCC Conference, Avignon, France.

Rubio, F. J. (1999). Metodología de Asignación de Costes de la Red de Transporte en un Contexto de Regulación Abierta a la Competencia. Doctoral Thesis. Escuela Técnica Superior de Ingeniería. Madrid (in Spanish), Universidad Pontificia Comillas.

Rubio, F. J. and J. I. Pérez-Arriaga (2000). "Marginal Pricing of Transmission Services: A Comparative Analysis of Network Cost Allocation Methods." IEEE Transactions on Power Systems **VOL. 15**(NO. 1): 448-454.

Schweppe, F., M. Caramanis, et al. (1988). Spot pricing of electricity, Kluwer Academic Publishers.

SNC/Statkraft (2004). Reglamento de Transmisión y Calidad de Servicio del MER, Report prepared by SNC Lavalin and Statkraft Groner for the Unidad Ejecutora del Proyecto SIEPAC, in Spanish: 120.

WB (2000). Interconexiones Regionales de Mercados Eléctricos en Sudamerica. Propuesta para Facilitar el Incremento de los Intercambios Energéticos en la Región. Buenos Aires, The World Bank, ESMAP, Department of Energy US, CIER.: 151.

Annex 1. Brief introduction to the concept and properties of nodal prices

References to nodal energy prices both in the introduction and in the remainder of this document are numerous. This is due to the importance they have in transmission pricing. The nodal price p_i at node i is the additional cost of operation that the system incurs when serving an extra unit of power at this node. Thus, it is a short-time marginal energy price with locational differentiation. The nodal price of node i may be obtained as a side result of the optimal dispatch problem, since this price is equal to the dual variable of the power balance equation for the corresponding node. Nodal prices were developed by Fred Schweppe and his team at MIT, see (Schweppe, Caramanis et al. 1988).

It can be easily demonstrated that nodal prices are optimal short term economic signals. In other words, when the energy produced or consumed at a certain node i is priced at p_i , market agents located at this node make optimal decisions, from the system point of view, on their level of production or consumption. The proof of this statement follows.

Mathematical Proof

The mathematical proof is comprised of two parts. In the first one we demonstrate that the operational decisions by generators are optimal. The second one focuses on the case of consumers.

The case of generators

We shall assume that marginal production costs monotonically increase with the amount of power produced. Consider the following formulation of the optimal dispatch problem:

$$\begin{aligned}
 & \max \sum_i \{B_i(d_i) - C_i(g_i)\} \\
 & \text{s.t.} \\
 & d_i - g_i + \sum_m m_{im} f_{im} = 0 \quad \forall i \quad p_i \\
 & m_{im} \frac{q_i - q_m}{x_{im}} = f_{im} \quad \forall i, m \quad x_{im} \\
 & f_{im} \leq \overline{f_{im}} \quad \forall i, m \quad g_{im} \\
 & q_{ref} = 0 \\
 & g_i \leq \overline{g_i} \quad \forall i \quad b_i \\
 & d_i \leq \overline{d_i} \quad \forall i \quad a_i
 \end{aligned} \tag{1.1}$$

where i is an index representing the set of nodes; m is an alias of i ; $B_i(d_i)$ is the benefit obtained by agents at node i from the power d_i they consume; $C_i(g_i)$ is the cost incurred by agents at node i when producing g_i units of power; m_{im} is a binary variable whose value is 1 when nodes i and m are connected through a line and is 0 otherwise; f_{im} is flow over the line between nodes i and m in the direction from i to m ; q_i is the phase angle at node i ; x_{im} is the reactance of the line between nodes i and m ; $\overline{f_{im}}$ is the maximum flow allowed over the line from i to m in the relevant direction; q_{ref} is the reference phase angle; $\overline{g_i}$ is the maximum output of power at node i and $\overline{d_i}$ is the maximum demand for power at the same node. Apart from that, p_i , x_{im} , g_{im} , b_i and a_i are the dual variables of the corresponding constraints.

If we express this problem as a minimization one, its first order optimality conditions include, among others:

$$\begin{aligned}\frac{\partial L_{sist}}{\partial g_i} &= \frac{\partial C_i^{opt}}{\partial g_i} - p_i + b_i = 0 \\ b_i(g_i - \overline{g_i}) &= 0 \\ b_i &\geq 0\end{aligned}\tag{1.2}$$

where $\frac{\partial C_i^{opt}}{\partial g_i}$ is the marginal production cost at node i computed at the optimum for the system and g_i is the most efficient production level at node i in order to maximize social welfare. Now, we formulate the optimization problem faced by the generator at node i when deciding how much to produce in order to maximize profits:

$$\begin{aligned}\max \quad & g_i^a r_i - C_i(g_i^a) \\ \text{st.} \quad & \\ g_i^a &\leq \overline{g_i} \quad b_i^*\end{aligned}\tag{1.3}$$

where g_i^a is the amount of power that agents at node i decide to produce, r_i is the unit price paid to these agents for their production and b_i^* is the shadow price associated with the restriction on the amount of power produced by the generator in i corresponding to the generator's problem. Again, when we express this problem as a minimization one and derive the optimality conditions, we obtain:

$$\begin{aligned}
\frac{\partial L_a}{\partial g_i} &= \frac{\partial C_i^a}{\partial g_i} - r_i + b_i^* = 0 \\
b_i^* (g_i^a - \bar{g}_i) &= 0 \\
b_i^* &\geq 0
\end{aligned} \tag{1.4}$$

If agents at this node are offered a unit price $r_i = p_i$ for the power they produce, two possibilities exist:

First option

Agents at node i (we shall assume from now on that there is only one agent) might decide to produce below their maximum capacity. Then:

$$g_i^a \neq \bar{g}_i \rightarrow b_i^* = 0 \rightarrow -p_i + \frac{\partial C_i^a}{\partial g_i} = 0 \rightarrow \frac{\partial C_i^a}{\partial g_i} = p_i \tag{1.5}$$

At the same time we know that $p_i = \frac{\partial C_i^{opt}}{\partial g_i} + b_i$ from the first order optimality conditions of the system dispatch. Then, we have:

$$\begin{aligned}
\frac{\partial C_i^a}{\partial g_i} &= \frac{\partial C_i^{opt}}{\partial g_i} + b_i \\
b_i (g_i - \bar{g}_i) &= 0 \\
b_i &\geq 0
\end{aligned} \tag{1.6}$$

Again, two possibilities exist theoretically speaking:

Case 1.A

If $g_i \neq \bar{g}_i$, i.e. if the optimal output decision from the system point of view is for the generator at node i to produce below its full capacity:

$$b_i = 0 \rightarrow \frac{\partial C_i^a}{\partial g_i} = \frac{\partial C_i^{opt}}{\partial g_i} \tag{1.7}$$

In other words, the agent decides to produce an amount of power such that its marginal cost is equal to the marginal production cost at the system optimum. Provided marginal production costs always increase with the power output, the power output from the generator at node i must be equal to the optimal output level from the system point of view, as we wanted to prove.

Case 1.B

Alternatively, the optimal output decision may be for the generator at node i to produce at its full capacity. Then, $g_i = \bar{g}_i \rightarrow b_i \geq 0$. Since marginal production costs increase monotonically with power output $\frac{\partial C_i^{opt}}{\partial g_i} = \frac{\partial C_i^{max}}{\partial g_i}$. Consequently, from equation (1.6) we have:

$$\left. \begin{array}{l} \frac{\partial C_i^a}{\partial g_i} = \frac{\partial C_i^{max}}{\partial g_i} + b_i \\ b_i \geq 0 \end{array} \right\} \rightarrow b_i = 0 \rightarrow \frac{\partial C_i^{max}}{\partial g_i} = \frac{\partial C_i^a}{\partial g_i} \quad (1.8)$$

Assuming that marginal production costs increase with the production level, equation (1.8) implies that $g_i = g_i^a = \bar{g}_i$, which contradicts the hypothesis under this option ($g_i^a \neq \bar{g}_i$). Then, this situation is impossible.

Second option

The generator at node i might decide to produce at its full capacity. In this case, we have:

$$g_i^a = \bar{g}_i \rightarrow \left\{ \begin{array}{l} b_i^* \geq 0 \\ p_i = \frac{\partial C_i^a}{\partial g_i} \Big|_{g_i^a = \bar{g}_i} + b_i^* \end{array} \right\} \rightarrow p_i \geq \frac{\partial C_i^{max}}{\partial g_i} \quad (1.9)$$

Therefore, we have:

$$\begin{aligned} p_i &\geq \frac{\partial C_i^{max}}{\partial g_i} \\ p_i &= \frac{\partial C_i^{opt}}{\partial g_i} + b_i \\ b_i &\geq 0 \\ b_i (g_i - \bar{g}_i) &= 0 \end{aligned} \quad (1.10)$$

Two possibilities exist:

Case 2.A

We may assume that $\frac{\partial C_i^{opt}}{\partial g_i} = \frac{\partial C_i^{max}}{\partial g_i}$. In this case, $g_i = \bar{g}_i = g_i^a$, as long as we accept the assumption that marginal production costs always increase with the power output.

Thus, the decision by the agent about how much to produce would coincide with the optimal one from the point of view of the system. As we aimed to prove.

Case 2.B

If we assume that $\frac{\partial C_i^{opt}}{\partial g_i} < \frac{\partial C_i^{max}}{\partial g_i}$, then:

$$b_i > 0 \rightarrow g_i = \bar{g}_i \rightarrow \frac{\partial C_i^{opt}}{\partial g_i} = \frac{\partial C_i^{max}}{\partial g_i} \quad (1.11)$$

which is against the hypothesis formulated in this case. Hence, this hypothesis is not possible.

The case of consumers

The mathematical formulation of the optimal dispatch problem is the same one considered when studying the optimal decisions made by generators.

The first order optimality conditions of this problem that are related to the amount of power consumer by the demand at node i are:

$$\begin{aligned} \frac{\partial L_{sist}}{\partial d_i} &= -\frac{\partial B_i^{opt}}{\partial d_i} + p_i + a_i = 0 \\ a_i (d_i - \bar{d}_i) &= 0 \\ a_i &\geq 0 \end{aligned} \quad (1.12)$$

where $\frac{\partial B_i^{opt}}{\partial d_i}$ is the benefit of consuming an extra unit of power for the consumer at node i computed at the optimum for the system and d_i is the most efficient consumption level at node i in order to maximize social welfare. Now, we formulate the optimization problem faced by the load at node i when deciding how much to consume in order to maximize profits:

$$\begin{aligned} \max B_i(d_i^a) - d_i^a r_i \\ \text{st.} \\ d_i^a \leq \bar{d}_i \quad a_i^* \end{aligned} \quad (1.13)$$

where d_i^a is the amount of power that agents at node i decide to consume, r_i is the unit price that these agents must pay for the energy they consume and a_i^* is the shadow price associated with the restriction on the amount of power consumed by the load at i corresponding to the demand's problem. Again, when we express this problem as a minimization one and derive the first order optimality conditions, we obtain:

$$\begin{aligned}
\frac{\partial L_a}{\partial d_i} &= -\frac{\partial B_i^a}{\partial d_i} + r_i + a_i^* = 0 \\
a_i^* (d_i^a - \bar{d}_i) &= 0 \\
a_i^* &\geq 0
\end{aligned} \tag{1.14}$$

If agents at this node are offered the energy they consume at a unit price $r_i = p_i$, two possibilities exist:

First option

Agents at node i (we shall assume from now on that there is only one agent) might decide to consume a smaller amount of power than the maximum amount they are able to consume. Then:

$$d_i^a \neq \bar{d}_i \rightarrow a_i^* = 0 \rightarrow p_i - \frac{\partial B_i^a}{\partial d_i} = 0 \rightarrow \frac{\partial B_i^a}{\partial d_i} = p_i \tag{1.15}$$

At the same time we know that $p_i = \frac{\partial B_i^{opt}}{\partial d_i} - a_i$ from the first order optimality conditions of the system dispatch. Then, we have:

$$\begin{aligned}
\frac{\partial B_i^a}{\partial d_i} &= \frac{\partial B_i^{opt}}{\partial d_i} - a_i \\
a_i (d_i - \bar{d}_i) &= 0 \\
a_i &\geq 0
\end{aligned} \tag{1.16}$$

Again, two possibilities exist theoretically speaking:

Case 1.A

If $d_i \neq \bar{d}_i$, i.e. if the optimal output decision from the system point of view is for the generator at node i to produce below its full capacity:

$$a_i = 0 \rightarrow \frac{\partial B_i^a}{\partial d_i} = \frac{\partial B_i^{opt}}{\partial d_i} \tag{1.17}$$

In other words, the agent decides to consume an amount of power such that its marginal benefit is equal to the marginal benefit obtained by the agent at the system optimum. Provided the marginal benefit the agents obtains decreases with the amount of power consumed, the power consumed by the load at node i must be equal to the optimal consumption level from the system point of view, as we wanted to prove.

Case 1.B

Alternatively, the optimal output decision may be for the load at node i to consume at its full capacity. Then, $d_i = \bar{d}_i \rightarrow a_i \geq 0$. Since the benefit obtained by the agent from an extra unit of power consumed decreases monotonically with power consumption,

$\frac{\partial B_i^{opt}}{\partial d_i} = \frac{\partial B_i^{min}}{\partial d_i}$. Consequently, from equation (1.16) we have:

$$\left. \begin{array}{l} \frac{\partial B_i^a}{\partial d_i} = \frac{\partial B_i^{min}}{\partial d_i} - a_i \\ a_i \geq 0 \end{array} \right\} \rightarrow a_i = 0 \rightarrow \frac{\partial B_i^a}{\partial d_i} = \frac{\partial B_i^{min}}{\partial d_i} \quad (1.18)$$

Assuming that marginal benefits decrease with the consumption level, equation (1.18) implies that $d_i = d_i^a = \bar{d}_i$, which contradicts the hypothesis under this option ($d_i^a \neq \bar{d}_i$). Then, this situation is impossible.

Second option

The generator at node i might decide to produce at its full capacity. In this case, we have:

$$d_i^a = \bar{d}_i \rightarrow \left\{ \begin{array}{l} a_i^* \geq 0 \\ p_i = \frac{\partial B_i^a}{\partial d_i} \Big|_{d_i^a = \bar{d}_i} - a_i^* \end{array} \right\} \rightarrow p_i \leq \frac{\partial B_i^{min}}{\partial d_i} \quad (1.19)$$

Therefore, we have:

$$\begin{aligned} p_i &\leq \frac{\partial B_i^{min}}{\partial d_i} \\ p_i &= \frac{\partial B_i^{opt}}{\partial d_i} - a_i \\ a_i &\geq 0 \\ a_i (d_i - \bar{d}_i) &= 0 \end{aligned} \quad (1.20)$$

Two possibilities exist:

Case 2.A

We may assume that $\frac{\partial B_i^{opt}}{\partial d_i} = \frac{\partial B_i^{min}}{\partial d_i}$. In this case, $d_i = \bar{d}_i = d_i^a$, as long as we accept the assumption that marginal benefits of consuming power always decrease with the

level of power consumption. Thus, the decision by the agent about how much to consume would coincide with the optimal one from the point of view of the system. As we aimed to prove.

Case 2.B

If we assume that $\frac{\partial B_i^{opt}}{\partial d_i} > \frac{\partial B_i^{min}}{\partial d_i}$, then:

$$a_i > 0 \rightarrow d_i = \bar{d}_i \rightarrow \frac{\partial B_i^{opt}}{\partial d_i} = \frac{\partial B_i^{min}}{\partial d_i} \quad (1.21)$$

assuming decreasing marginal benefits. This result is against the hypothesis formulated in this case. Hence, this hypothesis is not possible.

Comments on the proof

In the previous proof we have assumed that consumers decisions on the amount of power to buy exhibit some elasticity to the price of energy (the larger the amount of power purchased, the smaller value they place on consuming an extra unit of power). However, most consumers do not decide how much power to purchase at any given moment in time based on the price they will have to pay for it. Hence, the amount of power retrieved by consumers at the majority of nodes can be considered to be an input to the dispatch problem where prices are determined.

We have demonstrated that nodal or spot prices are optimal short term economic signals because they encourage agents to behave efficiently. However, throughout the proof we have assumed that the central dispatcher has access to the true cost function of the generator at node i (since we have worked under the hypothesis that the cost function $C_i(g_i)$ used in the dispatch and the agent problems is the same). This will not be true if a market is in place and generators are allowed to bid in order to maximize profits since, in this case, strategic generators will not bid according to their true cost function. In real life, some degree of market power always exists. What is more, we have assumed that the dispatcher has access to the true utility function of consumers (B_i for consumer at node i). This, may or may not be true. Therefore, unless a minimum cost regulated dispatch takes place, or agents are forced to bid their marginal production cost, and the dispatcher have access to complete information with regard to the value that consumers place on energy, nodal prices will not maximize social welfare in the short term. However, this pricing scheme will certainly be more efficient than others.

This proof is just a simplification of the one presented by Schweppe et al. in (Schweppe, Caramanis et al. 1988). However, it serves the purpose of illustrating the good properties that these prices have to drive the operation of the system.

Other properties of nodal prices

Decomposition of spot prices

If some conditions are fulfilled (see the previous point), nodal prices can be considered optimal short term signals. Rivier et al. carry out in (Rivier and Pérez-Arriaga 1993) an interesting decomposition of spot prices into several terms. One of these terms is normally called the *system lambda* and can be thought of as the system energy price. The other two terms are related to the effect that consuming an extra unit of power at a certain node has on the transmission losses and congestion in the system. Rivier also derives useful relationships between the expression of the spot prices at different nodes of the system. Thus, one can easily express the spot price at a certain node as a function of the price at any other node and the impact of a transaction between these two nodes in the system losses and congestion. Now we derive the most interesting expressions.

We must bear in mind that the mathematical expression obtained for nodal prices fundamentally depends on the formulation that is used for the economic dispatch problem. Consider the following formulation of this problem:

$$\begin{aligned}
 & \max \{ B_i(d_i) - C_i(g_i) \} \\
 & \text{s.t.} \\
 & \sum_i (d_i - g_i) + L(d_i, g_i) = 0 \quad l \\
 & \sum_i \{ PTDF_{i,l} (g_i - d_i) \} = f_l \quad \forall l \quad x_l \\
 & f_l \leq \bar{f}_l \quad \forall l \quad g_l \\
 & g_i \leq \bar{g}_i \quad \forall i \quad b_i \\
 & d_i \leq \bar{d}_i \quad \forall i \quad a_i
 \end{aligned} \tag{1.22}$$

where all the symbols have already been used in this section but $L(d_i, g_i)$, which represents the transmission losses in the system as a function of power injections and withdrawals; $PTDF_{i,l}$, which is the Power Transfer Distribution Factor of the flow over line l with respect to the power injection at node i , i.e. it is the sensitivity of the flow over this line with respect to the power injected at this node; f_l is the flow over line l and $\bar{f}_l$ represents the maximum amount of power allowed over line l in the direction in which the flow actually goes in the scenario considered. Finally, when used as an index, l refers to the set of lines in the system.

From the first order optimality conditions of this problem, we obtain:

$$\frac{\partial L}{\partial g_i} = \frac{\partial C_i}{\partial g_i} - l \left(1 - \frac{\partial L}{\partial g_i} \right) + b_i + \sum_l (PTDF_{i,l} x_l) = 0 \quad (1.23)$$

Then, we have the following expression for the spot price at node i :

$$p_i = \frac{\partial C_i}{\partial g_i} + b_i = l \left(1 - \frac{\partial L}{\partial g_i} \right) - \sum_l (PTDF_{i,l} x_l) \quad (1.24)$$

If now we derive the Lagrangian of problem (1.22) with respect to the flow over line l , f_l , we have:

$$\frac{\partial L}{\partial f_l} = x_l + g_l = 0 \rightarrow x_l = -g_l \quad (1.25)$$

Substituting equation (1.25) in equation (1.24) and operating, we obtain the following expression for the spot price at node i :

$$p_i = l + \sum_l (PTDF_{i,l} g_l) + l \frac{\partial L}{\partial d_i} \quad (1.26)$$

where l is the system energy price, $\sum_l (PTDF_{i,l} g_l)$ is the term associated with congestion in the system and $l \frac{\partial L}{\partial d_i}$ is the term associated with transmission losses.

Analogously, the expression of the spot price at node j is:

$$p_j = l + \sum_l (PTDF_{j,l} g_l) + l \frac{\partial L}{\partial d_j} \quad (1.27)$$

Therefore, we have that:

$$p_i = p_j + l + \sum_l \left\{ (PTDF_{i,l} - PTDF_{j,l}) g_l \right\} + l \left[\frac{\partial L}{\partial d_i} - \frac{\partial L}{\partial d_j} \right] \quad (1.28)$$

Considering that the flow over each line l is a linear function of the power injections and withdrawals in the system, and that the function of transmission losses is very approximately linear around the working point, we finally obtain:

$$p_i = p_j + \sum_l \left\{ PTDF_{j,i,l} g_l \right\} + l \frac{\partial L}{\partial t_{ij}} \quad (1.29)$$

This is the relationship between spot prices at nodes i and j that we were looking for. $PTDF_{jil}$ represents the impact on the flow over line l of a unit transaction between these two nodes (from j to i) while $\frac{\partial L}{\partial t_{ij}}$ is the impact on the system transmission losses of this same unit transaction.

Recovery of the cost of constructing, operating and maintaining the grid

As mentioned above, one first attempt to recover the regulated revenues of the transmission service would be applying nodal prices to the power injections and withdrawals in the system. Pérez-Arriaga and Rubio show in (Pérez-Arriaga and Rubio 1995) that spot prices (short term marginal costs) are equal to long term marginal costs under certain, very stringent, conditions that never occur but in stylized theoretical examples. These conditions are the following:

- Investment in new lines must be continuous as opposed to discrete. Thus, building 1.43 lines would be possible.
- The cost of a new line must be a linear function of its capacity.
- The dispatch problem must not include any reliability constraint.
- Demand must show some elasticity to energy prices or there must be some local marginal generation in every node of the system.

In reality, however, investments are always discrete: we can build 1 line or 2 but never 1.43 lines. In addition, economies of scale exist in the installation of lines: the more transmission capacity a line has, the lower is its cost per MW of capacity. Finally, energy prices do not affect the amount of power demanded by most consumers, i.e. the elasticity of demand is zero or close to zero in most cases (at least in the short to medium term), and there is not marginal generation in every node of the system.

If long term marginal costs were the same as short term marginal costs, applying spot prices would result in a complete recovery of the grid cost, unless certain very specific restrictions are added to the dispatch problem (for example, a restriction limiting the size of new grid investments or a security restriction limiting the amount of expected non-served energy per node). Pérez-Arriaga and Rubio formally proved in (Rubio 1999) what they had previously observed empirically in (Pérez-Arriaga and Rubio 1995).

Studies conducted in real power systems throughout the world show that, typically, spot prices applied to power injections and withdrawals only recover a minor fraction of the total cost of the transmission activity. Generally, this fraction is around 20% of the total grid costs, see (Pérez-Arriaga and Rubio 1995). However, this percentage may vary depending on several factors. First, revenues from the application of nodal prices in radial systems tend to differ significantly from those obtained in systems where the

grid is well meshed. Second, special restrictions considered in the system dispatch and planning may greatly impact the amount collected. Thus, security restrictions imposed on the system dispatch increase the revenues resulting from the application of nodal prices. The more restrictive these constraints are, the higher these revenues. We could increase the amount of congestion rents collected as much as we want by increasing the severity of the security restrictions that are taken into account when operating the system but not when planning the expansion of the grid. Many different methods exist to collect the fraction of the grid costs not recovered through the application of nodal prices. All these are explained in detail in Chapter 3 of this document.

2 Investments

2.1 Introduction

Classical optimization of the expansion of the transmission system was conceptually speaking an easy task regardless of the many practical difficulties in its actual implementation. All relevant decisions about the investment in new transmission and generation facilities were in the single hand of the vertically integrated utility or central planner (Pérez-Arriaga, Román et al. 1987; Latorre-Bayona and Pérez-Arriaga 1994; Joskow and Tirole 2002). Thus, typically, generation expansion planning took place first, normally assuming a perfect network without losses or capacity limits. Then the utility planned the transmission grid expansion, taking the location and characteristics of any new generation investment as input data and assuming a certain pattern of demand increase based on their best estimates. In some rare cases the expansion of both generation and transmission was planned jointly, with the single objective of minimizing the cost of electricity supply. Distribution network planning took place separately, where the main input data were the individual or aggregated demands of consumers—both load profile and location—and the substations connecting the high voltage distribution grid to the transmission network (Peco 2001).

In a liberalized regulatory environment a complete new paradigm for the expansion of power systems becomes necessary. Now, the planning of generation, transmission and distribution takes place independently from one another. Generation expansion and operation is now left to the initiative of private investors. The adoption of any regulatory instruments to promote adequacy of generation capacity and security of supply remains an open issue. Distribution expansion has presently become more sophisticated with new ideas, such as performance-based remuneration and benchmarking or the stimulating perspective of widespread distributed generation, but this change is not necessarily associated with the liberalization of the power sector. The accumulated experience in the restructuring of the sector shows that transmission expansion critically depends on the regulatory paradigm that is adopted for the activity. Designing a regulatory framework for the investment in new transmission facilities includes choosing the entity responsible of proposing the construction of new lines and the one that must give its approval to the investment projects, the decision on who will build the transmission facilities, the remuneration of the transmission services and the economic signals and incentives used to encourage any involved entity to perform these tasks efficiently.

There are different approaches to make compatible, in an efficient way, the two worlds of competitive generation and monopolistic transmission in the context of a liberalized regulatory framework. One possible option is leaving some specialized public-service type of entity to keep performing centralized transmission expansion in the classical manner. Under this scheme, there are several aspects that now become really important and deserve to be discussed specifically. A clear and objective process is needed—the so called “regulatory test”—to determine which new investments must

be made. This process must avoid being discriminatory, i.e. it should not benefit some agents at the expense of others. The advisability of establishing incentives for the System Operator to plan the expansion of the grid in the most efficient way must also be investigated.

Another option is to let market agents participate actively in the decision making process of transmission investment. There are several possibilities. One has to decide to which extent transmission investment must be left in private hands and how to make private investment compatible with some level of centralized decision-making or regulatory control. The role of transmission rights in this process is very relevant, see (Bushnell and Stoft 1996; Joskow 2005; Joskow and Tirole 2005).

The expansion of generation, transmission and distribution are intimately linked, as mentioned above. In order for the development of the system to be optimal investors in generation, transmission or distribution must take into account the investment and operational decisions made by other parties. Section 2.5 discusses the compatibility of investment signals in the three activities.

This chapter examines a number of issues related to the expansion of the grid that still do not have a conclusive answer in the new regulatory environment. Some of them are faced by all the investors in transmission facilities. Others are only faced by transmission planners:

- Whether the objective of transmission investment should be the same under the traditional and the market-based regulatory approaches.
- Once we know which is the objective to pursue, we must determine the most efficient regulation of the activity of investing in new transmission facilities.
- Within each existing regional market, this model must be adapted to the prevailing circumstances and organizational arrangements in it.
- Details must be provided about the criteria with respect to which each potential new investment should be assessed or the set of rules to be applied to ascertain whether this investment is optimal (or at least justified) or not.
- How to achieve the required level of coordination between the investments taking place in generation, transmission and distribution.

After this introduction, section 2.2 of this chapter reviews and classifies the several paradigms that have been proposed to regulate transmission investment. Section 2.3 proposes a model for the expansion of the grid and adapts this model to the particular cases of the most important existing regional markets. Section 2.4 introduces and discusses the regulatory test, a construct to provide conceptual support to any decision related to the regulatory approval of a new transmission investment. Section 2.5

discusses the economic signals for investment in generation, transmission and distribution and how they interrelate. Finally, section 2.6 provides the conclusions.

2.2 Regulatory paradigms in transmission investment

With the introduction of competition in generation and supply, the regulatory environment in which the planning of transmission takes place has changed radically. Even the organization model of the transmission activity is under scrutiny. Most systems maintain traditional planning approaches corresponding to a regulated monopoly, although some incentive schemes may have been added. And there are those who favor the restructuring of this activity by leaving some responsibility in the expansion of the grid to the private initiative. Conceptually, two possibilities exist in this regard. Coalitions of network users may be allowed to promote, or even decide, the construction of some lines, subject to some sort of regulatory approval and to the application of the same general rules of open access, remuneration and tariff setting used for any other lines. Alternatively, private investors may be allowed to build a line and sell its transmission services by buying electricity at one end of the line and selling it at the other.

Incentives to invest in transmission do not exist *per se*. They depend on the regulatory framework adopted for the transmission activity as a whole. Ideally, those agents that benefit from the expansion of the grid, i.e. the market agents, should be left to decide, or at least propose, the construction of new lines. However, due to the fact that market benefits are frequently very difficult to identify and allocate, or they are very widely distributed, such an approach may turn out to be problematic in many cases. This section describes the different regulatory schemes that exist for transmission investment. The classification presented here is based on the one included in (Pérez-Arriaga and Omos 2005).

The description of the different regulatory schemes includes both the institutional arrangements within each one of them (who proposes and/or makes the decision of investing in new lines, who –if anyone– supervises this process) and the remuneration of the owner of a transmission line, which also depends on the paradigm that has been adopted.

2.2.1 Centralized expansion of the transmission grid

The System Operator (SO) is responsible for proposing a reinforcement plan, which has to be authorized by the regulator. The System Operator may or may not be integrated with the transmission firm or firms. In the former case it is named Transmission System Operator or TSO. The construction of any reinforcement may be assigned by

competitive bidding or (more frequently) to the incumbent transmission company under cost-of-service remuneration.

Regularly, the System Operator must propose a plan to reinforce the transmission network after taking into consideration (or rejecting, with due justification) any proposals made by network users. Note that in a competitive environment network expansion may strongly affect the economic benefits of network users. The regulatory authorities approve the plan (or not) and, if it is approved, they authorize the construction of individual new facilities. Here the System Operator can be termed “passive” because it proposes plans for the expansion of the transmission network, but does not make the final decisions regarding the approval of the plan and the construction of the line (e.g. REE in Spain¹).

Both the System Operator, when proposing the construction of a line, and the regulator, when deciding upon the approval of the project, need some kind of rule to apply in order to determine whether the investment is justified under a technical and economic viewpoint. This conceptual rule is called the Regulatory Test. Section 2.4 introduces the concept of the regulatory test. Annex 1 presents the mathematical formulation of the problem of optimizing the expansion of the grid within the traditional framework. This annex also shows that, provided that negative benefits (losses) incurred by agents as a consequence of the construction of a line are taken into account, optimizing the expansion of the grid under a centralized planning type of regulation is conceptually equivalent to the regulatory test to be applied by the System Operator and the regulator in the new competitive environment.

Construction, operation and maintenance of each facility are here the responsibility of the incumbent transmission company (usually the TSO). An alternative is competitive bidding. If competitive bidding is adopted, the winner of the bidding process is paid as bid, and the bid becomes the regulated cost of the line or lines to be built. Since the contract will have a limited duration (the economic life of the asset, typically), the operation and maintenance of the facility may be auctioned again at the end of the period. Availability targets may be set for each facility and penalties or credits can be applied, depending on the actual performance.

Provided the SO is the owner of the transmission grid (then, it is a TSO), its remuneration must refer to the existing network. In principle, any incentives -either penalties or credits- should only depend on the availability record of the network equipment. Then, the TSO will typically have an incentive to over-invest in the transmission grid. If the TSO builds more lines, the size of the company will grow with assets that have a guaranteed and attractive remuneration (if the allowed rate of return is reasonable). Although the ultimate responsibility for transmission planning

¹ REE stands for Red Eléctrica de España, the Spanish Electricity Grid, in Spanish.

corresponds to the regulator it will be difficult for this institution to reject investments that have been proposed by the TSO.

If the SO is not the owner of the transmission grid its revenues should be independent of the volume of transmission assets. This entity should receive a fixed amount plus, maybe, some small incentives related to its performance. In any case, both the regulator and a “passive” SO have a natural inclination to accept overinvestment in grid facilities, since both can be held responsible for the overall security of the power system in one way or another. All in all, some degree of overinvestment is to be expected with respect to what would be optimal according to the assumed utility function of consumers. It is debatable whether this is a positive feature of this method or not.

Joskow concludes in (Joskow 2005) that most grid reinforcements need to be carried out by a central planner, preferably one of regional scope (the Federal Energy Regulatory Commission in the US for example). The same is claimed by Newbery in (Newbery 2003). Stoft states that centrally planning the construction of lines whose use will be spread widely among network users is necessary. He proposes defining clearly which lines must be built by private investors and which ones must be built by a central planner, see (Stoft 2005). The recently approved regulation for the Central American Electricity Market (MER) broadly coincides with this approach. Authors in (Dismukes, III et al. 1998) investigate the existence of economies of scale in the activity of the transmission of electricity. They find strong economies of scale, which dictates keeping this activity as a regulated one. Similarly, (Keller and Wild 2004) concludes that some coordination among the parties involved in the expansion of the grid is necessary even in the case where efficient transmission pricing mechanisms are in place throughout the system (nodal pricing). This coordination may materialize in an institution which looks after the optimal development of the grid. A coherent pricing system to recover total transmission costs will also be necessary.

However, Joskow also claims that regulated investment may not be free from some of the problems associated with merchant lines. Thus, difficulties to estimate the true congestion costs of the system, and the inability of short term marginal costs to reveal the market benefits that would be obtained from carrying out a grid investment, may lead the TSO not to invest optimally, see (Joskow 2005). Many authors argue that the TSO, or the entity in charge of the grid expansion and operation, should be made somehow responsible of the costs that its decisions cause the system to incur. Theoretically, this scheme would align the TSO’s interests with those of generators and consumers. See the next point for a discussion on this argument.

Regarding the organization of the central planning of the grid expansion, Baldick and Kahn propose the creation of an open planning process where market agents and regulators would have something to say. This could partially make up for the lack of coordination between the expansion of generation and transmission processes resulting from the liberalization of the electricity sector, see (Baldick and Kahn 1993).

2.2.2 The expansion of the grid is left to a private for-profit company

A for-profit (private, usually) company is in charge of deciding on the optimal expansion of the grid and implementing the resulting investment plan afterwards. This company is awarded the transmission license and it is regulated as a monopoly. The transmission company must follow prescribed design requirements (a mandatory grid code) and it is subject to incentives to meet performance targets for the transmission system. This is similar to the more customary regulation of distribution networks. This private company is what may be called an “active” TSO because it is ultimately responsible for the expansion of the network (e.g. NGC in England & Wales).

The remuneration of the transmission activity for these TSOs should refer to an efficient and well adapted network -not the actual one-. The scheme that is typically adopted to determine the global remuneration of the complete transmission network is of the type RPI-X². The SO, which typically is the owner of the transmission system, presents its transmission expansion plan for the next control period to the regulator. Based on this plan, as well as on the forecast of the construction, operation and maintenance costs that is also submitted by the SO, the regulator determines the remuneration of the transmission activity over this period. Finally, the regulator computes the parameters of the RPI-X scheme, namely the amount to be collected from transmission tariffs the first year of the control period and the adjustment factor X so that the net present value of the streams of expected costs and revenues coincide. Afterwards, the SO unilaterally decides the grid reinforcements to be carried out. Economic incentives could exist that are related to the actual performance of the transmission system regarding quality of supply. Obviously, in the next control period the regulator must take into consideration the grid investments that the SO has effectively undertaken during the current control period and the actual system performance.

Given that transmission network investments can be easily examined individually, the remuneration for a control period should take explicitly into account the planned new investments, the economic lives and the depreciation of existing investments, the financial situation of the transmission company and the expected improvements in efficiency. Combining the application of an RPI-X scheme to the entire network and the consideration of individual investments and economic incentives to promote better performance makes transmission remuneration more of an art than anything else.

² The alternative regulatory frameworks that are presented here are representative cases of the existing possibilities, since an exhaustive enumeration is out of question. The RPI-X scheme has been implemented in diverse forms although the one schematically described here appears to be the most sensible one.

Since the RPI-X remuneration scheme is such that the income of the TSO is not directly related to the actual investment, the System Operator may invest in transmission at least in order to comply with any mandatory grid codes, and only beyond this if the remuneration scheme (the adapted version of RPI-X plus performance-based incentives) appears to make the new investment profitable. Compliance with the minimum levels of quality of supply can be guaranteed by the use of administrative penalties. Overinvestment is unlikely to occur under this scheme. However, achieving an optimal level of investment seems difficult as well.

Some authors in the academic community are of the opinion that making TSOs responsible for those costs that are related to the operation and planning of the system is the only way to encourage these operators to carry out their function efficiently. Thus, Leautier proposes in (Léautier 2000) the definition of a regulatory contract for the TSO that would make this entity internalize the system costs related to congestion in the grid, losses and the construction of new lines. The difficulty with this scheme would be in computing an efficient revenue cap for the TSO. Similarly, Vogelsang proposes using two part tariffs forming a cap for the revenues to be earned by independent transmission companies. These tariffs would consist of a fixed part for the recovery of the cost of construction of lines and a variable part associated with congestion costs. However, he acknowledges that proper weights for these two parts should be found. Otherwise, the level of investment would not be optimal, see (Vogelsang 2001). Joskow proposes in (Joskow 2005) making the TSO responsible for the extra cost caused by this institution performing below certain standards. The TSO would not be paid if a collapse occurs and redispatching costs (assuming a nodal pricing scheme is in place) would be born by it as well. But these costs are not the sole responsibility of the TSO in many cases. In all these schemes the TSO would be subject to huge risks which are not in accordance with its potential profits.

Turvey concludes in (Turvey 2000) that signing a contract between the provider of a facility and network users may be necessary in order for beneficial reinforcements to be carried out. This contract could reconcile the revenue requirements on the part of the promoter and the long term value of the considered enhancement. This contract already exists in the case of regulated lines since, whenever a regulated project is deemed to be beneficial for the system (it passes the regulatory test), network users acquire the commitment to pay the TSO the cost of the investment plus a reasonable rate of return.

It is not a trivial task to develop remuneration mechanisms that encourage a for-profit System Operator to pursue those investments that are most beneficial for the system, since the incentives may have implications in different aspects of the security and economics of the power system. For instance, economic incentives to reduce congestion costs might result in reduced security margins. Or economic incentives to reduce network losses might result in operating measures that could be detrimental to the efficiency of the generation dispatch. A carefully designed combination of moderate economic incentives and regulatory supervision might be successful in practice.

Experiences in countries where this type of approach has been adopted show that certain well justified reinforcements -interconnections in particular- are not attractive to the planner because they would not earn any income for their construction, see (Pérez-Arriaga and Smeers 2003).

2.2.3 Network users form coalitions to promote the construction of new lines

As it has been mentioned before, the beneficiaries of the construction of a line are the network users. Therefore, those agents that would benefit from the construction of a line should be willing to cooperate to get it built. In order for a proposed reinforcement to be built, the regulator must authorize it. A regulatory test is needed for the regulator to be able to decide whether a new line is justified, see (Littlechild and Skerk 2004; Littlechild and Skerk 2004). As explained before, two possibilities exist here:

- The regulator may consider both the positive and negative effects of the line on the benefits of all the market agents.
- The regulator considers only the positive benefits brought about by the new line. Then, if the coalition of agents proposing the construction of a line agrees to pay its cost this line should always be built.

Section 2.4 includes a discussion on whether negative effects should be considered when applying the regulatory test to a potential grid reinforcement.

The construction of the approved reinforcements may be assigned by competitive bidding. This process would determine the total regulated cost to be recovered by network tariffs, as it is the case of alternative 1. Alternatively, coalitions of network users may also be in charge of building network reinforcements.

Consequently, several regulatory options are possible under this scheme. Some of them have been tried in the Argentinean system, see (Littlechild and Skerk 2004; Littlechild and Skerk 2004):

1. In a first option the coalition builds and pays for the reinforcement, which needs authorization from the regulator. This coalition has to provide open access to the facility and receives the proceeds from the transmission charges paid by its users. This option is commonly known as “investments at risk”.
2. In a second option coalitions pro and against the building of the line have to go through a quasi-judicial process, and the regulator decides whether the reinforcement is justified or not. If justified, its construction and maintenance is assigned by competitive bidding and paid-as-bid to the winner of the auction. Availability targets may be set and penalties or credits will be applied, according to the actual performance. The license will have limited duration

(typically, the economic life of the line) and will be auctioned again for operation and maintenance, and any additional costs needed for life extension, for the next period. The regulated cost of the reinforcement will be charged to all network users by means of transmission tariffs.

3. As in the second option, lines of the third type must go through a quasi-judicial process. If the reinforcement is finally found to be justified and the proposing coalition is properly defined, construction is assigned by competitive bidding and the cost of constructing and maintaining it is paid as bid to the winner of the tendering process. A fraction of this cost, which may be related to the fraction of the line that is actually used, is charged to all network users by means of regulated transmission tariffs. The remaining fraction is covered by the proprietary coalition, who also receives financial rights (firm transmission rights, FTRs) for the corresponding part of the capacity of the line.

Provided regulation allows agents to participate in the decision making process leading to the expansion of the grid, nodal energy prices and transmission tariffs with spatial differentiation, when they exist, will provide incentives for the network users to organize coalitions to promote new transmission investments that are beneficial for them. The success of this scheme critically depends on the existence and correctness of the locational signals in nodal prices and / or transmission tariffs.

Lines that are economically justified but whose benefit is widely dispersed (e.g. lines whose main purpose is to improve the overall reliability of the power system) will never get a coalition of users that is interested in promoting or building them. In a well meshed network, as it is generally the case in Europe, the economic benefits of new transmission reinforcements will probably be quite dispersed. Therefore, this scheme normally has to be supplemented by a centralized back-up scheme that makes sure that these lines are built.

The experience of utilization of these methods in Argentina is described in (Anderson and McCarthy 1999; Chisari, Da-Bó et al. 2001). According to these references, the mechanism, as it is applied, exhibits some important deficiencies. In addition to the already mentioned difficulty with the so-called reliability lines, small consumers are left out of the decision making process, the method substitutes electrical usage for benefits when allocating the cost of an investment and the use of a potential new line by network users is not accurately defined. All this indicates the need of introducing some modifications in this approach.

Leautier concludes in (Léautier 2001) that strategic generators, i.e. those with potential of abusing market power, will not be interested in expanding the transmission capacity because the benefits they obtain from exercising the market power they enjoy before the construction of a certain line overweigh the potential benefits resulting from their sales in other markets that are presently closed to them. On the other hand, (Gans 2001) claims that an efficiently designed set of transmission tariffs would encourage

market agents to invest in new transmission capacity at the right moment in order to obtain a competitive advantage over its competitors. However, conditions assumed in the latter work are far from being realistic.

It remains the controversial issue of whether a line should be accepted whenever it is proposed by a coalition whose benefits related to the construction of the line are higher than the cost of this line, i.e. whether the negative benefits caused by a new line should also be taken into account when deciding on its construction. According to this rule, some investments that harm other users and that, taking into consideration the profits and losses they create, are not profitable for the system as a whole, would be approved. If it is agreed that negative benefits should be taken into account when assessing grid investment projects, two possibilities exist: either the construction of a line is discarded whenever the net balance of profits and losses is negative, or transmission charges are designed so that beneficiaries not only pay for the cost of the line but also for the harm that the construction of the line is doing to other agents in the system.

Another alternative for the participation of network users in the process of development of the grid is the organization of long-term auctions where parties bid for capacity that has not been constructed yet. This mechanism would allow market agents (the ultimate beneficiaries of transmission grid enhancements) reveal their willingness to have new transmission capacity built and to pay for it. Long-term capacity auctions for transmission expansion have been long discussed in the academic world and economists find serious problems in exclusively relying on them for the construction of new lines, see (Helm 2003; McDaniel 2003; Newbery 2003; Stern and Turvey 2003; Yarrow 2003). Arguments found by economists and engineers for not relying on transmission capacity auctions for the expansion of the grid are coincident. These auctions may contribute partially to the construction of new lines but they cannot be believed to result in a significant development of the grid.

Most authors agree that capacity auctions are efficient to allocate existing capacity in the long term. However, implementing long-term auctions for capacity yet to be constructed seems far more problematic. The latter auctions will lead to a situation where necessary transmission upgrades will probably not be undertaken, see (Helm 2003; Newbery 2003). This may be related to the lack of liquidity in these auctions (McDaniel 2003) or to the incentives for the agents not to reveal their true expectations of future transmission capacity needs so as to drive up the price of the transmission rights they have already acquired. According to these authors, a regulated procedure for the centrally planned expansion of the grid should coexist with these auctions. In any case, if these auctions are implemented secondary markets for capacity would be crucial to facilitate the trade among agents of the previously auctioned capacity products, see (Yarrow 2003).

2.2.4 Merchant lines

Merchant lines are regulated as any other competitive business. Therefore, the owner of a merchant line could use it to buy energy at one end of the line (where the energy is cheaper) and sell it at the other (where it is more expensive) at market-determined prices. Network constraints –most frequently line congestions- are at the origin of the differences in energy prices between nodes.

Thus, the remuneration of a merchant line comes from the difference in energy prices between the end nodes of the line or from the revenues resulting from the sale of the corresponding transmission rights (some authors have proposed bidding the network capacity of some DC lines in a short term market). In other words, the expectation of these congestion rents is the only driver behind network investments by merchant grid promoters. These differences in energy prices may take place within a single power system or in between power systems. Obviously, open network access to both buyers and sellers is necessary.

Selling firm transmission rights (FTRs) over the capacity of a merchant line must be seen not only as a risk hedging mechanism but mainly as a way to finance the investment that is required to construct it, see (Bushnell and Stoft 1996; Chao and Peck 1996; Joskow and Tirole 2005). A promoter considering investing in a new line or corridor will have more certainty of recovering at least a fraction of the investment if he is able to sell transmission rights for part of the capacity of the line in advance of its construction.

Some of the difficulties in generalizing this approach to a significant number of lines, i.e. relying mainly on merchant projects for the expansion of the grid, are:

1. Insufficiency of market driven revenues to pay for the total cost of a well developed network³. Pérez-Arriaga and Rubio demonstrate in (Pérez-Arriaga and Rubio 1995) that under ideal conditions congestion rents would be enough to pay for the total cost of the grid, but they also show in (Pérez-Arriaga and Rubio 1995) that, when we consider real networks, and mainly due to the discrete nature of the investments, congestion rents fall significantly short of paying the whole cost of transmission lines except for a few cases. Therefore, only those lines where high congestion rents are expected will be built under this approach as Joskow states in (Joskow 2005). This statement has proven to be true in numerous actual networks.
2. The revenue stream from the operation of the line is subject to significant risks, mainly due to the possibility of new regulated lines being built, which would

³ This is a well known fact that has proved to be true in numerous actual networks and that also has a theoretical support, see (Pérez-Arriaga, J. I. and F. J. Rubio 1995)

reduce the amount of congestion happening in the line, see (Brunekreeft 2003; Joskow 2005) among others.

3. Given the highly discrete nature of transmission investments, underinvestment in a merchant line is a way of increasing the potential for the exercise of market power. Once a merchant line is built that is smaller than the optimal investment from a social point of view, the central planner may find installing a new line unprofitable for the system since the total investment in the corridor may then happen to be excessive.

Therefore, whenever merchant lines are allowed as a possibility to build reinforcements, they should always be just a complement to another approach, such as the ones proposed in 2.2.1 or 2.2.2 above, see (Joskow 2005). Contrary to what happens with merchant lines, the schemes in 2.2.1 or 2.2.2 may well result in most of the lines that are justified because of economic and security reasons being built. Bushnell and Stoft argue in (Bushnell and Stoft 1996) that most grid investments promoted by private investors would not be detrimental to the system, since transmission rights associated with detrimental investments would have a negative value. Despite all this, some parties may be willing to build a detrimental line if the value of the transmission rights they already own increases as a consequence of the extra congestion caused by the new line.

If some kind of centralized network planning and merchant investments coexist, rules for priority between both approaches have to be established. In some systems the installation of a merchant line is subject to the regulatory authority giving its consent to it. The consent might be subject to verification that the proposed merchant line is not detrimental to the system and, it seems also appropriate, that it is not included in the plan of regulated investments that has been prepared by the responsible centralized institution. This is related to the fact that, whenever the construction of a merchant line is justified from the point of view of the private promoter, congestion rents will be higher than the cost of the line. Then, network users will benefit from paying the construction of this line as a regulated one and retaining these congestion rents. See (CEER 2003) for details on this type of regulation in the context of the process of formation of the Internal Electricity Market of the European Union.

Broadly speaking, the main drawbacks of the proposed methods stem from the lack of agreement, to a smaller or larger extent, between the interest of the grid promoter, whoever it may be, and that of the system as a whole. The regulatory test, to be discussed in section 2.4, is aimed at determining whether any proposed transmission investment increases the global welfare or not.

2.3 Selection of the optimal scheme for the regulation of the grid expansion

2.3.1.1 General considerations

After having discussed different regulatory models for the activity of transmission expansion, we must conclude that regulated investment, which may adopt a variety of formats, must play a predominant role in the development of real world transmission networks. Centrally planning the expansion of the grid is especially important in those systems where the network is well meshed. This is the case of most European countries or states in the USA (at least within each country or state, respectively).

When TSOs are ultimately responsible for deciding on the construction of new facilities, some performance-based incentives may be placed on them so that network reinforcements go beyond the minimum that is required by any mandatory grid code. However, when TSOs only propose the reinforcements and the regulatory authorities make the final decision about the realization of these investments, the natural incentive of the TSO will be towards overinvestment, since this creates a comfortable security margin for the operation of the power system and also provides the opportunity for the company to grow in size. The regulator must be aware of this tendency, although the regulator itself is not free from the same impulse towards more security.

As mentioned in the previous paragraphs, if economic signals are provided by nodal prices and transmission tariffs with locational content, coalitions of network users receive the correct economic incentives to promote the construction of new transmission facilities. However, this sound mechanism can only be trusted for those lines with clearly defined and relatively few economic beneficiaries. Thus, generators or consumers (or associations of them) in radial systems (underdeveloped networks) may find it attractive to build their own transmission line or corridor to have access to consumers in the former case or to cheaper energy in the latter. These are normally called investments at risk by (associations of) market agents.

Also interesting is the idea of coalitions of TSOs who could propose the development of interconnections or other facilities of interest for cross-border trade. The facilities could be authorized by the corresponding regulators after using some kind of extended regulatory test. Regulators should make sure that these cross-border investments are properly remunerated. This scheme appears to be a pragmatic approach to getting some lines built when they are mainly justified in terms of their impact on cross-border trade. This is particularly true of those lines that are built within a country that does not happen to be the main beneficiary of their construction, see (CEER 2003; Pérez-Arriaga and Omos 2005).

Nodal prices and/or congestion rents are natural incentives for merchant investors who want to appropriate these price differences. However, the investment that would maximize the profits of a merchant investor is typically of a lower capacity than the

investment that the regulator would have chosen. A larger investment would normally reduce too much the remaining congestion rents. Besides, if both merchant and regulated network investments coexist, then the merchant investor has always the risk of losing his congestion rents because of a new regulated investment. This is why merchant investments can only contribute to the development of a transmission network in a few specific instances.

Merchant projects may play a more important role in the expansion of radial networks since significant systematic congestion may occur in these cases that private promoters may want to take advantage of. Regulators should make sure that the merchant lines built are optimal investments (or close to be optimal) from the point of view of the system as well. There are several variants on the basic type of merchant line to align the interests of the private promoter and those of network users, see (Vázquez, Rivier et al. 2005).

In the new regulatory environment there is uncertainty about generation investment and the difference in time scales between the construction time of transmission and generation no longer exists. This is one major problem faced by grid promoters in most liberalized systems and adds a new level of complexity to transmission network expansion that will be discussed when presenting the regulatory test to be applied to authorize grid investments.

2.3.1.2 Recommendations

Based on the considerations made in the previous paragraphs, we now outline what should be the features of a global regulatory framework for the construction of new lines in a regional market (as indicated before, the regional dimension adds the difficulty of having to deal simultaneously with different national regulations of transmission investment):

- The initiative to propose new investments will be mostly left to SOs, since they have the experience to do so and the technical knowledge, as well as the expertise to evaluate cross-border transmission investments. Depending on whether lines belong to the category of “domestic” or “regional” (cross-border) only one SO or several SOs will be involved in the process. Merchant lines, either promoted by “active” SOs or by other private entrepreneurs, are also possible, subject to some conditions and priority rules. Lastly, associations of network users may promote and even construct new lines that are of especial relevance for a limited number of them. In any case, the construction of these lines must be subject to the authorization of the relevant SO or SOs and/or regulatory authorities, according to the existing national regulation.
- When identifying the need for new network reinforcements, SOs must take into account the proposals by (associations of) network users, who, in principle, are the ones that directly receive any existing locational signals. SOs must incorporate them to the expansion plan or reject them in a justified manner.

- A clear and objective procedure must exist aimed at determining whether a potential network reinforcement is “domestic” or “regional”. Thus, regional authorities may set a threshold for the utilization of the line by agents located within areas (countries, states) other than the one or ones where this line would be built. If the utilization of the line by “external” agents is above this threshold the line should be considered a regional reinforcement. Otherwise, it should be considered a domestic line.
- Only the local authorities (regulator) and the local TSO can participate in the decision making process leading to the construction of a “domestic” line. On the other hand, regulators and TSOs of control areas that are affected by the construction of a new “regional” line in a third country or state should have something to say in this process. Alternatively, in some cases, it will be better that a regional regulatory body exists that determines whether the construction of any specific “regional” line is beneficial for the region as a whole. This body should also settle any disputes among SOs over the construction of cross-border lines.
- The incentives for SOs to promote the construction of a new network facility will vary depending on whether the approval of the investment project from the regulator or regulators is necessary or not. “Passive” SOs will not find it attractive to invest in a new line if the rate of return on the investment is too low. “Active” SOs will not build a new line if the performance index that is used to remunerate them does not reflect the cost of the investment. Thus, the essential recommendation here is that the involved regulators guarantee a satisfactory remuneration for cross-border investments. There would be no problem in leaving to each country or TSO the construction of domestic or cross-border lines when the involved TSOs find that the existing remuneration mechanisms provide enough incentive for them to do so. However, in many cases, these remuneration schemes are not able to encourage TSOs to promote efficient cross-border investments.

A high enough rate of return is necessary for cross-border network reinforcements to be built under the traditional cost-of-service scheme. Thus, we propose the introduction of a regulated scheme for cross-border lines that is based on public tendering of the construction of any new cross-border transmission facility that has been identified by the corresponding SO or SOs, with automatic recognition of the remuneration requested by the winning bid. This mechanism will achieve a drastic reduction of risk for the potential investors. This new regulated scheme is meant to overcome the difficulties of having to deal with the current diversity of regulatory schemes for transmission investment, some of which may not provide sufficient remuneration in some cases. In this way, the involved SO or SOs may resort to a procedure that will always guarantee an adequate remuneration for the new transmission investment. At the same time, incentive schemes should be used in the case of

“active” TSOs to try to encourage them to promote and build those cross-border lines (including the cross-border ones) that are most efficient. These incentives should be limited in scope and quantity. Thus, incentives might be associated with reliability objectives both in the operation and planning of the system. SOs could also be encouraged to reduce losses and/or congestion when operating the system. In any case, these incentives should be modest in order not to completely condition the operation of the system and the planning of the grid.

- Merchant lines have a higher remuneration risk for the investor and a higher cost for network users than the regulated option. However, in some cases regulated investments may not take place and merchant lines can be a viable scheme to promote new network investment. Therefore, as mentioned above, merchant lines may be allowed as a complement to the construction of regulated lines. The regulatory approval of a merchant project must, in general, be subject to the following specific conditions and priority rules:
 - o The proposed merchant line should not basically coincide with any line that has already been proposed by SOs and authorized by national regulators (if applicable) under the regulated investment track, as long as the regulated line has a reasonable probability of being built in due time. This rule tries to protect the network users from paying more than they would do under the regulated approach.
 - o Because of security of supply reasons, the actual physical operation of merchant network facilities needs to be undertaken by the corresponding SO or SOs.
 - o There will be open access to merchant network facilities under non-discriminatory conditions. Any network charges or collection of congestion rents must be determined by market mechanisms in a fair and transparent manner. Regulatory authorities have the responsibility for ensuring this. Initial long-term contracts may be authorized, in particular for DC links, where the flows can be controlled.
 - o Merchant lines can also be classified as domestic or cross-border investments. Only in the second case should the proposal of construction of the line be assessed by regulators in other areas that are also affected by the investment project. Alternatively, cross-border merchant lines could be assessed by a regional regulatory body - assuming it exists-.

Adequately remunerating new lines probably is the major factor that facilitates transmission investment. This will normally be achieved through an adequately designed cost-of-service scheme, or through a pay-as-bid one if there is an open-auction process. As for merchant lines, private entrepreneurs will only invest in grid reinforcements when the revenues expected from these projects sufficiently exceed

their costs. Finally, network users may associate to promote and pay for the construction of those lines that yield significant benefits for a reduced number of them.

The next subsections briefly discuss the implementation of the general regulatory framework proposed here in the specific cases of the regional markets in Europe, the US and Central America.

2.3.2 The European Electricity Market

2.3.2.1 The situation in the region

National regulatory regimes for the expansion of the transmission grid in the 27 countries participating in the EU Internal Electricity Market vary widely from one country to another. In most countries, an institution, normally the TSO, centrally plans the expansion of the grid subject to the oversight of the regulator or government. However, in other countries like England (and to some extent also in some Nordic countries and the Netherlands) the expansion of the grid is left in the hands of an “active” TSO. Finally, several countries have opted for the merchant model for the construction of lines with some specific features. For instance, Italy planned to have several merchant interconnectors constructed on the north of the country. Therefore, the model adopted in the IEM should be able to accommodate all the different national regulatory regimes.

Contrary to the situation in other regions, there is not a centralized institution in charge of promoting the development of the horizontal network. The Horizontal Network (HN from now on) is the set of lines and nodes that are significantly used by cross-border transactions. The HN is said to be comprised of cross-border lines. Therefore cross-border lines are those that are significantly used by agents in other areas than that where the line is located. Thus, the term “cross-border line” is approximately equivalent to the term “regional line” used in other regional markets like the one in Central America. Analogously, the HN in Europe is the equivalent of the regional grid in Central America. The European Commission carries out periodical studies to identify priority axes or corridors in the regional grid that should be developed to achieve an adequate level of integration of national systems. However, these studies are mainly indicative. In fact, European funds devoted to the implementation of the recommendations made are very limited. The development of national grids (comprised of those lines within a country that are not part of the HN) is the responsibility of national authorities, according to the deeply rooted principle of subsidiarity. The obligation to develop interconnectors or internal lines that may be required for cross-border transactions (cross-border lines) is loosely stated or is absent in national regulations.

2.3.2.2 The proposed regulatory mechanism for the development of the grid

After having briefly discussed the present situation in the region, we move on to describe how the expansion of the grid should be regulated in our view. Our proposal

shares some features with the regulatory mechanisms proposed by some other parties or institutions, especially with the proposal by the Council of European Energy Regulators (CEER), see (CEER 2003). In fact, we believe that, in order for any recommendation or mechanism to be implementable, it must agree with the basic regulatory arrangements that presently have a general consensus.

One should distinguish here between the case of purely domestic facilities (not belonging to the HN) and that of cross-border facilities. Again, regional lines could be identified based on their level of utilization by agents in countries other than that where the line is located.

The case of national transmission systems

As mentioned above, the construction of lines that do not belong to the HN should, in principle, be left in the hands of the competent national authorities. Given the diversity of regulatory regimes that exists, incentives for the construction of new lines depend on the country. In those countries, like Spain or France, where a passive TSO exists, the level of investment in the development of the transmission grid should normally be above the optimal one because of the natural incentives embedded in the regulatory scheme, as indicated above. The regulator should be aware of this natural tendency and check that the investments proposed by the TSO are well justified. However, other factors may hinder the construction of new lines. These include environmental considerations and the allocation of responsibilities in the decisions on the use of the territory.

On the other hand, in those systems where a single for-profit institution decides on the expansion of the grid (the aforementioned “active” TSO) this entity has incentives to invest in new lines below the optimal level in order to increase its profits. Incentives associated to the reliability in the operation of the system and the quality of service should be employed to counter to some extent this natural tendency. Merchant lines may be another option to get new lines built. Authorities should guarantee that these lines are, at least, not detrimental to the system and that they also abide by the common rule of open access.

The development of the regional grid

As mentioned above, no single entity is in charge of developing the HN. Countries in the region have not agreed yet on a process to be followed to identify and approve the construction of cross-border lines. However, most of these countries have shown their opposition to entitling a regional entity to identify the cross-border investments that must be built and promote their construction. Therefore, reinforcements to the HN must be promoted by national TSOs and/or regulators.

New cross-border lines would thus be proposed by at least one national SO or regulator. No upgrade to the HN should be unilaterally decided by one country since, by definition, these upgrades affect more than one country. Proposals to build new

cross-border lines should be jointly assessed by all the countries affected by the construction of the considered line. In practice, preventing a TSO from building a line in its own territory seems unlikely. The following is a pragmatic blueprint of the process that could be followed to get new cross-border lines built:

1. One or several TSOs propose the construction of a line.
2. The future external use of the line by other countries is assessed according to a standard procedure. If it is concluded that the line will be significantly used by external countries it qualifies as an reinforcement to the HN. Otherwise, the line is considered not to belong to the HN and the corresponding country is left to decide on its construction. Here, we are implicitly assuming that the economic benefits resulting from the new line can be estimated from the electrical usage that each agent (country) makes of it.
3. Two possibilities exist when the proposed network reinforcement belongs to the HN:
 - a. The country where the line would be located proposes its construction and agrees to pay its whole cost. Then, the project is given regulatory approval as long as the line is not found to be detrimental to the region as a whole. The set of concerned regulators and/or TSOs (those belonging to the countries affected by the investment) must be in charge of assessing whether the previous condition is fulfilled.
 - b. The reinforcement is proposed by a country other than that where the line would be located, or the country where the line would be located wants others to pay part of the cost of the investment:
 - i. The set of concerned regulators and/or TSOs must determine whether the line is beneficial for the system, i.e. they must determine whether the line passes some kind of regulatory test, as the one described in section 2.4.
 - ii. Assuming the line passes the regulatory test, the cost of this reinforcement is allocated among the different countries based on how much each country uses it. This is automatically done through the inter-TSO payment mechanism. If countries do not agree with the estimated share of the cost of the line that they will have to pay, some previously agreed mechanism must be applied in order to determine how much each country in the region would benefit from the reinforcement. Then, the cost of the line would be allocated based on the projected net aggregated benefit that the agents within each country will obtain from it. Other possible methods to deal with the conflicts

arising from the allocation of the cost of new network reinforcements are discussed in 2.3.5.

Once the process leading to the construction of new lines has been outlined, we discuss some specific aspects of it, namely the role of institutions of a European scope in the development of the regional grid.

The role of European institutions

A sub-committee of the CEER, comprised of representatives from those countries affected by the construction of a new line, must be in charge of determining whether the new cross-border line is beneficial for the whole region or not. This committee could request ETSO to perform any technical analysis related to the assessment of this regional investment project. Additionally, whenever countries affected by the construction of the cross-border line do not agree to pay the fraction of the cost of this line that they have been allocated using the inter-TSO payment mechanism, this committee would apply a predetermined procedure to assign the cost of the cross-border line according to how much each country will benefit from it. The composition of the committee appointed to examine different reinforcements to the HN would probably be different.

Allocation of the cost of regional reinforcements

Transmission tariffs should be aimed at recovering the total cost of the network, including the cost of any new lines once they enter into operation. Besides, transmission tariffs should send sound economic signals, whenever it is convenient and possible. Instead of defining a system of pan-European transmission tariffs, European countries have agreed to establish a system of compensations whereby countries are compensated for the use that external agents make of their grids. The net compensation a country has to pay is computed as the difference between the total compensations it is entitled to and the total payment it owes other countries (see chapter 3 of this document). Any inter-TSO compensation mechanism automatically assigns the cost of new network reinforcements among the different TSOs. Afterwards, authorities in each country should modify the local transmission tariffs according to the net compensation received by this country.

Provided the inter-TSO compensation (ITC) mechanism provides an accurate enough estimation of the use that agents within each country are making of a certain line, and assuming that economic benefits (and other type of benefits like the reliability ones) can be reasonably estimated using electrical usage, countries should not contest the allocation of the cost of new reinforcements to the HN resulting from the application of the ITC method. In spite of this, given that the electrical usage of a line by agents or countries is very difficult to assess, and electrical usage is only a proxy to economic benefits (which are the real driver behind network investments), a country might feel that the results of the ITC mechanism do not adequately reflect the benefits that this country obtains from a reinforcement. In this case, a predetermined, non-contestable,

procedure would have to be used to compute how much each country will benefit from this line and, therefore, the fraction of the cost of the line that this country will have to pay. Other methods exist to resolve conflicts of this type. They are discussed in section 2.3.5.

The role of merchant lines

Due to the meshed nature of national grids, private entrepreneurs may not find it profitable to invest in new lines that will not be part of the HN. Thus, most reinforcements to national grids will have to take place as regulated investments. On the other hand, interconnections between countries are often too weak to support the level of regional trade requested by market agents. Consequently, congestion occurs. Thus, private entrepreneurs may probably promote the construction of some cross-border lines. National and European regulatory authorities should assess the merchant projects to guarantee that the merchant lines whose construction is approved are in the best interest of the system. Specifically, merchant projects should comply with the following requirements:

- Market agents are provided with open-access to the facility. If congestion occurs, the capacity on the line is jointly assigned with that of others using an efficient coordinated market-based mechanism. This requirement is essential in the case of AC lines. Provided a merchant DC line complies with certain criteria (the ones described in the exemption procedure in article 7 of the electricity regulation could be used in this regard, see (EP 2003)), its owner could negotiate with the competent regulatory authorities (the CEER for cross-border lines) the exemption from the requirement to provide open access.
- As it has been explained above, the line should not be detrimental to the system.
- This merchant project does not interfere with any regulated investment already planned.

Despite the stringent conditions set out here for approving the construction of merchant projects, some merchant lines have already been built in Europe. These are mainly DC submarine cables between countries. For instance, the 'SwePol' link lies between the Swedish and Polish systems and the 'BalticCable' connects the Swedish and German systems. Local authorities in Italy have also proposed the construction of some merchant lines in the north of the country.

2.3.3 The regional markets in the US

2.3.3.1 The situation in the region

As mentioned in Chapter 1, when comparing the regional markets in the US to the IEM in Europe, one must be aware that the latter is, more or less, of the same size as the

whole US market. Some of the most important regional markets that have been created in the US are approximately of the size of a large national market in Europe. Therefore, the energy trade taking place between different countries in Europe is analogous to that between different regions in the US. The creation of a truly integrated national market in the US is probably even further from being a reality than the IEM in Europe.

To some extent, the situation in the US resembles that in Europe. Thus, the transmission network within many states was developed for the local generation to serve the local load. As a consequence of this, the transmission grid is prepared to accommodate local transactions but not a great deal of regional trade among areas traditionally served by different vertically integrated utilities. Additionally, regulation differs much from one state to another. Some areas have achieved a high level of deregulation in their markets whereas in others vertically integrated utilities are still responsible for all the activities from generation to retailing.

Similarities between the European and American cases end here. Contrary to what happens in Europe, the design of a common market model to be implemented in the different regions of the country has been in the agenda for a long time now. Relevant federal regulation exists on this subject issued by the Federal Energy Regulatory Commission (FERC). The FERC is a regulatory body with authority over inter-state trade. The EU Commission may be regarded as the FERC counterpart in Europe since it issues directives governing the functioning of the IEM that must be transposed into national laws. However, responsibilities of both institutions differ significantly. Thus, the FERC is in charge of setting regional transmission tariffs while the EC is not. Given the will of European countries to preserve their independence as far as possible, parties involved in the development of the IEM are unlikely to create a European regulatory body in the short to medium term.

Going back to the US, we can identify other institutions of a national scope. Thus, the NERC (National Electricity Reliability Council) develops standards aimed at achieving similar degrees of development of the system, and the utilization of common system operational procedures, in different parts of the country. ETSO performs a similar function within the IEM.

Traditionally, grid investments in the US have been planned by local utilities or SOs and licensed by State authorities. However, RTOs are responsible for planning the construction of those grid reinforcements that are necessary for the satisfactory functioning of regional markets. Additionally, as a result of disruptions of supply experienced recently by large areas of the country, there is much controversy over whether state or federal authorities should regulate the construction of large multi-state lines. Local authorities in each state are in charge of deciding over the construction of new lines internal to the state. However, the transmission service within many control areas is vertically integrated with the remaining activities including retailing, which is the sole responsibility of states. Hence, even if it were finally agreed to give the FERC authority to decide on the construction of transmission

facilities relevant for whole-sale markets, state authorities may have a significant influence in the development of regional markets in general and inter-state network investments in particular.

Most investments in new transmission facilities have taken place as regulated projects. Some markets, like PJM in the north east of the country, had initially decided to leave grid expansion mainly in the hands of private promoters. However, investment in the form of merchant projects has been very modest to date. Very few merchant lines have been constructed, see (Joskow 2005) for an example. Thus, FERC encouraged PJM to carry out periodical analysis to identify congestion in the grid that is not covered by existing capacity rights. Then, PJM must perform a cost/benefit analysis for different investment projects that are aimed at solving (easing) this congestion. Those projects that are found to be beneficial for the system are included in a list of potential grid reinforcements and private promoters are given some time to undertake them. If the proposed reinforcements are not built by the private initiative they are included in the centralized expansion plan of the grid.

2.3.3.2 Regulatory mechanism proposed for the development of the grid

After having discussed the situation in the US, we now propose an approach to regulate the expansion of the grid. We shall discuss separately the case of transmission lines that are necessary for the development of regional markets and those that allow the local generation to serve the local load. In any case, many features should be common to the approach implemented for lines that are relevant to the trade in each region and those that are not. We take into account the division of responsibilities that exist between state and federal authorities.

The approach proposed here aims to reproduce what would be the outcome of the reference approach described at the beginning of section 2.3. Thus, ideally, the approach proposed here and those corresponding to other regional markets would coincide. Differences between this and the reference regulation are due to the need to comply with the basic provisions in place in the US.

The development of local lines

Given the meshed nature of transmission grids within each control area, it is highly unlikely that the reinforcement that is required by local grids will be undertaken by private entrepreneurs. Similarly, benefits yielded by local reinforcements will probably be too much dispersed for network users to create associations to promote the construction of a line. Besides, there is no tradition in the US of the latter type of network investments. Consequently, most local grid reinforcements must take place as regulated projects. The construction of these lines should be promoted by the System Operator in the corresponding control area. This SO should consider the proposals by market agents and local authorities. In the end, the approval to build a new line must correspond to the state authorities.

In some control areas, the operation of the system is vertically integrated with the activities of generation, distribution and retailing within a local utility. In this case the SO has incentives to discriminate against local generation and power marketers that do not belong to this utility. Unbundling activities that are natural monopolies (transmission and distribution) from those where competition can be introduced (generation and retailing) may be an option. However, these measures may face opposition from local authorities.

In other cases the transmission service within the corresponding area has been unbundled. Finally, there are some cases where the considered control area has joined a regional market (like PJM or ERCOT), where a Regional System Operator (or an RTO) is in charge of the system and market operation. In these cases, local SOs should disappear and their functions should be taken over by the regional operator. If state regulators do not accept the disappearance of local SOs, these entities should be in charge of very specific local functions (local upgrades and maintenance of lines) and should not interfere with the efficient functioning of the market in the region.

As discussed in the description of the optimal regulatory model for the expansion of the grid in the reference case, which is provided at the beginning of section 2.3, some performance-based incentives for local SOs could be created so that they carry out their functions more efficiently.

The development of lines that are relevant to the trade in the region

From now on, we will use the term “regional lines” or “regional grid” to refer to the lines that are necessary to achieve integrated regional markets. A clear, objective, procedure must be used to determine whether a projected line is a regional reinforcement or not. The decision criterion may be related to the level of expected electrical usage of the reinforcement by agents external to the area where the line is going to be built. By definition, those lines that join two different states should be considered regional upgrades.

Both the private and the public initiative have a role to play in the construction of regional lines. Despite this, regulated grid augmentations should be given priority over other options. The next paragraphs discuss the functions to be performed by the different parties that may participate in the development of the regional grid.

Central planning is crucial for the development of the grid. As explained above, merchant projects and lines promoted by associations of network users will certainly represent a minor part of the grid upgrades required by regional markets. Here, we must make a distinction between those areas where a regional system operator exists and those where there is no one in place. In the former case, this regional entity should be in charge of promoting the construction of regional reinforcements. If no regional system operator exists, local SOs should be responsible for proposing the construction of those lines that they consider to be necessary for the local generation and load to engage in power transactions with agents from other areas. Obviously, this is unlikely

to occur if the operation of the system in an area is carried out by a vertically integrated utility. Thus, unbundling the transmission service is of utmost importance. Local and regional SOs should also promote the construction of those lines that allow power exchanges between different regions of the country to take place. Power exchanges among different regions may occur because of reliability reasons but also for economic ones. When developing an expansion plan or identifying individual reinforcements, local and regional SOs must take into account the proposals by other System Operators and authorities as well as those by network users. They should either incorporate these proposals into their expansion plans or justify why they have rejected them.

Regarding the approval of regional investment projects, we must distinguish between those lines that cross at least one inter-state border and those that are internal to a state. The decision on whether to approve the construction of lines of the former type must be made by federal authorities, i.e. by the FERC. Despite the fact that they are necessary for inter-state commerce, regional intra-state lines will certainly need the approval of state authorities for their construction. Thus, the remuneration mechanism employed should provide incentives for authorities in a state to approve the construction of those reinforcements internal to the state that are required for the development of the market in the region, even when these reinforcements are not of special interest for the state. A sensible regulatory test should be used to ascertain the advisability of installing a new line.

The cost of the new regional reinforcements must be paid by market agents in the different areas of the region (or the different regions of the country) according to the results of a cost-benefit analysis. This analysis should be carried out by the regional system operator and the resulting transmission rates should be approved by FERC. As explained for the European system, if electrical usage is accepted as a proxy to the economic benefits produced by a line, and transmission tariffs in the different areas within a region (and the different interconnected regions) are computed using a common usage-based method, this method will tell us how to allocate the cost of the whole regional grid, including any new lines, among grid users. Therefore, this same method will allocate the cost of regional reinforcements among the different areas and regions. Authorities in a state may oppose the construction of a regional line that is internal to the state if they feel the share they would have to pay of the cost of this reinforcement is not in line with the benefit they would obtain from it. Forcing authorities in a state to approve the construction of a regional line located within the state will probably not be possible. Therefore, a system of side payments between states may help get this line built. Once a line is built, FERC has authority to approve or not regional transmission tariffs aimed at recovering the cost of this line.

The construction of an approved reinforcement could be assigned through a competitive bidding process whose winner would be paid the amount he requested when making the offer. This mechanism is thought to be superior to the computation of standard unit construction costs based on the utilization of benchmarks.

As mentioned at the beginning of this point, there are alternatives to the construction of regulated lines by RTOs or SOs. Private entrepreneurs may wish to participate in the process of development of the grid through merchant projects. Additionally, associations of network users may promote and/or construct certain reinforcements that would be beneficial for them. Now, we discuss these options.

Due to the fact that transmission grids in the country were mainly built to allow the local generation serve the local load, interconnections among control areas are not developed enough to accommodate all the regional transactions that market agents want to engage in. In other words, interconnection capacity⁴ may be scarce and systematic congestion may occur. Thus, if energy prices in the different areas exhibit locational differentiation, private promoters may find it profitable to build a line and appropriate the corresponding congestion rents. However, merchant projects should be subject to strict regulatory oversight given the well-known tendency of their promoters to invest below the optimal level in order to maximize the resulting congestion rents.

Building a line may be profitable for a merchant investor in those very specific cases where systematic congestion exists which will persist even after the construction of new transmission capacity. Merchant investors should construct those lines that they consider would be profitable but are perceived as too risky by the central planner. If authorities want to leave the private initiative a more predominant role to play that is compatible with an adequate development of the system grid, they may give private promoters the opportunity to carry out those projects previously identified by SOs and RTOs. However, this would come at a cost since lines owned by private promoters would be more expensive for consumers than regulated lines (the congestion costs corresponding to these lines would be appropriated by the private investor instead of the system).

Associations of network users, or some entity acting on their behalf, may be interested in promoting the construction of some lines, and paying for them, even when they run the risk of not recovering their investment. For example, a local retailer representing the interest of its consumers could be interested in building a line to be able to import cheaper power from another region. In any case, experience suggests that non-regulated reinforcements will fall far short of the total amount of capacity needed (Joskow 2005). In other words, private initiative cannot be relied upon to promote the optimal development of the grid, as we have repeatedly stated.

⁴ With the term interconnection capacity we refer not only to the lines connecting two different control areas but also to those that are significantly used by power exchanges taking place between areas in a region or between regions.

2.3.4 The Central-American regional market

2.3.4.1 The situation in the region

The regional electricity market in Central America (MER) is much smaller than the IEM in Europe or the regional markets created in the US. The major objective of MER is the development of large and efficient generation facilities that can serve load in other countries than that where the generation plant is located. This is due to the fact that the total demand within each country is small and economies of scale exist in the construction of power plants. Thus, future power exchanges among countries are expected to be comparatively much larger than those in the two regional systems previously studied, though they are now very small because of the lack of enough interconnection capacity and the existence of transient rules for the MER that prevent it from exploiting its full potential.

Both the national transmission systems and the interconnections between them are not sufficiently developed. The topology of the grid in the region is radial, which conditions the features of the most adequate regulatory framework. A procedure is already in place to identify the regional grid, named RTR ('Red de Transporte Regional' in Spanish). The definition of the RTR is very relevant since regional and national lines are subject to different regulatory schemes, see (EOR 2005). The process followed to define the regional grid follows:

1. Interconnections among national systems are included in the RTR.
2. National systems choose a collection of nodes where agents submitting their bids to the regional market are supposed to be located. These nodes are included in the RTR as well.
3. The EOR determines the smallest interconnected network that includes the previous nodes and lines. All the transmission facilities in this network are also deemed to be part of the RTR.
4. The RTR also includes all the lines where regional transactions frequently cause the flow to change by more than 10% of the total volume of regional trade within the corresponding country. This must happen at least in 10% of the scenarios.
5. Finally, national system operators and the EOR may jointly decide to incorporate any other network component into the RTR.

The definition of the RTR is updated every three years or whenever the EOR requests it. Every offer to buy or sell energy in the regional market is assumed to be located in the closest node of the regional grid. The EOR must be able to oversee and control the system variables in all the nodes of the RTR. Congestion rents are defined between the

nodes of the RTR. Therefore, congestion rights are defined between these nodes as well.

The EOR is not the owner of the regional transmission grid. Part of the RTR is owned by the EPR (for 'Entidad Propietaria de la Red' in Spanish). The EPR is a regional institution that owns the SIEPAC line, which connects all the countries in the region (still under construction). In addition, the EPR may own other transmission facilities. National transmission systems also own part of the lines of the RTR. Lastly, private investors (mainly grid users) build and own lines in order to access other areas of the region and get the corresponding transmission rights. These are the so-called investments at risk. Thus, the operation and planning of the system and the ownership of transmission assets are independent from one another. This organizational model of transmission activity has its advantages and shortcomings. It reduces the incentives for the SO to over-invest in the development of the grid because, contrary to what happens with TSOs, the revenues of the SO are not related to the size of the transmission grid. In any case, incentives to over-invest are not completely eliminated. A SO will always be interested in operating a highly developed grid since security and reliability problems will be less frequent and significant. This organizational model also produces some inefficiencies resulting from the lack of coordination between the SO and the transmission owners.

From what has just been mentioned, one can easily conclude that reinforcements to the RTR may be of any of three types:

- Regulated investments included in the expansion plans developed by the EOR, which has the knowledge and expertise to propose new investments. In order to develop these plans, the EOR must take into account the proposals by market agents and regional institutions like the CRIE. Regulated lines belonging to the RTR earn the investment cost plus a reasonable return on it.
- Investments that belong to the RTR but are promoted by national TSOs. These lines are the result of the centralized planning of the expansion of the grid taking place within each country. According to the regional regulation, see (EOR 2005), the regional regulated revenue of these lines, which is collected from the application of nodal energy prices and regional transmission charges, is much smaller than the total cost of constructing, operating and maintaining the corresponding line and is related to the reliability provided by the line (it is the so-called *Valor de Indisponibilidad* in Spanish).
- Because of the special characteristics of this market, where the transmission grid is very weak and national markets are small and unable to obtain enough funds for the construction of all the grid reinforcements that are required, regulated investments in the grid must coexist with a significant volume of investments at risk. As explained above, these investments are promoted by market agents in order to access other areas of the region. Investments at risk

allow market agents to trade their energy in other areas. Additionally, they obtain the transmission rights created by the new line, which allows them to sign firm contracts with agents in these areas and earn the capacity payments that are in place there.

Thus, the development of the regional grid is the result both of regulated expansions and of those promoted and paid by market agents (mainly generators). According to the regional regulation that exists, see (EOR 2005), investments at risk must be given priority over regulated investments (proposed by the EOR). Only when the benefits produced by an efficient investment are very much dispersed, and no coalition of network users forms in order to construct the line, must this line be built as a regulated project. The CRIE must study the proposals by the EOR and market agents and give its approval to them if it considers that the proposed upgrades are justified on economic or reliability grounds. Criteria used to approve a regulated reinforcement are not coincident with those used for investments at risk, which are financed by market agents. The development of national transmission grids is left in the hands of the competent national authorities.

The method applied to allocate the regulated cost of new transmission facilities is the same as that applied for the already existing ones. This is a combination of two usage-based methods: the Dominant Flow Method (DFM) and the Average Participations method (AP). Therefore, contributions to the regulated cost of reinforcements are computed based on the electrical usage that the agents within each country make of these lines. DFM divides the flow over each line in two parts: one associated with the national transactions taking place within the country where the line is located and one corresponding to regional transactions. Then, both components of the line flow are tracked separately using the same algorithm that is applied in the AP method. When tracking each one of these components, only national or regional power injections and withdrawals are considered, depending on the case. Transmission charges obtained using DMF are transaction based; i.e. they depend up to a point on the commercial transactions taking place in the region. In addition, the fraction of the cost of a new line belonging to the RTR that is recovered through the application of regional transmission tariffs depends on whether this reinforcement was planned by the EOR or by the national transmission systems.

2.3.4.2 Regulatory mechanism proposed for the development of the grid

The set of rules presently applied to distinguish between the regional grid and the rest of existing transmission facilities has already been explained. We consider this to be a reasonable mechanism. Now, we move on to propose a regulatory framework for the development of both national and the regional transmission networks. Quite naturally, this framework is based on the reference regulatory framework outlined at the beginning of section 2.3. Differences between this and the reference scheme are due to the necessity to account for the specific characteristics and arrangements made in this region.

The construction of lines that do not belong to the RTR

The RTR that is defined following the process outlined above includes all the facilities that are essential to allow power transfers in the region. Therefore, the development of national transmission systems, which include all the transmission equipment but that belonging to the RTR, can remain in the hand of local authorities. Each country must be responsible of the expansion of the local grid in order to guarantee, at least, that the local load can be served either by local generation or by external agents. The procedure used within each country to carry out the local dispatch must determine the identity and volume of the national grid reinforcements that are considered to be needed.

In those countries where economic signals for investment in new transmission lines do not exist (energy prices in all the nodes of the national grid are the same), investments at risk or merchant projects for the development of national grids will probably be very rare. Thus, reinforcements to national transmission systems in these countries will have to be promoted by national SOs. The reinforcements proposed by SOs may need to get the approval of the corresponding national regulator or government depending on the paradigm adopted (passive versus active TSOs).

In those other countries where locational differences exist among the price paid for energy in different nodes of the national grid, investments in new “domestic” lines carried out by private investors may be more frequent. In any case, establishing efficient pricing mechanisms in some countries would be very difficult at the moment, given that, not only their grids, but also their national regulatory frameworks, are still very poorly developed.

The development of the regional grid

As the following paragraphs will explain, there is room both for the construction of regulated reinforcements to the regional grid and for the participation of the private initiative in the development of this grid. Due to the fact that national markets in the region are very small, power exchanges among countries will probably be very significant in the near future. In order for large power exchanges between countries to take place, strong enough interconnections must be built between countries in the region. Thus, the public electricity companies of the countries in the region, the Spanish company Endesa and the Colombian ISA joined forces to build a 230KV line running from Guatemala to Panama across the 6 countries in the region. This line has been named the SIEPAC line and the project to build the line has been named the SIEPAC project (for ‘Sistema de Interconexión Eléctrica entre los Paises de América Central’ in Spanish).

As we have already suggested, this regional market has some specific features which make it different from the two other ones that have been studied:

- The regional grid is very much underdeveloped. Interconnections among systems are very weak.

- It is expected that a number of power plants will be built to trade their energy regionally. These plants need access to other countries than that where they are located.
- Ownership of physical rights is necessary to sign firm contracts with agents from other countries, which allow generators to earn the corresponding capacity payment.

Consequently, market agents are interested in promoting the construction of their own transmission facilities, which would give them access to other systems and allow them to earn the corresponding capacity payments. Contrary to what happens in other markets, a significant number of investments at risk will take place. Investments by market agents are necessary because regulated network reinforcements proposed by the EOR will be limited in number and they may suffer delays. This is due to the difficulty of authorities in the region to obtain the funds necessary to undertake all the grid investment projects that are necessary. As a result of this, regulated augmentations are unlikely to result in a sufficient development of the regional grid. They need to be complemented by private investments. The next paragraphs discuss both types of projects.

Regulated projects for the construction of a line, whenever they exist, should take precedence over investments at risk since the private initiative cannot be believed to promote the optimal reinforcement from the point of view of the system (the size of the new line will probably be below the optimal level in order not to reduce congestion rents too much). Additionally, if the beneficiaries of a possible reinforcement are too much dispersed no coalition will form to build it. Regulated reinforcements must be promoted by the EOR. This institution must include them in the indicative expansion plans for the RTR it submits periodically to the CRIE. Alternatively, the EOR may propose an specific grid upgrade that is not part of a plan when this entity considers that the prompt construction of this line will render significant benefits. When developing an expansion plan, or promoting a specific upgrade, the EOR must take into consideration the proposals by market agents and other institutions in the region. Afterwards, the CRIE must decide on the construction of the reinforcements proposed by the EOR. Whenever modifications are made to the plan submitted by the EOR, this entity should be able to propose alternative investments to the ones rejected by the CRIE. Both the EOR and the CRIE must have in mind the interest of the whole region (generators and consumers). Designing a satisfactory regulatory test is necessary both for the EOR to identify those reinforcements that are most beneficial for the system and for the CRIE to determine whether the construction of the proposed new facilities is advisable or not.

National authorities and private parties could not contest the decision by the CRIE to construct new regulated lines since it is only the CRIE which is in charge of approving or rejecting the construction of regional grid upgrades. The allocation of the cost of the new facilities must be carried out according to the same method used to allocate the

cost of the already existing transmission facilities that belong to the regional grid. This contradicts the regulation that presently is in place in the region. The method applied to allocate the regulated cost of new and existing lines should not be the MFD method since it is transaction based. What is more, every regulated transmission reinforcement should recover the whole cost of the investment regardless of the institution that promoted the construction of this line.

The construction of regional regulated lines should be assigned through a concurrent bidding process where the winner is paid its bid. Thus, authorities would avoid having to compute a benchmark for the cost of construction of the different kinds of transmission facilities in the region. However, the bidding process would only be efficient if enough potential contractors compete to win the contest.

As explained above, there are some cases where relying on the private initiative to build some lines may be beneficial. First of all, new generators willing to sell their power to consumers in different countries of the region may be interested in building their connection lines as part of the overall investment project. Given the electrical size of national systems, most projects to build large generators only make sense when these generators plan to serve consumers on a regional basis. Since these projects have a regional scope the associated connection line may be considered a regional reinforcement. Waiting for the EOR to include the connection line in its regional grid investment plan may cause further unnecessary delays. One must be aware that these are not connection lines properly speaking, since their construction is caused by the location of both the generator and the load it serves. These lines might be built as investments at risk undertaken by the corresponding generator. Both the load served by the generator and other generators which end up using this line (probably generators that are installed in the same area) should appropriately compensate the promoter of the line.

But generators of regional scope will probably be interested in building (and paying for) other lines. The regulation on the organization of the regional market that is already in place considers the possibility of generators and consumers from anywhere in the region signing long-term firm supply contracts. In addition to earning the revenues corresponding to the sale of the amount of energy contracted, generators signing firm contracts have the right to get the capacity payment corresponding to the country 'C' where they are selling their output. However, in order for a generator to be allowed to sign a firm contract with a consumer in country 'C' it will be requested to physically reserve the transmission capacity required to access this area (country). Thus, this generator would have to buy the corresponding physical transmission capacity rights. If these rights are not available, the generator will have to construct some transmission lines to obtain them. This represents an incentive for generators to invest in new transmission lines. This incentive will be significant if the capacity payments are high enough.

Finally, merchant projects may be attractive for private promoters. Due to the topology of the network in the region, the current level of development of this grid and the fact that locationally differentiated energy prices are computed, congestion rents in the regional dispatch may be significant. Thus, private promoters may be interested in building certain lines and appropriating the corresponding congestion rents. However, merchant investors could only build those regional reinforcements that are not included in the centralized expansion plan of the grid submitted by the EOR. Additionally, these projects should be subject to certain conditions concerning the size of the investment. In general, the characteristics of the investment project may be jointly agreed by the private promoter and the CRIE. The CRIE should have available the tools necessary to negotiate with private entrepreneurs the construction of new lines. This includes being able to use a regulatory test aimed at assessing the convenience of authorizing merchant lines. This test should be different to that used for regulated investments.

2.3.5 Possible mechanisms to deal with conflicts among countries or states that are related to the construction of lines that belong to the regional grid

Several alternatives exist in order to settle those conflicts that may arise among countries or states that are associated with the construction of a new line and the allocation of its cost. One first possibility is implementing a mechanism that allocates the cost of the proposed reinforcement based on the benefits that each country or state obtains from it. Countries involved in this settling process would not be able to contest the results of the application of this mechanism. In any case, the construction of the line would be subject to the approval of the corresponding national regulators. In order for this mechanism to be applied, countries or states would have to commit themselves to abiding by the results of the method employed. Both in the IEM and in the US, countries or states are unlikely to agree to build and pay a major part of the cost of a line located within their territory if they do not feel that they will benefit from this reinforcement.

Another alternative involves establishing a negotiation process among the countries or states affected by the proposed reinforcement. Bilateral or multilateral compensations may be agreed in order to get the line built. Provided compensations are large enough, no area would oppose the construction of this reinforcement. Since benefits yielded by a justified investment would suffice to pay the compensations required to guarantee that every country is benefited by the construction of the line, these reinforcements would be approved. In other words, a system of side-payments like this should guarantee that any line beneficial for the region is constructed as long as property rights are well defined and the benefits produced by new lines are not spread over a large number of different countries, see (Coase 1990). However, this alternative may lead to countries taking advantage of their negotiating position to appropriate the

benefits that others would obtain from the construction of a line. One must bear in mind that ownership of transmission capacity, or the ability to influence the development of the transmission grid, may confer significant market power.

2.4 The regulatory test

The “regulatory test” is the set of rules that must be applied to determine whether the construction of a particular network reinforcement, or a set of them, is justified. This set of rules may vary depending on the identity of the party that is applying the test and on the nature of the investment being assessed. Therefore, different kinds of regulatory tests may be defined. Each one of them will be used in one of the following situations:

1. A passive System Operator must employ it to identify the network upgrades to be proposed for their construction. An active TSO may use the same regulatory test to show the regulator that the investments it has proposed are justified and should be included in the remuneration of the transmission activity for the next control period.
2. The regulator must decide on the construction of a line that is proposed by another party. Two possibilities exist here:
 - a. The investment proposal may come from a SO that is centrally planning the expansion of the grid. In this case, it corresponds to a regulated investment. Both the investments proposed by an active or a passive TSO are included here.
 - b. This proposal may come from a private entrepreneur or an association of network users willing to finance the upgrade. In this case it corresponds to a merchant line or an investment at risk.

The regulatory test to be applied by a SO must identify the most efficient investment among a set of possible ones. Strictly speaking, in order to identify the next ‘n’ reinforcements to be undertaken, the test should allow one to compute the optimal expansion plan of the grid where only a set of different grid upgrades are allowed. However, in practice, the SO could evaluate each one of the investment options separately making use of the expert knowledge he has got over time as well as any other tool he has available. The regulatory test applied should evaluate the opportunity cost corresponding to the alternatives to the reinforcement under consideration, see (ACCC 2004). In other words, when assessing an investment, we must take into account the cost associated with the fact that, if this upgrade is carried out, others will not. This is equivalent to comparing the increase in the social benefit associated with all the expansion plans that we have considered, and choosing the expansion plan that results in the highest increase.

On the other hand, a regulator should only check that the expected net social benefit of market agents, when the expansion plan proposed by the SO is followed, is higher than the net social benefit corresponding to a reference expansion plan. The regulator might as well ask the SO, or any other institution, to evaluate a set of alternatives to the proposed investment. In the case of investments proposed by private entrepreneurs, the regulator should only check that the considered reinforcement is not detrimental to the system.

Under the traditional cost-based regulation the rule to apply in order to compute the optimal expansion of the grid is as follows: “One should build new transmission network assets only while the additional network investment cost is still smaller than the additional saving in system operation costs resulting from further investing in new lines”. Obviously, -as it has just been mentioned and expressed in simplified terms-, when alternative investment options exist, the network planner should choose the combination of investments that maximizes the difference between the saving in the operation cost and the cost of building, operating and maintaining the new reinforcements.

There is not wide agreement yet on which should be the objective of transmission expansion in a context of competition. Specifically, controversy exists over whether the estimated negative benefits (losses) that might be incurred by some agents as a result of a network augmentation should be taken into account jointly with the positive ones when assessing the investment project, see section 2.4.3. If we decide to consider both the negative and positive benefits brought about by a new line, the optimal investment rule in the new regulatory context would be: “Invest while the additional network investment cost is smaller than the aggregated benefits obtained by all network users (i.e. all generators and all consumers) as a result of carrying out the investment. Network charges resulting from the investment must be excluded when computing the benefit of market agents”. Annex 1 proves that this definition of the optimal investment rule is consistent with the one applied in a traditional cost-of-service context. The net benefit of a generator, because of a new transmission investment, is equal to the increment in its margin of market revenues over operation costs. The net benefit of a consumer, because of a new transmission investment, is equal to the increment in the margin of its utility function over its cost of purchasing electricity.⁵

From the discussion in the previous paragraphs, one can easily conclude that the regulatory test is very difficult to apply in practice in strict terms. Different kinds of approximations and simplifications are used throughout the world by SOs and

⁵ This is how the Australian regulatory authority has expressed the regulatory test: “A new interconnector or transmission system augmentation satisfies this test if it maximizes the net present value of the market benefit, having regard to a number of alternative projects, timings and market development scenarios. Market benefit here means the total net benefits to all those who produce, distribute and consume electricity in the electricity market”.

regulators⁶. Note that, in order for the planner to strictly assess whether a new line is beneficial or not -or the best option to meet the new needs in the system-, one should try to compute, at least approximately, the true economic benefits that are produced by the line. Additionally, a huge amount of information is needed to be able to obtain any meaningful results. Section 2.4.2 discusses the cost factors involved in the computation of the net aggregated benefit brought about by a reinforcement or a collection of them.

Determining the net aggregated benefit obtained by network users as a result of a network augmentation (or the saving in operation costs that this augmentation produces) involves computing the system dispatch as accurately as possible. Provided some non-negligible degree of market power exists, this would involve computing the strategic market equilibrium over the period of time studied. Here, we are assuming that the market benefits produced by a network augmentation under oligopoly will not be the same than those existing under perfect competition. However, if we assume that generators bidding strategically, instead of according to their variable costs, do not significantly affect the total production costs in the system, we may estimate the total net aggregated benefit of generators and loads from the competitive outcome of the market. In other words, if strategic generators did not alter the dispatch of generation in the system, the difference between the competitive and the strategic market equilibriums would only be a wealth transfer from the load to generators, i.e. the social benefit would remain the same. Computing the increase in the social benefit corresponding to the competitive case instead of that in the real case may or may not be accurate enough depending on the system considered.

Theoretically speaking, the optimal expansion of generation and transmission over a period of time 't' should be jointly determined given a previously made estimation of the future evolution of demand in the system. In practice, and in order to make the process simpler, SOs usually estimate first the evolution of generation based on their knowledge and the available information. Then, they optimize the expansion of the grid taken the evolution of the generation as given.

As we have just mentioned, the single measure used to assess an expansion plan should be the expected net aggregated benefit that market agents will obtain over a period of time 't' if this grid expansion plan is followed. The different criteria related to the cost components that must be taken into account to compute the expected net aggregated benefit of market agents are explained in section 2.4.2. The period of time 't' considered may be the economic life of a typical reinforcement to the grid (40 years is a reasonable option).

⁶ For instance, "benefits" are frequently replaced by some measure of "electric usage" and very often the traditional version of the rule is used even in the context of competitive wholesale markets.

The computation of the future economic benefits associated with an investment plan is subject to a high degree of uncertainty. In order to deal with uncertainty, several scenarios may be considered corresponding to different levels of future energy prices, demand, etc. In this regard we could use Montecarlo simulation techniques. The expected net aggregated social benefit for an investment plan would be obtained as the weighted average of the social benefit computed for the different scenarios. Weights could be made proportional to the estimated probability of occurrence of the corresponding scenario.

The next paragraphs discuss the process to be followed in each one of the situations described above (1, 2.A, 2.B).

2.4.1 Description of the regulatory test to be applied in different situations

Next we shall describe the process to be followed to assess future grid reinforcements. As explained earlier, this process depends both on the identity of the institution applying the regulatory test and on the type of institution proposing the construction of new lines. Here we have assumed that the benefits obtained by all network users (and not only by those positively benefited by the construction of the line) should be considered. Again, section 2.4.3 deals with this issue in more detail.

2.4.1.1 Proposal of a (set of) reinforcement(-s) by a SO

In order to decide which reinforcements to propose, the System Operator must evaluate, for a set of grid expansion plans, the expected net aggregated benefits to be obtained by market agents over period 't'. Afterwards, it will propose the expansion plan for which the expected net social benefit is highest. In order to compute an expansion plan of the transmission grid the SO must consider both the future expansion of generation and demand in the system over the period of time 't'. As mentioned earlier, different scenarios corresponding to different possible evolutions of generation, demand and energy prices may be taken into account.

Obviously, not all the reinforcements included in the expansion plan will be undertaken if this plan is approved by the regulator. The SO will update periodically the proposed expansion plan. Therefore, once a plan is approved, only the reinforcements scheduled to be built between the present time and the time the SO submits the following plan will actually be carried out.

The last point of this section (subsection 2.4.5) provides a classification of the algorithms that are available to compute an expansion plan for the grid and briefly characterizes each of them. All the investments in generation and transmission that are already scheduled should be an input when computing a transmission expansion plan. The set of options studied may include the possibility of not installing any new line

over the period of time 't'. This may depend on the length of the considered period of time.

2.4.1.2 Assessment by the regulator of a proposal by another party

Here we must make a distinction between the case where the regulator assesses a proposal by the SO (a regulated reinforcement) and that where it decides on the construction of a reinforcement by a private entrepreneur or a coalition of network users.

Assessment of regulated investments

Here we assume that the regulator is not in the position to assess investment proposals itself. Therefore, the regulator must rely on the assessment provided by the SO or on that provided by any other institution that the regulator may choose to advise on this issue. When evaluating a proposal by the SO, the regulator will request this institution to submit all the information needed to show that the proposed reinforcements are beneficial for the system. An investment plan for a period of time 't' is beneficial for the system when the expected net aggregated benefit obtained by market agents over period 't' when this plan is implemented is higher than that obtained for a reference expansion plan.

Additionally, the regulator could ask the SO to evaluate any investment plan submitted by any other party and compare it to the one the SO has proposed. In order to assess a single reinforcement proposed by a third party, the SO must first determine the set of grid upgrades that should be undertaken during period 't' after the proposed one. In other words, the SO must compute an expansion plan for the grid that includes the reinforcement previously proposed by a third party as the first one to be undertaken, see 2.4.5 for a description of the different methods that exist to compute an expansion plan for the grid. Then, the SO must compute the expected net social benefit obtained, over period 't', for this plan. Finally, the regulator must compare the expected net social benefit obtained for the expansion plan just computed to the expected net social benefit for the best expansion plan found until that moment. Obviously, the regulator will decide to implement the expansion plan that results in the highest social benefit.

When evaluating a regulated reinforcement, computing the net social benefit obtained by market agents over a period of time 't' when only a particular grid reinforcement is built does not make sense, since this is not realistic. All the grid reinforcements that should be carried out over the period of time 't', once the considered one is built, must also be taken into account.

As mentioned above, an institution different from the SO may be commissioned this task or that of proposing alternative expansion plans.

The main difference between the investments resulting from the private initiative and the regulated ones lies in the fact that promoters of the former are willing to pay the cost of the proposed reinforcement. If a coalition of network users asks the SO to build a line that they do not want to pay entirely, this reinforcement should be treated as a regulated one. One may think of applying the same conditions described for the case of regulated investments to the investment projects resulting from the private initiative. However, this would unduly hinder the development of merchant lines and investment projects at risk. Conditions to approve merchant projects or projects at risk could thus be the following:

- The proposed reinforcement is not detrimental to the system. In other words, the net aggregated benefit of market agents over period 't' when only the proposed reinforcement is built, and agents do not have to pay this reinforcement, is higher than the net aggregated benefit obtained by market agents when no reinforcement to the grid is built.
- No regulated alternative to the proposed reinforcement has been included in a plan that has been approved by the regulator. We must bear in mind that regulated lines must be given priority over the investments by private promoters, as section 2.3 explains.

Lastly, one must be aware that the uncertainty about the outcome of the regulatory review of a proposed reinforcement, as well as an excessive duration of this process, decrease the value that the promoter places on the investment project. Therefore, as far as possible, the regulatory test should be easy to apply by the regulator so that reviewing a project does not take too much time. Additionally, the test should be transparent and objective in order to decrease the uncertainty of private promoters, see (Saphores, Gravel et al. 2004) for an interesting discussion on the subject.

2.4.2 Cost factors to be taken into account when assessing potential grid reinforcements

This sub-section discusses the different criteria that must be taken into account by the regulator or the competent authority when deciding on the construction of a line. A TSO proposing the construction of new lines should take into account all of this criteria when identifying the reinforcements that are most beneficial for the system. A private entrepreneur promoting the construction of a line will only consider part of them (the direct ones as we shall see later). In order to identify the main effects of a grid investment in the system welfare we have consulted a document by the Australian Competition Commission (ACCC 2004), among others.

We classify these criteria into two main categories according to whether they would automatically be internalized or not by a profit-seeking grid investor:

- Criteria not directly related to the investment project: those criteria that an entrepreneur will not internalize in its investment decision. They are external to the investment project itself and are related to the impact that the project will have on the rest of the system. Therefore, they need to be taken into account both when assessing regulated projects and merchant projects.
- Criteria directly related to the investment project itself: these criteria are internalized by private investors when deciding which line to promote. Therefore, they only need to be taken into account when deciding on the construction of regulated lines.

2.4.2.1 Criteria not directly related to the investment project

We may distinguish between two types of criteria within this group: technical and economic criteria and environmental criteria. If possible, the quantitative impact of every cost factor associated with the criteria outlined here should be computed. In this way, we would be able to compare the results of assessing an investment with respect to different criteria and obtain a single measure representative of the total benefits associated with the investment. This global measure corresponds to the expected net aggregated benefit of all the market agents in the system, see section 2.4.1.

Economic criteria

Impact of the investment on the transmission losses in the rest of the system: losses in the system may change because of two reasons:

- There will certainly be a redistribution of flows in the system associated with the change in the network topology.
- The load and generation pattern may also change due to this change in the network topology. If load and generation patterns change, transmission losses will also change.

Impact on the cost of providing the ancillary services required in the system: more efficient generators may be able to provide the ancillary services that are necessary in the operation of the system.

Impact on the cost of providing long-term security of supply: building a line may allow different systems or areas to share the generation reserves necessary to provide long-term security of supply. This will depend in any case on the characteristics of the system.

Impact on the costs of non-supplied energy: the better meshed a grid is the more reliable the operation of the system is and the fewer and shorter interruptions consumers will suffer.

Impact on the maintenance cost of the grid assets: building a new line will probably affect the number and magnitude of line outages because some overloads of lines may well be avoided.

Impact on the economic surplus of consumers and producers resulting from the dispatch of energy: this is probably the most important benefit associated with the construction of a line in many cases. As a consequence of the construction of a line, some more efficient generators will be dispatched instead of others whose energy is more expensive. Exports from areas where the energy is cheaper into areas where it is more expensive will increase. Thus, the generation and load pattern will change as a result of the investment. This will certainly affect both the revenues obtained by consumers and generators since it will change the distribution of energy prices in the region and the amount of energy produced and/or consumed by each agent. In those areas where the energy price decreases as a result of the grid reinforcement the benefit of generators will decrease while that of consumers will increase. In those areas where the energy price increases the benefit of generators and consumers will change in the opposite way. The exercise of market power may alter the net aggregated benefit of market agents as a result of the sale and purchase of energy.

Impact on congestion rents: the pattern of flows in the system will change as a result of the change in the topology of the grid. Apart from that, if the pattern of generation and demand also changes because of the installation of a new line, flows in the grid will vary as well. Congestion in the system may increase or decrease when line flows change. Changes in the amount of congestion rents in the system, which are initially collected by the market operator though they must be afterwards reinvested in the system, should also be taken into account when assessing a grid reinforcement.

Impact on the level of competition: if competition among generators increases and prices go down, the allocative efficiency of the market may increase, though this is already taken into account when assessing the economic benefit of consumers and producers resulting from the sale of energy. However, driving prices down may be an strategic objective in many markets as may be concluded from the statements of some governments over the functioning of the market. See the discussion in (Pérez-Arriaga 2005) on the strategy of the different Spanish governments when setting rules for the Spanish electricity market. In other words, reducing the level of market power in the system may be in itself an objective, even when the net aggregated economic surplus of generation and load does not change significantly (it will change anyhow, since the equilibrium point will be different because of the oligopolistic behaviour). This may be more of a qualitative objective than anything else. The Australian Competition and Consumer Commission considers this one of the major objectives of any network augmentation, see (ACCC 2004).

Environmental criteria

The aggregate level of emissions from generation plants will change if the network is reinforced. Pollution of any kind will also be affected. All these effects should be quantified (maybe using the prices of emission permits in organized markets) and taken into account. The amount of emissions may increase or decrease as a result of a network investment. If a price is set for emissions, environmental issues become also economic ones.

2.4.2.2 Criteria directly related to the investment project

These include the investment, operation and maintenance costs of the line as well as all the financial criteria to be considered by the investor. The investment cost should be the result of a public tendering in the case of a regulated line. Operation costs include the losses on the constructed line in the case of merchant lines. Transmission losses in a line represent an expense for a private entrepreneur buying energy at one end of the line and selling it at the other.

Financial criteria include the economic life of the line to be built, the date when the line will enter into operation -which will affect the net present value of the revenue stream associated with the investment- and the discount rate, which relates to the opportunity cost of investing in the construction of a line instead of investing in any other economic activity.

A risk analysis should also be carried out to weigh the risk associated with the uncertainty existing over the costs and revenues resulting from reinforcing the grid. The investor may consider taking some risk mitigation measures.

2.4.3 Discussion on the treatment to be given to the negative benefits brought about by the installation of a new line

Sub-section 2.4.1 assumes that both the negative and positive benefits resulting from the construction of a line are considered. On the other hand, some in the academic community and industry state that every investment project backed by one or several parties willing to finance it should be given regulatory approval. This thinking implies that losses incurred by agents must be neglected. This sub-section discusses the consideration of the negative benefits produced by a potential line when deciding on its construction. Here, we shall discuss the case of the expansion of the grid within a single, perfectly integrated, system. As mentioned in section 2.3, the institutional arrangements in each region may dictate that, regardless of the soundest mechanism to be used to identify new grid reinforcements and allocate their cost, compensations may need to be negotiated to get some lines built that are beneficial from a regional point of view.

We must distinguish between the regulated investments and those promoted by a private party. The following paragraphs discuss each one of the two cases separately.

2.4.3.1 Treatment to be given to the negative benefits associated with a regulated grid investment

A central planner should defend the interest of the system as a whole when proposing and/or deciding on the construction of a line. Therefore, both benefits and losses experienced by market agents because a regulated line is constructed should be considered. In this way, he would be maximizing the net social benefit of the system over the study horizon.

Neglecting the negative benefits (losses) incurred by some agents could result in conflictive situations where a line that reduces the level of competition in the system, thus benefiting few market agents at the expense of the interests of the region, could obtain the regulatory approval for its construction. This would be the case of a new transmission line that reduces the transfer capacity between an importing area and the rest of the system (admittedly, this will not typically be the case of new lines). Generators in the importing area will benefit from this investment and they may be willing to finance its construction. However, it is obvious that this reinforcement would be detrimental to the system. Considering negative benefits seems therefore unavoidable.

On the other hand, if the optimal investment rule for regulated lines that was presented in section 2.4 is employed, some lines which result in a decrease in the system production costs will not be built because the saving in operation costs they cause is lower than the cost of building, operating and maintaining them. This may be used as an argument to claim that not considering negative benefits would allow the system to evolve more rapidly towards a perfectly adapted one. However, it can be argued that the risk of being harmed by the construction of a new line may also discourage market agents from building a new plant or load centre. We will elaborate on this reasoning in the following paragraphs.

Strictly speaking, if negative benefits are considered when deciding on the expansion of the grid, market agents benefiting from the construction of a line should compensate those that are harmed. Compensations should be large enough so that no agent opposes the construction of the facility. Thus, the construction of a line that is beneficial for the system would represent a Pareto improvement with respect to the situation where this line is not built. No agent would be worse-off after the construction of the line. All of them will improve somewhat.

What is more, according to the allocation of benefits proposed here, agents benefiting from the construction of a line would not back its construction unless this line is beneficial for the system as a whole. If constructing a line were not efficient from a social point of view, the sum of the transmission line charges and the compensations that their beneficiaries would have to pay would outweigh the benefits they get from

it. Thus, adequately compensating agents for the losses they incur would result in a situation where the interest of individual market agents would be in line with the general interest (the interest of the system as a whole).

In any case, compensations associated with a certain line would have to be computed at the time the line is built and they would not be changed afterwards. Thus, these compensations and payments would not condition the operational decisions by agents on how much energy to produce or consume. Compensations and payments of this kind could result in a fixed uplift (for those agents benefited by the construction of the line) or reduction (for those harmed by it) in the transmission charge to be paid by each type of agent installed in each of several predefined areas. The areas would correspond to groups of nodes where the price of energy is approximately homogeneous (these are the Single Price Areas defined in chapter 4 when talking about congestion management).

But there are some difficulties associated with the computation of compensations to be received by those agents harmed by the construction of a line. If beneficiaries of grid investments compensate those who incur losses, market agents considering their installation will receive a locational signal corresponding, not only to the investment cost of new and existing lines, but also to the benefits and losses for the rest of agents resulting from the construction of these lines. One may think of only compensating those loads and generators that have been installed less than 'n' years before the time 't' when the line is built. Agents having installed a plant more than 'n' years before time 't' would be assumed to have already recovered their investment. Agents installing a plant after the construction of the line would know about the existence of this line when they decide whether to install it or not. In this way, the total amount of compensations to be paid by beneficiaries would be reduced.

As mentioned above, compensations and payments among agents that are related to the construction of a line would not alter the operational decisions by these agents as long as they are structured as a fixed transmission charge. However, these compensations and payments may have an impact on the long term decisions by agents on whether to install a new plant and where to install it. Some efficient generators might only be installed if they do not have to pay compensations to others. At the same time, some non-efficient generators (old and expensive) would retire if they are not paid any compensation but they would remain working if they receive a compensation. This would not be the case if compensations are only paid to the generators having installed shortly before the construction of the line (old inefficient generators would not qualify for receiving a compensation).

If computing suitable compensations for agents is deemed to be very problematic, one may think of considering negative compensations when planning the expansion of the grid but not compensating agents harmed by the construction of the approved lines. However, this alternative is not free from causing problems. As mentioned above, if agents are not compensated for their losses, they may perceive a strong risk associated

with the possible future installation of new lines, which could discourage some of them from installing a new generator or a load. If generator 'A' is not compensated because of the construction of a new line that allows more efficient generation to access his supply area, this agent will run the risk of not being able to recover his investment costs. Thus, not compensating agents when they are harmed may act as a deterrent of investment. Every potential new generator could see its future revenue stream threatened by the possibility of a more efficient generator being installed right after him in any part of the region and being dispatched instead of him.

In short, we propose taking into account negative benefits when deciding about the construction of a regulated line. Besides, agents harmed by a line being constructed should receive an adequate compensation. In any case, this is still an open issue that deserves further discussion.

2.4.3.2 Treatment to be given to the negative benefits associated with grid investments by private parties

Merchant lines are promoted and built by a private entrepreneur with the aim to maximize the profit resulting from the operation of these lines. The owner of a merchant line makes a profit by buying energy at one end of the line, where energy prices are lower, and selling it at the other, where energy prices are higher. Thus, merchant projects may be regarded as an alternative to investments in more efficient local generation. The installation of a merchant line may be accepted even when it causes the profit of generators in the expensive area to decrease for the same reason that efficient generators should not be banned from installing in this area to compete with the local ones. Consequently, merchant projects should get the green light as long as they do not cause an increase in the system operation costs and no alternative project exists that was proposed by the SO and has already been approved by the regulator.

If competition among agents increases, some efficient generators will be substituted for more expensive ones. Then, the aggregated production cost will decrease. Hence, any merchant project that increases the degree of competition will not be detrimental to the system as a whole. Thus, we conclude that conditions presented in section 2.4.1 for the acceptance of private investments in transmission result in an efficient development of the system.

Note that we are not neglecting the negative benefits incurred by agents when stating that every not detrimental merchant project should be approved. Rather, we claim that positive benefits should be weighed against the negative ones and the reinforcement should be built only when the former are greater than the latter (the cost of a merchant project should not be taken into account when deciding whether to approve its construction).

2.4.4 Some final remarks on the regulatory test

Note that, conceptually at least, the regulatory test should allow the justification of the so called “reliability lines”, i.e. lines that are justified because they contribute to the general improvement of reliability conditions in the power system. Reliability can be quantified in economic terms even if the beneficiaries whose reliability conditions are improved are widely dispersed. This may be done by making use of the utility function associated to the consumption of electricity. However, most frequently, transmission planning is subject to hard reliability constraints in network design that are imposed by some sort of national or international “grid code”, regardless of their economic justification.

Note, also, that reinforcements to the transmission network may bring substantial benefits to large aggregations of distributed generation that are already or will be connected to distribution networks. These reinforcements may facilitate –or even make possible- the secure injection into the lower voltage grid of significant amounts of power that otherwise could not be reliably absorbed by the local demand. These benefits should also be taken into account when computing the economic benefits produced by a line.

Most required network investments to be evaluated are not large ones, such as long lines or complex substations, but minor reinforcements, replacements or modifications. The nature of actual transmission investments is well explained in (Joskow 2005).

2.4.5 Classification of the existing algorithms to compute an expansion plan for the grid

This section provides an overview of the different methods that exist to develop an expansion plan of the transmission grid. The SO may use these algorithms as a tool when developing an investment plan to be submitted to the regulator. Additionally, these algorithms can be of help in the evaluation of individual grid reinforcements proposed by a third party. The SO should not evaluate a regulated reinforcement proposed by a third party separately from other reinforcements but jointly with the network upgrades that should be built after the proposed one.

The development of an investment plan for a large-scale system is a challenging and very complex task. It is a problem of great dimensions whose mathematical properties make it difficult to address. The planning problem is non-linear and non-convex in nature. In addition, much of the information required to obtain a plan is not available, since it refers to the future conditions in the system (load and generation pattern, etc). Sources of uncertainty already present in the traditional regulatory context include, among others: the future pattern of demand growth, the future weather and economic conditions, which may affect the efficiency of some generation technologies with respect to others, etc.

Planning the expansion of the transmission grid has become much more challenging in the new deregulated environment. Investment decisions in generation and transmission are closely interrelated. In fact, transmission and generation should be considered to be substitutes of one another in many cases. However, investment decisions in generation are no longer the responsibility of a central planner but of independent investors whose sitting alternatives depend on the availability of the required transmission facilities. What is more, strategic generators no longer bid their marginal costs but compute their offers in order to maximize their benefits. They may exploit the market power they hold to a lesser extent than they could because they fear the reaction of the regulatory authorities. Thus, estimating the behavior of market agents is not easy at all. In addition to all this, traditional sources of uncertainty remain.

Methods to identify the optimal grid investments may be classified according to different criteria. Here, three different classifications will be presented according to the article published by Latorre et al., see (Latorre, Cruz et al. 2003). First, we may characterize algorithms according to the solution method used. Second, we may classify approaches based on whether investments are assumed to take place all at the same time or in a sequence. Lastly, we will take into account the treatment that algorithms give to the new challenges posed by the restructuring of the electricity sector.

2.4.5.1 Solution method used

Regarding the solution method used to compute the expansion plan, algorithms can be of any of two types: those that make use of mathematical programming techniques and those that are heuristic. Some approaches combine features that are characteristic of both categories. These are called meta-heuristic. Methods based on mathematical programming usually rely on well-studied algorithms, based on the theory of operations research, to solve the optimization problem of expansion of the grid. This problem consists of an objective function to maximize or minimize and a set of constraints. The objective function measures either the increase in the global social welfare (deregulated environment) or the cost of supplying the electricity demand (traditional regulated environment). Constraints model technical, economic and reliability conditions imposed on the expansion of the grid.

Due to the complexity of the problem, this type of algorithms can only compute an optimum solution for a stylized representation of reality. Therefore, the outcome of a model of this type should be closely examined. Mathematical programming guarantees convergence and provides the global optimum of linear problems with continuous variables. Some non-linear problems with integer variables can also be solved though. Most important problems encountered by these type of algorithms are convergence problems related to the use of a DC load flow model or a more complex one and huge computational times associated with the use of integer variables or stochastic techniques to model uncertainty. Methods belonging to this group make use of different techniques like the classical optimization ones (linear and non-linear

programming, dynamic programming, mixed-integer programming, etc) or others like the Benders and hierarchical decompositions or the combinations of these with other approaches (the branch and bound algorithm). In (Alguacil, Motto et al. 2003) the computation of the optimal expansion of the grid including losses is formulated as a linear integer problem. Authors in (Binato, Pereira et al. 2001) decompose the optimization problem using Benders techniques while (Shrestha and Fonseka 2004) decomposes the same problem in a master problem that determines grid investments and a slave problem that optimizes the system operation. Examples of application of Branch-and-Bound techniques to compute the optimal expansion plan of the grid are provided in (Haffner, Monticelli et al. 2001; Choi, El-Keib et al. 2005). Authors in (Haffner, Monticelli et al. 2001) apply a Branch-and-Bound algorithm to solve the expansion problem when we consider a transport model of the grid while authors in (Choi, El-Keib et al. 2005) use a combination of fuzzy logic and Branch-and-Bound techniques.

Heuristic models are normally iterative and make use of non-proven methods to arrive at the optimum solution. On each stage of the process the method generates new expansion alternatives, evaluates them and selects the best according to the evaluation criteria used. They need a method to compute the benefit associated with a particular investment and may or may not use simulation in order to do so. Empirical rules or sensitivities are used to search for better options until the algorithm is not able to improve the best solution found.

Heuristic algorithms cannot guarantee obtaining the global optimal solution unless all the possible solutions are explored. However, high quality feasible solutions can be found with a relatively small computational effort. These methods employ many different techniques including Genetic Algorithms, Fuzzy Logic and many others. Authors in (Latorre-Bayona and Pérez-Arriaga 1994) propose an heuristic algorithm that takes advantage of the natural decomposition between the investment and operation sub-problems. (Escobar, Gallego et al. 2004) and (Silva, Gil et al. 2000) are examples of the application of genetic algorithms to the problem of the expansion of the grid. Faria et al. propose using an heuristic method named “Path Relinking” whereby new grid configurations are created by combining the characteristics of two good solutions, see (Faria, Binato et al. 2005). The underlying philosophy is therefore quite similar to that in genetic algorithms. Authors in (Kandil, El-Debeiky et al. 2001) use both a mathematical model and an expert system consisting of a set of rules to optimize the expansion of the grid. This algorithm determines both the optimal expansion of the grid and that of generation resources.

2.4.5.2 Timing of the investments proposed

Approaches to the planning of the grid expansion may be classified according to whether the considered algorithm determines when each upgrade must take place or it assumes that all the grid reinforcements are undertaken at the same time. The first type of algorithms is termed dynamic whereas the latter is termed static.

In a static algorithm, the grid planner is only concerned about the configuration of the grid at the end of the planning horizon. Therefore, he aims to find the configuration that best adapts to the situation of the system at the end of the period of time studied. Authors in (Silva, Gil et al. 2000) propose a genetic algorithm for the computation of a static investment plan. Most algorithms that make use of mathematical programming are also static.

On the other hand dynamic algorithms computes the optimal expansion strategy along the whole planning horizon. The planning horizon is divided into several time periods (years) and the algorithm takes into account the coupling among these periods. The planner is not only concerned with the identity of the grid connections to be built or reinforced, or the amount of transmission capacity to be built, but also with the time when each reinforcement should be undertaken. He aims to minimize the net present value of all the system costs incurred during the horizon of study.

Dynamic algorithms are much more complex in nature and therefore are far less developed than the static ones. The most important limitations are related to the system size and the complexity level of the system model. The dynamic problem must be greatly simplified in order to be able to solve it in a reasonable amount of time. One possible way of dealing with this challenge is dividing the dynamic problem into a collection of static problems that we solve one after the other. This is the strategy followed in the “backward” and “forward” procedures, which can be combined to obtain other methods, as in (Binato and Oliveira 1995). Romero et al. first proposed in (Romero, Rocha et al. 2003) a constructive heuristic algorithm to address the problem of identifying the optimal investments to be undertaken at each stage of the planning horizon. In this work they make use of a transportation model of the system grid. Afterwards, they presented in (Romero, Rocha et al. 2005) a constructive heuristic algorithm that employed a DC model of the transmission grid. This algorithm subsequently adds lines to the expansion plan or discards others that had previously been selected. Given the complexity of dynamic algorithms, heuristic algorithms are better adapted to solve large problems of this kind.

2.4.5.3 Ability to adapt to the situation emerging from the restructuring of the electricity sector

As mentioned above, the process of liberalization has affected most electricity systems in the world. Simultaneously, the creation of regional markets results in the size of the planning problem to be solved increasing notably.

Transmission expansion methods must adapt to deal with the special characteristics of the activity of transmission in the context of liberalized markets. Some aspects which require further research work to be properly addressed are the definition of the objective of transmission expansion planning in the new environment (the discussion about the consideration of the negative benefits incurred by network users is relevant here), the relationship between transmission planning and generation expansion (they are dependant one upon the other but take place separately), the ability to cope with

the high level of uncertainty that now exists or the reconsideration of the different criteria having to do with reliability, security and quality of service that have been used until now.

Among the different research results of interest regarding the expansion of the transmission system in the new environment, Pereira et al briefly analyze in (Pereira and Gorenstin 1996) the relationship of the generation and transmission expansion problems. Styczynski considers several objectives to be achieved simultaneously thus creating a multi-objective function (Styczynski 1999). Sun et al. apply the fuzzy theory to express several objectives as a single one (Sun and Yu 2000). Contreras et al. make use of the cooperative game theory to simulate the interactions among rational grid users, see (Contreras and Wu 1999; Contreras and Wu 2000). Pereira et al. aim to measure and hedge the risk inherent to competitive markets in (Pereira, McCoy et al. 2000).

Rau et al. address the problem of determining the optimal expansion of transmission and generation capacity in a regional context, see (Rau and Zeng 2004). They make use of mathematical programming to compute the optimal generation capacity within each area as well as the optimal transfer capacity between areas. Authors in (Fang and Hill 2003) explore different possible scenarios to deal with the increasing uncertainty in a deregulated environment.

2.5 Compatibility of investment signals in distribution, transmission and generation

In order for the development of the system to be efficient it is necessary that economic signals are sent to coordinate the investment decisions made by the institution in charge of the network expansion -or the coalitions of network users, depending on the adopted approach- and generation investors. A major unsolved problem is the interaction between transmission expansion and generation expansion.

The links between transmission and distribution expansion also need to be studied. First, the reasons for regulating distribution and transmission independently have to be examined and it must be also evaluated whether they are still justified in the new regulatory environment. One has to bring here the experience that has been obtained from existing sound regulatory schemes of transmission and distribution network planning, see (Khator and Leung 1997; Brown 2002).

In this point of the chapter we will examine two basic issues that still lack a satisfactory answer:

- The role that transmission-related economic signals and incentives -when separately applied to the activities of generation, distribution and transmission-

can fulfill in achieving joint efficiency in investment and operation of the complete power system.

- The relative strength that transmission-related locational signals can have when compared to other competing economic signals or to the distortions that are introduced by lack of regulatory harmonization in tariff setting. In particular it will be examined whether transmission-related locational signals –as they are currently used in some power systems- could be enhanced in an efficient and cost responsive way so that they may result in more efficient location decisions by the network users.

After this introduction, subsection 2.5.1 analyzes the compatibility of economic signals for generation and transmission and how these signals may complement one another. Subsection 2.5.2 examines this very same issue but now concerning distribution and transmission. Finally, subsection 2.5.3 presents several case examples, which are based on actual power systems, that help to illustrate the principles that have been presented in the rest of the section.

2.5.1 Compatibility of investment signals in transmission and generation

2.5.1.1 The difficulties

The objective now is to make compatible, in an efficient way, the two worlds of competitive generation and monopolistic transmission in the context of a liberalized regulatory framework. Strong evidence supports the claim that, despite being independently managed, transmission and generation (and to a lesser extent also demand) influence each other. As a consequence of this, the optimal development of the system can only be achieved if coordinating signals are employed that indicate network users what the impact –in terms of network costs of new generation or demand being installed at each point of the system will be and, vice versa, how the agents will be affected by the addition of a new line.

Contrary to what happened in the traditional regulated environment, new investment in generation does not result from a centralized process but from the non-coordinated decisions of a multiplicity of investors. Therefore, the central planner no longer knows with certainty what the development of generation in the system will be.

The construction of new lines may have a significant influence on the competitive position of some generators in the wholesale markets, therefore guiding the installation of new generators in each area. For instance, some reinforcement may eliminate a network constraint that was forcing the operation of an expensive plant that, otherwise, would remain idle most of the time. Or, on the contrary, it may eliminate a bottleneck and allow a generator to increase its output and export its power to another region. Thus, for example, the reinforcement of the corridors connecting the Iberian

system to the rest of Europe could allow less expensive generation in a very wide area to access the future Iberian market and the average energy price in this local region should presumably fall below its present level.

Locational transmission signals (a combination of losses, congestions and transmission network tariffs) may be influential in the efficient election of network connection sites for new investments in generation and, perhaps, large consumers as well. In order for the agents to make efficient investment options, they should face the real cost that the system will incur because of the installation of a new generation or consumption facility at a given point in the network. This can be achieved by sending locational signals in the form of nodal energy prices on the one hand, and transmission tariffs with longer-term locational content on the other. Although the existence of transmission-related charges is just one among several considerations when choosing a site for a new investment, in some cases it may have a decisive influence. This is particularly true when there are actual choices to be made, such as building a plant near a convenient fuel supply (a gas field or a high pressure pipeline) and shipping the power via the transmission network to the load center, or building the plant near the load center and bringing the gas to the plant via a pipeline. It would be clearly inefficient to apply the same transmission-related charges to two identical new power plants such that: a) one would be located near the major load center in the system, so that losses and congestions are small and no network reinforcements are needed and b) the other would be located far from the main load centers, so that losses will increase, congestion problems will appear and the network will need to be reinforced. A numerical case example will be presented in section 2.5.3.3 that shows some of the major considerations to be included when determining transmission-related locational signals.

Last but not least, and mainly due to growing environmental concerns raised by local and regional administrations and organizations in general, the process for obtaining the environmental permits and rights of way needed to begin the construction of a new line is significantly longer than what was usual before. Consequently, construction time for the prevalent generation technologies in most countries (combined-cycle gas turbines, CCGT, typically) is often shorter than the construction time of new transmission lines. As a consequence of all this, the transmission planner finds it increasingly difficult to predict where the new generation will be installed with a sufficient time margin. For instance, according to the Spanish System Operator, the average time to get a line built is about 5 years now, far more than the between one and one and a half years strictly needed to build the line and also more than the two or three years that the construction of a CCGT plant ideally takes (in practice at least three and sometimes even up to five or six years). In some cases the construction of highly controversial lines has been delayed by more than 20 years or even indefinitely.

2.5.1.2 Signals derived from transmission pricing

This point analyses the short and long term signals to generators and load resulting from the allocation of the costs of transmission. Chapter 3 of this document focuses on

the computation of transmission tariffs as well. In that chapter, long term transmission prices are discussed with much more detail. Prices used to recover the cost of the grid should also be used to let market agents know about the cost that their decision to install a power plant or load in one node of the network or another causes the system to incur. Therefore, some overlap between both chapters is unavoidable.

Short and long-term signals

Transmission pricing is the allocation of the regulated annual revenues of the transmission activity to the network users. If possible, transmission costs should be recovered from nodal prices since these are perfectly efficient short-term signals, i.e. geographically differentiated short-term marginal costs of energy, see (Schweppe, Caramanis et al. 1988). Nodal prices of energy implicitly include the effect on prices of losses and congestions in the network. They send adequate signals for decisions concerning the economic operation of generators and loads. See the properties of nodal prices in chapter 1 of this document.

Strict application of nodal prices to generators and loads results in a net amount of revenues, which should be applied to partly pay for the cost of the network. Under ideal circumstances, impossible to find in practice, these revenues would suffice to pay the total network costs fully. However, as indicated before, these revenues are usually very insufficient to cover the total network costs (cost recovery by nodal energy prices typically does not exceed 20% of total transmission costs), see (Pérez-Arriaga and Rubio 1995).

Thus, additional signals are needed to recover the remaining transmission network costs. These costs have to be assigned to the network users so that distortion of economic efficiency is minimized. Therefore, in the first place, these signals must be long-term signals, so that they do not interfere with nodal prices. This can be achieved by designing them as annual charges (although they may be distributed monthly, for instance). Ideally, these long-term signals should be consistent with the underlying cost function of the transmission activity, i.e. one should design them according to the driver behind transmission investment. This involves applying cost-causality, i.e. cost responsibility, as much as possible, see (Pérez-Arriaga and Smeers 2003; Smeers 2005). In the new competitive regulatory framework, investment in a new transmission line is justified whenever the present value of the aggregated benefits of all the network users (generators and consumers) is larger than the present value of the cost of the line. Here we are assuming that no market power exists.

Then, conceptually, the fraction of the network cost that is not recovered through nodal prices must be charged in proportion to the benefits that the transmission network (either globally or line by line) provides to each one of its users. The resulting long-term economic signals have no purpose in the operation (i.e. short-term) timeframe. They are only meant to provide locational signals to new generators and loads—or to those considering retirement—, i.e. to inform them about the transmission network

costs that are incurred because they locate or have located in one part of the network instead of another one. In the long-term, the response of the potential new network users to the transmission charges (i.e., whether they decide to carry out the investment or not) will depend on their expected profits, after transmission charges are duly included.

Some difficulties may be encountered when computing the economic beneficiaries of lines. Most of them arise from the lack of adequate information about generators' behavior in a competitive setting and the need to estimate the future evolution of the system. Additionally, evaluating the economic impact on the market agents of each individual line is difficult when the network is well developed and some level of reliability-driven redundancy exists. However, if the benefits produced by a grid reinforcement are not wide spread (few, well identified, market agents benefit from it), and taking into account the fact that benefits realized by market agents shortly after the construction of a line are much more important than those realized long time after its construction, we may be able to accurately estimate the economic benefits produced by a line. In this case, we would be able to allocate the cost of the line based on these benefits.

In most systems, parties have agreed to adopt some measure of electrical use as a pragmatic approximation to benefits (this is much easier to compute), see (Pérez-Arriaga 2002; Pérez-Arriaga and Smeers 2003). Thus, if we refer to the case of the Internal Electricity Market in Europe, this is the prevalent line of thought in the Florence Forum. Several algorithms have been proposed to compute the electrical use that each agent makes of the system grid, see (Rubio 1999). Presently, in most systems transmission network costs are socialized and charged uniformly to all network users, or to generators and consumers separately, according to some prescribed ratios of the total network cost.

2.5.1.3 Recommendations for practical implementation

Some recommendations can be derived from the basic principles that were presented in the preceding section and the discussion in Chapter 3 of the thesis. If the transmission network is well meshed and there are no clear locational signals to be sent, because generation and load are more or less evenly distributed and systematic congestion is unlikely to occur, then the beneficiaries (or major users) of the network cannot be clearly identified on the basis of their location. According to economic theory, in the absence of a clear indication from the underlying transmission cost function, the cost of the grid should be charged applying the inverse price elasticity rule (i.e. the concept behind Ramsey pricing) in order to minimize the loss of efficiency. This rule must provide an indication on how to split the global charge between generators and consumers on the one side and on how to compute transmission charges to individual consumers and generators on the other. Regarding the allocation to the individual consumers, the inverse elasticity rule would advise to charge more to the least elastic consumers. Note, however, that parties involved in the tariff setting

process may consider this an unacceptable discrimination, see (Pérez-Arriaga and Smeers 2003; Smeers 2005).

Regarding the split of charges between generators and loads, if we assume that there is strong competition on the generation side and that generation is perfectly adapted⁷ to the system, the rule advises to charge transmission costs mostly to consumers. Under these conditions, generation will be very elastic to prices and, in the long-run, a complete transfer of charges to consumers will take place. In other words, energy prices will rise or decrease as necessary in order for the generators to recover the transmission charges they are paying from consumers. Thus, generators will not make any profits in real terms, once the average rate of return for an investment in a sector with the same level of risk is deducted. Note that this is not a trivial or universal rule, contrary to the wide-spread belief that “consumers always pay all network charges in the end”. A simple example will illustrate this fact: Consider a project to build a new generator with non contested access to an inexpensive energy source in a remote location. Assume that transmission pricing rules are such that the generator is charged a large fraction of the cost of the transmission line that would connect it to the major load centers and that these charges do not turn the project unprofitable. In this case the network charge will not be transferred in the end to consumers since it will be fully absorbed as a cost of the new generation project.

If the transmission network is such that long-term locational signals are needed and they can be clearly identified, —because of systematic structural limitations of the network—, then the allocation of transmission costs should pay attention to location. Note that these long-term signals are no longer useful for existing generators and loads (except for those considering retirement because of economic reasons). They are meant to promote the installation of new facilities in the adequate sites and to fully recover network costs. However, for the sake of simplicity and avoidance of any appearance of discrimination, most regulators decide that both the existing and the new network users must be subject to the same charges. According to what has been said in the previous point, if generation is perfectly adapted due to free entry and exit, charging generators more or less is not important since consumers will end up paying the entire bill. However, locational differences among transmission charges to generators that are installed in different nodes or areas must provide them with the right incentive to locate in the network and, analogously for consumers. We must bear in mind that if the conditions presented above are not fulfilled, generators will not be able to completely transfer transmission charges to consumers. Then, the absolute value of transmission charges may affect the decision of a generator to install.

⁷ In a perfectly adapted system generators are totally free to enter and exit it. Therefore, whenever a generator is no longer profitable it will leave and, accordingly, generators will enter the market whenever there are business opportunities to take advantage of.

Both situations may take place at national level. In those countries where there is little need for long-term locational signals in transmission, network costs may be allocated to generators and consumers without any geographical differentiation. This seems to be the case in most European countries. On the other hand, in those countries where long-term location signals appear to be necessary (e.g. England & Wales, Norway or Sweden in Europe; Chile, Argentina or Australia are also good examples), transmission charges could have geographical differentiation.

These criteria are equally valid in a regional or multinational context. If geographical differentiation of the long-term signals is not a major concern, then uniform regional transmission charges for generators and consumers could be applied in strict application of the single system paradigm. However this would require a very high level of regulatory integration. Hence, letting each country compute the national tariffs to be paid by its network users may be a pragmatic alternative. In this way, agents would automatically gain access to the entire region.

However, the opposite situation is also possible. Authorities in a region may want to send long-term signals in order to let agents know which are the most appropriate and inappropriate zones to locate new generation and load. If the economic utilization of generation resources at regional level to meet the regional load causes much stress in the existing transmission network long-term locational signals are needed. A rigorous approach would consist of assigning the cost of each one of the lines in the region to those agents that use it (or benefit from it) while ignoring any political borders. However, implementing a regional tarification scheme may only be possible in markets with a very high level of integration. Another less intrusive alternative involves replacing the nodal allocation of transmission costs at regional level by a mechanism of compensations among countries, see the section of chapter 3 on the computation of transmission tariffs in Europe. These compensations would be based on how much each country uses (or benefits from) the networks of other countries.

As indicated above, complete socialization of transmission tariffs (i.e. the postage stamp tariff) does not contain any locational signal. Other more cost-reflective approaches, either based on the computation of electrical usage –as the “areas of influence” method in Chile or Argentina-, or on some estimate of the responsibility in network investment, as the “investment cost related pricing” (ICRP) method in the UK or Colombia, allocate the annual transmission network charge among all network users according to some procedure that typically involves a high level of cost socialization. Transmission tariffs that reflect the positive or negative incremental cost impact that new network users impose on the transmission network would have a stronger locational component. The next subsection develops a more “aggressive” set of transmission network charges, i.e. one with a stronger locational component.

As mentioned above, we have just presented some guidelines on the computation of transmission tariffs that encourage agents to locate in the right place in the grid. For a

comprehensive discussion on the development of tariffs, see chapter 3 of this document.

2.5.1.4 A proposal to compute incremental transmission charges for new network users

The basic concept

The emphasis here will be placed on identifying the incremental grid cost that a generator causes the system to incur when locating at a given node, while keeping in mind that the final driver behind the network charges should be cost causality, see (Pérez-Arriaga and Smeers 2003).

One should be aware of the frequent discrepancy that exists between network charges on the one hand and the volume of network usage by each agent, or the benefits that agents obtain from the network, on the other. The total amount of income that has to be collected to remunerate the owner of any of the lines must equal the regulated annualized cost of the line, which is the same for every year. However, due to the existence of economies of scale and the lumpiness of investments in the grid, the lines that are actually built may not coincide with the volume of investment that is strictly needed at a given time. In principle, one should avoid making new agents responsible for the transmission capacity in excess of that necessary just because of their connection. This excess capacity is actually meant for future network users that obviously cannot be charged now for it.

For the purpose of this analysis, we will make a distinction between:

- a) those lines that are built to meet the new needs of two or more network users and
- b) investments that are needed to extend the useful life of transmission facilities (since they are almost always replaced but not retired).

The allocation of the cost of connection lines (lines that are built for the exclusive use of a generator, or consumer, so that it can be connected to a suitable point of the common network) has been discussed separately and will not be addressed here.

Network reinforcements of type (b) are meant to replace existing facilities that arrive at the end of their physical lives. These lines are not built as a result of changes in the load and generation patterns. Consequently, the cost of these lines could be allocated just in proportion to some measure of the average use of the line by each agent. It is more difficult to assign the responsibility for lines of type (a). Here, it is argued that the allocation mechanism should reflect the fact that these lines are built in response to a change in the flow pattern related to the installation or retirement of some generators and/or loads. Additional transmission charges assigned to the owners of new generators and loads (those charges in excess of the “average” ones) should correspond

to the investment costs or savings that their decision to install or retire a power plant or load has caused the system to incur.

The decision by a generator to install in a mainly importing area (or a demand to install in a mostly exporting one) may reduce the need for new investment in the grid. In this case, a sound transmission pricing mechanism may result in a negative charge (i.e. a credit) for that generator or demand, thus making it attractive for new generators (loads) to install in importing (exporting) areas, since this will promote a more efficient use of the existing grid.

The proposal

The method that is proposed here employs the installation time of lines and generators and some measure of how much each agent uses –or benefits from- the network as the input to compute transmission charges. The method is applied only to lines of type (a) above. Any agent requesting connection to the grid should pay the cost of its direct connection line. In case the connection line has extra capacity so that future agents may connect to the same node, the cost of this extra capacity can be socialized to demand. As for those network reinforcements of type (b) that are built to extend the useful life of a given line, the same transmission pricing rules that were applicable to the original investment should be applied now.

Figure 2.5-1 graphically shows the process of computation of the locational signals that result from the allocation of the regulated revenue of the lines of type (a) to the network users. Several choices are offered to the regulator regarding the share of the cost to be jointly born by consumers and generators. General criteria to decide on how to make use of these choices in each particular case have been presented before, although in practice this has a certain degree of arbitrariness. A first choice concerns the possibility of socializing to demand the cost of the fraction of the transmission capacity of each line that is not actually used⁸, according to some agreed definition. In Figure 2.5-1, a is the used fraction of the line and $1-a$ the unused part.

Next, the cost of the used fraction of each line is divided into a fraction b to be paid by generators and the remaining fraction $1-b$ to be paid by demand. Again this is a choice to be made by the regulator using the general criteria that were discussed before. The cost of each one of these two parts is initially allocated to the corresponding market agents in proportion to their share in the utilization of the line (C_{Gi} for

⁸ Assume a line with a nominal transmission capacity (in MW) that is defined as $C1$ according to some agreed procedure. However, because of system security reasons, the maximum real power that the line can transmit under normal operating conditions is $C2$. Then, if the actual real power flowing through the line is P , the fraction a of the line that is actually used is simply $P/C2$.

generator i and C_{Dj} for load j). What is new in the method being proposed here, is that these participation factors are modified according to how long ago the line and the agents (generators and loads) were connected to the system. This is accomplished through the use of coefficients representing how old the line and each one of the agents are (k_l for the line, k_{Gx} for generators, where x is the generator number, and k_{Dy} for loads, where y is the load number). Thus, the participation factor of an agent in the recovery of the corresponding part of the cost of line l (where this corresponding part is $a \cdot (1-b) \cdot CT_l$ for generators and $a \cdot b \cdot CT_l$ for loads) comprises three components. The first one is simply the agent's share in the use made of the line: C_{Gi} for generator i , for instance. The second component is a function of the time for which the line, and the agent itself, have been in service (k_l and k_{Gi} , respectively, for line l and generator i , where these are input parameters whose value has to be defined a priori⁹). Finally, these participation factors are scaled down so that the sum of the participation factors of all the generators is 1 and the same for the loads. Therefore, the expression for the total usage charge, corresponding to line l , that generator i must pay is:

$$P_{Gi} = CT_l \cdot a \cdot (1-b) \cdot C_{Gi} \cdot (1+k_l k_{Gi}) \cdot CS \quad (0.1)$$

where CT_l is the total cost of line l , a is the used fraction of the line, b is the fraction of the cost of the used part of line l to be recovered from consumers (and therefore $1-b$ is the fraction to be recovered from generators), C_{Gi} is the usage factor corresponding to generator i , k_l is a factor representing how old line l is, k_{Gi} is a factor representing how old generator i is and CS is the scale factor used to modify participations so that all of them add up to 1. The expression for the scale factor is:

$$CS = \frac{1}{\sum_{i=1}^{ng} (1+k_l k_{Gi}) C_{Gi}} \quad (0.2)$$

We are proposing here that cost causality should be interpreted in an “incremental” rather than a plain “average” way, so that new agents are assigned more responsibility for the most recent network reinforcements. This results in the use of coefficients k_l and k_{Gi} so that new generators and loads pay a larger part of the cost of new lines than those market participants who have already been operating for a number of years. Thus, the factor used to compute generator i 's contribution to the recovery of the cost

⁹ For instance, we could use coefficients that are much larger than 1 (like 5, for instance) at the time the corresponding line (in the case of k_l) or generator (in the case of k_{Gi}) enters into operation. Then, these coefficients would gradually decrease with the passing of time until they become zero after 5 years for generators and 10 years for lines, for instance.

of a line l results from increasing generator i 's usage factor (C_{Gi}) by the proportion represented by the product of coefficients k_l and k_{Gi} .

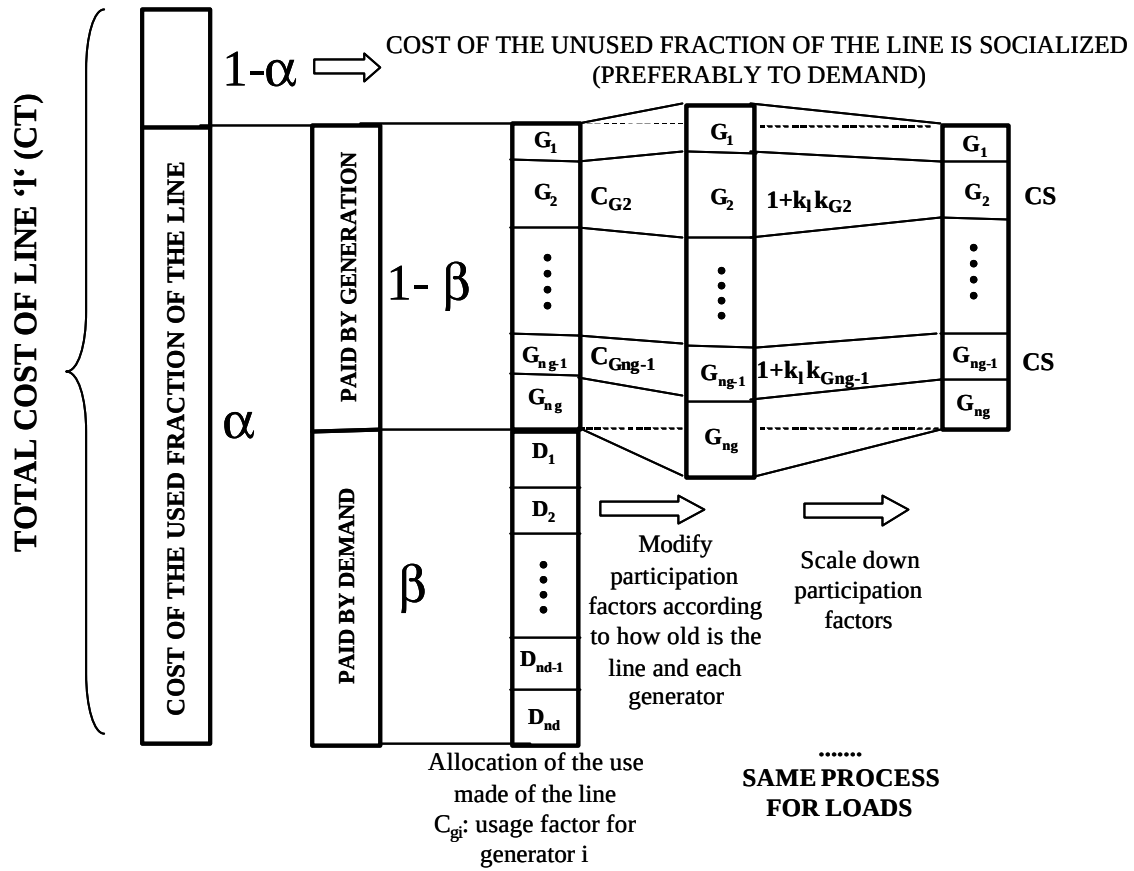


Figure 2.5-1: Process of computation of locational signals

Discussion

Chapter 3 of this document discusses the different methods that exist to compute transmission charges. The major difference between ‘marginal’ methods of transmission pricing and ‘average’ ones is that the former result in both positive charges (when the installation of the agent creates the need for new grid investments) and negative ones (when the agent reduces the amount of network investment that is needed). “Marginal” methods send stronger locational signals typically, although they require a higher level of arbitrary assumptions than “average” methods, see (Pérez-Arriaga 2002; Pérez-Arriaga, Olmos et al. 2002) for details. The approach that is proposed here can be used with any type of method for network cost allocation, either marginal or average, but it introduces the notion of temporality in the allocation, so that the link of causality between recently installed network facilities and new agents is emphasized. The method is supposed to produce stronger locational signals, where “strong” does not mean that the numerical value is larger, but that the relationship of causality has been reinforced. The method can be easily adjusted to emphasize more or less this temporal factor.

The method can be equally applied to compute transmission charges for new generators and consumers. However, consumers are not expected to be very responsive to these economic locational signals. Besides, many countries have decided to apply uniform electricity tariffs for every class of end consumer. If this is the case, flat charges may be applied to consumers and the proposed method could be applied only to generators.

The concept of introducing the temporal dimension in the causality relationship when computing transmission charges can be found in the regulation of PJM in the US, see (Joskow 2005). Transmission charges for generators corresponding to new grid investments may differ widely in PJM from one area to another thus sending important locational signals to new generators. However, unlike the scheme of signals proposed here, grid charges to generators in PJM are computed once (at the moment the generator is installed). Furthermore, charges paid by a new generator only cover the cost of facilities built because of the construction of the generator as well as those lines built shortly before the installation of the generator from which it is benefiting. In other words, new generators are not held responsible for the cost of lines built after the installation of these generators unless these lines were already projected when the generators decided to install. Responsibilities in network cost are evaluated by computing the incremental impact of the installation of each new generator on the flow over congested lines or corridors that must be reinforced.

Temporal differentiation plays an important role as well in the proposal for the computation of grid charges for the Peruvian power system by a group of international consultants that included the director of this PhD thesis, see (ME 2005). An interesting feature of the approach that has been proposed for Peru is that it employs two different network cost allocation schemes, one for the existing lines and the existing network users (method A) and a second method for the new lines and network users (method B). The point here is that there is little use in sending economic signals to consumers and generators that are already connected to the grid. Hence, there is not much interest in perfecting any current transmission pricing method A that is already in use for existing lines and network users. However, one could aim even for the most efficient method B –allocation to the economic beneficiaries- when it only concerns the new lines and the new network users. The major difficulty with the method of allocation to beneficiaries -going deep into the past and the future guessing what the economic benefits were or will be- disappears if there is no need to go into the past and if the exploration into the future is limited in time. Limiting the exploration into the future can be achieved by returning to method A once some years have passed after the construction of a line. Then, what were new lines and network users can be considered to be existing ones. Another approach -the one that was adopted in the Peruvian study- consists of computing the future transmission charges of the new lines for the new users once for all, no subject to future revision. This also has the positive effect of eliminating any uncertainty of the new users with respect to any possible changes in transmission charges in the future.

Locational signals for retiring generators

In order to encourage market agents to make the right decisions in the long term we must send them economic signals so that market agents realize the cost that the system incurs as a consequence of either their installation or retirement. The previous section discusses a system of economic signals aimed at encouraging market agents to install generation and load facilities in the best location from the system point of view. This section presents a scheme of signals that is meant to induce those agents considering the possibility of retiring a power plant to take into account the grid costs that may result from this decision.

Computing the effect that the decision of retiring a plant has on the number and nature of future grid investments is very difficult. Besides, the application of economic signals that are based on costs not yet incurred by the system is very problematic from a regulatory point of view. Therefore, it seems more appropriate to derive signals to market agents from the impact that their decision to leave the system would have on their current electrical usage of the existing grid.

Here we propose following for retiring plants exactly the same basic approach that has been proposed for new entrants. There are two major possibilities. The first one is to hold the agent responsible only for the change in the use of any congested lines, which will probably have to be reinforced in the near future if the plant retires. The second one is to consider that the agent is responsible for the change in the use of any line in the system, whether congested or not. It seems more adequate to compute the grid charges taking only into account the congested lines, since there is no evidence that other lines will have to be reinforced because of the decision of the agent to leave.

In summary, the same method presented above to send “strong” locational signals to new generators is also proposed to compute the grid charges associated with the retirement of an existing plant. The retiring generator would be regarded as new as well as the congested lines and the lines that have actually been installed recently. The signal should be implemented as a single payment to be made when the agent retires the plant.

2.5.1.5 Other locational signals for generators and network planners

The previous sections have discussed the design of a system of transmission charges that could send meaningful economic signals to existing and future network users, while recovering the total network costs. However, there are other locational signals of a different nature that generators and also system planners could send and receive. These signals are mostly related with information that helps removing the uncertainty that accompanies investment decisions in both network and generation.

As we have seen, the development of the grid must be planned in conjunction with that of the generation and the demand in the system. The location of new power plants and loads influences the decision by grid planners on which are the most interesting

reinforcements. Likewise, market agents should take into account the limits that are imposed by the existing grid and the investments in new lines or other transmission facilities that are planned for the future. Both the network planners and the potential investors in new generation need to know what the other party will do with as much certainty as possible.

On the one hand, it is important for the network planner to have a reliable indication of the level of commitment of new agents that ask for connection to the grid, since network expansion should be specially addressed to meet the needs that result from new network users. Consider as an example the case of the Spanish power system, with a peak load of 43.000 MW in 2005 and where the System Operator has standing requests for new connection of more than 50.000 MW of wind generators and more than 50.000 MW of CCGT power plants. Most of these requests will not materialize, since their purpose is just to make sure that a particular site has met all the administrative requirements to install new capacity -a procedure that usually takes several years-. Requiring from the potential investor a financial guarantee can help the System Operator sort out the more serious requests from the merely speculative ones. This guarantee would be necessary in order to start the administrative process of authorization of the connection to the grid plus any environmental or other kind of required permits. The guarantee will be returned only if the power plant is actually built within a prescribed time limit.

On the other hand, the System Operator should provide as much information as possible regarding the future development of the grid to the market agents that are considering the installation of a new plant. This information should include the projected importing and exporting capacity with neighboring systems, the expected value of transmission charges at each node as well as the loss factors and the nature, the location and severity of the restrictions that the grid is anticipated to impose on the operation of the system. An adequate time horizon for these estimations is between five and ten years. This information should be made available in some kind of “N-year-ahead transmission report” to be issued annually.

Additionally, the regulator and the System Operator must make clear the network access rules to all potential generation investors. Access rules encompass the management of network constraints as well as the criteria to grant connection to the transmission network. Ill conceived access rules may deter new generation investment by creating too much uncertainty or, at the other extreme, they may provide excessive advantages to the first new entrants with respect to future ones. For instance, a reasonable access rule for generators may specify that connecting first to a given node of the transmission network does not provide any advantage regarding priority of access to the network. This rule is relevant when the network cannot accommodate all the power that the generators that are connected to that node are willing to inject. An investor that considers building a new power plant and to connect it to this particular node should expect competition with the generators already connected to this node just for the use of the limited capacity of injection at the node. Firm transmission rights,

FTRs, may be used by the existing and the future power plants to manage the economic risk of being displaced from producing at a node by more competitive generators, therefore creating more favorable conditions for investment. Note that, in general, if there is a bottleneck, the network planner will decide to reinforce the grid -if it is economically and technically justified- so that the congestion is conveniently reduced or even eliminated. Therefore one should not expect this type of situation to be frequent, unless some of the existing generators at the node are very inefficient and network reinforcement in this case is not justified. Each power system may choose different access rules from the ones that have been presented here as an example. Access rules are an essential ingredient of the information that potential generation investors need in order to make sensible decisions on whether to invest or not in a given power system and whether to connect or not at a given node of the transmission network. For a comprehensive discussion on access rules to the transmission network go to chapter 4 of this document.

Access rules and transmission charges can be used to guide the process of connection to the grid of new loads in a meaningful way. The applicable legislation in many countries allows agents to apply for connection at any voltage they choose and the network planner is then forced to comply. Some agents may take advantage of this situation to get connected to voltage levels where they know they will pay fewer network charges or have other types of advantages, but which may not be convenient for the overall design of the distribution or transmission network. One possibility to guide the location of new agents is to charge them according to the voltage level where they should be connected -depending on their peak demand and total consumption, for instance- instead of the voltage level where they are actually located.

2.5.2 Compatibility of investment signals in transmission and distribution

The separation between transmission and distribution networks is not a clean one. Most European countries have adopted the simple and, up to a point, arbitrary decision to define transmission as encompassing all network facilities operated at 220 KV and higher (400 KV, typically) and, also frequently, all interconnection lines with other countries. Distribution would correspond to all other networks at lower voltages. Obviously, there are some interactions at the interface between both networks. For instance, depending on the physical configuration of each particular network, some further development of the 220 KV grid may help the underlying high voltage distribution network to avoid some investments. Similarly, a strong high voltage 132 KV sub-transmission grid may replace some investments at 220 KV level.

Traditionally, the regulatory approaches to the transmission and distribution activities have been very different. There are some good reasons for them to remain so:

- Transmission and distribution grids perform different functions. Transmission makes sure that major generators are well connected to major load centers. Flows over the lines in the transmission grid change direction frequently. Transmission has a great impact on the implementation of the wholesale market. On the contrary, if we leave aside the existence of distributed generation -which in some areas may outweigh local demand so that the corresponding distribution network is a net exporter- power over the distribution grid always flows in the same direction: downstream from transmission to the end consumers. The operation and planning of distribution networks as a whole are mostly influenced by the location and load profiles of final consumers. On the other hand, the internal configuration and operation of distribution networks has a negligible impact on the wholesale market and the configuration of the transmission network.

However, at higher voltage levels distribution grids tend to be meshed and their configuration can be adapted to the particular operating conditions by opening and closing the corresponding devices. In some cases the power flows in the transmission network cannot be fully understood without considering the impact of the sub-transmission network. Besides, more and more generation is connecting to the distribution grid at all voltage levels. Therefore, one can easily conclude that, in many cases, those parts of the traditionally considered distribution grids that are immediately below transmission substations are really performing transmission grid functions. The difference between transmission and distribution grids, especially at the border between them, is becoming increasingly blurred.

- The number of lines and other network facilities in distribution grids is much higher than that of transmission grids. In transmission each individual investment can be separately considered for authorization and remuneration whereas in distribution, particularly at lower voltages, this would be an extremely tedious and meaningless task. Performance-based regulation is a preferable choice for distribution networks.
- Transmission has an important role in keeping the overall security of the power system. Failures at transmission level are very infrequent. However, when they occur, there is usually a large disruption in power supply. Failures at distribution level are much more frequent. The quality of supply of most consumers mostly depends on the distribution network. Therefore, the regulation of distribution should heavily relate to quality of supply considerations.

The above mentioned broad differences that exist between transmission and distribution in many aspects should result in a specific treatment of remuneration, planning of the expansion, operation and, quite importantly, the short and long term signals to be sent in each one of the two cases:

- Regarding the remuneration of the activity, there are implications both for the computation of the regulated revenues of the owner of the grid and for the split of these revenues into the part to be collected from consumers and the one to be collected from generators. In the absence of distributed generation, consumers should pay 100% of the distribution network costs. According to cost causality, generators connected to transmission should not pay distribution network costs, since the design of distribution networks is not influenced at all by the nature or location of the generation power plants. On the other hand, the costs of the transmission network should be shared by generators and consumers (here we mean both large consumers that are directly connected to the transmission network and smaller consumers that obtain their power from the distribution networks). When distributed generation (i.e. generation that feeds into the distribution network) cannot be absorbed locally and feeds into the transmission network, it should also pay transmission network charges.
- Given the ambiguity that exists when defining both the border between transmission and distribution and the responsibilities in expanding the network, the approaches used to compute the remuneration of the transmission and distribution companies should at least be consistent with each other.

As said before, performance-based regulation should be employed to determine the revenues of the owner and developer of the distribution grid. In the case of transmission, currently most of the countries compute the income of the transmission company or companies through some sort of cost of service regulation. This lack of coherence may result in a suboptimal development of both grids. We have already mentioned the natural inclination of System Operators and regulators to overbuild the transmission network. Moreover, when faced with a remuneration scheme that is based on the cost of individual transmission facilities or on the winning bids in a public auction, transmission companies or TSOs will be happy that as many lines as possible are built. On the contrary, provided the remuneration of the high voltage distribution company is some kind of RPI-X, this company prefers leaving the transmission company to meet the needs of reinforcement of the sub-transmission network by reinforcing the transmission one. Thus, the distribution company could cut costs without hurting the overall network service. As a consequence of this, too much transmission grid may be built that performs a function that would have been better accomplished by the distribution grid. Regulators should be aware of these distortions in order to minimize them by adjusting the corresponding remuneration mechanisms. However, a complete elimination of these distortions does not seem to be possible because of the discontinuity of the regulatory treatment at the border between the two types of network.

- Ideally, the planning of transmission and distribution grids should be coordinated. For instance, the decision on the location and number of the transmission/distribution substations at the border between both grids is

critical in this regard. An iterative process, similar to the method that is used in some cases to determine the location of substations and transformers between different distribution grid levels, could provide the location and number of the substations between transmission and distribution as well, see (Peco 2001).

In the absence of any joint planning in practice, coordinating economic signals might be useful. These locational signals could be applied to the end consumers and also to the distribution utilities. As shown before, network charges for consumers in the distribution grid should in general cover both transmission and distribution costs, whereas those aimed at consumers and generators at transmission level could ignore the distribution grid. This is true as long as distribution grids do not contribute to transporting the power output from distributed generation to load centers located in other distribution networks. Locational signals in relation to transmission investment have been already discussed in section 2.5.1 of this chapter. The next paragraphs explain how nodal prices could be employed to send incentives for investment both to end consumers and to the distribution companies. Since the distribution and retailing activities should be unbundled in a sound regulatory framework, the distribution utility does not purchase electricity. However, the point here is that regulatory incentives for the distribution utility should be designed so that the price of acquiring the power that is needed to physically supply the end consumers of the distribution company is as low as possible. A regulatory mechanism to accomplish this is to make the distribution company responsible for the difference between the actual cost of purchasing electricity at hourly nodal prices from the corresponding transmission nodes and the amount of electricity consumed by the end consumers of this distribution company multiplied by some average loss factor and some average hourly energy system price. This is a generalization of the regulatory mechanism that is frequently used to provide incentives to distribution utilities to reduce network losses.

- The first level of implementation of this scheme of locational signals, with effects over the development of the distribution grid (and also the operation, as we shall see below), would consist of letting the distribution company be responsible for the economic difference between some reference purchasing hourly price for the energy demanded by the consumers of the distribution company (affected by some correction loss factor) and the pretended purchase (since the power is actually bought by the retailing company) of the energy from the transmission system at the corresponding nodal price, which would be dependent on the point of interface with the transmission grid. Under this scheme, the distribution company would react planning the development of the grid in order to minimize, not only the cost of losses over the grid, which has traditionally been its major objective, but also the cost of purchasing wholesale power.

Nodal prices can also encourage efficient operation of the distribution network. Similarly to what happens with the planning of the grid, leaving the distribution company to buy the energy for its customers at the corresponding transmission hourly nodal prices would lead the distribution utility to operate the grid in order to minimize the cost of purchasing the wholesale power at any given time.

By operation of the grid we mean the possibility of modifying the configuration of the sub-transmission grid. Normally, this grid is well meshed because of security, quality of supply and stability reasons. However, at least part of this grid is operated as a radial network by suitable switching. Changing the network configuration during the operation allows the distribution company to modify the incurred losses and, if they would buy power from the transmission grid at nodal prices, to modify also the amount paid for this power as well. Some indication of the quantitative relevance of these locational signals can be found in (Carrillo-Cacedo and Pérez-Arriaga 1995).

- A second level of implementation of a scheme of nodal prices would be to pass-through the wholesale power prices to the end consumers. Computation of nodal prices at the end consumer nodes of the distribution grid has to be discarded because of the high volatility these prices may experience, which would make them unacceptable, see (Outhred and Kaye 1996). It would be possible, however, to charge each end consumer a zonal price based on the nodal prices at the node or nodes of the transmission grid where the distribution company buys its power. These zonal prices could be obtained by averaging, over the period of time for which the prices are computed, the nodal prices at the different nodes of the transmission grid weighted by the amount of energy that the considered distribution zone is obtaining from each transmission node.

As mentioned before, the massive introduction of distributed generation will likely blur the differences between transmission and distribution regulation, especially if this generation ends up being partially fed into the transmission network. Then, the development of transmission and distribution grids will become much more dependent on one another. The planning of the transmission and high voltage distribution grids will probably have to be carried out jointly or, at least, will involve a much higher level of coordination. Transmission and distribution grids will perform much more similar functions and, therefore, the regulatory treatment of transmission and distribution will have to evolve accordingly.

2.5.3 Case examples

This subsection presents numerical examples of the most relevant concepts that have been presented in the preceding sections. First, we explore the application of the

method proposed above to compute strong locational signals to new generators and loads. Afterwards, we assess the potential of transmission charges to guide transmission and generation investment in the context of the European electricity market. Lastly, a simple numerical example shows how nodal transmission tariffs are just part of an ensemble of other locational signals and what the expected weight of each one of them is.

2.5.3.1 Numerical results resulting from the application of the method to compute incremental transmission charges for new network users

As explained in section 2.5.2, the method proposed there, hereinafter referred to as the Average and Load Factors (ALF) method, takes into account the different degree of responsibility that agents have in the development of the grid, depending on how much time has passed since they began to operate. This method has been used in combination with one of the usage-based methods discussed in chapter 3, the Average Participations method (AP). The AP method produces the usage based factors taken by our method as an input. Results have been computed for the Spanish system so as to provide economic signals that may guide, together with others, the location of agents in the grid.

Locational signals are only computed for generators. Apart from sending economic signals, transmission charges must achieve the total recovery of the grid costs. Charges to generators in the Spanish system have traditionally been zero (except for the connection lines some of them have had to pay). This situation is unlikely to change in the future. This means that the average transmission charges for generators will probably continue being zero. Therefore, transmission charges to load will have to pay the total cost of the grid. This will certainly condition the grid locational signals that can be sent to the load. Apart from that, the fact remains that load is typically far less elastic than generation to the economic signals in a country like Spain. All this has led us to focus on locational signals to generators. Computing locational signals to load as well is always possible. However, consumers in Spain pay a uniform transmission charge. Therefore, computing locational signals to load in this particular case would not make sense.

Description of the data used in the study

Application of the ALF method requires having available data on the time when the different generators and lines entered into operation (at least for those generators and lines recently installed). Since this information was not available for present or past scenarios corresponding to the actual operation of the system, we decided to apply the ALF method to future snapshots of the Spanish system provided by the Spanish system operator (REE). We have used input data corresponding to two snapshots of the Spanish system for the year 2008. These snapshots are an estimation of the summer and winter peak load operating conditions that may take place in the Spanish system during the year 2008. We also employed data on the lines and thermal generators scheduled to begin to function between 2004 and 2008. As we have just mentioned, all

these data are just an estimate corresponding to two feasible scenarios for the year 2008 which may not materialize in reality.

Presentation of the results

The ALF method, as explained in section 2.5.1, was applied to the above mentioned data. Our objectives were two. First, we wanted to determine whether transmission charges computed in this way were comparable with other locational signals, namely the loss factors. Therefore, we computed both transmission charges and marginal loss factors for a set of CCGTs planned to be installed between 2004 and 2008. Table 2.5-1 allows us to compare both types of signals. Second, we wanted to compare transmission charges obtained for ‘new’ generators (those recently installed) and ‘old’ generators (those that have been operating for a while) and see how much they differed. Figure 2.5-2 shows the geographical distribution of transmission locational signals to generators in the Spanish system for the year 2008.

Table 2.5-1 shows the values corresponding to the year 2008 of the marginal loss factor, the total annual grid payment and the corresponding energy transmission tariff for some generators expected to be built between the years 2004 and 2008. The marginal loss factor and the transmission tariff are expressed in €/MWh (losses have been valued at 36 €/MWh) while the annual grid payment is expressed in million Euros. The last two columns in the table represent the deviation from the average of the loss factor and the transmission tariff at each node. They are also expressed in €/MWh.

As already mentioned, the loss factor at a node represents the increment in the system transmission losses produced by a unit transaction between this node and the one that has been taken as the slack bus. Thus, negative transmission charges mean that injecting a unit of power at the corresponding node, and consuming it at the slack bus, results in a decrease in the system transmission losses. The slack bus chosen is a node where a large generator is operating (Villarino). Loss factors computed for all the nodes included in Table 2.5-1 are negative, which means that substituting a MW produced by the generator in Villarino by one produced by a generator located at any of these nodes would reduce transmission losses. If we had chosen a different slack node, loss factors would be the result of adding a common term to the corresponding ones obtained here. Therefore, differences among the loss factors at different nodes would not change. In terms of the numbers provided in Table 2.5-1, and assuming that loss factors and transmission charges are constant throughout the year, a generator located at node ‘ACECA’ which produced 1 kWh of power every hour of the year would be paid 22.06 € for the decrease in transmission losses it causes, while it would have to pay 4.81 € for the fraction of the annual transmission costs it is assigned. On the other hand, a generator located at node ‘SOTORIBERA’ that produced 1 kWh of power every hour of the year would be paid 3.58 € for the decrease in transmission losses it causes, while it would have to pay 18.70 € for the fraction of the annual transmission costs it is assigned. Last two columns correspond to the different between the loss and transmission charges paid at each node and the average ones.

MODE NAME	LF_50G_50L	TOTAL_GRID_PAYMENT	TRANS_TARIFF	DEV_FFROM_AVER_LF	DEV_FFROM_AVER_TARIFF
ACECA-2005	-2.518	3.6	0.549	-0.968	-0.724
PALOS-2005	-1.360	11.6	1.754	0.189	0.481
NESCOMBERA S-2005	-1.570	5.8	0.890	-0.020	-0.424
NESCOMBERA S-2005	-1.570	5.8	0.890	-0.020	-0.424
NESCOMBERA S-2005	-1.706	5.8	0.890	-0.196	-0.424
PALOS FRONTERA-2005	-1.357	11.9	1.795	0.193	0.521
PALOS FRONTERA-2005	-1.126	4.3	1.854	0.424	0.581
AMOREBIETA-2005	-1.191	4.3	0.990	0.359	-0.293
AMOREBIETA-2005	-1.170	4.9	0.991	0.379	-0.282
AMOREBIETA-2005	-1.191	4.3	0.990	0.359	-0.293
NESCOMBERA S-2006	-1.706	6.4	0.936	-0.196	-0.338
NESCOMBERA S-2006	-1.706	6.4	0.936	-0.196	-0.338
A RCOS FRONTERA-2006	-1.736	11.4	1.662	-0.187	0.389
A RCOS FRONTERA-2006	-1.736	11.4	1.662	-0.187	0.389
FAUSITA-2006	-1.503	3.5	0.795	0.046	-0.478
FAUSITA-2006	-1.503	3.5	0.795	0.046	-0.478
FAUSITA-2006	-1.681	3.5	0.795	-0.132	-0.478
MORVEDRE-2006	-1.964	4.4	0.664	-0.414	-0.610
ARA GON-2007	-1.731	18.4	2.778	-0.182	1.504
ARA GON-2007	-1.731	18.4	2.778	-0.182	1.504
C. COLON-2007	-1.517	6.2	0.938	0.032	-0.335
MORVEDRE-2007	-2.120	4.5	0.672	-0.571	-0.602
MORVEDRE-2007	-2.120	4.5	0.672	-0.571	-0.602
SOTORIBERA-2007	-0.409	13.1	2.135	1.141	0.862
ACECA-2008	-2.518	4.0	0.604	-0.968	-0.669
VANDELLOS-2008	-1.713	10.5	1.566	-0.164	0.312
VANDELLOS-2008	-1.713	10.5	1.566	-0.164	0.312
SABON-2008	-1.493	1.5	0.251	0.057	-1.022
TMBERO-2008	-1.404	14.2	2.181	0.145	0.908
P. G. RODRIGUEZ-2008	-0.717	9.5	1.546	0.832	0.272
P. G. RODRIGUEZ-2008	-0.717	9.5	1.546	0.832	0.272
A RUBAL-2008	-1.388	8.8	1.782	0.161	0.509
AVERAGE	-1.550	7.7	1.273		

Table 2.5-1: locational signals for new generators in the Spanish system coming from grid charges and transmission losses. Values estimated for the year 2008.

From the comparison of these two last columns we can draw the conclusion that locational signals for new generators resulting from the application of these transmission charges are of the same order of magnitude as the locational signals that are related to the marginal losses that can be attributed to these new generators. If our aim is to devise a suitable scheme of locational signals both types of signals would be equally important to implement.

As said before, Figure 2.5-2 presents the distribution over the Spanish system and for the year 2008 of the locational signals to generation that are related to the cost of the transmission grid. These transmission locational signals have been computed using the ALF method. Their average value is zero and they have been obtained from an initial set of transmission tariffs resulting from the ALF method by deducting from them their average value. Signals are represented using dots. The level of the signal is represented by means of a color code. Typical values range between -1.25 €/MWh for dark blue and 1.60 €/MWh for dark red.

Those charges whose value is above 1.6 €/MWh (most of which are well above 2 €/MWh) have been considered outliers since there are only a few of them and there is a large difference between their values and those of the signals obtained for the rest of the generators. The existence of outliers may be related to the fact that the method assigns the whole cost of each line to the agents that use it. Given that some small generators or consumers make a non-negligible use of large lines whose level of utilization is very low, the unit transmission charges computed for these agents tend to be much higher than those obtained for the majority of agents. Outliers have been represented by squares in dark red. The size of the square marker varies with the total amount of energy produced by the agent over the year. The more power it produces the larger the square marker is. Locational signals for 'new' generators have been represented by using a circle instead of a dot in order to be able to tell them apart from those markers corresponding to generators installed before 2004 ('old' generators). As mentioned above, a generator is considered to be new when it is expected to begin to function between 2004 and 2008.

Apart from allowing us to check whether the distribution of transmission grid locational signals makes sense, Figure 2.5-2 helps in determining whether locational signals for new generators are significantly stronger than those for old ones. In general, there does not seem to be a significant difference between the signal sent to a new generator and that sent to an old one located close to the former. This is due to the fact that the majority of new generators are not using new lines to a significant extent. Therefore, the participations of new generators in the recovery of the cost of these lines, despite increasing when we use the ALF method to compute them, is still quite modest in most of the cases. Besides, the number of new lines being considered here is very small compared to the total number of lines in the system. Different numbers would have been obtained if a higher number of lines were expected to be built in the coming years.

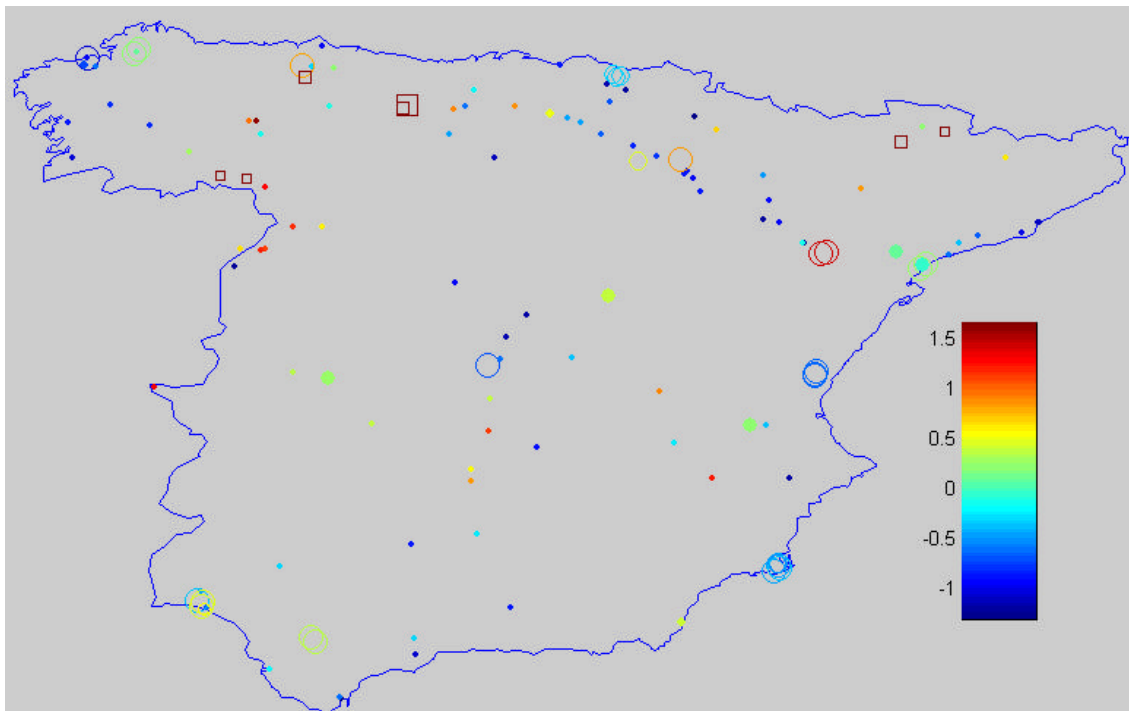


Figure 2.5-2: Geographical distribution of locational signals related to the cost of transmission lines when computed using the ALF method. Values obtained for the year 2008 in the Spanish power system

2.5.3.2 Comparison of transmission charges to other locational signals to market agents in the IEM

The objective now is to assess the relative importance of transmission charges when compared to other locational signals to generators and loads, namely energy prices. This comparison is accomplished in the context of the Internal Electricity Market of the European Union.

As outlined above, the application of a system of regional transmission charges has been discarded by parties in the IEM market, at least initially. Instead they have decided to implement a system of compensations among countries for the use that agents in each country make of the grid of others. In this example, we have applied the AP method to compute compensations among countries for the year 2002.

Data used in the study

We have used 24 different snapshots corresponding to the real operation of 13 countries throughout the year 2002 to compute the total net payment for each country. The data have been provided by the association of European Transmission System Operators (ETSO).

Some other data have been used to obtain the results presented here. We have had to estimate the annual cost per km of 400KV line in order to express compensations

among countries in monetary terms. First, we computed the number of kilometers of equivalent 400KV lines within the Horizontal Network of 11 European countries. These countries were Austria, Belgium, France, Germany, Italy, Portugal, Spain, Switzerland, the Czech Republic, The Netherlands and Slovenia. The figures obtained were 3896 km for Austria, 3930 for Belgium, 4477 for Switzerland, 5116 for the Czech Republic, 34811 for Germany, 21511 for Spain, 27916 for France, 13019 for Italy, 3737 for the Netherlands, 3627 for Portugal and 696 for Slovenia. Second, we divided the regulated cost of the Horizontal Network of each country (values were those actually used to compute compensations in the year 2002) by the number of kilometers of equivalent 400 KV lines within the network of the country. Thus, a regulated cost per kilometer of equivalent 400KV line was obtained for each country. Finally, the standard unit cost of a 400 KV line used to express all the results in monetary terms resulted from computing the average of the regulated unit costs obtained for the different countries. An annual value of 0.0352 million Euros per km of 400 KV line was obtained.

Presentation of the results

Figure 2.5-3 compares, for 17 European countries, the average transmission tariff of a typical industrial customer to the net inter-TSO payment resulting from the application of the AP method to the 24 scenarios referred to above, see (Pérez-Arriaga, Olmos et al. 2002; Pérez-Arriaga and Olmos 2003). Numbers are expressed in €/MWh. The total net payment faced by each country is obtained as the amount the country must pay others minus the compensations others must pay it. Figure 2.5-4 shows the evolution of the wholesale energy price in several European power exchanges between 2000 and 2004. Figure 2.5-4 has been provided by the Spanish power exchange OMEL.

According to these figures, there can be a difference of up to 7 €/MWh between the transmission tariffs in two different countries (see the difference between the transmission tariff in Sweden and Spain). This difference is of the same order of magnitude as that between the energy prices in two power exchanges, see Figure 2.5-4. Since transmission tariffs are long-term signals that should not interfere with system operation, it is most advisable to give transmission tariffs the format of a capacity (€/MW per year) charge, or even a lump sum charge (€/year), but never an energy charge (€/MWh), since this could introduce non negligible distortions in the generation dispatch in the region.

Figure 2.5-3 also shows that the net inter-TSO payment each country must receive (expressed again in €/MWh) is much smaller than the typical transmission tariff. Therefore, it is very unlikely that these inter-TSO payments will be able to send any significant locational signal in the presence of widely different transmission tariffs in the considered countries.

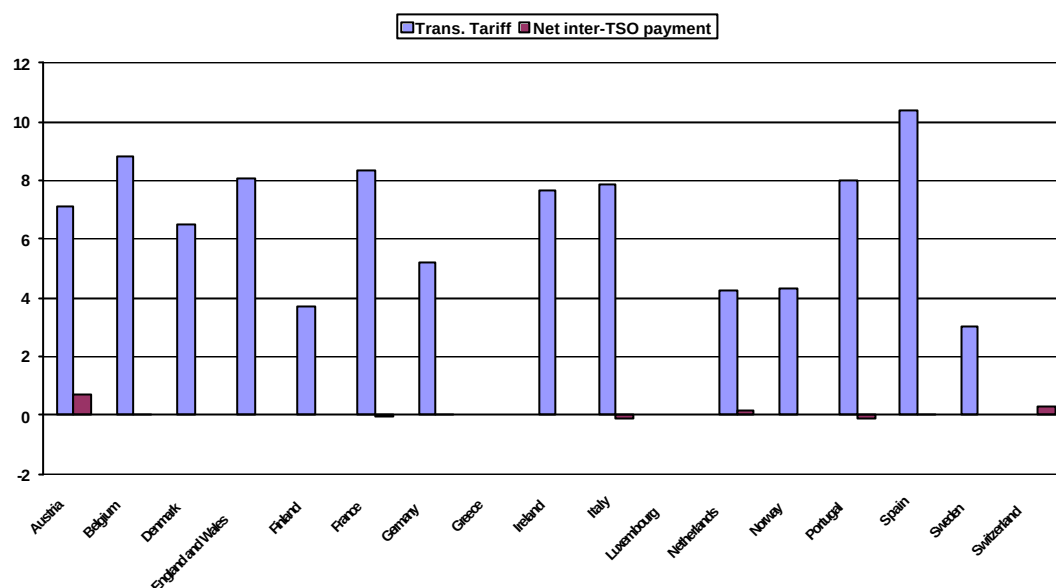


Figure 2.5-3: Comparison between the transmission tariff and the net inter-TSO payment for 17 European countries¹⁰.

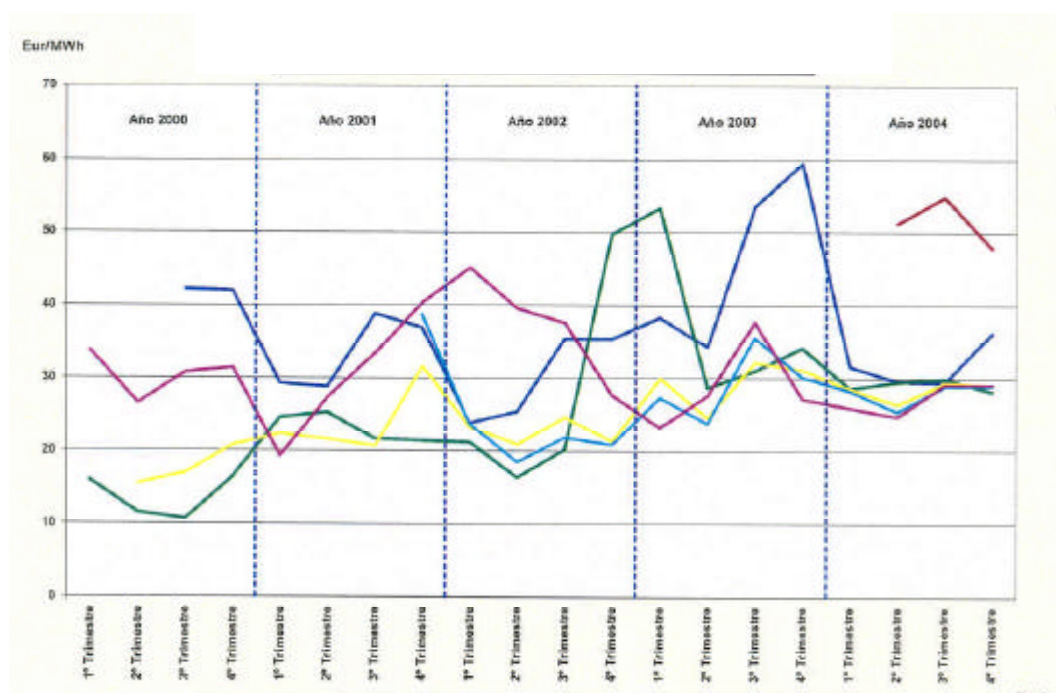


Figure 2.5-4: Evolution of average quarterly energy prices of several Power Exchanges belonging to Europex from January 2000 to November 2004.

¹⁰ Figures are expressed in €/MWh. Nordel Countries, Ireland and the UK did not participate in the study. Unit inter-TSO payments for Italy and France are quite small because these are large countries. Despite having to pay large net inter-TSO payments in absolute terms, the net inter-TSO payment per MWh of energy produced or consumed in these two countries is very small. No data were available for Greece and Luxembourg.

2.5.3.3 Locational Signals affecting the decision by a CCGT on where to install

This subsection compares the locational economic signals corresponding to two different areas in the transmission network of a fictive system where a potential investor might consider installing a new combined cycle gas turbine (CCGT) power plant. The first location L1 is close to a liquefied natural gas (LNG) regasification facility on the coast. The second location L2 is close to a main load centre. In both cases, the power plant is meant to meet the demand growth in the aforementioned load centre.

The most important locational factors can be grouped into three categories:

1. Locational signals related to the impact of losses, both in the electricity and the gas networks. Network constraints are ignored here since systematic congestions are infrequent in well developed networks. Most systematic congestions are usually removed by new network reinforcements in the short or medium term since they typically have a significant impact on the system operating costs.
2. Those signals related to the cost of new network facilities -both for gas and electricity- that are built because of the installation of new agents in the system.
3. Other locational signals that may exist because of various reasons. In this case we have to consider signals derived from the effect of the altitude over sea level on the useful generation capacity of the CCGT power plants.

With the purpose of facilitating the comparison of both options, the power plant has been assumed to provide the same amount of usable power to the system regardless of whether it is installed in one place or the other. Table 2.5-2 shows the effect of transmission losses, the provision of ancillary services and the altitude over sea level on the total generation capacity that is needed to supply an increment in demand of 384 MW in the load centre. In both cases, the power plant must have the same useful generation capacity (see first row of Table 2.5-2, 'Useful Power Delivered'), which equals the assumed increase in the power demand in the consumption centre. This figure is then modified to take into account the different increments in electricity transmission losses that the system incurs when the same extra amount of usable power is produced in the two locations (see second row of Table 2.5-2, 'Net Power Output'). Including the ancillary services that each power plant must provide leads to a 4% increase in its generation capacity (see the third row in Table 2.5-2 'Nominal Capacity'). Finally, we consider the effect of the altitude of each location. Whereas the plant near the LNG facility is located at sea level, the power plant close to the main load centre is supposed to be located considerably higher thus experiencing a reduction of 5% in its total generation capacity (20 MW for a plant with a nameplate generation capacity of 400MW). Therefore, in order to be able to produce 401.1 MW, i.e. the required nominal capacity of the plant in L2, the nameplate capacity of the power plant must be 422.2 MW. On the other hand, the power plant in L1 does not experience

any efficiency losses related to its location because the effect of the altitude is negligible there (see the last row in Table 2.5-2 'Nameplate Capacity').

Table 2.5-3 shows, for both locations, the different cost components involved in meeting an increase of 384 MW in the demand for electricity of the consumption centre close to L2. The useful life of the power plant has been assumed to be 20 years and the cost of capital per annum 6% real. The price of natural gas has been assumed to be 17 €/MWh for the entire time horizon of the study (20 years) and the thermal efficiency of the power plant has been assumed to be 50% in both cases. We estimate that in both locations the power plant would be functioning 5000 hours a year at its full capacity. The per-unit investment cost of the considered CCGT power plant is 1,208 M€ per megawatt of installed capacity. This figure includes the cost of some dedicated facilities assumed to be needed to connect the power plant to the main electricity and gas systems (both the electrical transmission line and the pipeline connecting the plant to the substation and the main pipeline system, respectively, have been supposed to be 40 km long).

Electricity transmission losses and the provision of ancillary services have an impact on the amount of energy produced as well as on the cost of construction of the power plant. The power plant must produce some extra amount of energy in order to supply the transmission losses and ancillary services allocated to it. This amount varies with the location of the plant. Moreover, the required capacity of the generation facility also depends on the loss of capacity because of transmission losses and ancillary services, see Table 2.5-2.

The altitude over sea level has an effect on the required nameplate capacity of a CCGT power plant because it limits the maximum power output of the generator. Therefore, if this power output is prescribed from the outset, the higher the power plant is located above sea level the larger the nameplate capacity of the power plant must be. Consequently, all other factors being equal, generation investment costs increase with the altitude over sea level. Since the final amount of energy produced and the fuel consumption are the same regardless of the altitude, production costs are therefore not affected by this factor.

Regarding the losses incurred when transporting the gas from the point of injection to the place where it is consumed, it has been assumed that a fraction of the transported gas is lost. The assumed percentage can only be regarded as an estimate for the purpose of this example, since we do not know about consistent statistics on the relative importance of losses in gas pipeline systems. No gas losses are considered for the plant located at L1, since the plant is very close to the LNG regasification facility. On the other hand, losses incurred when the power plant is located near the load center at L2 are assumed to account for 0.7% of the total gas consumed. We have taken into account both the price of gas and the thermal efficiency of the power plant in order to compute the cost of these losses, which affect the production cost of the power plant.

As for the locational signals related to the cost of those electric transmission lines that the new power plant must be held responsible for, we have only considered new lines here. The cost of those lines that already exist will probably be allocated according to some ‘average’ method which results in a socialization of their cost thus not giving raise to any locational signal. In order to assign the cost of new lines the novel method presented in section 2.5.1 has been used, with the following simplifying assumptions:

- The flow pattern in the transmission grid is such that area L1 exports power to area L2 where the large consumption centre is located.
- Lines connecting areas L1 and L2 are already congested or close to being congested.

Then, installing a new power plant in L1 to serve the load located near L2 would imply that a new line reinforcing the connection between both areas must be built. If the new CCGT plant is the only new generator using the line and if it is assumed a split of the cost of transmission lines of 50% to generation and 50% to load, application of the method explained in 4.4 would result in the new plant at L1 having to pay half the cost of the line (the entire fraction of the cost to be recovered from generation). The line has been assumed to be a typical 400 KV line with a capacity of 100MW and a cost per kilometer of line of 0.35 M€. If the length of the line is assumed to be 550 km, the annual transmission charge to the plant at L1 would be 7.9 M€. This figure is of the same order of magnitude as the locational signals that result from the existence of losses in the transmission grid or from the effect of the altitude on the generation capacity of the power plant. Therefore transmission charges in this example could influence the decision on whether to install the plant in one place or the other.

Alternatively, we could have used some ‘average use’ method like AP to allocate the cost of both the old and the new lines. In this case, the difference between total transmission charges in both locations (L1 and L2) would probably not have been very large and locational signals related to the allocation of the cost of electric transmission facilities would have been weaker than those considered in this case example.

Likewise, the costs of the infrastructures of the gas grid that are needed by the power plant at L2 have to be accounted for. Here we shall assume that the gas consumed by the CCGT plant at L2 is coming from L1 through a single pipeline whose cost is shared with gas consumers. It has been supposed that the infrastructures that are used to transport gas from L1 to L2 include 90 km of a 20’ pipeline, 450km of a 26’ pipeline and two compression stations, both in the 26’ pipeline. Given the amount of gas consumed by the power plant and the typical amount of gas transported annually through the different gas pipelines, it has been estimated that a CCGT power plant at L2 would be using between 10 and 15% of these infrastructures on average. This results in an annual charge of 3 M€ for the use of gas grid infrastructures by the plant at L2. Per-unit construction costs have been taken from (SMI 2005).

	Close to Main Load Centre (Loc. 1)	Close to an LNG facility on the Coast (Loc. 2)
Useful Power Delivered [MW]	384	384
Net Power Output [MW] <i>(Transmission Losses)</i>	385.1	407.8
Nominal Capacity [MW] <i>(Ancillary Services)</i>	401.1	424.8
Nameplate Capacity [MW] <i>(Altitude over Sea Level)</i>	422.2	424.8

Table 2.5-2: Impact of different factors on the total generation capacity needed to supply a 384 MW load, located close to a main consumption centre, from two different locations, one close to the load centre and the other one close to an entry point for LNG

	Market Value of Energy Produced [M€]		Investment in Generation Facilities [M€]		Investment in Transmission Facilities [M€]	
	Loc. 2	Loc. 1	Loc. 2	Loc. 1	Loc. 2	Loc. 1
Annual Impact of Electricity Transmission Losses	0.2	4.3	0.1	2.4	-	-
Annual Impact of the Provision of Ancillary Services	2.9	3.1	1.6	1.7	-	-
Annual Impact of the Altitude over Sea Level	-	-	2.1	0	-	-
Annual Impact of Losses in the Gas Pipeline System	0.5	0	-	-	-	-
Annual Impact of the Use of the Electricity Grid Infrastructures	-	-	-	-	0	7.9
Annual Impact of the Use of the Gas Grid Infrastructures	-	-	-	-	3	0

Overall Annual Value						
	72.7	76.5	42	42.3	3	7.9
(total annual value of the energy produced or total annual cost of the facilities built)						

Table 2.5-3: Comparison of the cost savings involved in supplying a 384 MW load located close to a main load center

Locational signals related to the existence of congestion in the grid have not been taken into account in this example since congestion is typically infrequent within any system where the grid is sufficiently meshed. Assuming the lines between L1 and L2 become congested in the future and different energy prices are computed for both areas, locational signals resulting from this congestion would amount to the difference in energy prices between both areas times the total power production of the power plant over the period of time during which the lines are congested. Considering a difference in energy prices of 10% of the average system price (supposed to be 36 €/MWh) and assuming the line is congested 30% of the time, annual revenues from the sale of energy would be about 2 M€ higher at L2 than at L1. This figure is comparable to the difference between the annualised effect of the altitude on the investment cost of the power plant in both locations (which is also about 2 M€) but clearly smaller than the difference between the annual value of the losses assigned to the generator in both cases (which is about 6.4 M€). However, numbers depend critically on the assumptions that have been made.

From the numbers shown in Table 2.5-3 one can conclude that the net present value of the cost savings that are obtained annually when installing the power plant close to the consumption centre, instead of installing it close to the LNG regasification facility, would amount to 9 M€. This figure is the difference between the total annual costs associated with each one of the two locations, see last row in Table 2.5-3. Numbers in this row correspond to either the total annual value of the energy produced in the two locations or to the total annual cost of the facilities built in the two cases ($9 = (76.5 - 72.7) + (42.3 - 42) + (7.9 - 3)$). Note that, when computing these numbers, we have assumed a certain pattern of electricity flows in the grid that remains the same during the entire time horizon of the study. Therefore, transmission charges and losses also remain the same. The breakdown of costs in Table 2.5-3 shows the relative importance of each one of the locational factors for this case example. Obviously, these results may be different for other case examples, but the approach to be followed in the evaluation should be basically the one that has been shown here.

2.6 Conclusions

After liberalization, investment decisions for generation and network infrastructures are no longer in the same hands. Therefore, the uncertainty of the network planner

when computing the optimal expansion of the grid increases. In addition to this, private grid promoters face the risk of not being able to recover the investment cost of the line or lines that they build.

We have discussed the different paradigms that exist for regulating the expansion of the grid. Based on this discussion, we have characterized the best approach to this problem. This approach may be termed “hybrid” since it leaves room both to the public and the private initiative. However, centralized planning must be given priority over the initiatives by private investors since only a centralized regulated institution may actually promote an adequate development of the transmission system.

Both the public and the private initiative must be given some role in the expansion of the grid. However, centralized expansion planning must take precedence over merchant lines and investments at risk.

Obviously, the features of the chosen approach must adapt to the particular situation in each regional market. Here, we have analyzed how the hybrid approach can be adapted to best fit the characteristics of three relevant regional markets, namely the IEM of the European Union, the regional markets being created in the USA and the Regional Electricity Market in Central America. In order to attract investment in new transmission facilities, the set of rules implemented in any regional market to regulate the expansion of the grid must reduce, to the extent possible, the regulatory risk perceived by promoters.

In those regions where the transmission grid is clearly underdeveloped and central planning is unlikely to succeed in getting all the required lines built, private promoters should be encouraged to undertake some of these investments by implementing efficient pricing schemes (nodal or zonal marginal pricing). Providing owners of (physical) transmission rights with extra benefits, in addition to the collection of congestion rents, should also be considered.

Accepting or rejecting a particular investment or a set of them (investment plan) requires applying a set of rules in order to determine whether the proposed reinforcement is beneficial for the system and how it compares to alternative reinforcements. This set of rules is commonly called Regulatory Test. Adequately designing the Regulatory Test is central to the successful implementation of any mechanism for the expansion of the grid. The regulatory test must be able to evaluate the benefits that each one of the investment proposals will yield in the time horizon considered. The test to be applied to regulated reinforcements must be different from that used for merchant projects or investments at risk since private promoters and not consumers pay the cost of the latter ones. In any case, negative benefits or losses incurred by grid users as a result of the construction of a line or lines should be taken into account when deciding about their construction.

Under the new regulatory paradigm, investments in grid facilities should be aimed at maximizing the net aggregate benefit of market agents in the region. Both positive and negative benefits (losses) incurred by market agents as a consequence of the construction of each line must

be taken into account. Contrary to the case of regulated investments, the regulatory test to be applied in order to decide on the construction of lines financed by private investors must not consider the cost of the investment. Starting from a set of possible reinforcements previously identified, the SO must propose those that are expected to result in the highest increase in the social welfare. The regulator must check that the investment proposals by the SO are beneficial for the system. This entity may also ask the SO to evaluate regulated reinforcements proposed by third parties, which should be compared to the ones already proposed by the SO itself. Investments at risk should be approved as long as they are not detrimental to the system and no alternative regulated project has been already proposed.

The ensuing loss of efficiency due to the unbundling of investment decisions in generation, transmission and distribution may be partly compensated by coordinating locational economic signals at the different interfaces. Efficient investment in generation can be encouraged by the use of nodal energy prices and transmission tariffs with locational content. These signals should be designed on the basis of responsibility on the incurred network costs. The usefulness of network-related locational signals for transmission investment critically depends on the adopted regulatory approach for transmission and, in particular, on which is the entity in charge of the expansion of the grid: the regulatory authority, the transmission system operator, merchant investors or the network users themselves. Different regulatory paradigms for the expansion of the grid should result in different remuneration schemes. Hence, the economic incentives for the construction of new lines will probably differ from one organizational structure to another. Other useful locational signals of non economic nature, whose purpose is to decrease the level of uncertainty that both potential generation investors and transmission network planners must face, involve providing reliable information on the future investment decisions (generators) and the expected network conditions (System Operator).

Economic locational signals that are based on cost causality are always useful to guide the siting in the network of new generation and load facilities. The issue here is whether these locational signals may or may not be strong enough to have any actual impact on the location of new network users or even on the decision to retire a power plant or a load from a given site. The chapter has provided criteria and numerical examples showing the nature of these signals and what can and cannot be expected from their application. A novel scheme to design transmission charges with a special emphasis on cost causality has been proposed. Locational signals may be also useful at the interface between transmission and distribution networks in order to minimize the distortions that result from having separate network operators under widely different regulatory rules.

Transmission facilities that are built under central regulated planning must be paid according to the incurred costs or the winning bid in public auctions. Some moderate performance-based incentives related to losses and reliability may be included in the remuneration scheme. Private owners of grid facilities must earn the marginal value of their lines in the dispatch.

In order to encourage efficient investment in new generation, energy prices must be efficient (they must include the marginal effect of losses and congestion). Additionally, transmission charges to be paid by new network users must vary, from one point of the grid to another, according to the extra grid costs that the system incurs when a new generator or load is installed in each node or area. Relevant information must be continuously exchanged between market agents and the central planner.

Coordinating signals should be sent so that the joint development of transmission and distribution grids is efficient. These signals are specially important at the interface between both types of grids.

2.6.1 Interaction between the regulation of the expansion of the grid and that of other aspects of transmission regulation

From the discussion above and the arguments laid out in this chapter one can immediately conclude that the regulation of investments in new transmission facilities cannot be considered in isolation of other aspects of the regulation of electricity markets in general and the regulation of transmission in particular. As it has just been mentioned, different regulatory paradigms for the expansion of the grid should result in different remuneration schemes. Thus, the revenues of grid owners, and therefore the total amount to be collected from grid users, must be determined as a function of the existing network assets and the regulatory paradigm adopted in each case. Chapter 3 is concerned with the computation of the grid charges to be paid by network users where the total amount to be recovered from network charges must be equal to the allowed remuneration of the transmission activity.

At the same time, network charges to be paid by market agents may condition their decision to install a new load or generator in one part of the grid or another if these charges exhibit some kind of locational differentiation. The pattern of generation and demand in the system will in turn be one of the main inputs to the process of computation of the optimal development of the transmission grid. Thus, we can see how the expansion of the grid and the tariff setting process are interdependent and cannot be considered in a completely isolated manner.

The expansion of the grid is also related to the third major aspect of transmission regulation, namely congestion management. In fact, if an adequate mechanism to manage congestion in the grid is in place, locational differences in energy prices among different nodes or areas should be one of the major indicators of the need to reinforce the existing grid. In this sense, chapter 4 proves that, under certain conditions, the marginal value of increasing the capacity of a line is a function on the amount of flow on the line and the difference in energy prices between both ends of this line. For merchant investors, the sole driver behind the decision to invest in a new line is the amount of congestion rents that the promoter expects to extract from the market.

The decision to build a new line or reinforce an existing one may probably alter the pattern of flows in the grid and, hopefully, reduce congestion. In fact, the construction of a new line represents a major threat to private investors, who may see their revenues from the congestion rents they earn (those corresponding to the merchant lines they own) sharply reduced. The optimal development of the grid achieves the appropriate balance between the system operational costs (including the congestion cost) and the cost of investing in new lines, see the optimal investment rule in section 2.4 above.

Bibliography

- ACCC (2004). Review of the Regulatory Test for Network Augmentations, Australian Competition and Consumer Commission (ACCC). Web page: <http://www.accc.gov.au/content/index.phtml/itemId/532002/fromItemId/54368>.
- Alguacil, N., A. L. Motto, et al. (2003). "Transmission Expansion Planning: A Mixed-Integer LP Approach." IEEE Transactions on Power Systems **18**(3): 1070-1077.
- Anderson, K. P. and A. McCarthy (1999). "Transmission pricing and expansion methodology: lessons from Argentina." Utilities Policy **Vol. 8**(No. 4): 199-211.
- Baldick, R. and E. Kahn (1993). "Network Costs and the Regulation of Wholesale Competition in Electric Power." Journal of Regulatory Economics **5**: 367-384.
- Binato, S. and G. C. Oliveira (1995). A heuristic procedure to cope with multi-year transmission expansion planning. IEEE/KTH Stockholm Power Tech Conference, Stockholm, Sweden.
- Binato, S., M. V. F. Pereira, et al. (2001). "A New Benders Decomposition Approach to Solve Power Transmission Network Design Problems." IEEE Transactions on Power Systems **16**(2): 235-240.
- Brown, R. E. (2002). Electric Power Distribution Reliability. New York, Marcel Dekker, Inc.
- Bruneekreeft, G. (2003). "Market-Based Investment in Electricity Transmission Networks: Controllable Flow." CMI electricity project working paper. Web page: <http://www.econ.cam.ac.uk/electricity/publications/wp/index.htm>.
- Bushnell, J. and S. Stoft (1996). "Grid Investment: Can a Market do the Job?" The Electricity Journal **9**(1): 74-79.
- Bushnell, J. B. and S. E. Stoft (1996). "Electric Grid Investment Under a Contract Network Regime." Journal of Regulatory Economics **10**: 61-79.
- Carrillo-Caicedo, G. and I. Pérez-Arriaga (1995). Optimal reconfiguration of distribution networks for a diversity of regulatory frameworks. IEEE/KTH Stockholm Power Tech Conference, Stockholm, Sweden.
- CEER (2003). CEER Guidelines on Regulatory Control and Financial Reward for Infrastructure, Council of European Energy Regulators (CEER). Web page: <http://www.ceer-eu.org/>.

- Chao, H.-P. and S. Peck (1996). "A Market Mechanism for Electric Power Transmission." Journal Of Regulatory Economics **10**: 25-29.
- Chisari, O. O., P. Dal-Bó, et al. (2001). "Hight-tension electricity network expansions in Argentina: decision mechanisms and willingness-to-pay revelation." Energy Economics **23**: 697-715.
- Choi, J., A. R. A. El-Keib, et al. (2005). "A Fuzzy Branch and Bound-Based Transmission System Expansion Planning for the Highest Satisfaction Level of the Decision Maker." IEEE Transactions on Power Systems **20**(1): 476-484.
- Coase, R. H. (1990). The firm, the market and the law, The University of Chicago Press.
- Contreras, J. and F. F. Wu (1999). "Coalition Formation in Transmission Expansion Planning." IEEE Transactions on Power Systems **14**(3): 1144-1152.
- Contreras, J. and F. F. Wu (2000). "A Kernel-Oriented Algorithm for Transmission Expansion Planning." IEEE Transactions on Power Systems **15**(4): 1434-1440.
- Dismukes, D. E., R. F. C. III, et al. (1998). "Capacity and economies of scale in electric power transmission." Utilities Policy **7**(3): 155-162.
- EOR (2005). Reglamento del Mercado Eléctrico Regional (RMER). Libro III del RMER. De la Transmisión., Ente Operador Regional (EOR).
- EP (2003). Regulation (EC) No 1228/2003 of the European Parliament and of the council of 26 June 2003 on conditions for access to the network for cross-border exchanges in electricity. Web page: http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_176/l_17620030715en00010010.pdf: 10.
- Escobar, A. H., R. A. Gallego, et al. (2004). "Multistage and Coordinated Planning of the Expansion of Transmission Systems." IEEE Transactions on Power Systems **19**(2): 735-744.
- Fang, R. and D. J. Hill (2003). "A New Strategy for Transmission Expansion in Competitive Electricity Markets." IEEE Transactions on Power Systems **18**(1): 374-380.
- Faria, H., S. Binato, et al. (2005). "Power Transmission Network Design by Greedy Randomized Adaptive Path Relinking." IEEE Transactions on Power Systems **20**(1): 43-49.
- Gans, J. S. (2001). "Regulating Private Infrastructure Investment: Optimal Pricing for Access to Essential Facilities." Journal of Regulatory Economics **20**(2): 167-189.

- Haffner, S., A. Monticelli, et al. (2001). "Specialised branch-and-bound algorithm for transmission network expansion planning." IEE Proceedings on Generation, Transmission and Distribution **148**(5): 482-488.
- Helm, D. (2003). "Auctions and Energy Networks." Utilities Policy **11**(1): 21-25.
- Joskow, P. and J. Tirole (2002). Transmission Investment: Alternative Institutional Frameworks. Wholesale Markets for Electricity. November 22-23., Toulouse, France.
- Joskow, P. L. (2005). Making Transmission Owners Accountable. Roundtable on "Investments in Transmission", Conference on "The Economics of Electricity Markets", Toulouse.
- Joskow, P. L. (2005). Patterns of Transmission Investment. Electricity Infrastructure Investment Workshop. To be published as a chapter in the book 'Investments in Generation and Transmission in Electricity Liberalized Regimes', CRE, Paris, France.
- Joskow, P. L. (2005). "Transmission Policy in the United States." Utilities Policy **Vol. 13**: 95-115.
- Joskow, P. L. and J. Tirole (2005). "Merchant Transmission Investment." Journal of Industrial Economics **Vol. 53**(2): 233-264.
- Kandil, M. S., S. M. El-Debeiky, et al. (2001). "Hybrid mathematical and rule-based system for transmission network planning in open access schemes." IEE Proceedings on Generation, Transmission and Distribution **148**(5): 455-462.
- Keller, K. and J. Wild (2004). "Long-term investment in electricity: a trade-off between co-ordination and competition?" Utilities Policy **12**(4): 243-251.
- Khator, S. K. and L. C. Leung (1997). "Power Distribution Planning: A Review of Models and Issues." IEEE Transactions on Power Systems **12**(3): 1151-1158.
- Latorre, G., R. D. Cruz, et al. (2003). "Classification of Publications and Models on Transmission Expansion Planning." IEEE Transactions on Power Systems **18**(2): 938-946.
- Latorre-Bayona, G. and I. Pérez-Arriaga (1994). "CHOPIN, A heuristic Model for Long Term Transmission Expansion Planning." IEEE Transactions on Power Systems **9**(4): 1886-1894.
- Léautier, T.-O. (2000). "Regulation of an Electric Power Transmission Company." The Energy Journal **21**(4): 61-92.

- Léautier, T.-O. (2001). "Transmission Constraints and Imperfect Markets for Power." Journal of Regulatory Economics **19**(1): 27-54.
- Littlechild, S. C. and C. J. Skerk (2004). Regulation of transmission expansion in Argentina Part I: State ownership, reform and the Fourth Line, Working paper ep61. CMI electricity project. Department of Applied Economics. University of Cambridge. Web page: <http://www.econ.cam.ac.uk/electricity/publications/wp/:75>.
- Littlechild, S. C. and C. J. Skerk (2004). Regulation of transmission expansion in Argentina Part II: Developments since the Fourth Line, Working paper ep62. CMI electricity project. Department of Applied Economics, University of Cambridge. Web page: <http://www.econ.cam.ac.uk/electricity/publications/wp:75>.
- McDaniel, T. (2003). "Auctioning access to networks: evidence and expectations." Utilities Policy **11**(1): 33-38.
- ME (2005). Plan Estratégico para Modernización del Marco Regulatorio, Final report prepared by Mercados Energéticos for Osinerg (Peruvian Energy Regulatory Agency). Web page: <http://www.osinerg.gob.pe/>.
- Newbery, D. M. (2003). "Network capacity auctions: promise and problems." Utilities Policy **11**(1): 27-32.
- Outhred, H. R. and R. J. Kaye (1996). Incorporating Network Effects in a Competitive Electricity Industry: an Australian Perspective. Issues in Transmission Pricing and Technology. M. Einhorn and R. Siddiqui, Kluwer Academic Publishers: 207-228.
- Peco, J. P. (2001). Modelo de Cobertura Geográfica de una Red de Distribución Eléctrica. Doctoral Thesis. Instituto de Investigación Tecnológica. Madrid (in Spanish), Universidad Pontificia Comillas.
- Pereira, M. V. F. and B. G. Gorenstin (1996). Transmission pricing methods. Seminar: Remuneration and Pricing of Power Services, Bogotá, Colombia.
- Pereira, M. V. F., M. F. McCoy, et al. (2000). "Managing risk in the new power business." IEEE Comput. Applicat. Power: 18-24.
- Pérez-Arriaga, J. I. (2002). Methodology for cross-border tarification in the Internal Electricity Market of the European Union. Power Systems Computing Conference (PSCC), Seville.
- Pérez-Arriaga, J. I. (2005). Libro Blanco sobre la reforma del marco regulatorio de la generación eléctrica en España. Madrid, Ministry of Industry, Tourism and Commerce 579.

- Pérez-Arriaga, J. I. and L. Olmos (2003). Extension of the project on cost components of cross border exchanges of electricity, Directorate-General for Energy and Transport / European Commission. Web page: [http:// europa.eu.int/ comm/ energy/ electricity/ publications/ index_en.htm](http://europa.eu.int/comm/energy/electricity/publications/index_en.htm).
- Pérez-Arriaga, J. I., L. Olmos, et al. (2002). Cost components of cross border exchanges of electricity, Directorate-General for Energy and Transport / European Commission. Web page: http:// europa.eu.int/ comm/ energy/ electricity/ publications/ index_en.htm.
- Pérez-Arriaga, J. I. and L. Omos (2005). Compatibility of Investment Signals in Distribution, Transmission and Generation. Electricity Infrastructure Investment Workshop. To be published as a chapter in the book 'Investments in Generation and Transmission in Electricity Liberalized Regimes', CRE, Paris, France.
- Pérez-Arriaga, J. I., T. G. S. Román, et al. (1987). Situación del estado del arte en la planificación de redes de transporte de energía eléctrica, Report prepared for Red Eléctrica de España (in Spanish).
- Pérez-Arriaga, J. I. and F. J. Rubio (1995). "Marginal pricing of transmission services: An analysis of cost recovery." IEEE Transactions on Power Systems **Vol. 10**(No. 1): 546-553.
- Pérez-Arriaga, J. I. and Y. Smeers (2003). Guidelines on tariff setting. Transport Pricing of Electricity Networks. F. Levesque, Kluwer Academic Publishers.
- Rau, N. S. and F. Zeng (2004). "Adequacy and Responsibility of Locational Generation and Transmission- Optimization Procedures." IEEE Transactions on Power Systems **19**(4): 2093-2101.
- Romero, R., C. Rocha, et al. (2005). "Constructive heuristic algorithm for the DC model in network transmission expansion planning." IEE Proceedings on Generation, Transmission and Distribution **152**(2): 277-282.
- Romero, R., C. Rocha, et al. (2003). "Analysis of heuristic algorithms for the transportation model in static and multistage planning in network expansion systems." IEE Proceedings of Generation, Transmission and Distribution **150**(5): 521-526.
- Rubio, F. J. (1999). Metodología de Asignación de Costes de la Red de Transporte en un Contexto de Regulación Abierta a la Competencia. Doctoral Thesis. Escuela Técnica Superior de Ingeniería. Madrid (in Spanish), Universidad Pontificia Comillas.

- Saphores, J.-D., E. Gravel, et al. (2004). "Regulation and Investment under Uncertainty: An Application to Power Grid Interconnection." Journal of Regulatory Economics **25**(2): 169-186.
- Schweppe, F., M. Caramanis, et al. (1988). Spot pricing of electricity, Kluwer Academic Publishers.
- Shrestha, G. B. and P. A. J. Fonseka (2004). "Congestion-Driven Transmission Expansion in Competitive Power Markets." IEEE Transactions on Power Systems **19**(3): 1658-1665.
- Silva, E. L. d., H. A. Gil, et al. (2000). "Transmission Network Expansion Planning Under an Improved Genetic Algorithm." IEEE Transactions on Power Systems **15**(3): 1168-1175.
- Smeers, Y. (2005). Long term locational prices and investment incentives in the transmission of electricity. Electricity Infrastructure Investment Workshop. To be published as a chapter in the book 'Investments in Generation and Transmission in Electricity Liberalized Regimes', CRE, Paris, France.
- SMI (2005). Orden ITC/102/2005 por la que se establece la retribución de las actividades reguladas del sector gasista., Spanish Ministry of Industry, Turism and Trade. Web page: <http://www6.mityc.es/energia/archivos/Orden%20Retribuciones%202005.pdf>.
- Stern, J. and R. Turvey (2003). "Auctions of capacity in network industries." Utilities Policy **11**(1): 1-8.
- Stoft, S. (2005). Problems of Transmission Investment in a Deregulated Power Market. Electricity Infrastructure Investment Workshop, CRE, Paris, France.
- Styczynski, Z. A. (1999). Power network planning using game theory. 13th Power Systems Computation Conference, Trondheim.
- Sun, H. and D. C. Yu (2000). A multiple-objective optimization model of transmission enhancement planning for independent transmission company (ITC). Proceedings IEEE Power Engineering Society Summer Meeting.
- Turvey, R. (2000). "Infrastructure access pricing and lumpy investments." Utilities Policy **Vol. 9**(No. 4): 207-218.
- Vázquez, C., M. Rivier, et al. (2005). Adecuación del marco regulatorio para la integración regional. Madrid, Report prepared for the Peruvian energy regulator (OSINERG-GART) by the Instituto de Investigación Tecnológica from the Universidad Pontificia Comillas. In Spanish: 123.

- Vogelsang, I. (2001). "Price Regulation for Independent Transmission Companies." Journal of Regulatory Economics **20**(2): 141-165.
- Yarrow, G. (2003). "Capacity auctions in the UK energy sector." Utilities Policy **11**(1): 9-20.

Annex 1: Mathematical formulation of the problem of centrally planning the expansion of the grid

Centralized planning of the expansion of the grid in a traditional regulatory setting

The objective of transmission planning in the traditionally regulated environment is to maximize the global social welfare resulting from the expansion of the grid. Given that generation was traditionally planned together with transmission, maximizing the social welfare was equivalent to minimizing the cost incurred when supplying the system load while complying at the same time with certain technical and reliability criteria. In order to do so one has to take into account the investment, maintenance and fuel costs of the power plants, as well as the investment, operation and maintenance costs of the grid.

When deciding how to supply the demand, it was possible to consider the joint cost of the transmission and generation involved in a number of different possible expansion plans and to choose the most efficient alternative. The central planner had to consider a number of uncertain inputs to the planning process, which basically were the demand forecast, the estimation of hydrological conditions, the costs of fuel and capital, and the expected availability of the already existing and new generation and transmission equipment. From all this information the central planner had to decide where and how much to invest in new transmission and generation facilities. In other words, contrary to what happens now that competition has been introduced, the central planner knew about the generation expansion plans and could modify them.

It is possible to write the mathematical formulation of this optimization problem. For the sake of simplicity, a “static” optimization model is presented, where the optimal mix of generation and transmission is determined for a single given future year, ignoring the trajectory of investments along the planning horizon. The objective function to be minimized is:

$$\sum_{l \in \Omega} \sum_{k=1}^{K_l} h_{lk} \sum_{m=1}^{M_{lk}} X_{lkm} + \sum_{n \in \Psi} \sum_{w=1}^{W_n} h_{nw} Y_{nw} + D \sum_{s \in \Phi} p_s \sum_{i=1}^N \sum_{c=1}^{E_i} c_{ie} g_{sie} + CEN \sum_{i=1}^N r_{s,i} \quad (0.3)$$

where Ω is the subset of branches (i.e., corridors) where new investments are allowed, h_{lk} the investment cost of a line of type k in corridor l , K_l is the number of different types of line (220 or 400 KV, single or double circuit, number of conductors per phase, overhead or underground) that are allowed as expansion options in corridor l , M_{lk} is

the maximum number of new lines of type k that can be added to corridor l , X_{lkm} is a binary variable whose value is 1 if line m of type k is built in corridor l and it is 0 otherwise, Ψ is the subset of nodes where new investment in generation can be located, h_{nw} is the investment cost of a power plant of type w at node n , Y_{nw} is a binary variable that is 1 if a plant of type w is built at node n and it is 0 otherwise, D is the duration of the planning horizon, Φ the subset of considered production cost scenarios, p_s is the probability of occurrence of scenario s , N represents the total number of areas, E_i is the total number of generation blocks in area i , c_{ie} corresponds to the variable production cost of generation block e in area i , g_{sie} is the power output of generation block e in area i for scenario s , CEN is the cost of unserved energy and $r_{s,i}$ is the unserved power in area i for scenario s .

The objective here is to minimize the total cost of supplying the demand, while taking into account the fixed costs of the newly installed generation, the variable power production costs and the transmission costs. Ohmic losses can, for instance, be taken into account by adding a fictitious load at both ends of each line with every load being equal to one half of the ohmic loss of the line.

A number of constraints have to be considered: maximum total number of lines that can be installed in each corridor (equation (0.4)), maximum number of lines of a given type that can be installed in each corridor (intrinsic to the formulation), maximum total investment per corridor (equation (0.5)), or for the entire network (equation (0.6)), maximum output for every thermal generation block and for the equivalent hydroelectric unit in each area (equation (0.7)), maximum unserved energy in every area (equation (0.8)), electric network equations (equations (0.9) and (0.10)) and the transmission capacity limit for every corridor (equation (0.11)). The mathematical expression of these constraints follow:

$$\sum_{k=1}^{K_l} \sum_{m=1}^{M_{lk}} X_{lkm} \leq L_l \quad l \in \Omega \quad (0.4)$$

$$\sum_{k=1}^{K_l} h_{lk} \sum_{m=1}^{M_{lk}} X_{lkm} \leq I_l \quad l \in \Omega \quad (0.5)$$

$$\sum_{l \in \Omega} \sum_{k=1}^{K_l} h_{lk} \sum_{m=1}^{M_{lk}} X_{lkm} \leq TI \quad (0.6)$$

$$0 \leq g_s \leq \overline{g_s} \quad (0.7)$$

$$0 \leq r_s \leq d_s \quad (0.8)$$

$$-Sf_s + g_s + r_s = d_s + \Lambda_s \quad (0.9)$$

$$f_s - DS^T q_s = 0 \quad (0.10)$$

$$|f_s| \leq \overline{f_s} \quad (0.11)$$

where I_l is the maximum investment in corridor l , L_l is the maximum number of lines to be installed in corridor l , TI means the maximum total investment in the network, $\Lambda = \frac{1}{2} S^T l$ is the vector representing the ohmic losses for each one of the areas, l represents the vector of ohmic losses for each one of the corridors, g is the vector of active outputs from every thermal generation block and every hydroelectric equivalent unit in each area, r is the vector of area load curtailments and $\bar{g}$ is the vector of maximum active power outputs from every thermal generation block and every hydroelectric equivalent unit in each area. This last vector is a function of the investment variables in generation Y_{nw} .

Given that the construction times and the investment costs of power plants used to be much higher than those for transmission lines, generation planning normally took precedence over transmission planning. Once the expansion of generation had been planned, investments in the transmission grid were decided considering the investments in generation as given. In this case, the mathematical statement of the problem can be simplified since the variables for investment in generation Y_{nw} are now input data.

As for the planning of the distribution grid, the objective remains the same as that used for transmission. The planner must minimize the cost incurred by supplying the forecasted demand while complying with certain reliability criteria. Reliability can be incorporated into the objective function to be minimized, it can be taken into account through reliability constraints or it can be included both in the objective function and the set of constraints. Traditionally, planners developed the distribution grid independently from the transmission grid. Despite the fact that the planning of transmission and distribution grids deserve a specific treatment, some level of coordination between them is also advisable, as it has been explained. Here we leave aside the fact that the volume of distributed generation is growing very fast in many countries. The cost factors that play a role in this optimization process are mainly investment in distribution grids, ohmic losses and the cost of non supplied energy, see (Peco 2001) for details.

Differences and similarities in centralized transmission planning under the traditional and competitive regulatory frameworks

When competition in the generation activity is introduced, the entity in charge of optimizing the development of the grid no longer knows what the evolution of generation in the system will be. This fact adds much more complexity to the task of deciding which lines to build. It is generally agreed that, also within a competitive framework, the ultimate objective of the process of planning the expansion of the grid should be to maximize the global social welfare (see the discussion in section 2.4). At

the same time, generators are competing among them with the purpose of maximizing their profit, like consumers do. Therefore, assuming there is a central network planner, his objective will be to maximize the joint net profit made by generators and consumers because of the utilization of the grid. This means that both the benefits and the losses of any type that would be incurred because of the construction of a line must influence the decision by the planner on whether or not to construct it.

If it is accepted that generation has precedence over transmission -both in the traditional and in the liberalized setting-, then it can be proved the theoretical equivalence of the outcome of transmission planning under both regulations, when the former objective function is replaced by the new one. It is therefore possible to provide a conceptually sound answer to the problem of determining the objective function to be pursued by transmission planning in the new competitive environment. This has practical implications in the definition of the regulatory test, see section 2.4. The simple proof of this statement follows.

Within the traditional approach, investment in transmission and generation is jointly optimized. Since generators are paid their cost of service, the objective is just to maximize the consumers' welfare, i.e. the utility they obtain from the use of electricity minus the cost of producing and delivering electricity. Therefore, the objective function can be simply expressed as:

$$Max \{U(D) - CFG - CVG - CT\} \quad (0.12)$$

where $U(D)$ is the utility function of consuming a demand D , CFG represents the generation fixed costs, CVG are the generation variable costs and CT are the transmission costs (capital plus operation and maintenance costs, which altogether can be considered fixed costs).

Assuming that demand is inelastic in the range of prices that may result from changes in transmission investment, and if generation planning is also prescribed from the outset, transmission planning becomes the typical minimization of generation operation costs via network reinforcements:

$$Min \{CVG + CT\} \quad (0.13)$$

Within the competitive regulatory framework, the entity in charge of transmission planning (the System Operator, typically, under regulatory supervision) must apply the following optimization criterion in order to identify the network reinforcements that must be proposed to the regulator for authorization:

$$Max \{Net\ benefit\ of\ consumers + Net\ benefit\ of\ generators\} \quad (0.14)$$

where the total cost of any justified investment is implicit in the net benefits of market agents since the latter are computed taking into account network charges to be paid by these agents. The following proof shows that the ideal outcome in competitive markets coincides with the one that the traditional approach would render under the same circumstances. Typically, this indicates that the new regulatory setting is efficient.

In a competitive wholesale market, the following expression holds:

$$PD - IG - IVT - RNC = 0 \quad (0.15)$$

where PD is the total payment (at wholesale level) of consumers, IG is the total income of generators (net of any network payments), IVT is the global variable income of the transmission network (based on the application of nodal prices to both consumers and generators) and RNC is the residual network charge of the transmission network (i.e. the part of the total network cost CT that is not recovered by IVT).

The preceding expression allows one to replace the objective function of the maximization problem in the traditional approach by this one that is entirely equivalent:

$$\{U(D) - PD\} + \{IG - CVG - CFG\} + \{IVT + RNC - CT\} \quad (0.16)$$

which shows that the maximization problem in the traditional approach can be replaced by the following equivalent problem in the context of the competitive approach:

$$\text{Max}\{\text{Net benefit of consumers} + \text{Net benefit of generators}\} \quad (0.17)$$

Transmission facilities are assumed to be regulated. Thus, $CT = IVT + RNC$. Note that, embedded in the net benefit of consumers and generators, are the complete payments for any justified investment in transmission facilities.

One must be aware that, in order for this proof to be valid, the entity in charge of the centralized development of the grid in a competitive environment must pursue those grid expansions that maximize the global surplus. The objective of the promoters of transmission reinforcements under other regulatory schemes than the centralized regulated expansion of the grid is not the same as that under traditional cost-based regulation.

3 Transmission tariffs

3.1 Introduction

Transmission tariffs result from the allocation of the regulated revenues of the transmission activity among the network users. Chapter 2 discusses the computation of the total amount to be recovered from network charges (the aforementioned regulated revenue), which depends on the regulatory paradigm that is adopted for the expansion of the grid. Transmission charges to network users must comply with the following basic requirements:

- They must recover the cost of construction, operation and maintenance of the regulated network facilities.
- They must promote the efficient functioning of the system both in the long and the short term, see (Rubio 1999; Rubio and Pérez -Arriaga 2000):
 - o In the long term they must encourage market agents and grid promoters to make optimal investment decisions in generation, load and transmission facilities.
 - o In the short term they must not interfere with efficient short term signals¹ that encourage market agents to make optimal operational decisions.
- They must be perceived as fair (non-discriminatory) by market agents .

Because transmission tariffs must complete the recovery of the cost of any regulated network facility, they must be also used to allocate the cost of network reinforcements. Satisfactorily allocating the cost of new transmission lines may be essential to getting these lines built, especially in regional markets, where several control areas may jointly decide on the construction of lines that are relevant to the trade in the region.

In practice, the cost components considered in transmission tariffs vary much from system to system, as a recent report from the European association of Transmission System Operators (ETSO) highlights, see (Fernández and Remacha 2005). Consequently, the level of transmission tariffs varies widely as well, see (Fernández and Remacha 2005) and (Pérez-Arriaga, Montero et al. 2002). In any case, certain aspects of the computation of transmission tariffs, to be paid by agents within the same market, must be harmonized in order to achieve a level playing field. Agents from different areas within a market should be able to compete under equal conditions.

¹ Efficient short term signals may result either from locationally and temporally differentiated energy prices (nodal prices) or from the combination of a single energy price and separate signals corresponding to losses and congestion in the grid.

Regardless of the factors taken into account in the computation of transmission tariffs and the total amount of regulated revenues, the level of grid access charges should not be left at the discretion of an entity that has interests in non-regulated activities such as generation or retailing. Otherwise, this entity will set inefficiently high charges. High charges may not be able to prevent competitors from entering the market, but these charges may help in creating cross-subsidies by allowing the regulated monopoly to lower access charges in other segments where it does not face tough competition, see (Riechmann 2000).

To sum up, the definition of a method to allocate the cost of the grid is a very delicate matter due to its implications in many different aspects of the functioning of the system. The following sections in this chapter explain the problem of computing transmission tariffs, propose solutions to this problem and discuss the implementation of a tariff system in different regional markets. Section 3.2 carries out a conceptual analysis of the computation of transmission tariffs. Section 3.3 discusses the different methods that exist for the allocation of the cost of the grid. Section 3.4 lays out the solution proposed to this problem in a general context. Section 3.5 extends the solution derived for the reference case to that of several regional markets with a focus on the EU market. Finally, section 3.6 concludes.

3.2 Conceptual analysis of the computation of transmission tariffs

As already explained in appendix 1, nodal prices are optimal short-term economic signals, i.e. they encourage market agents to make optimal decisions in the operation time-frame. Nodal prices are unable to recover more than a small fraction of the total cost of the grid (Pérez-Arriaga and Rubio 1995; Rubio and Pérez-Arriaga 2000). Therefore, we must resort to some other charges so that, once the construction of a regulated line is approved, the owner of the line recovers the whole cost of the investment plus some reasonable rate of return on it. Therefore, the cost of transmission must be recovered from:

1. The net revenue resulting from the application of energy charges in either of the following cases:
 - a. nodal prices are computed,
 - b. a single energy price is computed for the whole system and short-term economic signals related to losses and congestion are sent separately.
2. Additional charges aimed at completing the recovery of the grid cost and, maybe, sending locational signals. These additional charges are commonly referred to as the complementary charge.

Two main approaches to the computation of the complementary charge exist: allocating the fraction of the grid cost that is not recovered through energy prices according to the responsibility of each agent in the development of the grid, and computing the complementary charges corresponding to each line according to the benefit that each agent obtains from the existence of the line. The following paragraphs explain each one of the two approaches.

As mentioned in the previous paragraph, we may allocate the complementary charge corresponding to each line to those agents for whom it has been built. Then, the charge to be paid by an agent should be based on the impact that installing additional generation or load at its node would have on the system grid costs. One possible implementation of this principle is to compute unit transmission charges that are equal to the Long Run Marginal Cost at the corresponding node. In other words, unit transmission charges that are equal to their corresponding Long Run Marginal Costs. Long Run Marginal Costs represent the optimal solution from a theoretical point of view. Under ideal conditions, they coincide with short term marginal costs. Therefore, besides optimizing the system dispatch, they would lead to an optimal expansion of the system and achieve recovery of the whole cost of the grid. Unfortunately, Long Run Marginal Costs do not exist in reality because of the indivisibility of investments in transmission. Even assuming continuous investments in transmission, reliability constraints, if not included in the planning of the expansion of the grid, and the lack of elasticity on the demand side would result in short term marginal costs differing widely from the long run ones. Marangon and Oliveira compare other pricing schemes with LRMCs aiming to find some tariffs that optimize the system expansion, see (Marangon and Oliveira 1998). They illustrate the close relation that exists between transmission tariffs and system expansion and planning. A method that is based on these principles has been applied in England and Wales and Colombia to compute transmission charges (the ICRP method), see (Green 1997; Bakirtzis, Biskas et al. 2001).

Alternatively, complementary charges may be computed according to the benefit that each agent obtains from the existence of the grid. If the cost of every line were allocated proportionately to how much each agent benefits from its existence, no agent would oppose the construction of a line that is beneficial for the system. In principle, one may think of computing the benefit that each agent obtains from one or several lines by comparing two situations, one where these lines exist and the other one where they do not. In this regard, comparing the present situation to the one where no transmission line exists, and everything else remains the same, would not be realistic (at least when this results in loss of supply), since regulatory authorities, the SO and the existing agents would have taken some measures in order to guarantee that the load is supplied by the generation in the system.

Transmission lines may produce benefits of three types:

- economic benefits related to the replacement in the dispatch of less efficient generators by more efficient ones,

- reliability benefits which result in a decrease in the number and duration of the supply disruptions and system failures,
- and competition benefits which occur when consumers benefit from a decrease in the sale price of electricity as a consequence of greater competition among generators.

Unfortunately, it happens that allocation of transmission costs to the economic beneficiaries is plagued with difficulties in practice. Most of the problems arise from the lack of adequate information about the generators in a competitive setting and the need to estimate the future behavior of the system. But also because it is difficult to evaluate the economic impact on the market agents of each individual line in a well developed network with some level of reliability-driven redundancy.

Both methodologies to compute complementary charges are equivalent, as we shall explain now. The cost function of the grid consists of the rule or rules applied to decide upon the construction of new grid facilities (ACCC 1999; ACCC 2004). These rules determine whether a new investment is justified. As explained in chapter 2, in the new deregulated environment, a line is deemed to be beneficial for the system whenever the aggregated net benefit that market agents get from it is higher than the cost of constructing, operating and maintaining it. The benefits that network users obtain from a line are the only decision factor, or driver, behind its construction. Therefore, the more an agent benefits from the existence of a line, the more it is responsible for its construction and the more it should pay for it compared to others.

From what has been said, we must draw the conclusion that the responsibility in the cost of constructing, operating and maintaining a line is not related to the commercial transactions in which each agent engages (Pérez-Arriaga 2002). Assuming perfect information, the set of market agents that are dispatched do not depend on the commercial transactions taking place. Let us imagine that we were able to repeat many times the trading situation whereby the agents in a region agree on the commercial transactions they want to engage in. Every time we let the market agents trade their energy they would come up with a different set of market transactions. However, the group of generators selling power and the group of consumers buying it would always be the same and they would always be the most efficient ones. For every possible set of operating conditions there is only one optimal set of generators that should produce power most economically. Therefore, both the most efficient use that can be made of the network and the corresponding profits that are made by each agent are completely determined.

Complementary charges have two purposes: a) they must complete the recovery of the cost of the grid, and b) these charges must make network users internalize, in their investment decisions, the effect that installing a new generator or load in one point of the grid or another has on the system grid costs. In the long-term, the response of the potential new network users to the transmission charges (i.e., whether they will decide

to install or not) will depend on their expected profits, after transmission charges are duly included. Consequently, complementary charges must be long-term economic signals with no purpose in the operation (i.e. short-term) timeframe. In other words, they should not interfere with efficient short-term signals (short term marginal costs of electricity), which are designed to promote the optimal operation of the system. Thus, energy prices should be independent from charges aimed at paying network infrastructure costs already incurred. This problem is common to every network industry, where market agents compete to gain access to some essential facility, see (Turvey 2000). Kim et. al study in (Kim and Baughman 1997) the influence that different price modifying methods to achieve revenue reconciliation have on market results. They conclude that market results are significantly affected by the fact that short term energy prices are modified. Therefore, leaving energy prices unchanged and using some additional method to recover the complementary charge from system users seems preferable.

Benefits resulting from the construction of a line may be positive or negative (losses). Chapter 2 concludes that both positive and negative benefits should be taken into account when deciding on its construction. Analogously, charges resulting from the allocation of the cost of a line could be positive or negative, depending on whether the corresponding agent is benefited or harmed by the construction of a line. If network users are appropriately compensated for the construction of those lines that harm their interests, no agent would oppose the construction of a line that is beneficial for the system. However, applying negative transmission charges seems rather difficult in practical terms.

According to the discussion above, the following is a blueprint of the method that should be applied to allocate the complementary charge corresponding to a line or a set of lines:

- In principle, the cost of each line should be assigned to those agents for whom the line has been built in proportion to the benefit that each of these agents obtains from it. Since computing the benefits produced by a line may be rather difficult, most of the times, though not always, some measure of electrical use should be adopted as a pragmatic approximation to benefits (electrical use is also much easier to compute), see (Pérez-Arriaga 2002; Pérez-Arriaga and Smeers 2003). In (Pan, Teklu et al. 2000; Rubio and Pérez-Arriaga 2000) several usage-based methods to compute the complementary transmission charge are evaluated. (Rubio and Pérez-Arriaga 2000) compares three of the main methods, the so called Average Participations method, the Marginal Participations method and the Allocation-to-Beneficiaries method. All of them are briefly presented in section 3.3 of this chapter.

It is generally accepted that the responsibility of an agent in the construction of a line depends both on the location of the agent and its profile (when and how much electricity it consumes or produces). Consequently, transmission tariffs

used to recover the cost of the grid should be spatially and temporarily differentiated:

- o Spatial differentiation: the benefits an agent obtains from the grid depend upon its location on the grid. Thus, a generator located within an exporting area will, in general, be benefited by the building of a new line linking this area to the rest of the system, whereas a generator located within an importing area could be harmed by the reinforcement of the connection between this area and the surrounding ones. We could identify a number of areas in the system and compute uniform per unit transmission charges within these areas. Thus, the charge to be paid by each agent would be less dependent on the unilateral commercial decisions he makes.
- o Temporal differentiation: reinforcements to the grid reduce the number and duration of service disruptions but they also contribute to the reduction of the system fuel costs and the losses in the grid. These last two benefits happen during the whole year. The operation of the system during all the hours of the year must therefore be considered. Since every hour market agents get different profits from each line (in the same way as the use of the line varies with the time of the day and the year) the responsibility in the incurred grid costs varies with time as well. Introducing some temporal differentiation in the transmission tariffs makes sense. One possibility is to compute transmission charges according to the electrical usage that each type of agent is expected to make of the grid during each time of the year (season or any other division that makes sense). Again, this has the advantage that it makes the charge levied on each agent less dependent on the operation time frame decisions of this particular agent. Computing transmission charges for a number of blocks of hours that we have divided the year in is easier to implement than considering each hour separately.

In order to make transmission charges as independent as possible from unilateral operating decisions by agents, we could combine the proposals in the previous two bullets. Thus, we should explore the possibility of computing transmission charges based on the electrical usage that agents of the same type that are located within each one of a set of areas previously defined are expected to make of the grid.

- If the benefits produced by a line are very much dispersed, or there is no need to send locational signals (since these signals would be very weak anyhow, due to the fact that a transmission grid is well meshed) and we only need to allocate the sunk cost of the grid, socializing the cost of the line may be preferable. In this case, the only objective would be to minimize the distortion of the dispatch resulting from the application of grid charges. Ramsey proved that access

charges should in this case be higher the lower the demand elasticity is, see (Ramsey 1927). This may only be right for perfect competition, however, see (Poletti 2005).

3.2.1 Discussion of several issues related to the implementation of a method to allocate the grid costs

Now, several issues related to the implementation of a network cost allocation method are discussed: these are the treatment to be given to the cost of the unused capacity of each line (how it should be recovered) and the consideration of the generation and load within each node.

3.2.1.1 Treatment to be given to the unused fraction of the capacity of a line

Regarding the recovery of the cost of the unused fraction of a line, one first possibility is to allocate the total cost of the line independently of its level of utilization. In principle, this appears to be correct, since the purpose of the network cost allocation exercise is completing the recovery of the cost of the transmission network through the application of network tariffs.

However, one could claim that the principle of cost allocation based on responsibility in network development should not be taken that far. Assuming the cost allocation method employed is based on the computation of the electrical usage that agents make of lines, we are substituting “electric network use” for “responsibility in network development”. Therefore, we should leave some margin for error. Additionally, it is not clear how the unused fraction of the line can be attributed to any network user on the basis of “use”. Therefore, leaving the cost corresponding to the unused part of the lines to be assigned via a less engineering-like procedure seems to be a sensible approach. Ramsey pricing principles may have something to say here, as well as more pragmatic considerations, such as charging consumers a flat fee for the cost of the entire unused part. Then, specific ad-hoc rules may be used to avoid assigning the grid cost 50% to generation and 50% to load. As we shall see later, a smaller G charge can be accepted under most commonly encountered market conditions.

Lastly, there is the problem of determining the actual “capacity” of a line. The amount of flow that a given line can take depends on the operating conditions of the power system. For the sake of simplicity, we might initially assume by default that the capacity of any line is the one given by its thermal rating at some agreed temperature.

3.2.1.2 Treatment of the generation and load that are located within the same node

Generation and demand in each node of the network must be separately considered. They must not be netted out as a single net demand or a net generator. This appears to be the most consistent choice with the principle that any network cost allocation

method should be non-transaction-based, as stated in the most advanced regulation of some regional markets (see, for example, the conclusions of the Florence Forum and the EU Regulation applicable to the Internal Electricity Market of the European Union). Netting generation and load according to the representation that is made of the grid would be a discriminatory measure. The generation located in other nodes than that where a load is located would be discriminated against when determining the generators serving this load. Assume that the aggregated generation in node k is G_{Ak} , the aggregated demand D_{Ak} and also that $G_{Ak} > D_{Ak}$. Regarding the node as having only a generator with an output of $G_{Ak} - D_{Ak}$ would amount to assuming that a fraction of the generation equal to D_{Ak} serves locally the load D_{Ak} , without using the transmission network, and, therefore, without paying for it. This should certainly be considered a transaction based decision. If this approach were chosen, network users that are in a minority position at any given node k (less generation than load, or the opposite) would not have to pay any transmission network charges.

3.3 Review of network cost allocation methods

Grid costs may be classified according to whether or not those agents for whom the corresponding network facilities have been built can be easily identified. Costs that can be easily attributed to the connection of an agent (or a group of agents) to the grid are direct connection charges. Direct connection costs must be recovered through connection charges. On the other hand, some grid costs cannot be directly attributed to any agent. They must be recovered through “use of the system charges”.

Connection charges could in turn be divided into shallow and deep charges. Shallow connection charges include the transmission facilities strictly needed for the connection of the agent to the grid. Deep connection charges correspond to those network facilities that are built as a consequence of the installation of a generator or load in the system, but whose construction is not required, strictly speaking, for the connection of the agent to the grid. Thus, the cost of a line that allows an agent to trade a larger amount of power in other areas of the system could be charged to this agent in the form of a deep connection charge.

This section discusses the application of connection and use-of-the-system charges throughout the world. Section 3.3.1 focuses on connection charges, whereas section 3.3.2 focuses on use of the system charges. Section 3.3.3 describes the tariff setting process that is applied in several regional markets. Finally, section 3.3.4 discusses the treatment given to other aspects of the implementation of transmission tariffs.

3.3.1 Connection charges

Focusing on the charges that may be applied upon the connection of an agent to the grid, and as explained in the introduction to this section, we may distinguish between shallow and deep connection charges. Shallow connection charges strictly result from the recovery of the cost of connection network facilities (those that are used to connect a specific agent or group of agents to the rest of the system). Deep connection charges must recover the cost of those lines that are partly or totally built as a result of the installation of an agent, whose construction is not strictly necessary for this agent to get connected to the grid.

Most countries apply only shallow connection charges, others compute also deep connection charges, and still others do not consider any kind of connection charge. Here, we shall limit the discussion on connection charges to the shallow ones. We think it is more coherent to discuss deep connection charges together with the rest of transmission charges.

Talking about shallow connection charges, they have been structured in many different ways (CER 2004). There are systems like England and Wales which apply these payments both to generators and consumers. Others like Australia charge only generators. It is far less common to charge just consumers because generators are more likely to have specific connection lines dedicated to them. There are systems such as Ontario where a shallow connection charge is only levied on new consumers or generators, see (CER 2004). Referring to the way in which the charge is levied on the agent, it is possible to make the agent pay the whole charge at the same time or to distribute it among several years. In the latter case, market agents must pay an annual transmission rate by way of connection charge. However, most of the systems use a one-time shallow connection charge.

Another option is to consider adjustments to the charge, according to the use others end up making of a line initially devoted to the connection of an specific agent. In this way, this agent will share with others the cost of these facilities. This adjustment system is applied to the generators in Ontario, for example.

3.3.2 Network use of the system charges

This section provides a review of the different mechanisms that have been applied (or at least proposed) for the computation of transmission use of the system charges in the context of a single, perfectly integrated, system. Section 3.3.3 deals with the case of different regional markets that presently exist. We shall begin by describing those methods that are conceptually sound, i.e. those methods that are in line with the rationale presented in section 3.2. Afterwards, we shall briefly comment on those other methods that we consider to be flawed. Lastly, we focus on some methods recently

proposed in the technical literature on this subject, which have not been actually applied yet.

3.3.2.1 Conceptually sound methods

First, we present those methods that apply the cost causality principle in order to compute grid charges. In other words, these methods aim to assign the cost of each line to those agents for whom the lines have been built. Charges to be paid by these agents are proportional to the estimated benefits that each one of them will obtain from the construction of the line. Conceptually speaking, these are the most efficient methods. Methods of this group must be applied to new lines built for very specific network users and, in some cases, to any new line when the number of years that is considered in the computation of grid charges is low (around 5).

- Beneficiaries: this method allocates the cost of each line based on the increase in the operational profit of each agent that is produced by the existence of the line (Pérez-Arriaga and Rubio 1995). Conceptually speaking, it is the best method and therefore it should be used as a reference to assess the performance of the rest of the methods. Unfortunately, benefits may be very difficult, if not impossible, to compute, except for those new lines whose cost is allocated considering a small number of years. In fact, this method has never been properly applied. It was adopted for application in Argentina and also later in California, as well as in Peru and other places, but a full implementation has not been achieved yet.
- Methods that are based on the computation of long run transmission costs: several methods exist, which differ in the underlying assumptions.
 - o Investment Cost Related Pricing (ICRP): it is based on the computation of the long run marginal transmission costs (LRMTCs) (Green 1997). The method obtains the minimum grid cost that is necessary to supply the demand with the generation that is available in the system. Each generator is assumed to respond according to its capacity. LRMTCs are the sensitivity of the cost of this grid with respect to the power injected at each node. In order to apply the method many simplifications are introduced such as the assumption that continuous investments in transmission and generation are possible. There are several variants of the ICRP method, depending on whether they take or not into account the electrical parameters of the lines when computing the minimum cost grid. In some cases, these electrical parameters are considered constant, i.e. they do not vary when we increase the capacity of the line. Other methods assume that the transmission system behaves like a pipeline system and electricity flows like a liquid through it (they consider a transport model of the transmission grid). Both methods are explored in (Bakirtzis, Biskas et al. 2001). These simplifications allow the method to

be easily applied but the results are not very realistic. It has been applied in England & Wales and Colombia to compute network tariffs.

In (Lima, Pereira et al. 1995), Marangon et. al propose basing the computation of grid complementary charges on the sensitivity of transmission costs with respect to the existence of each particular agent (generator or load). This is a concept quite similar to that of ICRP, though the authors establish a close relation between their method and the MW-mile one.

- o Alternative to ICRP: the method is addressed to encourage the loads to make an utilization of the grid that corresponds to the optimal level of expansion of this grid, see (Farmer, Cory et al. 1995; Farmer, Perera et al. 1995). First, the problem that optimizes the network expansion is formulated. From this problem, the method computes the optimal flows over the lines. Then, these flows are used in the market agent problem where each agent tries to maximize its benefits. The optimal transmission charge that each consumer must pay for each line is computed from this second problem. The H matrix relates the flows over the lines in the system to the net injection of power by each consumer.
- Methods that are based on cooperative game theory: cooperative game theory is an approach to solve the problem of allocating a given amount of resources among a number of agents. Here, the method allocates the cost of the grid among the market agents that make use of it. Many different cooperative games can be formulated. In some of them, coalitions of network users are formed. Within each coalition, the agents pay the cost of the minimum grid that is able to support the transactions that would take place among those agents if they were considered in isolation. Coalitions are such that every agent is better off than if he belonged to any other coalition (or no coalition at all). For more information on this type of methods, see (Contreras and Wu 1999; Contreras and Wu 2000), among others. Methods based on game theory have very interesting properties from the point of view of economics and equity.

One interesting concept is that of the Shapley value. We may compute how much of the whole grid cost should be paid by each agent as the Shapley value corresponding to the agent in the formulated game. In the game where the grid costs are allocated to market agents, each agent can form coalitions with others in order to pay a smaller charge. Each coalition must pay the cost of the grid that is strictly necessary for the agents within the coalition to trade among themselves the energy they consume or produce. The incremental contribution of an agent to the total transmission charge to be borne by a coalition can be computed as the difference between the cost of the grid that the coalition must pay and the transmission charge this same coalition would have to pay if the

considered agent did not belong to it. The Shapley value of an agent in a game is the average incremental contribution of the agent to the costs to be born by all the coalitions in which it may take part (in this case, the grid costs). In (Tan and Lie 2002) a rather simple and intuitive explanation and application of the Shapley Value (SV) concept is provided.

In (Contreras and Wu 1999) the cost of new grid investments is allocated among grid users based on the Bilateral Shapley Value (BSV) concept, which is an approximation to the Shapley value, where coalitions between two players or groups of players are considered in turn until the grand coalition containing all the players is formed, if possible. The BSV is simpler to compute than the SV. One major problem of this method and others within this group is that they do not necessarily allocate the whole cost of a grid investment but only the fraction that is used by the grand coalition of players. In (Zolezzi and Rudnick 2002) the responsibility of each agent is measured in peak conditions, which, from the point of view of beneficiaries theory, is not right since economic benefits are obtained at any time of the year and not only when the flow through the line is highest. Stamtsis et al. apply in (Stamtsis and Erlich 2004) the Shapley value method and the Nucleolus method, both to allocate the system grid costs as a whole and the cost of each line on a separate basis. The analysis is carried out for different organizational structures of the market. See (Evans, Zolezzi et al. 2003) for a review of different methods based on cooperative game theory.

When benefits resulting from the construction of a line cannot be estimated, but we still aim to send long term locational signals, we may compute the use of the system charges based on the electrical usage that each agent makes of each line, as a reasonable proxy, see (Rubio and Pérez-Arriaga 2000). As already mentioned, electrical usage tends to be easier to compute than network benefits. A list of the main usage-based methods follows:

- Average Participations (AP): the AP method takes the set of network flows as an input. It makes use of the actual pattern of these flows to determine how much of the power flow on each line is produced by each network user. A heuristic rule determines the path followed by the power flow injected or withdrawn by each agent. The AP method neglects the existence of counter-flows since they cannot be measured. Only the existing (net) flow over a line can. If we accept this assumption (which seems to be very reasonable), the path followed by the total power flow entering a node is perfectly determined. The AP method assumes that any entry flow to a node distributes among the exit flows from the node proportionately to the way the total flow entering the node does. When applying this method, the generation and load located at any given node must be separately considered. Considering only the net demand (or generation) at each node is equivalent to assuming that the demand within a node is served, to the extent possible, by the generation located at the same node. Thus, only the excess of demand over generation at the node (or the

excess of generation over demand) would be served (consumed) by the generation (load) at other nodes. When determining the sources of the power consumed by a load, there is no reason for discriminating between the generation located at the same node as the load and the generation located at other nodes. Apart from this, the AP method may either be used to allocate the cost of the used fraction of each line or that of the entire line. This method was originally applied in New Zealand (TP-NZ 1990) and was then considered for its application in Poland and other countries. It is presently used in the Central American Regional Electricity market. This thesis proposes that this method be implemented in the context of the IEM of the EU. Other sections of the chapter discuss the method and some of its more interesting properties.

Some variants on the basic method exist, see (1990; Bialek 1996; Strbac, Kirchen et al. 1998; Kattuman, Green et al. 2001; Pérez-Arriaga 2002; Pérez-Arriaga, Olmos et al. 2002; Vazquez, Olmos et al. 2002; Bialek and Ziemianek 2003; Pérez-Arriaga and Olmos 2003). Kirchen et al. apply AP to an equivalent system made up of commons, or groups of nodes supplied by the same set of generators, and links, or groups of transmission lines between the same two commons, see (Kirschen, Allan et al. 1997; Strbac, Kirchen et al. 1998). In (Strbac, Kirchen et al. 1998) authors focus on the usage made of each line under peak load conditions. In (Kirschen, Allan et al. 1997) they suggest that the method can be applied to determine the responsibility of market agents in creating both active and reactive power flows. AP may also be applied to allocate transmission losses. In (Kirschen and Strbac 1999), Kirchen et al. propose some modifications to the basic AP method in order to allocate both active and reactive power flows in a satisfactory way. These authors transform power injections into current injections (both real and imaginary) and track currents on the grid. Once they have allocated the current on each line among grid users, they obtain the power flow attributable to each user by transforming line currents back into power flows.

In (Bialek 1996; Bialek 1997), Bialek employs a matrix algorithm that makes the computation of participations very compact. In (Bialek 1997) he applies AP to the problem of allocating the complementary charge corresponding to each line. Cost responsibilities are assigned in proportion to the fraction of the active power flow on each line that each agent is responsible for. In (Bialek 1996) he applies AP to allocate active and reactive power flows as well as transmission losses. The loss allocation system makes use of a loss-apportioning algorithm.

In (Wu, Ni et al. 2000) Wu et al. propose the creation of directed graphs that always begin or end in a pure source or a pure power sink. This article proves the existence of pure power sources and sinks. In (Pantos and Gubina 2004), the authors apply the AP method to a set of snapshots and then build a regression model that expresses the flow on each line as a function of the net power injection at each node. In another article, Wei et al. modify the settings of phase

shifters and voltage regulators in order to eliminate the loop flows that may appear. Then, they apply the basic AP method using the graph methodology, see (Wei, Ni et al. 2002). Finally, in the 8th Florence Forum, Glavitsch et al. presented an alternative implementation of the same basic AP algorithm presented above. In the implementation proposed by professor Glavitsch, each line flow has an associated cost corresponding to all the lines it has gone through either upstream or downstream. Glavitsch proposed some small and irrelevant modifications to the basic AP algorithm which result in slightly different numbers.

- Marginal participations (MP, also called Areas of Influence): here, the participation of an agent in the use of the grid results from simulating the marginal impact on the flow over every line of a unit transaction between the node where the agent is located and a node 'R', taken as the reference node in the system. The impact of a marginal transaction between nodes 'A' and 'R' on the flow over line 'l' is commonly known as the Power Transfer Distribution Factor (PTDF) of the flow over this line with respect to the injection at node A. The participation of an agent in the use of a line results from multiplying the corresponding PTDF by the amount of power traded by the agent. This participation can be either positive or negative. A negative participation implies that the power injection or withdrawal by the agent increases instead of decreasing the amount of capacity available to others. Thus he should, in principle, be paid for this. Other possibilities exist nevertheless. Normally, payments depend on the results obtained for several scenarios.

The Marginal Participations method may provide reasonable results in a the context of a single, perfectly integrated system with a clear dominant node where most demand and/or generation is located. However, this cost allocation method is conceptually flawed. This is due to the fact that MP implicitly assumes that the same node would marginally respond to a power injection or withdrawal taking place at any node of the system, regardless of the location of the later within the grid. Consequently, the election of the balancing or reference node critically affects the amount to be paid/earned by each agent. For more information on the MP method, see (Agurto 1992; Pérez-Arriaga and Rubio 1995; Rudnick, Varela et al. 1996; Pérez-Arriaga 2002; Pérez-Arriaga, Olmos et al. 2002; Vazquez, Olmos et al. 2002).

Similarly to what happens with the AP method, different implementations of the MP method are possible, see (Agurto 1992; Rudnick, Varela et al. 1996; Park, Park et al. 1998). In (Anderson and McCarthy 1999) authors favor the substitution of the MP method, which was in place at that moment in the Argentinean system, with another method based on the computation of the benefits that market agents obtain from each line. In (Rudnick, Palma et al. 1995), Rudnick favors the utilization of nodal energy prices complemented by marginal transmission prices that are based on generalized load distribution

factors obtained for generation and load. Application of this system of prices in the Chilean system is illustrated in this paper.

According to some variants on the basic MP method, a different responding node must be considered depending on the type of agent. This is the case of (Galiana, Conejo et al. 2003) where generators respond pro rata when computing the participation of each load in the use of the lines of the system and loads respond pro-rata when computing the participation of each generator. Even in this case, the same node, or group of nodes, would respond to all power injections (or withdrawals) in the system. Therefore, the basic problem of the MP method persists.

In those systems where the transmission grid is developed enough, we may not be interested in sending long term locational signals. Then, use of the system charges should be aimed at recovering the sunk cost of the grid without interfering with the economic signals that are sent in the short term. Now, we briefly describe the more significant methods that have been proposed within this category.

- Ramsey like cost allocation methods (Ramsey 1927): these methods try to minimize the distortion in the operation of the system that is caused by the allocation of network charges. Ramsey pricing is used to compute energy charges aimed at minimizing the interference with the operation decisions by market agents. It computes higher transmission tariffs for those grid users whose decision on how much to produce or consume depends less on the transmission tariff they have to pay. Most of the transmission tariff schemes that result in a different unit transmission charge for each type of transmission user are actually applying the same idea, though maybe they are not making use of Ramsey Pricing, strictly speaking. In developed countries, these schemes typically result in higher tariffs for households than for industrial customers, while in less developed countries the opposite occurs. This method is very much advisable when we lack the information required to apply cost causality (how much each agent benefits from a particular line or piece of line) or when there is no need to introduce locational signals because the transmission grid is developed enough. In this case the sole objective is to compute network charges that fully cover total transmission costs without compromising the efficiency of the system operation. Additionally, applying this method requires having access to information that is difficult to obtain (the elasticity of market agents' operating decisions with respect to the level of transmission charges).
- Postage Stamp: these are uniform charges throughout the system. They can adopt any type of structure: they may be made dependent on the contracted capacity; they could be structured as a fixed rate per energy unit produced or consumed; lastly, charges may be completely fixed in absolute terms, so that any network user pays the same with independence of how much it uses the grid. Postage stamp rates are applied in many systems (they are probably the

most common tarification method). See (Pérez-Arriaga, Montero et al. 2002; 2004) for examples of application of transmission tariffs of this type. Transmission charges paid by generators in the Netherlands include a component that does not depend on the energy or capacity contracted, see (CER 2004). Postage stamp tariffs have also been extensively used in the US (Baughman 1997; Bushnell and Oren 1997; Green 1997). In fact, many new designs and initiatives advocate the use of this type of charges (2003; Jong 2004).

- License Plate Rate: according to this method, all the generators or loads within each one of a predefined set of areas pay the same per unit transmission charge. Therefore, this method results in G and L per unit transmission charges that are characteristic of each area. G and L transmission charges vary from one area to another but are uniform within each one of them. Typically, the total cost of the local grid in each area is split into two parts, one to be paid by the local generation and one to be paid by the local load. Afterwards, G (or L) per unit transmission charges result from dividing the total amount to be recovered from generation (load) by the total amount of generation (load) in the area (we may use either the total generation/load capacity installed or the total amount of energy produced/consumed). Transmission charges in the Netherlands have this format, though they also vary from one voltage level to another (CER 2004). License Plate rates have also been proposed as part of the Standard Market Design in the US (2003; Jong 2004; Joskow 2005).

3.3.2.2 Other methods

This section briefly describes some methods that have been proposed, or actually used, for the computation of use of the system charges despite being conceptually wrong. The theoretical foundation of all these methods is flawed, since the transmission charge to be paid by each agent depends on the commercial transactions that the agent takes part in. Therefore, these methods are transaction based. It can be easily proved that, provided perfect information exists, the use made of the grid of a system, and consequently the benefits obtained from it, do not depend on the commercial transactions.

Most of the methods that are presented in this section belong to the so-called wheeling philosophy. We make reference to them mainly because of historical reasons. Wheeling methods were developed during the 80's and the early 90's and they are aimed at computing the cost incurred by a system because of the existence of wheeling transactions that crossed its grid. With the passing of time, methods based on this philosophy were abandoned and others that determine the responsibility in grid costs of each individual agent or group of agents took their place. Some of these methods are described in (Rau 1989; Shirmohammadi, Rajagopalan et al. 1991; Happ 1994; Shirmohammadi, Vieira et al. 1996; Yu and David 1997; Silva, Mesa et al. 1998).

Methods within this group can be classified into those that are based on the electrical usage that system transactions made of the grid, and those that do not adopt a measure of electrical usage in order to compute transmission charges.

Methods based on electrical usage

We begin by listing those methods that are based on the electrical usage of commercial transactions:

- Contract Path: parties involved in a transaction, together with the regulator, the System Operator or the competent authority, agree on the path followed by the transaction from the point where the power is injected to the point where it is retrieved. For each line, the transaction will pay for the amount of capacity it uses according to the contracted path. This method is very simple but it is certainly deficient. To begin with, the path actually followed by a certain amount of power injected at a node and retrieved at another one will not coincide in general with the one previously agreed. Physics laws rule the way flows spread over the grid and not any kind of agreement. Besides, this method favors transactions that take place between agents located close to one another (mainly local transactions) over those that take place between agents of different countries or systems. In other words, it is openly discriminatory, which is against the regulation approved for the Internal Electricity Market of the European Union and many others. This method has been widely used in many US systems (Happ 1994).
- MW-mile method: it was originally aimed at computing the cost attributable to wheeling transactions. The method computes the fraction of the grid cost that each transaction crossing a system must pay by comparing the total use made of the grid in two different situations (Shirmohammadi, Gribick et al. 1989; Shirmohammadi, Rajagopalan et al. 1991). First, the actual use that is made of the grid in reality is obtained. Then, the concerned transaction is removed and the system grid use is computed again. The electrical usage attributable to the transaction results from computing the difference between the use made of the grid in the two cases. The method computes the electrical usage of the transmission grid as the sum, over all the lines in the system, of the product of the length of each line and its flow. Each transaction must pay a fraction of the cost of the existing grid equal to the ratio of the electrical usage allocated to this transaction to the total usage made of the grid. Several variants exist of the method depending on how they take into account the negative incremental impact that the transaction has on the flow through certain lines. We may either take these impacts as positive, we may not take them into account or we can take them as negative impacts, which are deducted from the positive impact that the transaction has on the flow over other lines. This method was proposed by Shirmohammadi et al in (Shirmohammadi, Gribick et al. 1989) and presents

the problems common to any transaction based method. This and the Contract Path methods are compared with others of different nature in (Shirmohammadi, Vieira et al. 1996).

- General Agreement on Parallel Paths: it is used to determine how much a transaction that crosses the grid of a system must pay this system (Consultants and Putnam 1995). It was commissioned by the Interregional Transmission Coordinating Forum in the US. A matrix specifies the fraction of the amount of power transacted that goes through the grid of each one of the systems in the region. The transaction is charged according to these percentages, which are computed running power flows for different operating conditions. Contrary to what happens with other methods, this one takes into account how the flows actually spread over the grid. However, it is only valid for market arrangements based on the establishment of bilateral contracts. Like all the previous methods, it makes the compensation due to each system dependent on the set of transactions happening in the market, which is not conceptually sound.
- Analytical methods based on the physical laws governing the functioning of the system: these methods make use of the Kirchoff laws to derive more or less complex relationships among the different electrical variables of the system (flows and power injections, for example). They are not only used to allocate the cost of the lines in the system but also to allocate the cost of different system services, such as reactive power support by generators, see (Chicco, Gross et al. 2002). In (Chang and Lu 2001), line currents are decomposed into components associated with each generator or load. Authors obtain an equivalent system of voltages and currents for each grid user and compute the flow on each line that is caused by each agent. Participations of agents in the flow of the lines are employed to allocate line losses. In (Fang and Ngan 2002) authors express network losses as a function of the active and reactive power injected at each node. A method quite similar to the one proposed by Fang is compared in (Mutale, Strbac et al. 2000) to the traditional method for the allocation of losses based on marginal loss factors. Despite being a method for the allocation of losses, the marginal loss factors method is not comparable at all with other methods discussed in this chapter, since it treats transmission losses as a short-term problem (associated with the energy price) rather than as a long term problem, as methods discussed in this chapter do.

Most of these methods cannot avoid obtaining results that are dependent on the election of the slack bus used when running the power flow. The election of a slack bus, despite being seldom discussed, is implicit in most of the methods of this kind. See (Skokljek and Maksimovic 2004) for an example. An alternative to overcome this problem is to compute the responsibility of balanced transactions instead of individual power injections or withdrawals. Some methods employ some form of tracking rule, instead, which is based on the electrical properties

of the system. Thus, (Reta and Vargas 2001) substitutes generators with equivalent current sources and tracks the current that each generator causes on each line using the electrical equivalent of each node in the system. The method tracks electrical currents on the grid starting from those nodes that are pure sources (nodes from which the power flows outwards through all the lines connecting it to the rest of the system). This method is used to allocate transmission losses. In (Daniel, Salgado et al. 2005) generators are represented as impedances when transmission losses are allocated among loads and vice versa when they are allocated among generators. Thus modified Y_{bus} matrices are used to express transmission losses as a function of current injections at generator and load nodes.

Methods that are not based on electrical usage

Now we discuss those flawed methods that are not based on the computation of the electrical usage that transactions make of the transmission grid. These may be called incremental methods. Incremental methods directly estimate the costs incurred by a system that are the responsibility of the transactions that cross the system. In order to do so, these methods compare the total grid cost in the real case to the cost corresponding to the situation where these transactions have been eliminated. There are several ways of doing this, see (Shirmohammadi, Rajagopalan et al. 1991). Methods that apply the superposition principle assign a different unit charge to each transaction depending on the order in which the existing transactions are considered. Baran et. al propose assigning each incremental transaction a participation which is somewhere between the payment that this transaction would get if it were considered the first of all and what it would get if it were considered the last transaction. They aimed to allocate the flows on the lines, the transmission losses and the reactive power support from generators without having to assume an order for wheeling transactions, see (Baran, Banunarayanan et al. 1999). Shirmohammadi et. al compare in (Shirmohammadi, Vieira et al. 1996) incremental methods, such as LRIC (Long Run Incremental Cost) or SRIC (Short Run Incremental Cost), to some basic methods based on electrical usage, such as the MW-mile or the contract path methods. Marginal methods, such as LRMC (long Run Marginal Cost) or SRMC (Short Run Marginal Cost), and the postage stamp method are also considered in this article. A comparison between incremental methods and sunk cost methods is carried out in (Happ 1994) as well. In (Lo and Zhu 1993), Lo proposes charging wheeling transactions according to the marginal system cost associated with each one of these transactions. It is therefore a method that belongs to the wheeling philosophy but makes use of marginal instead of incremental costs.

3.3.2.3 Methods recently proposed in the technical literature

Methods described in this section have been proposed in the technical literature by academics but they have never been actually implemented. Many of them are based on

the computation of the sensitivities of line flows with respect to other variables. See for example (Bakirtzis, Biskas et al. 2001; Fradi, Brignone et al. 2001; Danitz, Rudnick et al. 2002; Yu 2002; Llaens 2004) among others.

The main drawback of these methods lies in the fact that defining these sensitivities involves choosing a reference node in most cases. Some methods try to avoid this difficulty by computing the economic dispatch corresponding to the situation where the generation or load in a node have been marginally modified, and comparing it to the original economic dispatch. The latter methods may produce reasonable results if the computed dispatch covers the whole system. This is normally the case within a national or local market but it is less common in the context of regional markets. Additionally, collecting the information required to compute the dispatch tends to be very difficult. (Yu 2002) presents an allocation mechanism whereby the cost of the lines is assigned to the commercial transactions in the system according to the contribution of the power injections and withdrawals involved in each transaction to the flow on each line. The contribution of each power injection to each line flow is computed from the sensitivity of the line flow, as resulting from the solution of an Optimal Power Flow, with respect to the power injection. Some methods apply sensitivity indices previously modified in order to comply with some desirable requirements, see (Vlachogiannis 2000).

Other methods compute sensitivities to allocate different system variables (losses, for example) among balanced power transactions. In this case, the effect of the slack node is irrelevant as it cancels out for each one of the defined transactions. However, it becomes necessary to determine which agent responds to the power injected or withdrawn by each other, i.e. we need to identify the generators and loads involved in each balanced transaction.

Finally, there are some methods that combine the features of several others. Thus, (Bjorndal, Stamtsis et al. 2005) proposes a method where the solution provided by basic usage-based methods such as the MW-mile method or postage stamp are combined to produce a solution that is within (or at least as close as possible to) the core of the game where the cost of each line is allocated among grid users. In this way, this method aims to reconcile usage-based methods with game theory.

Another example of a method that combines features of others is provided in (Usola 2005), where the author decomposes the system dispatch in a set of balanced transactions. Average Participations is applied to determine the generators and loads involved in each of these balanced transactions. Afterwards, the article computes the use that each one of these transactions makes of the grid as the incremental impact of this transaction on the grid flows. The author applies this method to allocate the cost both of the losses in each line and of the construction, operation and maintenance costs of the grid infrastructures. In the first case, both positive participations (payments) and negative participations (compensations to be received by the corresponding transaction(-s)) in the use of the lines are considered, whereas the construction,

operation and maintenance costs of each line are assigned to transactions proportionately to the absolute value of the participation of each transaction in the use of the line. The transmission charge computed for each transaction must be somehow divided into two parts: one to be paid by consumers, the other one by generators.

3.3.3 Application to regional markets

Originally, transmission tariffs in most regional markets resulted from voluntary agreements between vertically integrated utilities operating in different areas. In many cases tariffs to be paid by regional transactions were computed as the cumulative sum of the tariffs corresponding to all the systems that were “crossed” by the transaction. Therefore, if a Dutch generator wanted to sell its power to a consumer in Spain these agents had to pay, not only the Dutch and Spanish transmission tariffs, but also the Belgian, German and French ones. This is the problem that is commonly known as tariff pancaking (Pérez-Arriaga, Olmos et al. 2002; Jong 2004). Tariff pancaking is based on the idea that each system should treat an external transaction in the same manner as it treats the national ones. Thus, a transaction crossing the grid of a system was considered as a generator located at the node where the transaction entered the system and as a load located at the node where the transaction exited it. Presently, this philosophy is acknowledged as incorrect, since, according to it, tariffs paid by agents trading their energy regionally are clearly dependent on the commercial transactions that these agents engage in and on the shape of the political borders. As we have mentioned above, transmission tariffs must be non-transaction based also in regional markets. Different approaches have been developed since these early days. We now focus on the case of some of the most representative regional markets that exist and explain the solution presently adopted in each of them.

3.3.3.1 Regional markets in the US

First, we shall describe the solution proposed by FERC as part of its Standard Market Design (SMD). The SMD model has not been implemented in any region of the country since it was proposed in 2002. In fact, this comprehensive model had to be watered down to a collection of recommendations on the design of regional markets. Finally, we also briefly comment on the special characteristics of the transmission pricing scheme used in the PJM regional market.

According to the Standard Market Design (SMD), which was developed by the Federal Energy Regulatory Commission (the federal body in charge of setting the guidelines that rule interstate power trade in the US), only consumers should pay transmission use-of-the-system charges (FERC 2002; Jong 2004). Apart from this, both generators and consumers might pay connection charges. Use-of-the-system charges proposed by FERC are of the type entry/exit, i.e. each agent would have to pay according to the place where it is located and not depending on the commercial transactions it engages in. In this case, only exit tariffs would be considered, since only consumers would

have to pay use-of-the-system charges. Paying these exit tariffs would allow consumers to use all the regional grid. The SMD proposes two methods to compute use-of-the-system charges: the Postage Stamp and the License Plate Rate. The final decision on which method to apply would correspond to the competent authorities in each region.

The pricing scheme considered in PJM includes connection charges and use-of-the-system charges. Connection charges are aimed at making agents pay the lines that are built as a result of these agents installing a new generator or load. Use-of-the-system charges are of the type License Plate Rate and are only paid by customers (Load Serving Entities (LSE) on behalf of their end consumers). The charge paid by a LSE depends on its peak demand throughout a certain period.

As mentioned in the case of local markets, Postage Stamp (PS) and License Plate Rate (LPR) tariffs are well suited to those systems where sending locational signals is not necessary, which is not the case of regional markets in the US, since significant new network investments are needed.

3.3.3.2 The IEM in the European Union

Participants in the Florence Regulatory Forum (where regulators, TSOs, consumers, power exchanges, producers, the European Commission and all the remaining stakeholders meet to discuss the development of the IEM) have opted for a tarification system based on the concept of subsidiarity among countries. According to this principle, countries are free to set internal transmission tariffs as long as they comply with some common guidelines. These guidelines are requirements to be fulfilled by tariffs so that implementing a competitive market at European level is possible. Each country is responsible for the design of local tariffs, which should not hamper international trade. Therefore, the situation in Europe resembles that of an LPR system where each one of the areas (countries) applies a different unit G and L transmission charge. However, transmission charges within some countries exhibit some spatial and temporal differentiation. As we have just mentioned, a minimum level of harmonization of the local charges paid in different countries is necessary in order not to introduce distortions in the functioning of the regional market.

At the same time, a system of inter-TSO compensations has been established to complement local G and L charges. These compensations account for the use that market agents in a country make of the grid of others (Pérez-Arriaga, Olmos et al. 2002; 2004; Jong 2004). Therefore, all the considered methods take electrical usage as a proxy to benefits. At the moment, a transitory method determines how much each country must pay others. Several methods (AP, APT, WWT,...) have been considered to replace the transitory one. The final decision on the ITC method to use is still pending. Now we briefly review these methods. A more detailed discussion of the different ITC methods that have been considered can be found in section 3.5 of this chapter.

ITC methods

We can distinguish between those methods that compute the compensation to be paid by each country from the electrical usage that each agent makes of the grid in the region, and those others that directly compute the external usage of the grid of each country as a whole, instead of computing it as the sum of the participations of all the external agents in the use of this grid. Many of the latter methods determine the compensation to be received by a country from the electrical usage that the transit through the country makes of its grid. The transit through a country is defined as the set of power flows that are deemed not to be related to the activity of the agents within the country.

Methods based on the electrical usage by individual agents

These are mainly the AP method and the MP method. The AP method produces results that are quite reasonable. Thus, we may conclude that it is well suited to a regional market like the IEM. On the contrary, MP does not provide reasonable results. This has to do with the fact that this method assumes that the same node would marginally respond to a power injection or withdrawal taking place at any node of the system, as we have previously explained. It neglects the existence of constraints within the system that isolate the operation in some regions from that in others. Thus, the MP method may be acceptable in a local context but not in a regional one.

Methods that directly compute the external usage of the grid of a country as a whole

Several methods can be included within this category. They are listed below:

- With and Without Transits method (WWT): it compares the electrical usage made of the grid in two different situations: the one actually taking place and that where transits through the grid of the country have been removed. It is presently being applied in the NORDEL system to compute compensations among countries for the losses incurred due to cross-border transactions.
- Average Participations applied to Transits (APT): it is a variant of the method of Average Participations (AP), which is explained later in detail. It applies the AP method to the cross-border flows deemed to be part of the transit through each country.
- Modified Average participations (MAP): like the APT method, this is a variant of the AP method. It considers a different model of the system grid when computing the use that each country makes of the grid of others. The grid of the concerned country C is represented using a single node and its net demand or generation. The rest of the system is represented in detail. The AP method is applied to this modified system model. The same process is repeated for every country in the system.

Within the context of the IEM, a regional market has been operating in the Nordic region since the late 1990s. See (Kristiansen 2004) for a description of different aspects related to the regulation of transmission in this market.

3.3.3.3 The market in Australia

Here, we must make a distinction between connection charges and use of the system charges, see (CER 2004). A description of both types of charges follows. More information on the tariff setting process can be found in (CER 2004):

- Connection charges: generators pay a shallow connection charge. Specifically, they pay a fixed amount per day in order to recover the cost of the connection equipment they needed to begin to operate. Distribution companies do not have to pay a connection charge on behalf of their customers: the cost of the connection equipment is paid by means of use of the system charges.
- Use-of-the-system charges: consumers pay an exit tariff and generators an entry tariff. Consumer tariffs are non-linear and vary from one node to another. The energy part of this two-part tariff does not have temporal differentiation whereas the capacity part is a function of the maximum monthly power consumption. Consumers also pay a common general services tariff which varies with the connection point. Normally, the general services tariff is a function of the use that the agent has made of the system during the two previous years. Alternatively, general services tariffs may be made dependent on the contracted capacity. However, large penalties are applied if the power consumption exceeds this contracted capacity at any time. Therefore, the system includes charges associated with the electrical usage made of the grid and others based on the contracted capacity or the maximum consumption. Some temporal and spatial differentiation exists as well.

3.3.3.4 The MER in Central America

The regulation adopted at regional level establishes in the first place the regulated revenue that corresponds to each line. The regulated revenue in the regional market that is associated to each line depends on the type of the line. Thus, the regulated revenue associated to the facilities that belong to the primary regional transmission system (the SIEPAC line) is equal to the debt contracted by the EPR in order to finance these investments (there is a limit to the amount that can be collected annually by the EPR). The regulated revenue associated with the facilities belonging to the RTR that have been centrally planned by the EOR is equal to the total cost of the line including a reasonable rate of return on the investment. Existing transmission lines and investments at risks belonging to the RTR earn a regulated revenue related to the reliability value of the line, which represents a minor part of the cost of the line. Finally, the owners of investments at risk that are beneficial for the regional system can negotiate with authorities earning a certain regulated R . This revenue R is equal to

the fraction of the cost of the line represented by the ratio of the aggregated benefit that agents other than the owner obtain from the line to the total benefits produced by the line. In return for earning this revenue, the private promoter would have to transfer the property of the corresponding fraction of the transmission rights over the capacity of the line. If private promoters do not negotiate the collection of a regulated revenue, their revenue is equal to the market value of the lines they own, i.e. these lines earn the corresponding congestion rents in the regional dispatch.

The regulated revenue to be perceived by the owner of a line is broken down into several parts, see (PHB-HAGLER/SYNEX 2000; CRIE 2002; EOR 2005). A list of these parts follows:

- Variable transmission revenues (CVT for the Spanish *Costes Variables de Transmisión*): they result from the application of nodal prices. Every line belonging to the regional grid is entitled to this revenue. In principle, every line earns a congestion revenue that equals the difference between the energy price at both ends of the line times the flow over the line. The national regulation may devote the congestion revenue corresponding to regulated lines to the partial recovery of the cost of the line. Alternatively, it could be used to recover part of the total cost of the national grid as a whole. Lastly, these congestion rents could finance the construction of new lines in the region or in the country where these congestion rents have been collected.
- Use-of-the-System charges (also called tolls): they complete the recovery of the part of the regulated revenue of a line that is represented by the used fraction of the line. Thus, if 40% of the capacity of a line is used, the sum of the CVT and the Use of the System charges corresponding to the line must be equal to 40% of the regulated revenue of the line. The amount of use-of-the-system charges corresponding to a line that must be paid by each country is proportional to the use that the agents in this country have made of the line. A combination of the Dominant Flow and the AP methods is used to determine how much the agents within each country have used each line (the amount of flow through the line that they are responsible for).
- “Complementary Charge” (CC): it is the difference between the allowed revenue for a line and the amount recovered through congestion rents and use-of-the-system charges. The complementary charge corresponding to each line is allocated among countries in the region using a variant of the method applied to compute Use of the System charges.

The Dominant Flow method (PHB-HAGLER/SYNEX 2000) considers one national super-transaction per country and a regional super-transaction. It decomposes the flow over each line into the one caused by the corresponding national super-transaction and the flow produced by the regional super-transaction. Each one of the national super-transactions comprises all the power injections and withdrawals accepted in the

corresponding national market. Apart from them, there is one regional super-transaction that comprises all the injections and withdrawals accepted in the regional market (modifications to national dispatches). The flows corresponding to a super-transaction are determined, by applying the superposition principle, as the joint incremental impact on the flows over all the lines in the system of all the injections and withdrawals that are part of the super-transaction concerned. Once the flows of the regional super-transaction have been determined, the AP method is applied to these flows in order to compute how much each agent involved in this super-transaction is using each line. Two different variants of the method presented here are used to allocate the fraction of the regulated revenue of a line to be collected from Use-of-the-System charges and “complementary charges”. When computing Use-of-the-System charges, only those super-transactions causing a flow over a line that goes in the same direction as the net flow over the line have to pay for the use of this line. If there is a super-transaction that causes a flow over a line that goes in the opposite direction to the net flow over the line, agents involved in this super-transaction do not pay for the use they make of the line. On the other hand, both the super-transactions that create a flow over a line that goes in the same direction as the net flow, and those that create a flow that goes in the opposite direction, have to pay part of the “complementary charge” (or revenue) the line is entitled to. Scenarios corresponding to the last three years of operation are employed to calculate transmission charges.

It can be easily concluded that the method described here is a transaction-based method. Transmission charges resulting from the application of this method discriminate between agents trading their energy regionally and those that trade it locally. Obviously, this is not a desirable feature of the method.

Application of the method that has been just described to all the injections and withdrawals that are part of the regional super-transaction results in a net amount to be received or paid by each country. This amount must be allocated internally to the generation and load within the country. Local authorities are responsible for determining how to assign the aggregated charge to the generators and load in each country. However, we propose here a methodology to assign this charge efficiently. If the global charge is positive (the country must pay), it may be levied on generators if the country is exporting more than it is importing and on loads if it is importing more than it is exporting. Conversely, if the net amount to be recovered from a country is negative (the country must be compensated), then this amount may be distributed among generators when the country is importing more power than it exports and among loads if it is exporting more. As for the national super-transactions, the regional bodies can only make general recommendations but they could not interfere with the national policy.

3.3.3.5 The MERCOSUR market in South America

As far as we now, no comprehensive transmission pricing scheme has been adopted in this region yet. Here, we shall describe the scheme that is proposed for the region by the World Bank, see (WB 2000).

The total transmission charge is divided into variable transmission costs, use of the system charges and a specific charge related to the benefits yielded by the existence of interconnections among the different countries. Variable transmission costs result from the application of nodal prices. Use of the system charges comprise two different payments: tolls, which are obtained through a sensitivity analysis, and a complementary charge aimed at completing the recovery of the regulated revenue of each installation. The toll can be computed as the long run marginal cost (LRMC) for each agent or by applying MP to allocate the use made of the lines (50% to generation and 50% to loads). The complementary charge is necessary to complete the recovery of the regulated revenue of those installations whose flow is below the capacity of the line (those that are not used 100%). It is a postage stamp charge. The net balance for each country modifies the transmission charges to be paid by national generators and loads. Inter-TSO compensations take place annually. A more detailed description of the market and the computation of transmission tariffs is provided in section 3.5.3.1 and (WB 2000).

The assessment of the tarification methods included in the approach proposed here is the same as that provided in section 3.3.2 when discussing these methods in the context of a perfectly integrated market. In any case, combining several different mechanisms is unlikely to result in efficient price signals.

3.3.4 Different aspects of the implementation of any transmission tariff. Structure of the tariff and others

This section discusses some additional aspects of the tariff setting process, namely the structure of transmission charges, the method used to recover the cost of the unused fraction of the capacity of transmission lines and the number of scenarios considered when determining the level of the tariffs. These three issues are discussed separately.

3.3.4.1 Structure of transmission charges

We must distinguish between the method applied to compute transmission charges and the structure of these charges. Sections 3.3.1, 3.3.2 and 3.3.3 have described and discussed several methods to compute connection and use-of –the-system transmission charges. This section discusses the different structures of transmission charges applied in real-life systems. These include:

- Energy charges: each agent must pay a certain charge per unit of energy produced or consumed. Unit charges may vary from node to node, they may be

uniform within each one of a set of areas or a single energy charge may be computed for the whole system. Energy charges are known as linear tariffs as well.

- Capacity charges: each agent must pay a certain charge per unit of capacity contracted, or per unit of power consumed or produced at a certain time. Unit charges may vary from node to node, or from area to area, according to different considerations. Alternatively, a single capacity charge may be computed for the whole system.
- Fixed charge: the amount paid by each agent does not depend on how much capacity it has contracted or how much electricity it has traded.
- Non-linear tariffs: these are combinations of the former ones. They are also called multiple part tariffs (two part tariffs with an energy and a capacity charge for example, etc).

The method used to compute transmission charges may not determine how much each agent must pay, but the total amount to be paid by each group of agents (those belonging to the same country or area, for example). In other cases, the method used in a system only establishes some general rules to be applied in order to compute tariffs (like the postage stamp method, which establishes that transmission tariffs must be structured as a uniform charge but does not say anything about the structure of this charge). Then, the structure of tariffs may determine how the total transmission charge to be paid by a group of agents (maybe all of the agents in the region) is distributed among these agents. Consequently, the annual charge to be paid by each agent would depend on the structure chosen for the tariff. Thus, for example, if network users in a system have to pay a uniform capacity charge, an agent that consumes a large amount of power at very specific times will pay twice more than another one who consumes half this amount of power every hour of the year. On the other hand, if they have to pay a uniform energy charge, the latter agent will certainly pay much more than the former.

Alternatively, tariff structures may become irrelevant if the method applied to compute transmission charges determines the level of the transmission charge itself. If the annual regulated cost of the grid is allocated to network users in proportion to the total electrical usage that each one of them has made of the grid, in 'n' snapshots that are deemed to be representative of the operation throughout the year, the structure adopted by charges will not affect the total amount to be paid by each agent. Thus, for example, the unit charge corresponding to an agent might result from dividing the total amount to be recovered from this agent (known before computing this unit charge) by the amount of capacity it has contracted.

In (Perán, Pérez-Arriaga et al. 2001; Pérez-Arriaga, Montero et al. 2002) the structure of the transmission tariffs in the different EU countries is explained. (Fernández and Remacha 2005) provides a more recent review of the structure and level exhibited by

transmission tariffs in 19 European countries belonging to ETSO. Authors show the importance of non-transmission cost factors that are included in the transmission tariff applicable in each system (an example is the cost of decommissioning nuclear power plants in some countries). There are reports, like (CER 2004), which describe the transmission tariffs used in countries all around the world.

3.3.4.2 Recovery of the cost of the unused fraction of the capacity of a line

Another aspect to be taken into account when designing transmission tariffs is the treatment to be given to the unused capacity of each line. In principle, the owner of a regulated line should recover the whole cost incurred when constructing, operating and maintaining it. Lines that are not regulated (investments at risk or merchant lines) should not receive any regulated payment. Their owners should only earn the corresponding congestion rents. Thus, this issue should not be controversial.

However, in the context of regional markets, an area may oppose paying the unused fraction of a line whose construction and size has been decided by another area, even when this line is significantly used by agents within the former area. Several possibilities exist in this regard. Presently, the inter-TSO compensation mechanism applied in the IEM of the EU allocates the whole cost of each line belonging to the HN among the different TSOs. Therefore, agents in a country may pay part of the unused capacity of a line located outside this country. A document by the Florence School of Regulation on the ITC method to be applied permanently proposes to distinguish between new and existing lines, see (FSR 2005). According to this work, agents in an area should only pay the cost of external lines to the extent that they use these lines. Therefore, the non-used fraction of the capacity of exiting lines would always have to be paid by the country where the line is located (which is the one that decided to build it). Regarding new lines, if their construction is jointly decided by several TSOs, these TSOs should pay the whole cost of the line in proportion to the electrical usage that each one of them makes of this line (regardless of where the line is located). On the other hand, if the construction of a new line is decided unilaterally by the country where this line is located, other countries would only pay the first one the cost of the fraction of the line that they use.

In the context of the MER in Central America, a distinction is made between those regional regulated lines that were promoted by the EOR and those that were promoted by countries. Regional transmission charges must pay the whole cost of the former while they only recover the cost of the used fraction of the latter.

The academic community has also provided some mechanisms to deal with this issue. Thus, in (Hur, Yoo et al. 2004) the cost of the unused fraction of the capacity of each line 'l' is allocated based on: a) the participation of each agent in the electrical usage of the remaining lines under normal conditions and b) the impact that an outage affecting these other lines has on the flow through line 'l' (n-1 security criterion).

3.3.4.3 Number of snapshots considered when computing transmission tariffs

Again, this issue may be relevant or not depending on the method used to compute transmission charges. In general, when transmission charges depend on the economic benefits that network users obtain from the grid (or on the electrical usage they make of this grid), the operation of the system in several snapshots, which are deemed to be representative of the system functioning throughout an entire year, is considered. On the other hand, when transmission charges do not depend on the economic benefits that the transmission grid yields, nor do they depend on the electrical usage that agents make of this grid, considering more than one snapshot generally is unnecessary. Section 3.5.1.3 discusses this issue in the context of the IEM of the EU.

3.3.4.4 Global split of transmission charges between generation and demand

The split of the grid costs between G and L is also relevant (50% to generation and 50% to demand, 100% to generation, 100% to demand, etc). In most systems consumers pay a larger fraction of the total grid costs than generators. This is based on the fact that operating decisions by load (especially households) tend to be less sensitive to energy prices (and, in particular to the level of transmission charges) than those of generators. (Pérez-Arriaga, Montero et al. 2002; Fernández and Remacha 2005) provide an overview of the solution adopted in this regard by most EU countries.

3.4 Proposed solution to the problem of allocating the grid costs

As section 3.2 of this chapter explains, the cost of building, operating and maintaining a line should ideally be allocated based on the market and reliability benefits that each agent obtains from the existence of the line. Section 3.2 also explains that economic and reliability benefits are very difficult to assess, with the exception of connection lines (this is typically straightforward) and new lines (whose cost may be allocated considering the estimated operation of the system during a small number of years following their construction). Previous work presented in (Rubio 1999; Rubio and Pérez-Arriaga 2000) shows that, in most cases, trying to estimate benefits is meaningless, given the large amount of sensitive information required and the extraordinary difficulty of the task.

Thus, many tariffication methods employ electrical usage as a proxy to economic benefits. The electrical usage of a line by an agent seems to be a reasonable approximation of the benefit that the agent obtains from the line, see (Rubio and Pérez-Arriaga 2000). Despite acknowledging that economic and electrical usage do not need to coincide, it is clear that both of them are related. However, estimating the electrical usage made of the grid by market agents is not possible, strictly speaking. We only know how much power is consumed or produced in each node and how much flows

over each line. However, we will never know how much of the power physically produced by a generator is consumed by each load.

Despite all this, some measure of electrical usage might be useful in determining which fraction of the cost of the line an agent should be deemed responsible for. The work presented in (Rubio 1999) shows that there are significant similarities between the numerical results produced by two usage-based methods (Average Participations and Marginal Participations or Areas of Influence) and the results produced by the beneficiaries method. However, results yielded by usage-based methods are, by no means, identical to those produced by the beneficiaries method. What is more, results produced by different usage-based methods may be significantly different. Therefore, the decision on which usage-based method to apply in order to compute network use-of-the-system charges is relevant. Now, we revisit the most important methods for the computation of the electrical usage of lines by generators and loads. These are, as explained in section 3.3.2.1, Average Participations (AP) and Marginal Participations (MP).

3.4.1 Main usage-based methods

This section discusses in more detail the different methods for the allocation of the cost of the transmission facilities² that are based on the concept of electrical usage. Aspects like the international experience in the application of each method are also discussed. Here we shall only focus on the two most promising methods, namely AP and MP. The next paragraphs discuss the reasons why the remaining methods have been discarded from the outset, though arguments have been already provided in sections 3.2 and 3.3.2.

As explained before, the responsibility in the grid costs incurred by the system bears no relation to the commercial transactions that the agents are involved in. We shall assume that perfect information exists, i.e. every agent is aware of the business opportunities present in the market and they can perfectly arbitrage prices among different locations. Let us imagine that we were able to repeat many times the trading situation whereby the agents in a region agree on the commercial transactions they want to engage in. Then, every time we let the market agents trade their energy they would probably come up with a different set of market transactions. However, the group of generators selling power and the group of consumers buying it would always be the same. Only the most efficient agents whose power injections and withdrawals are compatible with the physical operating constraints imposed by the system would operate.

² We use the terms ‘lines’ and ‘transmission facilities’ indistinctly in this context. In any case, we always refer to the infrastructure needed to supply loads with the power produced by generators.

Therefore, those methods which make transmission charges dependent on the commercial transactions arranged in the market place or within the corresponding political borders are fundamentally flawed and must be discarded. The market outcome should not be affected by political borders in the region but only by the cost structure of the generators offering to produce electricity. Transmission charges should vary by agent, according to the location of these agents in the grid and their profile of production or consumption of electricity. Transaction-based methods are not an option for the allocation of grid charges.

Within the non-transaction-based methods, some of them are clearly outperformed by others. The Kirchen's method is based on the AP method. This method identifies the groups of nodes supplied by the same set of generators, the domains, and applies to them the AP method instead of applying it at node level. This makes the method more complex than necessary. Furthermore, the classification of the system nodes in domains lacks any theoretical justification and adds an extra level of arbitrariness to it.

As for the simplified version of AP, the use of this method only makes sense when the results obtained are similar to those yielded by the AP method or when we do not have available the information required to apply AP in full scale. These premises are not true for most of the real life systems we may encounter.

Consequently, we shall focus on the only two methods that compute individual transmission charges for network users based on the estimated electrical usage that these agents make of the grid: Average Participations (AP) and Marginal Participations (MP).

3.4.1.1 Average Participations

Background

The basic intuition behind the Average Participations method is that the sources of the supply to loads and the destination of the power injected by generators, as well as the responsibility for causing the power flows, can be assigned by employing very simple heuristic rules that only make use of the actual pattern of network flows. One could argue that, in an electricity market that works reasonably well, the power flows from nodes where it is less expensive towards nodes where it is more expensive. Thus, using the actual network flow pattern may be a way of assigning sources and sinks to loads and generators, respectively, in a reasonable and economically-meaningful manner. It is not the only possible way, but it is intuitive and simple to explain and to compute. More details on the justification of the AP method from the point of view of the physics of power flows on lines are provided in section 3.4.3.2.

The algorithm of Average Participations is simple and robust. The basic inputs to the model are the actual network topology and a comprehensive set of historical or computed network flows. In order to track the flows through the grid the method

assumes that inflows to a node are distributed proportionally among the outflows. This is the so-called proportionality assumption. Causality, or attribution of responsibility, directly results from this assumption that allows one to trace the flows backward or forward. Robustness, i.e. the ability of the method to produce results that are not sensitive to small variations in the input data, as well as the absence of arbitrary decisions, such as the choice of a slack bus, which may critically influence results, is a very important property. This is even more so when the application of the method has significant economic implications for network users. Having said that, it must be acknowledged that the method of Average Participations is not free from assumptions. The validity of the proportionality rule used to track flows cannot be proved using engineering principles. However, this rule is of a physically intuitive nature, which is very much associated with the robustness qualities of the method. See section 3.4.3.2. Furthermore, it basically guarantees the reasonability of its results, which cannot be easily challenged.

The algorithm of Average Participations has been used, with minor variations, in New Zealand. Additionally, it was considered for its implementation in Poland and South Africa, as well as in specific transmission pricing studies in, at least, Chile, Central America, Romania and Spain. It is also partly used in the Central American Regional Electricity Market. It can be easily programmed and it has been studied thoroughly in the technical literature. Reference to some additional work on the application of AP can be found in section 3.3 of this chapter. It is not free from objections, as it will be further discussed below, but it appears to be a reasonable procedure whose level of sophistication is well adapted to the problem under consideration.

Detailed description of the method

The method requires as its basic input data a collection of snapshots of the network power flows corresponding to the specific system conditions of interest. The algorithm is based on the assumption that electricity flows can be traced—or the responsibilities for causing them can be assigned—by supposing that, at any network node, the inflows are distributed proportionally among the outflows. Implicit in this rule is the assumption that electricity does not flow in the opposite direction to that of the prevalent flow over each line, which according to this assumption is the only existing flow (AP dismisses the existence of counterflows). Under these assumptions, the method traces the flow of electricity from individual sources to individual sinks. In other words, the model identifies, for each generator injecting power into the network, physical paths starting at the generator that go through the grid until they reach certain loads where they end. The method also works the other way around, thus identifying paths from loads to generators. Those paths are coincident with those starting from generators and therefore yield exactly the same results of allocation of responsibility in network flows to generators and loads. Then, the cost of each line is allocated to the different grid users according to how much power produced or consumed by each

agent flows along the corresponding line. Figure 3.4-1 shows a simple example of application of the method.

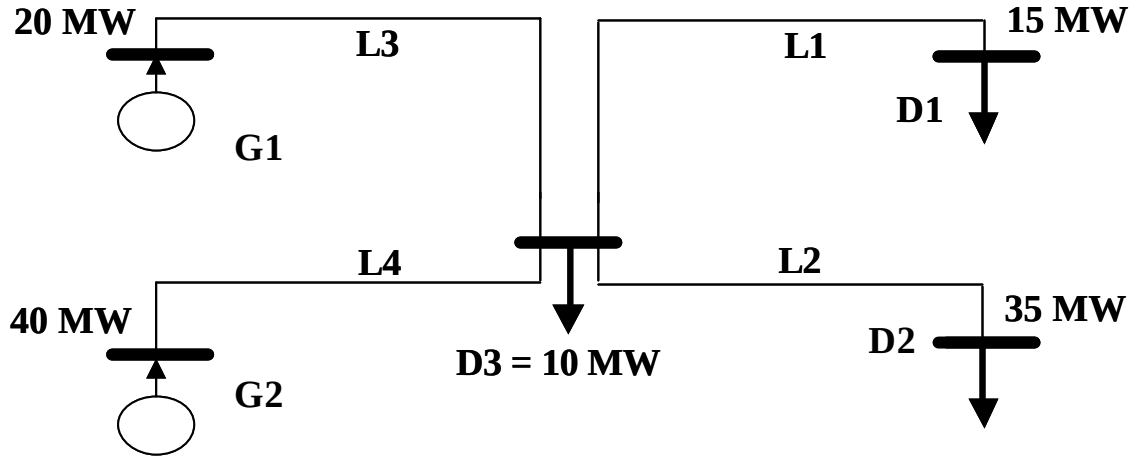


Figure 3.4-1: Proportionality principle in average participations

According to the proportional distribution rule, generator G1 would contribute $15 \times 20 / (20 + 40)$ MW to the flow over line 1, $35 \times 20 / (20 + 40)$ MW to the flow over line 2, 20 MW to the flow over line 3 and nothing to the flow over line 4. Similarly, demand D2 would contribute $20 \times 35 / (10 + 15 + 35)$ to the flow over line 3, $40 \times 35 / (10 + 15 + 35)$ to the flow over line 4, all 35 MW to the flow over line 2 and nothing to the flow over line 1. Demand D3 has a contribution of $20 \times 10 / (10 + 15 + 35)$ to the flow over line 3, $40 \times 10 / (10 + 15 + 35)$ to the flow over line 4 and no contribution to the flows over lines 1 and 2.

The participation of agent i in the flow through line j is obtained as the part of the flows starting at agent i that passes through this line. The method implicitly results in a 50/50 global allocation of costs to generation and load. However, if desired, an ad hoc weighting factor could be used to modify this percentage. Other methods let us change the global split of transmission charges between generation and load as well. For instance, in the MP method this decision is linked to the choice of the slack node, as we shall see in the next subsection.

3.4.1.2 Marginal Participations (MP)

Background

The method of Marginal Participations (often named “areas of influence”, after Chile) has been used in Chile and Argentina since these systems were restructured in 1981 and 1992, respectively, as well as in other Latin American countries. See (Rudnick, Palma et al. 1995) for a study on the application of the MP method in Chile. This method calculates how much the flow over line j would increase if generation (or load)

located at node i increased by 1 MW. Due to Kirchhoff's laws, any 1 MW increase in generation (or load) at node i needs to be accompanied by a corresponding 1 MW increase in load (or generation) at some other node or nodes (ignoring losses). Thus, the impact on the network flows of a marginal increase in the net injection at a certain bus depends on the decision about which is the responding node. The MP method assumes that the responding node or nodes, which are used to compute the participation of agents in the use of lines, are the same regardless of where the concerned agent is located. Using the same slack node for every agent in the system results in very elegant mathematical properties of the method, see (Rivier and Pérez-Arriaga 1993). However, this assumption does not seem to be in accordance with the reality of the functioning of large systems where systematic congestion exists. The decision on the location of the responding node, which is an input to the method, heavily conditions the results obtained. The impact of a unit transaction between node i and the responding node on the flow over line j is called the unit marginal participation of an agent located at i in the flow over j . The total participation of the agent in the flow through line j results from multiplying its unit marginal participation by the total amount of power produced or consumed by the agent. This calculation is carried out for every node, for all the lines in the grid and for a representative number of load flow scenarios. Then, the cost of each line is allocated to the different users according to their participations in the line.

Different choices are possible regarding the election of the “slack bus” (the responding node or nodes in power systems terminology). It could be chosen near the major load center (as it is done in Argentina or Chile). Alternatively, we could define a distributed virtual node so that all the demand or generation or both responds pro rata. Another option is making the slack bus coincident with the marginal generator in the market or in the centralized economic dispatch, etc. Choosing different responding nodes leads to widely different results. Despite the fact that the selection of the slack bus has essential practical consequences, there is no criteria to decide which slack bus to use.

Description

Any usage-based methodology tries to identify how much of the power that flows through each of the lines in the system is due to the existence of a certain network user. This measure is then used in order to charge the agent according to the adopted measure of utilization. The MP method analyzes how the flows in the grid change when minor modifications are introduced in the production (or consumption) of agent i . It assumes that the relationship between the flow through line j and the behavior of agent i can be considered to be linear. The following is a description of the procedure to be carried out for each of the considered scenarios:

- a) Marginal participation sensitivities $A_{i,j}$ are obtained. These sensitivities represent how much the flow through line j increases when the injection at

bus i is increased by 1 MW, and the net injection at the slack bus decreases by 1 MW.

- b) Total participations for each agent are calculated as the product of its net injection by its marginal participation. The net injection is positive for generators and negative for demands. Therefore, the total participation of a generic agent i in the flow over line j is obtained as $A_{i,j}(g_i - d_i)$.
- c) The cost of each line is allocated pro rata to the different agents according to their total participation in the flow over the corresponding line.

The linearity assumption does not introduce significant errors. The critical task is the computation of the sensitivity factors $A_{i,j}$.

As mentioned above, it is a fundamental technical characteristic of power systems that generation and demand must always be balanced. Therefore, if the generation at node i is increased by 1 MW in order to compute marginal sensitivities, then some other nodes in the grid must increment their demand or reduce their generation in order to keep the system in balance. This is the reason why sensitivity $A_{i,j}$ actually expresses how much the flow through line j increases when generation at bus i is increased by 1 MW and the demand at the slack bus is increased by 1 MW (ignoring losses). In other words, $A_{i,j}$ tells us how much of a hypothetical transaction starting at node i and ending at the slack bus would go through line j .

Variants of the method

DC load flow models allow one to define any node or combination of nodes as the slack bus using some easy numerical manipulation. However, one must be aware that choosing the nodes that respond critically determines the final results. This is a decision not to be made arbitrarily. A common choice for the slack node is a major load center. Chile and other Latin-American countries use some network node that is as close as possible to the largest demand (typically, the capital of the nation). In this case, network users that are close to the slack bus (most of them consumers) pay reduced network charges, while distant market participants (most of them generators) tend to pay high transmission charges.

Using a distributed virtual node is another option. In this case all the generators (or all the loads, or both) respond jointly to any imbalance in proportion to their level of production (or consumption, in the case of loads). According to this alternative, in order to compute the marginal participations of an agent located at i in the line flows, we must assess the impact on these flows of a transaction that starts at node i and reaches practically every node in the system. Consequently, agent i participates in the flow on almost every line in the grid. Of course, participations are higher in those lines that are closer to the considered agent, but the area where these participations are not

negligible is relatively wide. In fact, depending on the generation and load pattern, agent i may have an important participation in the flow over very distant lines.

Finally, it is also possible to take as the slack bus the node where the marginal generator in the dispatch is located. This is the most efficient solution, conceptually speaking. Each agent would be found responsible for the increment (decrement) in the total use made of the grid that would result from an increase in the power this agent trades, i.e. each agent would be charged according to the true marginal use that he makes of the grid. Notice that, in order for this variant of the method to be properly applied, the same marginal generator should respond to an increase in the power traded by every agent in the system. This is not true in large systems, where transmission losses are significant, and those systems where congestion occurs, thus creating several areas within which the operation is independent from the rest of the system.

3.4.2 Assessment of the main usage-based methods

The success of a network cost allocation method in promoting, or at least facilitating, cross-border electricity trade in a multinational market depends on a wide variety of factors. The two methods selected before, AP and MP, are compared here with regard to the main aspects to be considered when assessing their performance, namely: the soundness of the method from a technical point of view, its compatibility with the remaining aspects of the regulation of the transmission service, its compliance with the Single System Paradigm principle, the soundness of the method from an economic point of view, the data requirements, the simplicity and transparency of the method and, finally, the soundness of the results provided by each method in a regional context. When talking about the MP method, we refer to the variant where the slack node is chosen to be located at the same node as the marginal generator in the dispatch. If different areas with different marginal generators exist (as it will probably be the case in regional markets like the IEM), choosing as the slack bus a single node (or a set of nodes) that would respond in the dispatch to a marginal increase in the power produced or consumed by every agent in the system is not possible. Different generators would respond to marginal variations in the dispatch in different areas. Thus, whatever the slack node we choose, it will never be the marginal generator considered in every part of the system. Choosing a different slack bus for each area is not possible since this is not consistent with the MP method. Choosing as the slack bus the marginal generator in one of the areas would not be conceptually sound because of the following reasons:

- It would not be coherent with the fact that congestion in the system and transmission losses result in the marginal generator in different areas being different.

- It would not either be coherent with the fact that the each country or area in the region may carry out a separate energy dispatch.

3.4.2.1 Soundness of the method from a technical point of view

Both AP and MP take electrical use as a proxy to responsibility in network investments or the economic benefits produced by each line, which are the two reference principles that should be taken into account to compute network charges. Any usage based method relies on some assumptions or arbitrary decisions, since the electrical usage cannot be computed in an indisputable way. Then, the difference between both methods lies in the assumptions made when determining the usage that each agent makes of the grid. Now, we recall the main assumptions behind the AP and MP methods:

- **Average Participations:** Accepting the technical impossibility to assign line flows unambiguously to sinks or sources, it is assumed that this responsibility can be simply obtained by tracking each flow upstream & downstream following the same paths & branching proportions of the actually existing line flows. By construction, the use of each line is assigned 50/50 to demand & generation, although this proportion may be easily modified afterwards if the user of the algorithm has a reason to do so.
- **Marginal Participations:** Any power inflow or outflow in any node of the system has as source or sink the same (slack) node, which has to be chosen somehow before the method can be applied. Here, we consider the case where the slack node is the one where the marginal generator in the dispatch is located. This seems to be a reasonable assumption in local markets (national markets in the IEM, states in the US). However, in regional markets, even if a single dispatch is carried out for the whole region, several price areas may appear if there are congestions, each one with a different (set of) marginal generator(s). As explained above, the MP method cannot provide any satisfactory solution in this case.

Given a slack node (whichever we take):

- o The responsibility of a given inflow (outflow) in the use of each line of the network is obtained by injecting (retrieving) power at the considered node & retrieving (injecting) it at the slack node, & examining which is the impact on each line of the entire network
- o Obviously, participations obtained critically depend on the choice of the slack node
- o The choice of the slack node also determines the split $X / 100-X$ between the total charges to all demands & to all generations, see the discussion in Annex 1 of this chapter.

When evaluating AP and MP methods, one should do it based on the underlying assumptions behind each one of them, since this is what determines if a method is correct or not. However, a frequent mistake is to ignore the assumptions & to focus on the computational algorithms that are needed to make the calculations that each method requires. Note that if the assumptions behind a method are wrong, this cannot be fixed by a sophisticated numerical algorithm that just makes calculations based on the mistaken assumptions. Also note that if the assumptions behind a method are reasonable, the results will be also reasonable, even if the required computations are simple to perform. Given that the underlying assumptions behind the AP method seem far more reasonable than those under behind the application of the MP method, we must conclude that the former definitely ranks higher than the latter in this respect.

Besides assessing the rationale behind each method, we must determine whether the results that AP and MP provide are technically sound with respect to several dimensions. These include the effect that the election of a slack node has on the results produced by the method, the dispersion of the participations of market agents in the electrical usage of lines, the results obtained for balanced groups of market agents (those where total generation within the group equals total demand) and the coherence of the results with the objective devised for the method.

Effect of the selection of the slack bus

The selection of the slack bus has a major impact on the process of allocation among market agents of the electrical usage that is made of each line. In the AP method the selection of the slack node is implicit. The method itself determines who is consuming the power produced by each generator and vice-versa. As for the Marginal Participations method, Appendix 1 shows that modifying the slack node boils down to modifying the overall allocation of grid costs between generation and load. Choosing one slack node or another does not alter the differences that exist among the locational signals sent to market agents, which are independent from the slack node considered. However, the fact that differences among transmission charges produced by MP do not depend on the choice of slack bus does not mean that these differences are efficient. In other words, despite exhibiting this elegant property, these charges may be completely nonsensical.

Dispersion of the participations

The Marginal Participations method results in participation factors that are very dispersed. According to this method, and regardless of the flow pattern in the grid, many agents are found to be responsible for a non-negligible part of the flow over very distant lines. This is a direct consequence of the fact that the slack node is the same for every node in the system. This method assumes that the power injected or retrieved at each node flows through the grid according to Kirchoff laws until it finally reaches the slack node. All those agents located far away from the chosen reference node will be

deemed responsible for non-negligible flows over lines that are close to the slack node and, consequently, far from them. Again, the dispersion of the results obtained with the MP method may only be acceptable if we compute marginal participations at local (national) level, and the slack node chosen is the one where the marginal generator is located. However, if the method is applied at regional level, results will be wrong.

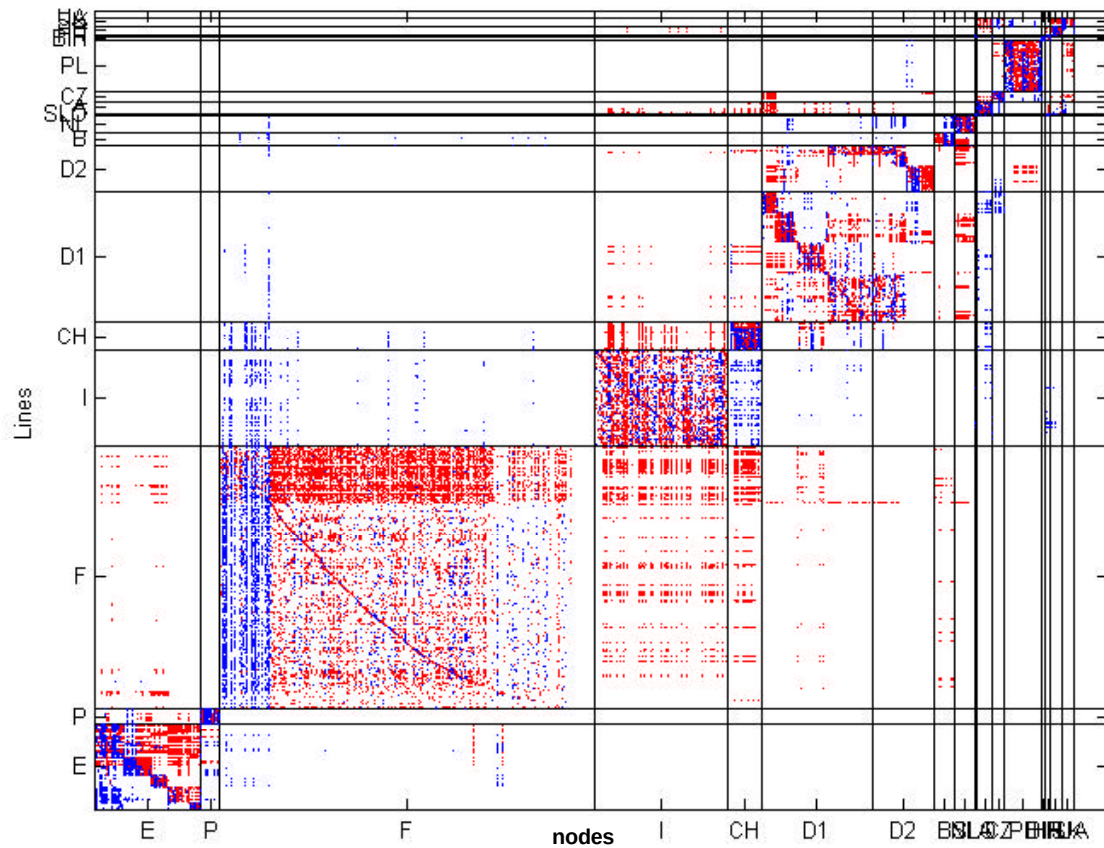


Figure 3.4-2: Participation of generation and demand nodes in the lines of the UCTE system according to AP

As for the Average Participations method, it results in participations that are far less dispersed. The electricity grid usage by each agent is restricted to the area where it is located except for those cases where there are prevalent flows crossing large system areas. Figure 3.4-2 provides information for the UCTE system on how far from each line the nodes that participate in the use of the line are according to AP. All 3,383 nodes in the network of the considered system are represented in the horizontal axis and all 3,655 lines in the vertical axis. Nodes and lines that are located within the same country have been represented together. Eurostat codes have been used to indicate the country that each set of nodes or lines belongs to. Nodes and lines within each country have been sorted by proximity. There is a dot in the figure when a node has a non zero participation in a line. Dots are blue for generation nodes and red for demand nodes. With France being an exporting country, it makes sense that red dots do not extend outside France while the blue ones reach as far as Spain, Italy, Switzerland, Germany or Belgium. From the figure it is immediate to conclude that there is a higher density of dots in the nodes that are closest to the lines.

Results for balanced groups of agents

The net participations assigned by the MP method to a balanced group of network users, —obtained by adding up the participations of all the individuals forming the group—, are independent from the choice of slack node. Application of the superposition principle shows that the joint effect on the system flows of a 1 MW transaction between a generator A and the slack node, on the one hand, and a 1 MW transaction between a demand B and the same slack node, on the other, is the same as the effect of a unitary transaction between generator A and demand B, this with independence of the slack node considered.

It is possible to compute the flows caused by a balanced set of injections and withdrawals just by analyzing how flows change when all these injections and withdrawals are removed from the system. This can be done without having to define any slack node. Assuming that the superposition principle holds when allocating responsibilities in grid costs, the total network usage yielded by the Marginal Participations method for a balanced group of agents coincides with the flows actually caused by this set of agents. For instance, a set of agents consisting of a generator and a demand of the same size and located at the same node will be assigned no network usage by the MP method. This feature is an indicator of the soundness of a method only when this method is examined from the point of view of the superposition principle. Thus, it should not be considered as a strong criterion to determine the value of a usage-based method. In fact, the MP method is unable to capture several effects, such as congestions and other non-linearities, that characterize the functioning of the system. This may weaken the reasoning above.

The AP method tracks the power flows that actually exist. Thus, it tracks the power injected by a generator until it reaches the loads where it is consumed. Consequently, according to the AP method, the power produced (consumed) by each generator (load) is consumed (produced) by a different set of loads (generators). This feature of the method still applies when we compute the total net usage that a balanced group of agents makes of the grid. It is the AP method itself which determines where the power injected by a generator is consumed, and vice versa. AP does not leave this decision in the hands of the person who applies it.

Coherence of the obtained results with the expected transmission tariffs

The results provided by the AP method are more in line with the expected transmission tariffs than those provided by the MP method. This is especially true in regional markets, where participations produced by AP are far less scattered over the regional grid. As we have already mentioned, results for MP do not seem to adapt well to the case of large regional markets where significant and systematic congestion takes place.

3.4.2.2 Compatibility with the remaining aspects of the regulation of the transmission service

Here we analyze the compatibility of the results provided by the studied methods for the allocation of grid costs with the desirable treatment to be given to other aspects of the regulation of the transmission activity in the context of multinational markets. These aspects are the design of transmission locational signals and the allocation of the cost of new network reinforcements.

Locational signals

Locational signals must drive, together with other factors, the installation of agents in the most efficient place in the grid. The two considered methods send some type of locational signals to generators and consumers since transmission charges for both present some form of spatial and temporal differentiation. However, locational signals computed with AP are significantly different from those produced by MP. As a matter of fact, whereas MP can give rise to both positive and negative locational signals, usage factors obtained with AP are always positive.

The two considered methods allocate already incurred costs. They do not take into account the cost of lines that have not been constructed yet. Locational signals must make market agents internalize the cost that the system incurs as a consequence of their decision to install in one point of the grid or another. Therefore, we should assess whether locational signals produced by each one of these methods comply with the aforementioned objective.

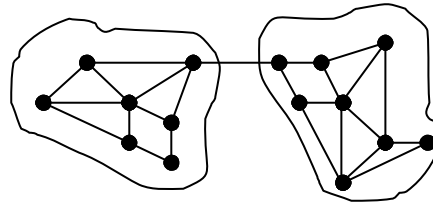


Figure 3.4-3: Two areas example

It can be easily checked that the locational signals produced by MP are not cost reflective in some cases. Consider the system in Figure 3.4-3. The system comprises two well-meshed areas linked by a very weak interconnection. If a fully distributed slack node is chosen, application of marginal participations results in a generator located in any of the two areas participating in the use made of lines on both sides of the interconnection. Then, this generator has to pay a relevant part of the network in the neighboring area. This is true even though the power flow through the interconnection might be close to zero. Common sense suggests that this generator should basically pay transmission charges for the network in the area where it is located. Hence, the allocation of network costs that is obtained by MP appears to be incorrect.

It is much more difficult to find counterexamples that challenge the validity of the results provided by the AP method. One may argue that the installation of those market agents that contribute to balancing generation and demand within an area may reduce, in some cases, the amount of future grid investments that will be required. Therefore, transmission charges for these agents should be negative. Contrary to what should happen according to this reasoning, transmission charges produced by AP are always positive. This is due to the fact that the AP method is unable to consider the incremental effect that the installation of a new agent may have on the number and size of the grid investments needed. AP just computes the average use that each agent makes of the grid.

Allocation of the cost of new network reinforcements

The two considered methods separately allocate the cost of each line in the system to individual agents. Therefore, AP and MP automatically allocate the cost of new network investments. Given that the two methods only consider already sunk costs, they will only allocate the cost of new lines from the time they enter into operation. Provided we accept the cost allocation process followed in a method, we should deem this method appropriate to assign the cost of any line in the system, including the new ones. Participation factors both in MP and AP add up to 1. These participations represent the fraction of the cost of the line that the corresponding agent must pay. We must bear in mind that, whereas participations produced by AP are always positive, those produced by MP may be either positive or negative. In other words, according to the results produced by the MP method, some agents (or areas) may have to be compensated for the construction of a new line. This does not seem to be acceptable in reality.

If applied at regional level, employing an efficient and non-disputable method to allocate the cost of network investments that are relevant to the trade in the region is critical in order to get them built. Otherwise, these investment projects will face the opposition of countries arguing that the fraction of the cost of the new line they would have to pay is disproportionately high when compared to the benefits they would obtain from it.

3.4.2.3 Compliance with the Single System Paradigm principle

The single system paradigm establishes that agents in a multinational market must not be treated in a discriminatory way based on the country or area where they belong. They should all be considered as if they belonged to the same single system.

The two methods allocate the cost of each line to market agents. Then, aggregation of these charges results in the total transmission charge that each agent must pay for the use of the entire grid. If applied in a regional context, aggregation of the latter charges at country level yields the total grid payment faced by the agents within each country.

Political borders are totally ignored in the network cost allocation process. They are only used when adding up the charges of the agents within each country.

Again, the two methods are coherent with this principle since they result in charges for individual agents which, provided they are not aggregated at control area or country level, do not depend on the political borders or the considered subsystems.

3.4.2.4 Soundness of the method from an economic point of view

The economic properties displayed by a method play a key role in the assessment of its overall performance. Network cost allocation methods must be efficient from an economic point of view. They should encourage agents to behave optimally in the long run. Besides, they should not interfere with the efficient short term signals.

Both AP and MP rely on a basic assumption: electrical usage is tantamount to responsibility in investment or economic benefit, which are the sound criteria that should be applied for assigning network costs. Once the utilization of “network electrical usage” has been accepted as the criterion to allocate network costs, the discussion about whether AP is better suited than MP to carry out this task boils down to determining which one of the two methods is better to compute electrical usage. This is an issue that is mainly related with engineering.

The only possible advantage of ‘marginal’ methods over ‘average’ ones is that the former may produce either positive charges (when the installation of the agent contributes to grid investments) or negative ones (when the agent reduces the amount of network investments required).

Therefore, typical “average use” methods seem to be more appropriate than “marginal use” ones when we do not particularly want to send strong locational signals, but we are primarily concerned about recovering the sunk cost of the grid. In order to send stronger locational signals, we may alter the average participations so that new agents pay a larger part of the cost of new lines. See the Agent and Line Factors method (ALF) proposed in section 2.5.1 of chapter 2. We cannot prove that either AP or MP promotes allocative efficiency (efficiency in the long run decisions made by agents). In fact, none of the methods used in reality can claim to produce efficient long term signals.

In order not to interfere with the short run decisions made by agents, transmission tariffs must be structured as capacity charges, i.e. network users must pay a fixed charge per unit of installed capacity (or per MW of power consumed or produced at the time of their peak load / production). Additionally, given that transmission charges produced by AP and MP depend on the electrical usage that agents make of the grid, we may think of computing average unit charges for all the network users located within the same area that have the same load or generation profile. Thus, these unit charges would be very approximately independent of the unilateral operating decisions by market agents.

Game theory has been repeatedly used to allocate the cost of the grid among the agents operating in the system. Thus, for example, (PSR 2002) presents a cooperative game whose solution lies somewhere in between those yielded by the AP and MP methods. In this method, each generator supplies different loads (and each load is supplied by different generators), as in the AP method. Generators choose the loads they supply, and loads choose the generators that supply them, in order to minimize the transmission charge they have to pay. However, a load-flow is run to compute the participations of agents in the electrical usage of lines, as in the MP method.

Cooperative games have also been used to justify other methods, such as AP, from the point of view of economic theory. Many of these proofs take advantage of the commodity nature of electricity to propose some form of proportional sharing of a system variable (the inflows among the outflows for every node, the load among the output of generators, etc), see (PSR 2002; Bialek and Kattuman 2004). Section 3.4.3 provides a theoretical justification for the AP method based on cooperative game theory.

3.4.2.5 Data and algorithmic requirements of the different methods

Given that both AP and MP calculate the electrical usage of each installation by each agent, part of their input data are common to the two methods. A description of the data requirements of each one of the methods follows:

- **Average Participations:** application of AP requires a description of the network configuration (grid topology), as well as the pattern of generation and demand in the system for each one of the considered scenarios. Additionally, we need either the pattern of flows over the grid or an electrical model of the system so that we can run a load flow using the power injections and withdrawals and the network topology. If the system is comprised of several areas or countries, results may be computed separately for each of them. Thus, we may use separate data sets for each country or TSO. Developing the computational tools that are required to handling separately the data corresponding to each country has been a non-negligible contribution of the thesis.
- **Marginal Participations:** as in the AP method, the pattern of generation and demand in the system are input data. Additionally, an electrical model of the whole system is also necessary in order to compute the impact on the line flows of a unit transaction between each node and the reference one. Given that running a load flow of the whole system is necessary, we must use a single model for the whole system, even when it is comprised of several areas or countries. Therefore, this is a more stringent requirement than for AP.

Once computed the fraction of the cost of each line to be paid by each agent, participations need to be expressed in monetary terms. Thus, we need to use an

estimation of the cost of each line or an estimated average cost per km of line and information on how to compute the length of a line.

Due to the fact that the final objective is computing annual transmission rates, we need to consider scenarios that are representative of the different possible operating conditions during the year. Thus, all the above mentioned data must be available for all the scenarios though some of these data are common to all or, at least, some of them. Finally, transmission rates are obtained as the weighted average of the charges for all the scenarios. Results for each scenario are weighted with the estimated duration of the scenario.

3.4.2.6 Complexity of the method, dependence of the method on commonly agreed definitions and experience with its utilization

The underlying idea behind both the AP and the MP methods is easy to understand. Computing charges using AP involves applying some simple heuristic rules to the map of flows that actually exist. Thus, running a load flow is not necessary in this case. Application of MP, however, requires computing the transmission line flows in a fictive situation where the power produced or consumed by an agent is increased with respect to the existing situation. Therefore, in the MP method it is necessary to run a load flow, which is, conceptually speaking, much more complex than applying the tracing rule used by AP, despite the fact that load flows are common practice among power system engineers. In any case, this should not be a major issue when choosing between the two methods.

Regarding the experience in the utilization of each method, both AP and MP have been used in some countries to actually compute network tariffs. Therefore, there would not be significant differences between the two in this regard.

3.4.2.7 Assessment of the results of application of the method to a representative regional system

Given that one main focus of this thesis is the development of regulatory designs to be applied in regional markets, applying the AP and MP methods to the case example of a stylized simple regional system may be useful in order to draw some interesting conclusions that are based on the obtained results. Many of the conclusions derived in this point have already been mentioned in the previous ones. However, it is much more illustrative to see how numbers obtained corroborate them. Additionally, some new remarks are made here.

Participations by agents in the system (the aforementioned equivalent generators and loads) are aggregated at country level and, finally, compensations among countries are derived for the use that agents in a country would make of the grid of others. This is due to the fact that the example was created to help clarify the merits of each method in the context of the computation of inter-TSO compensations in the IEM of the

European Union. In any case, conclusions derived here are extensible to the case of single perfectly integrated systems in most cases.

The case example used here is a fictive one. Its is presented in Figure 3.4-4. This system is comprised of 6 areas (countries): P,E, F, D, I and H. Load and generation within each country have been represented using a pair of equivalent agents: one generator and one load. The system includes 13 equivalent lines, whose electrical parameters are provided in the figure, and 12 nodes (two per country) where either an equivalent generator or an equivalent load is located. Similarities between this system and Western Europe are obvious. However, we do not claim that the results obtained here are representative of the ones that would be obtained if the corresponding method would be actually applied in this region.

Figure 3.4-4 includes the generation or load located at each node, the power flow on every line and the reactance of every line as well.

The design options taken when applying the methods are:

- Half of each cross-border line is located within each one of the countries sharing the border.
- The natural 50/50 split of electrical usage between generators and loads in AP is not modified.

Now, the results obtained for each method are presented and commented on separately. Finally, results obtained for AP and MP are considered jointly and compared. Numbers corresponding to the allocation of the cost of individual lines (which may happen to be new lines) are also taken into account.

Results for the MP method

We shall present the results corresponding to two possible choices of the slack bus. In case 1 we have taken node 8 as the slack bus. In case 2, we take node 2 as the slack bus. Finally, we draw some conclusions from the numbers computed in both cases.

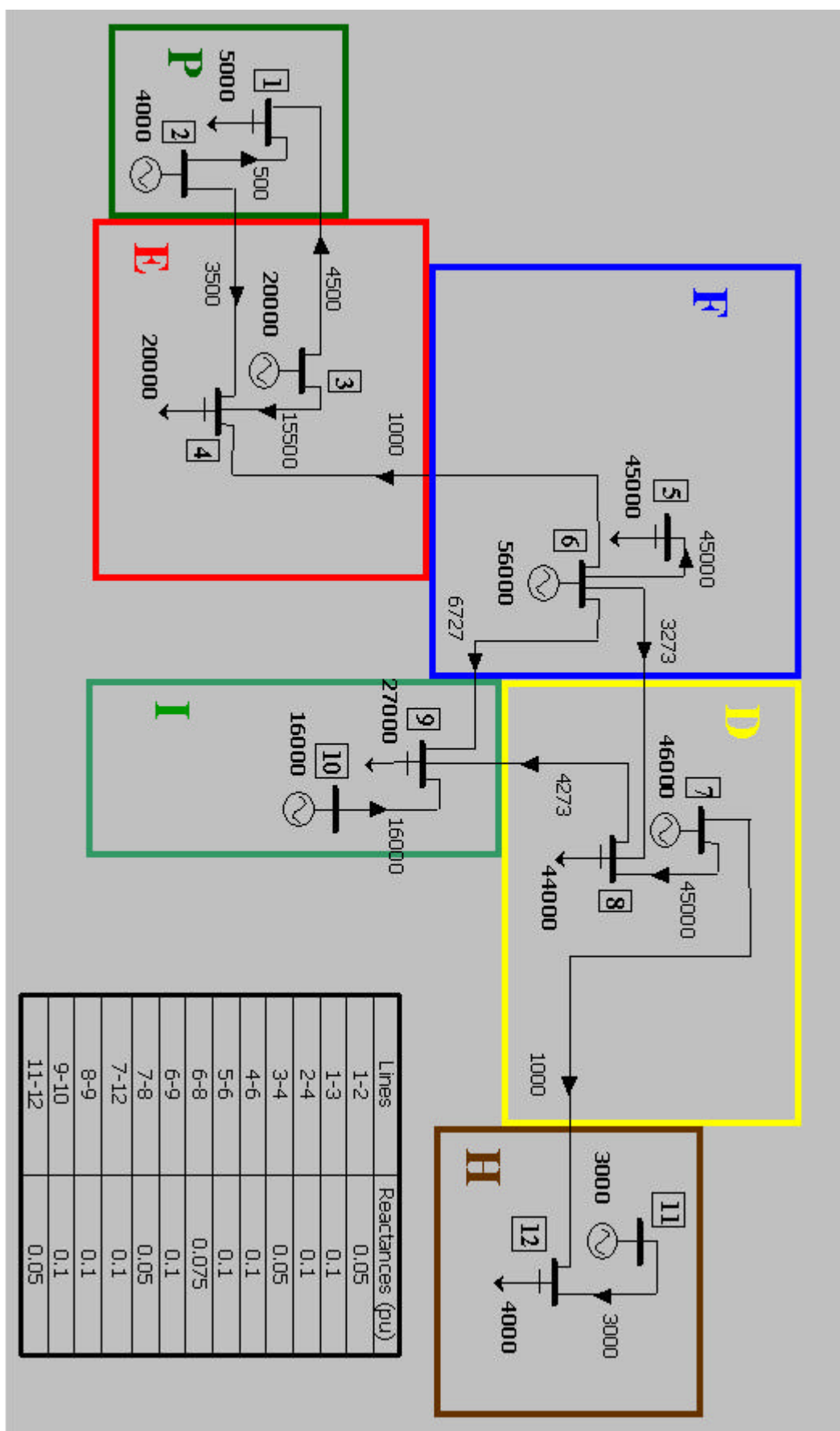


Figure 3.4-4: Stylized system used to show the properties of the results produced by the AP and MP methods

In both cases, numbers are presented in three different formats: first, Table 3.4-1 and Table 3.4-4 provide the participations of agents in the usage made of transmission lines in case 1 and 2, respectively; Table 3.4-2 and Table 3.4-5 provide, also for cases 1 and 2, the participation of all the agents within a country (generation and load considered together) in the usage made of the same lines. Numbers in all these tables are expressed in MW.

	<i>L1-2</i>	<i>L1-3</i>	<i>L2-4</i>	<i>L3-4</i>	<i>L4-6</i>	<i>L5-6</i>	<i>L6-8</i>	<i>L6-9</i>	<i>L7-8</i>	<i>L7-12</i>	<i>L8-9</i>	<i>L9-10</i>	<i>L11-12</i>
<i>N1</i>	2500	2500	-2500	-2500	5000	0	-3636	-1364	0	0	1364	0	0
<i>N2</i>	1333	-1333	2667	1333	-4000	0	2909	1091	0	0	-1091	0	0
<i>N3</i>	-3333	3333	3333	16667	-20000	0	14546	5455	0	0	-5455	0	0
<i>N4</i>	0	0	0	0	20000	0	-14546	-5455	0	0	5455	0	0
<i>N5</i>	0	0	0	0	0	45000	-32727	-12273	0	0	12273	0	0
<i>N6</i>	0	0	0	0	0	0	40727	15273	0	0	-15273	0	0
<i>N7</i>	0	0	0	0	0	0	0	0	46000	0	0	0	0
<i>N8</i>	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>N9</i>	0	0	0	0	0	0	-9818	9818	0	0	17182	0	0
<i>N10</i>	0	0	0	0	0	0	5818	-5818	0	0	-10182	16000	0
<i>N11</i>	0	0	0	0	0	0	0	0	3000	-3000	0	0	3000
<i>N12</i>	0	0	0	0	0	0	0	0	-4000	4000	0	0	0

Table 3.4-1: Participations by agents in the usage made of lines, resulting from the MP method, when we take node 8 as the slack bus

	<i>L1-2</i>	<i>L1-3</i>	<i>L2-4</i>	<i>L3-4</i>	<i>L4-6</i>	<i>L5-6</i>	<i>L6-8</i>	<i>L6-9</i>	<i>L7-8</i>	<i>L7-12</i>	<i>L8-9</i>	<i>L9-10</i>	<i>L11-12</i>
<i>P</i>	3833	1167	167	-1167	1000	0	-727	-273	0	0	273	0	0
<i>E</i>	-3333	3333	3333	16667	0	0	0	0	0	0	0	0	0
<i>F</i>	0	0	0	0	0	45000	8000	3000	0	0	-3000	0	0
<i>D</i>	0	0	0	0	0	0	0	0	46000	0	0	0	0
<i>I</i>	0	0	0	0	0	0	-4000	4000	0	0	7000	16000	0
<i>H</i>	0	0	0	0	0	0	0	0	-1000	1000	0	0	3000

Table 3.4-2: Participations of countries in the use made of lines, resulting from the MP method, when we take node 8 as the slack bus

	<i>P</i>	<i>E</i>	<i>F</i>	<i>D</i>	<i>I</i>	<i>H</i>
<i>P</i>	1292	292	45	-68	0	0
<i>E</i>	833	5833	0	0	0	0
<i>F</i>	0	0	24750	750	0	0
<i>D</i>	0	0	0	11500	0	0
<i>I</i>	0	0	250	1000	10750	0
<i>H</i>	0	0	0	0	0	1000
	833	292	295	1682	0	0
	269	833	750	0	1250	0
	564	-542	-455	1682	-1250	0

Table 3.4-3: Compensations among countries produced by the MP method corresponding to the usage that external agents make of the grid in each country. We take node 8 as the slack bus

Finally, Table 3.4-3 and Table 3.4-6 provide the compensations among countries for the electrical use that external agents make of the grid of each country. Table 3.4-3

corresponds to case 1 and Table 3.4-6 corresponds to case 2. In these two tables, the rows labelled with the name of a country show the use that the corresponding country makes of the grid of all the countries in the system. The row labelled 'Use by others' provides the use of the grid of each country by others. The row labelled 'Use of others' provides the use that each country makes of others' grids. Finally, the row labelled 'Net use' provides, for each country, the difference between the two previous results. Numbers in these two tables are expressed in '000 MWkm.

	L1-2	L1-3	L2-4	L3-4	L4-6	L5-6	L6-8	L6-9	L7-8	L7-12	L8-9	L9-10	L11-12
N1	4167	833	833	-833	0	0	0	0	0	0	0	0	0
N2	0	0	0	0	0	0	0	0	0	0	0	0	0
N3	-10000	10000	-10000	10000	0	0	0	0	0	0	0	0	0
N4	6667	-6667	13333	6667	0	0	0	0	0	0	0	0	0
N5	15000	-15000	30000	15000	-45000	45000	0	0	0	0	0	0	0
N6	-18667	18667	-37333	-18667	56000	0	0	0	0	0	0	0	0
N7	-15333	15333	-30667	-15333	46000	0	-33455	-12545	46000	0	12545	0	0
N8	14667	-14667	29333	14667	-44000	0	32000	12000	0	0	-12000	0	0
N9	9000	-9000	18000	9000	-27000	0	9818	17182	0	0	9818	0	0
N10	-5333	5333	-10667	-5333	16000	0	-5818	-10182	0	0	-5818	16000	0
N11	-1000	1000	-2000	-1000	3000	0	-2182	-818	3000	-3000	818	0	3000
N12	1333	-1333	2667	1333	-4000	0	2910	1090	-4000	4000	-1090	0	0

Table 3.4-4: Participations by agents in the usage made of electrical lines, resulting from the MP method, when we take node 2 as the slack bus

	L1-2	L1-3	L2-4	L3-4	L4-6	L5-6	L6-8	L6-9	L7-8	L7-12	L8-9	L9-10	L11-12
P	4167	833	833	-833	0	0	0	0	0	0	0	0	0
E	-3333	3333	3333	16667	0	0	0	0	0	0	0	0	0
F	-3667	3667	-7333	-3667	11000	45000	0	0	0	0	0	0	0
D	-666	666	-1334	-666	2000	0	-1455	-545	46000	0	545	0	0
I	3667	-3667	7333	3667	-11000	0	4000	7000	0	0	4000	16000	0
H	333	-333	667	333	-1000	0	728	272	-1000	1000	-272	0	3000

Table 3.4-5: Participations of countries in the use made of lines, resulting from the MP method, when we take node 2 as the slack bus

The results we have obtained seem to corroborate some of the main conclusions already obtained. Specifically:

- Consumers and generators in E and P are using lines in third countries to a very large extent (well above the transfer capacity with these systems assuming cross-border lines are heavily used). These participations do not seem to be compatible with the topology of the grid, since the use that agents in a country make of the grid of others should, in principle, be limited to the transfer capacities between these countries;
- P turns out to be using lines in F, I and D. This also looks surprising, since the latter countries are very distant from P;
- F and I must be compensated for the use they make of some lines in other countries (lines L8-9 and L6-8 respectively). In the context of the IEM, several

countries have expressed their opposition to compensating other countries for the use the latter make of the transmission grid of the former. ETSO has indicated in recent documents that counterflows that are due to transits do not reduce the need for new investments in the network. Therefore, this result may be unacceptable from a technical and political point of view;

- F is using 8000 MW of the interconnection capacity with D (more than twice the actual capacity of the inter-connector, provided if it is close to be congested). Again, this appears to be incompatible with the constraints imposed by the transmission system;
- I is using 7000 MW of the interconnection capacity with D (again, well above the capacity of the inter-connector provided it is close to be congested). The same comment.
- Results critically depend on the choice of the slack bus. One may argue that changing the slack node results in a different split of transmission charges between generation and load in the region, see (FSR 2005). Thus, we could compute where the slack node must be located so that the generation in the system pays a certain fraction of the grid costs. In other words, once we have decided how to split grid charges between G and L, the location of the slack node would be given. However, this decision will affect the transmission charge each agent will have to pay, or the net amount of money each country must contribute. Therefore, given that it is an arbitrary decision, any option will probably face opposition from some agent (in the regional context) or group of agents (in a single, perfectly integrated system).

	<i>P</i>	<i>E</i>	<i>F</i>	<i>D</i>	<i>I</i>	<i>H</i>
<i>P</i>	1458	208	0	0	0	0
<i>E</i>	833	5833	0	0	0	0
<i>F</i>	-1833	917	25250	0	0	0
<i>D</i>	-334	167	91	11363	0	0
<i>I</i>	1833	-917	-250	1750	10750	0
<i>H</i>	167	-83	-46	69	0	1000
	667	292	-205	1819	0	0
	208	833	-917	-76	2417	107
	458	-542	712	1895	-2417	-107

Table 3.4-6: Compensations among countries produced by the MP method corresponding to the usage that external agents make of the grid in each country. We take node 2 as the slack bus

Finally, we have represented the path followed by the flow injected by the generator in node 2 when the slack node is located in node 8. This is shown in Figure 3.4-5. As we can see, the flow injected in node 2 crosses the grid of most countries in the system

regardless of their proximity to this node and the capacity of the inter-connectors between the different areas.

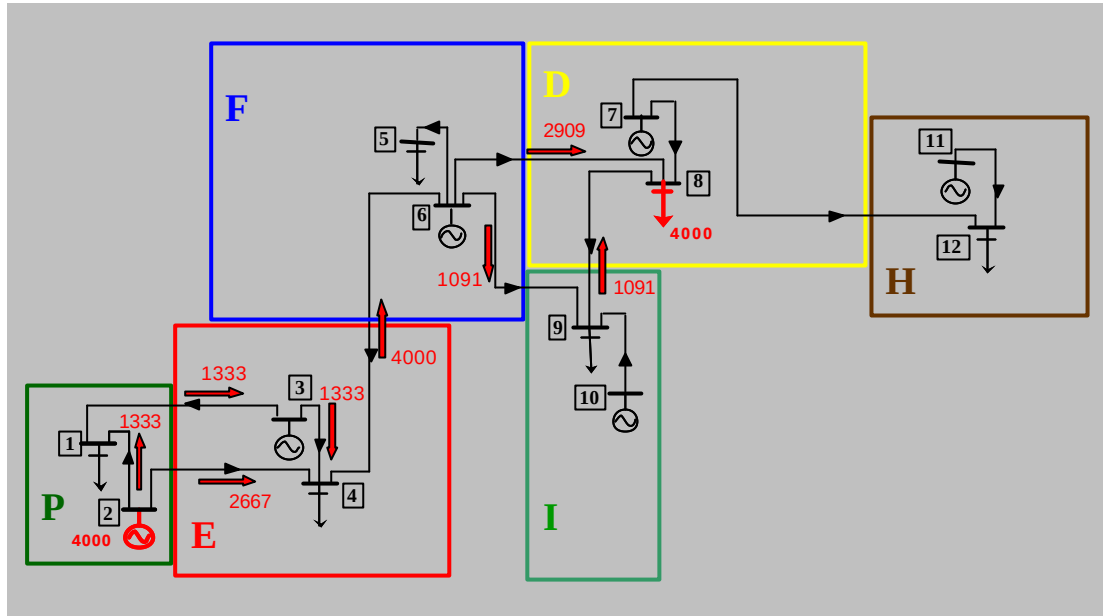


Figure 3.4-5: Path followed by the flow injected by the generator in node 2 when the slack node is in node 8.

Results for the AP method

Table 3.4-7, Table 3.4-8 and Table 3.4-9 show the results obtained when applying the AP method to the system in Figure 3.4-4. Table 3.4-7 provide the participations of agents in the usage made of transmission lines. Table 3.4-8 provides the participation of all the agents within each country (generation and load considered together) in the usage made of the same lines. Numbers of this two tables are expressed in MW.

	L1-2	L1-3	L2-4	L3-4	L4-6	L5-6	L6-8	L6-9	L7-8	L7-12	L8-9	L9-10	L11-12
N1	250	2250	0	0	0	0	0	0	0	0	0	0	0
N2	250	0	1750	0	0	0	0	0	0	0	0	0	0
N3	0	2250	0	7750	0	0	0	0	0	0	0	0	0
N4	0	0	1750	7750	500	0	0	0	0	0	0	0	0
N5	0	0	0	0	0	22500	0	0	0	0	0	0	0
N6	0	0	0	0	500	22500	1637	3364	0	0	145	0	0
N7	0	0	0	0	0	0	0	0	22500	500	1992	0	0
N8	0	0	0	0	0	0	1492	0	20508	0	0	0	0
N9	0	0	0	0	0	0	145	3364	1992	0	2137	8000	0
N10	0	0	0	0	0	0	0	0	0	0	0	8000	0
N11	0	0	0	0	0	0	0	0	0	0	0	0	1500
N12	0	0	0	0	0	0	0	0	0	500	0	0	1500

Table 3.4-7: Participations by agents in the usage made of electrical lines resulting from the AP method

Finally, Table 3.4-9 provides the compensations among countries for the electrical use that external agents make of the grid of each country. The meaning of the numbers in

the last table is completely analogous to that of number in Table 3.4-6. Numbers in this table are expressed in '000 MWkm.

	<i>L1-2</i>	<i>L1-3</i>	<i>L2-4</i>	<i>L3-4</i>	<i>L4-6</i>	<i>L5-6</i>	<i>L6-8</i>	<i>L6-9</i>	<i>L7-8</i>	<i>L7-12</i>	<i>L8-9</i>	<i>L9-10</i>	<i>L11-12</i>
<i>P</i>	500	2250	1750	0	0	0	0	0	0	0	0	0	0
<i>E</i>	0	2250	1750	15500	500	0	0	0	0	0	0	0	0
<i>F</i>	0	0	0	0	500	45000	1637	3364	0	0	145	0	0
<i>D</i>	0	0	0	0	0	0	1492	0	43008	500	1992	0	0
<i>I</i>	0	0	0	0	0	0	145	3364	1992	0	2137	16000	0
<i>H</i>	0	0	0	0	0	0	0	0	0	500	0	0	3000

Table 3.4-8: Participations of countries in the use made of lines resulting from the AP method

	<i>P</i>	<i>E</i>	<i>F</i>	<i>D</i>	<i>I</i>	<i>H</i>
<i>P</i>	1125	1000	0	0	0	0
<i>E</i>	1000	5000	125	0	0	0
<i>F</i>	0	125	23773	343	877	0
<i>D</i>	0	0	280	11655	498	125
<i>I</i>	0	0	868	1059	9375	0
<i>H</i>	0	0	0	125	0	875
<i>Use by others</i>	1000	1125	1273	1527	1375	125
<i>Use of others</i>	1000	1125	1345	903	1927	125
<i>Net use</i>	0	0	-72	625	-552	0

Table 3.4-9: Compensations among countries produced by the AP method corresponding to the usage that external agents make of the grid in each country

From the results shown in Table 3.4-7, Table 3.4-8 and Table 3.4-9, we can draw some conclusions which are in line with the assessment of the AP method previously made:

- Numbers seem to be in line with basic intuition: countries use the grid of neighboring countries only. Similarly, agents make use of neighboring lines only;
- participations of countries in the grid of others are consistent with the capacity of the inter-connectors. The same is true for the participations of agents, since all these participations are positive;
- as a consequence of what we have previously said, individual participations by agents make sense as well.

Analogously to the case of the MP method, Figure 3.4-6 represents the path followed by the flow injected by the generator in node 2. In other words, it shows the use that the agent in node 2 makes of the grid. Contrary to what happens for the MP method, participations of this agent in the use of the grid are restricted to the neighboring lines, which is quite reasonable taking into account the topology of the grid and the pattern of flows in the system.

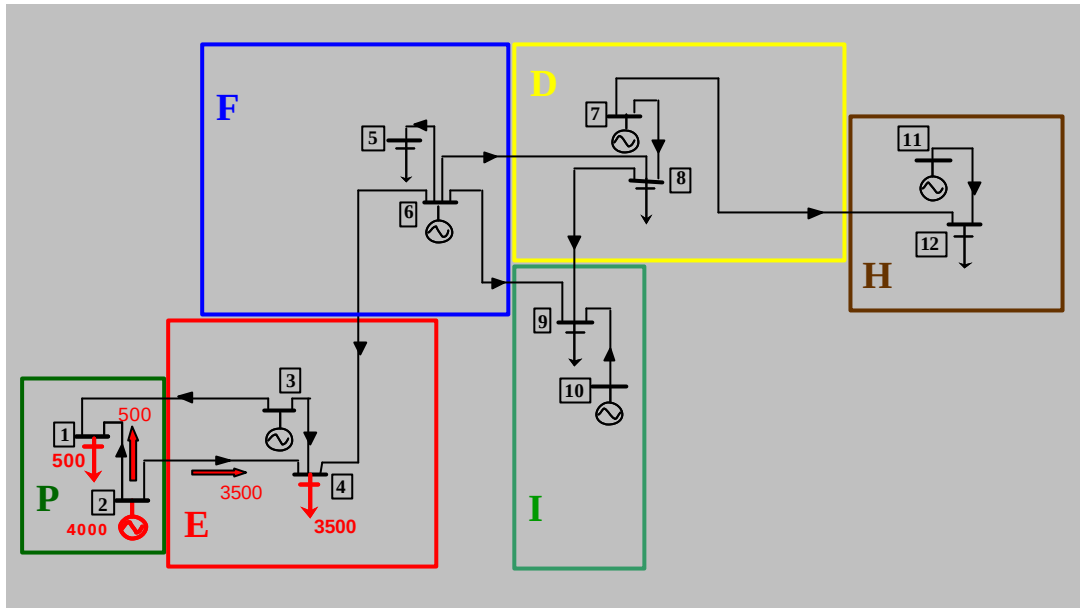


Figure 3.4-6: Path followed by the flow injected by the generator in node 2.

3.4.2.8 Conclusion

Based on the qualitative and quantitative analysis carried out, one may conclude that the Average Participations method is clearly preferable to the Marginal Participations method. The main reason for choosing the AP method is the fact that the underlying assumption behind the AP method makes much more sense than that behind MP. In AP, the flow injected / retrieved by each agent is traced until it reaches the nodes where it is consumed / produced. The tracing algorithm applied takes into account the power flows that actually exist. On the other hand, MP assumes that the same node responds to the flow injected (retrieved) by every agent in the system. This results in the following advantages of AP over MP:

- Participations of agents (countries) in the use of the grid are compatible with the restrictions imposed by the grid, unlike those obtained with MP.
- Participations of an agent (country), according to AP, are restricted to the neighboring lines (areas), unless dominant flows cross the system. On the other hand, results produced by MP show most market agents being held responsible for a non-negligible part of the use of a significant number of very distant lines.
- Results produced by MP critically depend on the choice of the responding (slack) node. On the other hand, application of the AP method is free from any arbitrary decisions (the method automatically decides how much each agent is using each line).

Having said that, we must acknowledge that the results obtained with MP for a single country (area), when the slack node is chosen to be located in the same node as the marginal generator in the dispatch and no significant congestion exist within this

system, may be reasonable. Applying the MP method at regional level seems to be much more problematic.

3.4.3 Foundation of the method of Average Participations

Acknowledging the fact that the origin of the AP method is simply a heuristic rule, this section aims to provide a theoretical foundation of the AP method. Section 3.4.3.1 discusses the rationale of the method from an economic point view. Section 3.4.3.2 provides a justification to the method based on the physics of the problem.

3.4.3.1 Rationale of the method from the point of view of economics theory

This section discusses the justification of the method based in economics theory that was originally published by Bialek et al. see (Bialek and Kattuman 2004). Authors in this article prove that transmission charges resulting from the application of AP coincide with the Shapley value corresponding to each agent in a certain game. We make explicit the critical assumption underlying the cooperative game designed by Bialek et al. The mathematical proof presented is reproduced here using an alternative formulation of the same game.

Statement of the problem

Our purpose is to allocate the cost of a jointly used facility, each line in the network, among its users. Cooperative game theory provides solutions to cost allocation problems that behave well in terms of fairness, efficiency and stability, see (Zolezzi and Rudnick 2002). These qualities are essential for an allocation of the cost of the transmission grid to be considered satisfactory. Game theory comprises both cooperative and non cooperative games. Players in a game may cooperate in order to get higher pay-offs (smaller transmission charges). Cooperation among agents and the formation of coalitions is the essence of cooperative games. Agents form coalitions in order to reduce (or augment) their share of the cost of the common resource they have to pay. All the agents within a coalition cooperate so that the aggregate cost they share is as small as possible. Cooperative games are regarded as a fair scheme for the allocation of joint costs or benefits among a group of agents.

Many applications of cooperative game theory to the allocation of the sunk costs of the grid, or the planning of the grid expansion, exist, see (Contreras and Wu 1999; Contreras and Wu 2000; Evans, Zolezzi et al. 2003; Stamtsis and Erlich 2004) among others. For example, (Stamtsis and Erlich 2004) compares two different methods of this type: the Shapley Value method and the Nucleolus method. Both methods are applied at system level and separately for every individual line. Additionally, two different organizational forms are considered for the market: a pool and a market based on bilateral arrangements. This paper concludes that the Shapley Value is superior to the

Nucleolus method as long as the solution obtained when computing the Shapley value belongs to the core of the game.

The game to be explained here is aimed at justifying the proportional sharing principle, which is the basis of the AP method. Once the tracing rule is accepted, determining the path followed by the power that each generator produces or the power that each load consumes directly follows, thus obtaining how much each agent uses each line. Then, provided electrical usage is deemed a reasonable proxy to economic profit, we can allocate the cost of each line in direct proportion to the use that each agent makes of it.

Our game is based on the one proposed by Bialek and Kattuman in (Bialek and Kattuman 2004), though the formulation of the game differs in some aspects from the one published by Bialek. The proportional sharing principle of the AP method states that flows entering a node are distributed among outflows according to the size of the latter ones. Here we allocate the megawatts of power flow entering a node among the ones leaving it according to the average value that each megawatt leaving the node contributes to a coalition. This average value is computed over all the possible coalitions the agent may take part in. This concept is called Shapley Value. We demonstrate that the resulting allocation coincides with that provided by Average Participations.

We must bear in mind the fact that each megawatt entering a node 'A' may have followed a different path before reaching node 'A' and analogously for those megawatts leaving node 'A'. Therefore each one of these megawatts may have used each line in the system to a different extent. When we allocate each megawatt entering node 'A' to a megawatt leaving this node we are really assigning the cost of the grid used by the former megawatt to the later megawatt or vice-versa. Therefore the problem of distributing megawatts flowing into a node among the megawatts flowing out of it is tantamount to assigning the corresponding use of the system charges. The following paragraphs describe the game.

The players in the game are the megawatts of power that flow outwards from each node. They form coalitions with the remaining megawatts leaving the node in order to establish the correspondence between the megawatts that enter the node and the ones that leave it. Each megawatt entering a node through one line (or produced by a generator located at this node) must leave it through another node (or supply a load also located at the node). A one to one correspondence between megawatts of one type and the other follows from the first Kirchoff's law. Each megawatt leaving a node is assigned a label that corresponds to one megawatt entering the node. If required, megawatts can be considered to be divided into more elementary units in order to preserve the validity of the method.

Coalitions are associations among players that cooperate to obtain some benefit. Not every coalition needs to be possible. In the game we define that coalitions are arrangements or agreements among megawatts. Exiting megawatts may agree on a

certain distribution of the labels, thus establishing through which line each one of them has reached the node. Since every megawatt reaching a node must leave it, one agreement or arrangement is not valid unless it assigns a label to every megawatt that flows through the node. Therefore, coalitions are one-to-one correspondences between exiting megawatts and labels (and implicitly between megawatts reaching a node and those leaving the node). If we consider labels from '1' to 'R', one for each entering megawatt, where 'R' is the number of megawatts passing through the node, and we assume that these labels remain always in the same order, then each possible correspondence between exiting megawatts and labels is actually a permutation of these megawatts.

Mathematical formulation of the problem of computing the Shapley value for this game

Now we compute how the flow reaching node 'A' through a particular line is distributed among the outflows. Repeating the process for each inflow to the node results in a complete allocation of inflows among outflows. Figure 3.4-7 shows a node connected to several lines. The flow over the lines to the left of the node is towards it, while the flow over the lines to the right of the node goes outwards from it. We aim to allocate the inflow f among the outflows $f_1, f_2, \dots, f_n$, i.e. we aim to compute how much power flow f contributes to flows $f_1, f_2, \dots, f_n$.

Consider megawatt ' i ' leaving the node. Now, we are going to compute how much of flow f must be allocated to megawatt ' i '. In order to do so, we have to compute the average contribution of megawatt ' i ' to flow f over all possible coalitions. As explained before, every coalition is a permutation of the outflow megawatts that results in a specific correspondence between outflow megawatts and labels, i.e. inflow megawatts. Megawatt ' i ' contributes 1 to flow f when it is assigned one of the labels corresponding to flow f and it contributes 0 when it is not assigned one of these labels.

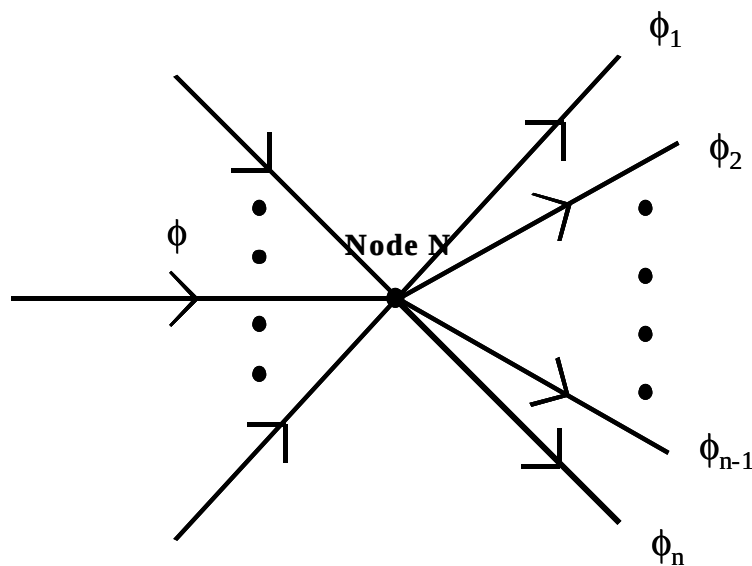


Figure 3.4-7: allocation of an inflow f to a node among the outflows

Therefore, the average contribution of megawatt ‘i’ is:

$$AC_{i,f} = \frac{\left(C_{f_{total}-1}^{|f|-1} P_{|f|} P_{f_{total}-|f|}\right) \cdot 1 + \left(C_{f_{total}-1}^{f_{total}-|f|-1} P_{f_{total}-|f|} P_{|f|}\right) \cdot 0}{P_{f_{total}}} \quad (3.1)$$

where $AC_{i,f}$ is the average contribution of megawatt ‘i’ to inflow ‘f’; f_{total} is the total amount of flow that goes through the node, i.e. $f_{total} = |f_1| + |f_2| + \dots + |f_{n-1}| + |f_n|$; $C_{f_{total}-1}^{|f|-1}$ is the number of combinations of the $f_{total}-1$ remaining outflow megawatts taken $|f|-1$ at a time (in groups of $|f|-1$). This represents, for those permutations where ‘i’ is assigned a label that corresponds to outflow f , the possible ways of choosing the remaining $|f|-1$ megawatts that are assigned a label corresponding to outflow f ; $P_{|f|}$ is the number of permutations of $|f|$ elements and represents, for each combination of megawatts that are assigned a label corresponding to flow f , the different orderings of these megawatts; and $P_{f_{total}-|f|}$ represents the different orderings of the remaining $f_{total}-f$ megawatts. In other words, $C_{f_{total}-1}^{|f|-1} P_{|f|} P_{f_{total}-|f|}$ represents the number of permutations or coalitions for which megawatt ‘i’ contributes 1 to inflow f . Analogously, $C_{f_{total}-1}^{f_{total}-|f|-1} P_{f_{total}-|f|} P_{|f|}$ represents the number of permutations for which it contributes 0. $P_{f_{total}}$ is the total number of permutations or coalitions. If we operate the expression in equation (3.1) we obtain:

$$AC_{i,f} = \frac{\frac{(f_{total}-1)!}{(f-1)!(f_{total}-f)!} f!(f_{total}-f)!}{f_{total}!} = \frac{f}{f_{total}} \quad (3.2)$$

We have already computed the Shapley Value for any outflow megawatt ‘i’ when allocating an inflow f . Any outflow megawatt must have an equal contribution to the inflow f . There is no reason to think otherwise since one cannot make any distinction between two megawatts leaving a node. Therefore, the total contribution of outflow f_0 to inflow f is:

$$C_{f_0,f} = \frac{f_0}{f_{total}} f \quad (3.3)$$

Equation (3.3) is precisely the same mathematical expression as that of the sharing principle applied in the Average Participations method. This result directly follows from the logical assumption that each megawatt that enters a node has the same probability of being any of the megawatts that leave the node.

Properties of the method

Given that this allocation coincides with the Shapley Value of the game just described, it has all the desirable properties in a cost allocation procedure that is based on game theory. The solution found belongs to the core of the game since all the possible coalitions attain the same cost. Therefore it is stable and the allocation is acceptable for all the players. It is fair and symmetric. It agrees with the intuitive idea that, given that all the megawatts are equal in nature, each megawatt has equal probability of being assigned to a particular inflow. Besides, it is monotonic and additive. Monotonicity results in the fact that all the charges are positive. As mentioned before, this is not necessarily a good property. Additivity implies that we shall always obtain the same allocation of inflows to outflows, regardless of the order in which we perform such allocation.

Objections to the method

Implicit in the mathematical proof that has just been presented is the assumption that power over a line only flows in the direction of the net or dominant flow. According to this assumption, counterflows do not exist. Actually, this is what happens. In reality, only the total flows exist as a result of the inputs and withdrawals of power in all the nodes of the system. Thinking of counterflows is a mathematical artifact. Power injections and withdrawals can be grouped into physical transactions where power is injected at some nodes and retrieved at others. When considered separately from the rest of the transactions in the system, some transactions produce a power flow over a certain line in the direction of the prevalent flow while others produce a power flow in the opposite direction. Given that the value of the power flow over each line can be very approximately represented as a linear function of the power injections and withdrawals in the system, the net aggregate contribution of all the transactions add up to the amount of power actually flowing over the line (this is commonly known as the superposition principle). Despite the fact that the physical or commercial transactions that we can define verify the superposition principle, AP implicitly neglects the existence of counter-flows caused by some of these transactions. The method considers the flow over each line as a whole and not as the sum of the separate contributions of the different transactions. As mentioned above, decomposing a flow into the contributions of several transactions is a mathematical artifact).

Analogously, throughout the proof we have assumed that, whenever a megawatt enters a node N, this megawatt leaves the node through the lines for which the net power flow is in this same leaving direction, i.e. those lines for which the net flow leaves node N as well. Thus, we are accepting that the method must be based only on the actual existing flows, and the conceptual scheme of superposition and counterflows is ignored.

The idea that power only flows in one direction at a time is in line with the physical reality of the grid³ as a guiding system for the flow of electromagnetic energy. According to this conception, the power flow over each line is the result of a certain configuration of the grid and a certain pattern of load and generation in the system. Had the pattern of load and generation been different, the flows over the lines of the system would also be different. In other words, according to the AP method, the use that is made of each line depends at the same time on all the power injections and withdrawals. We cannot compute how much each agent uses each line independently from the determination of the usage that other agents make of it. The usage made of the line by the former depends on the usage made of it by the latter ones. This is in accordance with the currently accepted laws of physics.

3.4.3.2 Rationale of the method from the point of view of the physics of power systems

Justification of the Average Participations method

From electromagnetic theory, we know that generators connected among themselves and to the load in a system provide the conditions for fields $\vec{E}$ and $\vec{B}$ to develop. Fields $\vec{E}$ and $\vec{B}$ produce an energy flow $\vec{S} = \vec{E} \times \vec{B}$. Energy flows from generators to consumers guided by wires (lines) though it does not flow within the transmission lines properly speaking but around them. Despite what many people claim when searching a method to trace energy flows, the way electrons move is not related to these flows.

We know where energy is produced and the path it follows until it reaches the load. At some nodes of the grid the power flows split and join others. Determining the path followed by the energy that arrives at a certain node flowing over a line is not possible. It may be guided out of the node by any of the remaining lines or it may be consumed in the node by the local load (provided it exists). Thus, it is reasonable to assume that the energy that flows into a node over a certain line leaves it through the lines that carry energy in the same direction, and that this energy flow splits into parts that are proportional to the size of the outflows. In fact, we do not know how energy actually flows and, probably, we will never know. Applying this simple rule, we can identify the loads that are consuming the energy produced by a certain generator, and vice versa. We can also determine the path followed by energy when flowing from a generator to the loads where it is consumed. Lines are not pipelines or roads but every line sets a limit on the amount of energy that it can guide according to the voltage of the line, the number of circuits it has and other physical and geometrical

³ It agrees with the accepted principles of electromagnetic field theory: the Maxwell's equations.

characteristics. The solution provided by the AP method complies with these constraints.

Once we have identified the set of loads $D_{k1}, D_{k2}, D_{k3}, \dots, D_{kj}$ supplied by a generator G_k , we could compute the line flows that would be produced by this set of agents if they were considered in isolation from the rest of the system. We would only have to compute the solution of a power flow assuming that generator G_k is the only one in the system and loads $D_{k1}, D_{k2}, D_{k3}, \dots, D_{kj}$ are the only ones as well. Flows resulting from the power flow do not coincide with the line flows that the AP method would allocate to these agents. However, this situation is not real. In reality, generator G_k and loads $D_{k1}, D_{k2}, D_{k3}, \dots, D_{kj}$ are not isolated from the rest of the system. They make use of transmission lines together with many other generators and loads. When all generators and loads are jointly taken into account, the line flows are the ones considered by AP to allocate responsibilities in network use, i.e. the actual ones. We are not interested in determining the use that generator G_k and loads $D_{k1}, D_{k2}, D_{k3}, \dots, D_{kj}$ would have made of the transmission grid if they were alone in the system. We only want to compute the use they make of lines in reality, and, in reality, they are operating together with the rest of agents.

Rationale of methods based on the computation of marginal use

On the contrary, the Marginal Participations method determines how much each agent would have used the transmission grid if only this agent (a generator or load), and its counterpart at a reference node (a load or generator of the same size), were operating in the system. The reference node that responds to the power injected/consumed by each agent is the same for all the network users. Obviously, this bears no resemblance to the electricity flows that actually exist.

Instead of assuming that the same node would respond to the injections and withdrawals by all agents, we might consider that the marginal generator in each area responds to the injections and withdrawals by all the network users in this area. Again, power flows resulting from a transaction between an agent and the corresponding marginal generator bear no relation to the flows that actually exist. Besides, results obtained with this method would also have to take into account the limits imposed by the transmission system, plus the fact that, when simulating a transaction of X MW between an agent and the corresponding marginal generator, we may have to consider a different marginal generator depending on the area where the agent is located instead of always the same marginal generator. In any case, implementing this method in the context of a regional market like the IEM, where there is no such a thing as a central economic dispatch, does not make any sense.

3.4.4 Structure of the tariffs

For a start, it is useful to distinguish between access transmission charges and connection charges. As explained before, the latter ones refer to the costs incurred by the system that can be related to the installation of a new agent. These may include both shallow charges (corresponding just to the line or lines that connect the agent to the rest of the system) and deep charges (corresponding to the lines necessary to transport to other parts of the grid the power produced by a generator or bring from other parts the power consumed by a load).

3.4.4.1 Connection charges

The regulator must specify, somewhat arbitrarily, the maximum length of a line to be considered a connection line (e.g. 25 km). It is commonly agreed that only those agents that need these lines to be connected to the grid should cover their cost. We do not need to resort to electrical usage in order to compute connection charges, since we can identify those agents that are responsible for the construction of a connection line. A fixed annual charge may be established to recover the cost of these lines. This connection charge may be modified afterwards, to take into account the fact that other agents end up using the line that was initially meant for the concerned generator or load. In this way, this agent may avoid paying part of the connection charge that he was initially allocated.

3.4.4.2 Access charges

Focusing now on use-of-the-system tariffs, these charges should encourage agents to make the correct investment decisions. One option would involve making each agent pay according to the actual grid usage that this agent has made of the grid. These economic signals will certainly interfere with the efficient short term behavior of the network users. Each agent would be able to modify its transmission charge by changing the amount of power produced or consumed at each time of the year. Then, market agents would consider transmission charges like a variable cost and they will take these charges into account when making an offer on a power exchange or signing a bilateral contract with other party. Therefore, the operation will not be optimal. This problem cannot be totally eliminated, although it can be mitigated.

In order not to interfere with short-term decisions by agents, one may think of computing uniform per-unit transmission charges to be paid by all the agents of the same type located within each of a set of predefined areas. These tariffs should be based on the estimated future usage that agents of the same type that are located within the same area will make of the grid (instead of the electrical usage they have actually made of it). First, each generator and load is assigned one among several possible generation or load profiles. Afterwards, we simulate the operation of the system using the aforementioned profiles in a set of scenarios that are considered representative of what is going to happen during the next year. Transmission tariffs

result from the electrical usage that is attributed in these scenarios to the agents of each profile that are located within each area.

Depending on the adopted classification of agents into different profiles, the decision on the structure of the transmission tariff to be applied may be relevant, since the total transmission charge to be paid by each agent may depend on this decision. Several options are possible:

1. Fixed charge (€/year): it does not affect the operational decisions by agents but it may affect negatively those agents or technologies that use the grid less than others. This effect may be corrected by applying a different charge depending on the type of consumer or technology of the generator.
2. Capacity charge (€/Mw $\times$ year): it encourages agents to reduce the capacity that they contract (consumers) or the nominal capacity (generators) and may discriminate against those agents with less hours of network utilization. Again, this effect may be corrected by applying a different charge depending on the type of consumer or technology of the generator.
3. Energy charge (€/Mwh $\times$ year): it may affect the dispatch if transmission charges are not negligible compared to differences among variable production costs for different technologies.

The characteristics of the system can help to decide which is the best tariff structure to use. Transmission access tariffs must have an structure such that these charges do not interfere with the short term signals meant for the operation of the system. At the same time they must complete the recovery of the grid cost and send efficient locational signals. If transmission charges were small enough they could be structured as energy charges. Employing energy charges is not conceptually correct but implementing them is easy and, provided charges are small enough, it would not distort the functioning of the market in the operation time-frame. Additionally, more experience exists of their application. If transmission charges are comparable in size to the differences among the variable production costs of agents, then transmission tariffs should be implemented as capacity charges, which are more difficult to compute but conceptually sounder than energy charges.

3.4.5 Numerical results

Section 3.4.2 has concluded that the AP method is the best suited, among the considered ones, to compute transmission tariffs. If sending locational signals is a priority, we may strengthen the relation between responsibility in grid costs and the applied transmission tariffs, see the method proposed in section 2.5.1 of chapter 2 of this thesis. Here we present the results of the application of this method to some load flow scenarios of the Spanish power system.

The following subsections provide numerical results of nodal transmission tariffs computed using the AP method. First, we apply the AP method in its basic form. The resulting nodal tariffs are displayed in various formats and analyzed. Afterwards, application of a smoothing function helps in correcting some problems that have been detected in the AP results. Once corrected, the AP results are pictured again.

3.4.5.1 Results for the AP method

Description of the data

This section presents the results of the application of the AP method to a system comprising 3965 nodes of 13 different countries that belong to the Internal Electricity Market of the EU. The 24 scenarios considered here correspond to the actual operation of this system in the year 2002. Data have been provided by the association of European Transmission System Operators (ETSO).

Computation of nodal transmission tariffs

Application of AP yields transmission charges or participations in the electrical usage of the grid for each agent and each one of the 24 scenarios. However, market agents will pay an annual charge, either as a single payment or divided into several identical payments distributed along the year. Therefore, computing annual tariffs is necessary. Annual tariffs result from averaging the charges corresponding to the different scenarios. The charge for an agent in each scenario is weighted by the amount of power produced or consumed by the agent in this scenario. By default, the AP method equally allocates the total flow over each line to the aggregate of generators and to the aggregate of loads in the system. Therefore 50% of the use of each line (and consequently of its cost) is allocated to generators and 50% to loads.

Section 3.5.1 later presents the results corresponding to the entire IEM system. Here, we focus on those corresponding to Spain so as to be able to analyze in more detail the results of the application of the method. Figure 3.4-8 and Figure 3.4-9 show the detail of the original G and L annual tariffs produced by AP for a selection of the nodes in the Spanish transmission system. In these figures, nodes are grouped by province. This is represented with a thin pink line. Figure 3.4-10 and Figure 3.4-11 represent the same results as Figure 3.4-8 and Figure 3.4-9 but now showing the geographical distribution of the Spanish L and G transmission tariffs over the Iberian peninsula.

Annual transmission tariffs have been represented using dots, whose color varies according to the numerical value of the charge. The color code that has been used goes from dark blue for very low tariffs to dark brown for very high ones. We have not taken into account outliers when defining the color code in order not to blur the representation of the tariffs for the majority of nodes, which are more or less close to the average. Thus, we have used different symbols to represent extremely low or high

transmission tariffs. Outliers above the average are always designated using a square. Outliers below the average are designated with a dark blue cross. The size of the symbols used to represent tariffs varies according to the amount of energy traded by the corresponding agent. The larger the amount of energy the bigger the symbol is. All the numbers are expressed in €/MWh.

Figure 3.4-8, Figure 3.4-9, Figure 3.4-10 and Figure 3.4-11 indicate that the distribution of average transmission tariffs within each province is not too disperse. Significant differences among transmission charges in different system areas are mainly due to the fact that different patterns of flow exist in different parts of the grid. Some general patterns in the spatial distribution of transmission tariffs can be identified.

Regarding L tariffs, charges in the north western part of the country are low, they are medium sized along the eastern coast and moderately high in the north central part of the peninsula and Madrid. The highest charges can be found in the south. This can be easily explained since the north western part of the country has an excess of generation because of the great amount of generation facilities that exist there. Consequently, it is an exporting area. On the other hand, in the south of Spain a deficit of power exists. In fact, this area has to import power from very distant places. Madrid and the Basque region are major load centers and therefore they also have to import power. The east coast has both major load centers such as Barcelona or Valencia and large generation facilities in Tarragona, Valencia and Castellón.

Regarding generation transmission tariffs, the geographical distribution of tariffs is somewhat the opposite, though fewer differences exist among tariffs in the different areas. Transmission charges are highest in the north west, specifically in the south of Galicia, followed by those in Asturias, Cantabria and León (which are located next to Galicia in the north). Tariffs in the major load centers (Madrid, Barcelona, the south west, etc.) are among the lowest. This agrees with the intuitive idea that, while L charges in exporting regions tend to be low because a demand located in these areas will probably consume the power produced locally, G charges may need to be high if the power production needs to be transported over a long distance before reaching the loads where it is consumed. The opposite applies to importing areas.

One major concern is the existence of some outliers, i.e. extremely low or high transmission tariffs at certain nodes that bear no relation to those for the neighboring nodes. In other words, the tariffs directly provided by the model make sense in general terms but transmission charges at some nodes differ quite much from those in the same area. Thus, whereas G transmission tariffs in the central part of the east coast tend to be reasonably low, there is one node where the tariff is very high, see Figure 3.4-11 (it is represented as an outlier with a square). This issue requires an explanation.

Analysis and explanation of the large dispersion or volatility of transmission tariffs computed with AP

In some cases, tariffs resulting from AP seem to be too volatile. Thus, for example, nodal G tariffs within the Spanish province of Orense range between less than 2 and more than 8 €/MWh. Some agents have to pay very high transmission charges while they are located at nodes within an area where charges are low in general. Analogously, some agents located in areas where charges are high would have to pay very small transmission tariffs.

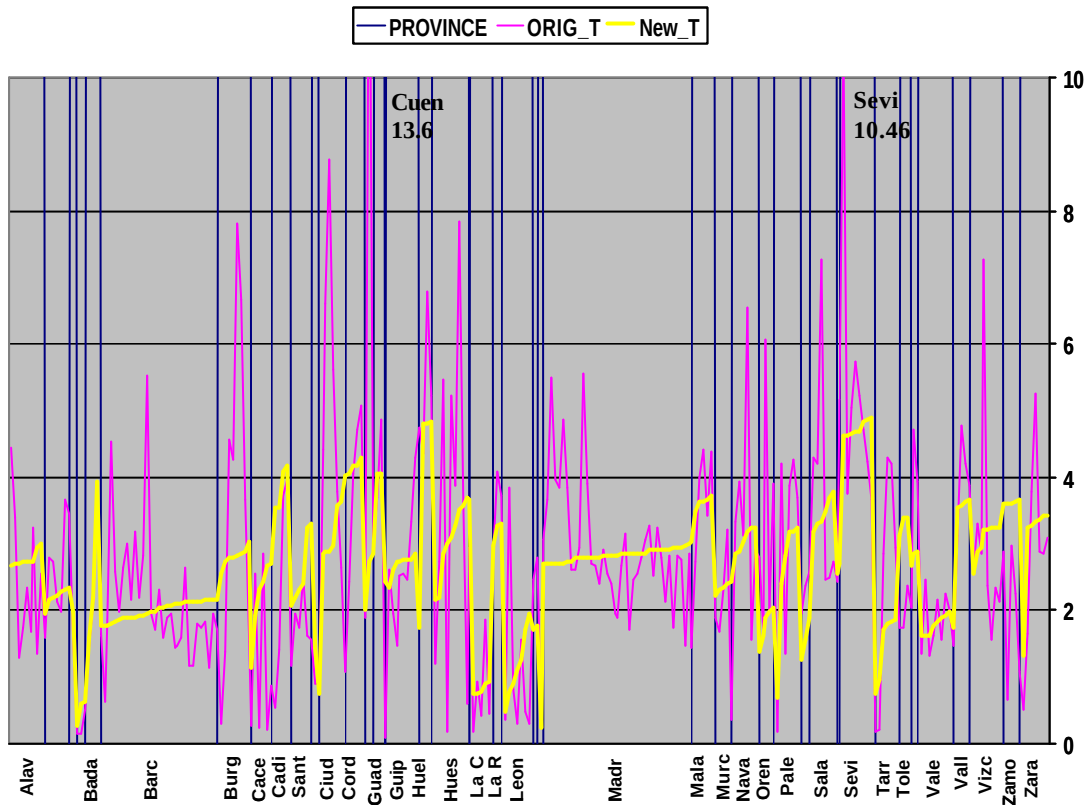


Figure 3.4-8: Original and new L tariffs within Spain for the IEM-13 system

This is related to the fact that, according to AP, the same pattern of flows might result in some special circumstances in very different transmission charges for two agents located not very far one from another. Due to the non-linear nature of the method, the unit participation of an agent in its use of the grid may critically depend on the amount of generation and load in the node where the agent is located as well as in the surrounding ones.

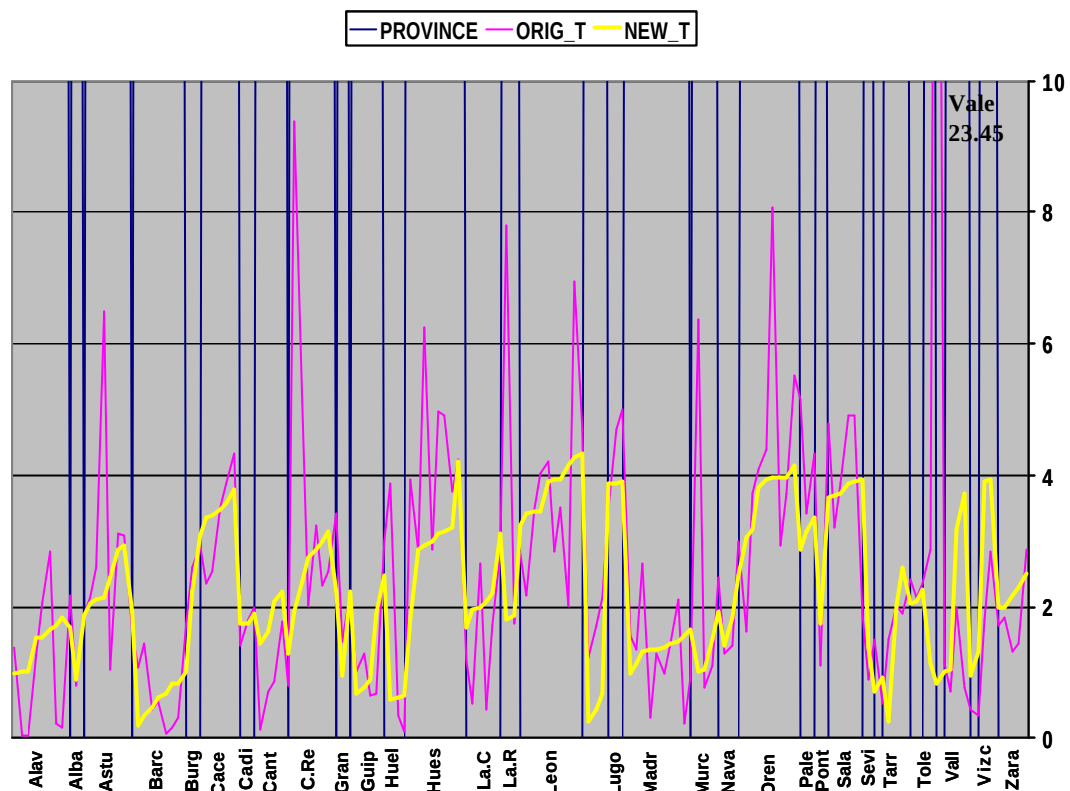


Figure 3.4-9: Original and new G tariffs within Spain for the IEM-13 system

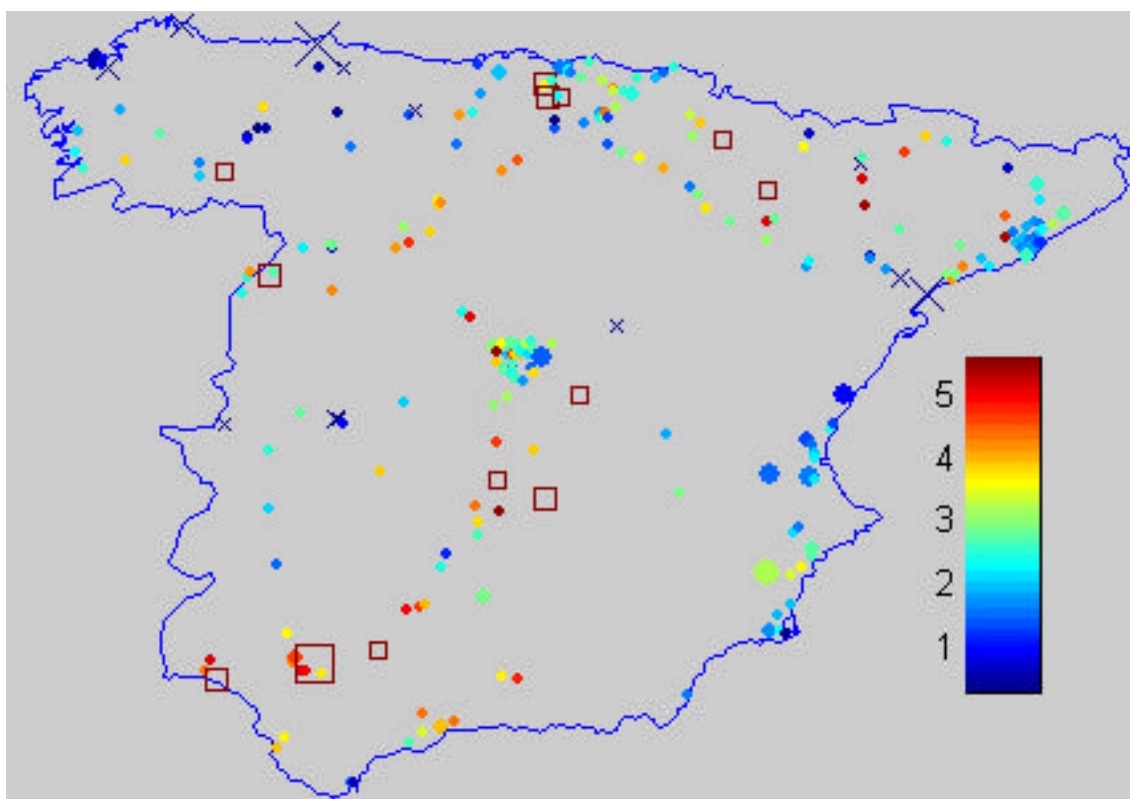


Figure 3.4-10: Geographical distribution over the Spanish system of the nodal L transmission tariffs computed using AP for the IEM-13 system

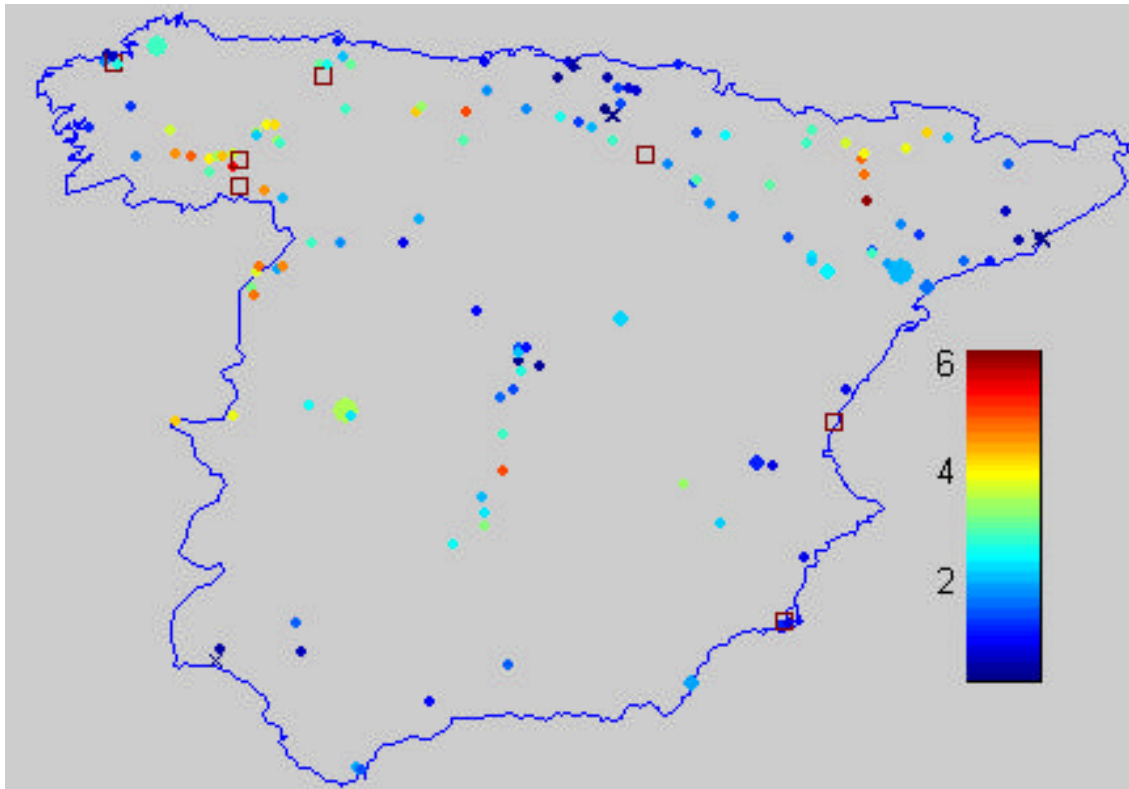


Figure 3.4-11: Geographical distribution over the Spanish system of the nodal G transmission tariffs computed using AP for the IEM-13 system

There may be different reasons for this to happen. A small generator may be located close to a large load within an exporting area. If the load is large enough, the flow over most of the lines connected to the node where the generator is located may always be towards this node. Then, according to the AP method, this generator would face a very low transmission charge. Given that the generator is located within an exporting area most of the remaining generators in this area will probably have to pay high transmission charges. In the same way, loads located close to a large generator in a mainly importing area may have to pay low transmission charges while the surrounding loads have to pay high tariffs.

Extremely high values of transmission charges may also occur when small loads or generators make use of long and therefore expensive lines whose flow to capacity ratio is very low. This is a consequence of the fact that AP, in its basic form, allocates the whole cost of each line to its users no matter how much of the total capacity of the line is being used. This effect could be corrected by employing the electrical usage just to allocate the cost of the used fraction of each line. The cost of the remaining part could be socialized (preferably to demand). Application of the AP method to allocate just the used fraction of each line is discussed in the next subsection.

Additionally, the power produced (respectively consumed) by generators (respectively loads) in areas where there is an excess (respectively deficit) of generation may have to flow through many and / or long lines, thus resulting in very high tariffs as well.

All in all, the AP method happens to provide very reasonable results at macro level (for example, inter-TSO payments among countries obtained with AP seem to be quite reasonable). However, nodal transmission tariffs sometimes are quite volatile as it has been just indicated, thus requiring some interpretation and further treatment.

How to deal with outliers resulting from the application of AP

Different options exist in order to deal with abnormally low or high values of transmission tariffs. Some of them have already been presented. Others are mentioned here for the first time. Avoiding the excessive dispersion of tariffs may be achieved by computing and applying zonal tariffs. Zonal tariffs can be obtained as the weighed average over predefined geographical regions of the nodal tariffs resulting from AP or any other method. Each nodal tariff in a given area would be weighted with the amount of power either produced or consumed by the corresponding agent. Extreme values would thus have a small influence in the resulting tariff to be applied to every generator or load in the area.

But other methods exist that help to deal with outliers. Smoothing out individual nodal tariffs by taking into account the values of nearby nodes is another possibility. The next paragraphs explain how to get rid of outliers and make the geographic distribution of transmission tariffs smoother.

Our aim is to produce smoother transmission tariffs for each market agent by combining the original tariff of the agent and those corresponding to market agents of the same type (either generation or load) that are located close to this agent. Therefore, only tariffs of generators influence the computation of the new tariff of a generator and the same happens with loads. If important imbalances exist between generation and load in the different regions, results of the allocation of network costs produced by any sound method differ significantly from one region to another. G tariffs in a mainly importing area should be smaller than those in a exporting area and vice versa for load, thus signaling the advisability of installing the new generation in the importing areas and the new load in the exporting ones. Our objective is not to distort these signals by mixing G and L tariffs.

The new tariff for a certain market agent A results from averaging the original tariffs of all the nodes of the same type as A. In order to compute this average, tariffs are weighed using a function that decreases with the distance to the node where A is located. The weighing factor used for the original tariff of agent A is 1. Weighing factors are smaller the further away we move from this node. Therefore, nodes that are closer to agent A will have a stronger influence in the new tariff of this agent than those located farther.

Figure 3.4-8 and Figure 3.4-9 show the new tariffs within the different provinces in Spain once the smoothing process is applied. New tariffs are represented by a thick yellow line. Figure 3.4-12 and Figure 3.4-13 show the geographical distribution over the

Iberian peninsula of these same transmission tariffs. All these figures show how the evolution of the new tariffs is much smoother. New tariffs present less dispersion and they change gradually with the geographical location of the corresponding agent.

Finally, Figure 3.4-14 and Figure 3.4-15 show the geographical distribution of transmission tariffs obtained when we employ electrical usage just to allocate the cost of the used fraction of each line while the cost of the unused fraction is socialized to demand. Due to the fact that, on average, only 20% of the capacity of lines is used, the resulting tariffs for generators are significantly lower than those corresponding to loads. Thus, the average L transmission tariff is 4.64 €/MWh whereas the average G transmission tariff is 0.43 €/MWh. The geographical dispersion of tariffs is now far less important since only a small part of the cost of the grid is allocated using geographically differentiated transmission tariffs (around 20% of the total cost of the grid corresponding to the aggregated cost of the used fraction of the lines in the system). The remaining part is recovered by means of a uniform charge to be levied on all the load.

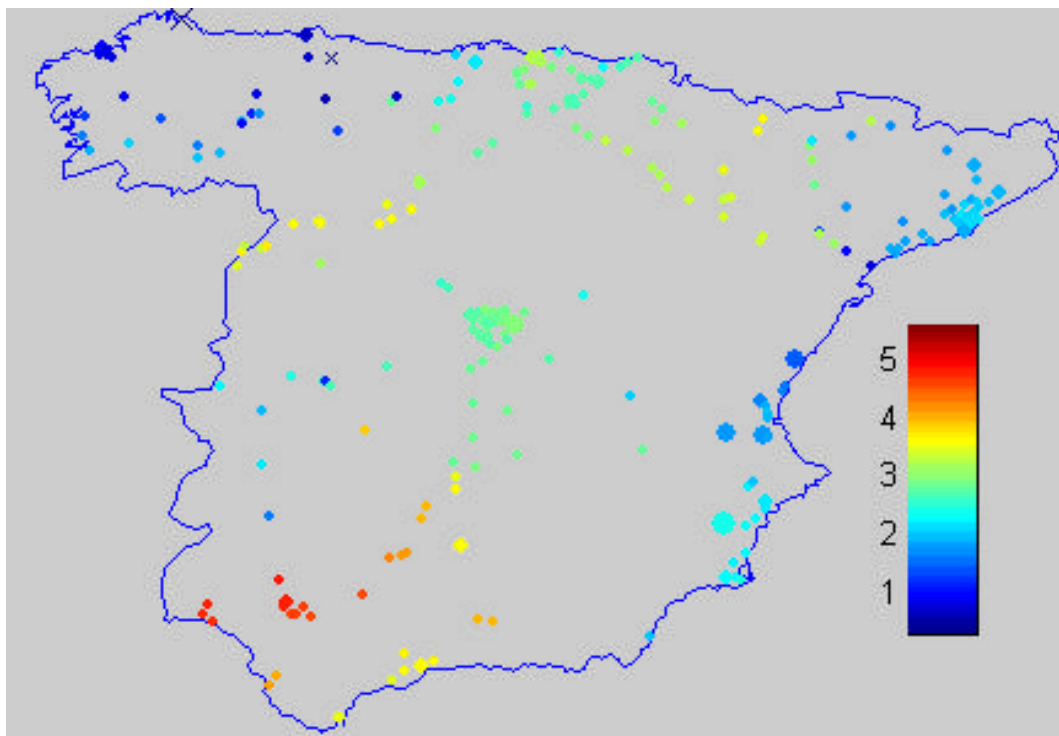


Figure 3.4-12: Geographical distribution over the Spanish system of the nodal L transmission tariffs computed for the IEM-13 system using AP and applying the smoothing process afterwards

Here we can distinguish between locational signals, whose secondary purpose (besides recovering the total transmission cost) is to promote the installation of market agents in those places that are most beneficial from the point of view of the system grid, and transmission charges just aimed at completing the recovery of the whole cost of the

grid. Meaningful efficient locational signals are compatible with any desired average level of transmission charges to be levied on generators or loads. For instance, if we want generators to pay X% of the cost of the grid we may compute optimum locational signals for them and then add a constant amount to these locational signals so as to modify the fraction of the system transmission costs recovered from generators. Note that by doing this we are not modifying the initial locational content of the resulting transmission charges. However, modifying the average transmission charge they have to pay may have an impact on their investment decisions. This has to do with the fact that the amount of expenses that generators (and demands) will face when they decide to invest will vary according to the decision by authorities on how to split transmission costs between generation and load.

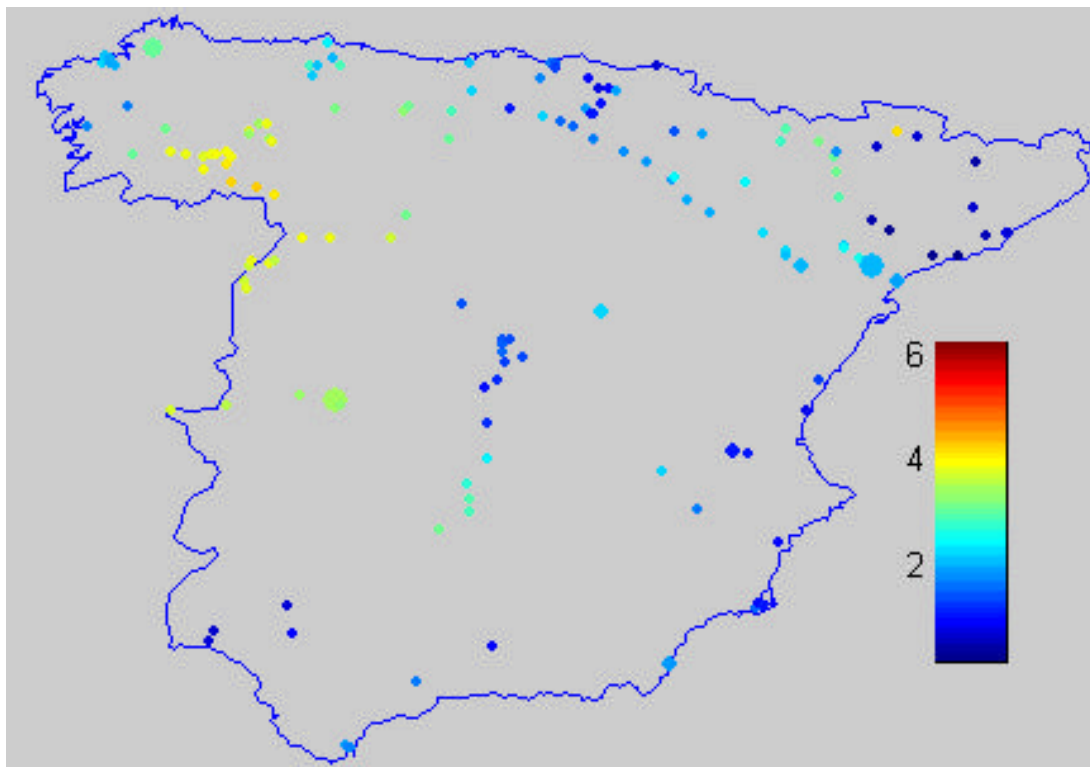


Figure 3.4-13: Geographical distribution over the Spanish system of the nodal G transmission tariffs computed for the IEM-13 system using AP and applying the smoothing process afterwards

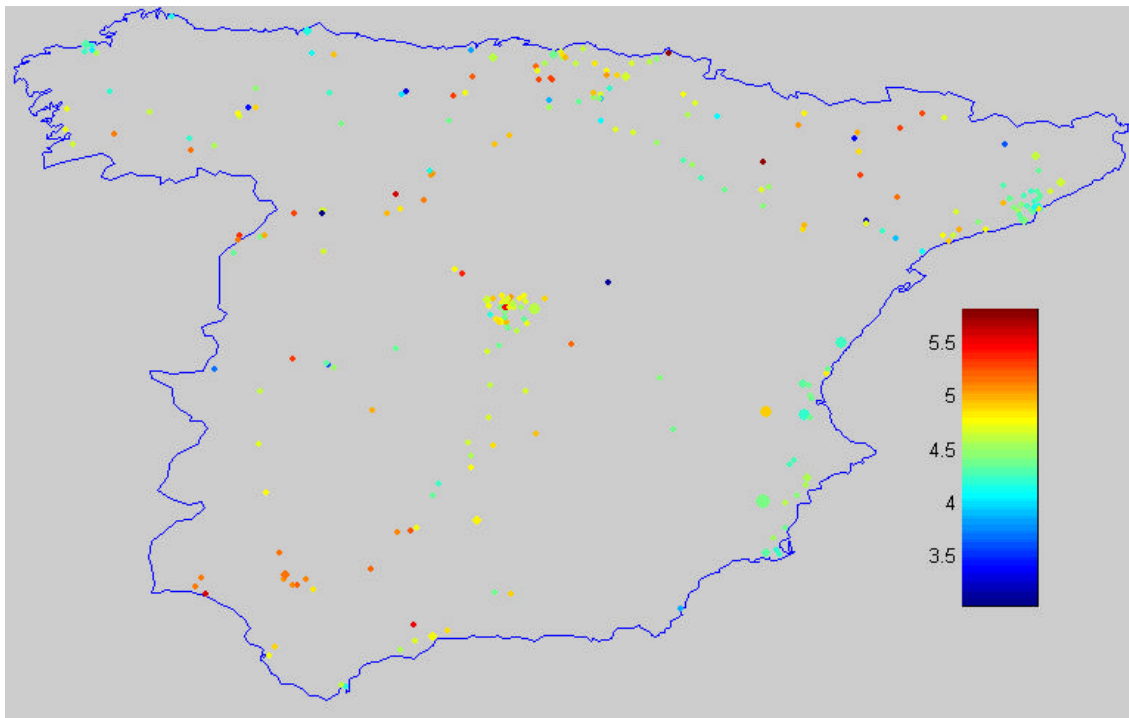


Figure 3.4-14: Geographical distribution of nodal L tariffs obtained using AP to allocate the cost of the used fraction of each line. The remaining part is socialized.

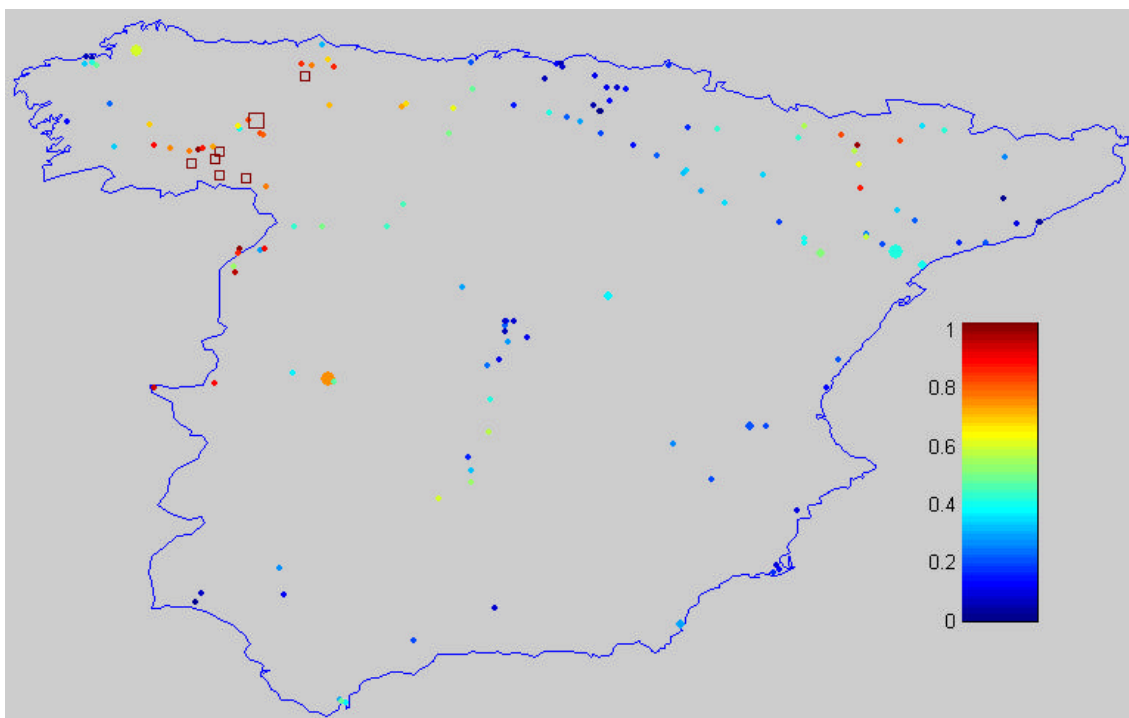


Figure 3.4-15: Geographical distribution of nodal G tariffs obtained using AP to allocate the cost of the used fraction of each line. The remaining part is socialized.

3.5 Extension of the computation of transmission tariffs to the case of regional systems

Circumstances existing in the different regions where a market is developing condition the practical implementation of the methodology proposed in the first part of this chapter for the reference case of a perfectly integrated market. When designing a regional market, circumstances of various types must be taken into account: political, institutional, level of implementation of national or local markets in the region, even environmental in some cases (in order to regulate the construction of new transmission lines), etc. Given the circumstances that presently exist in the different regions that we have studied, we describe here the best possible implementation of a system of transmission charges in each one of these cases.

3.5.1 The European system

3.5.1.1 The market context

The European electricity industry has undergone major changes since the late 90s. Most of the countries within the European Union, and other European countries as well, have moved from a system comprised of one or several regulated utilities to a situation where several companies compete in the market to provide their services. Furthermore, in many cases the global service of providing electricity to customers has been unbundled into different activities, namely generation, transmission, distribution and retailing.

However, important obstacles lie in the way towards the creation of a competitive market at regional level. The degree of liberalization attained varies widely from one country to another, see (EC 2005). This fact, together with other aspects such as the historical evolution of the industry in each country results in widely different approaches towards the regulation and organization of the electricity sector. Additionally, the local grid within each country is well meshed in general and interconnections between some countries are strong, but the level of interconnection between sub-regions (Iberian peninsula, England and Wales and Ireland, Italy, the NORDEL region and the continental plate) is very low.

All this has led parties involved in the Florence process to opt for a gradual approach to the optimal solution. In terms of cross-border tariffication, all parties have agreed to adopt a system of compensations among countries instead of implementing a system of pan-European nodal tariffs. This compensation system could evolve in the future into an integrated European system of transmission tariffs. Compensations among

countries should take into account the use that agents in one system make of the grid of other systems.

This chapter focuses on the principles affecting cross-border tarification that have been agreed at the Florence Forum. Other chapters focus on the remaining aspects of transmission regulation. Based on the agreements reached at the Forum, a provisional mechanism of cross-border tarification was implemented on March 1st 2002. It finally succeeded in eliminating tariff “pancaking” in the IEM. However, this provisional mechanism contained serious regulatory imperfections. It was therefore agreed at the 8th Florence Forum in February 2002 that the provisional mechanism should not last more than the year 2002 and that it should be replaced by a long-term mechanism where these major imperfections would be eliminated.

ETSO presented a revision of this mechanism that came into effect from January 1st 2003. This revised version is not the long term mechanism yet but it succeeds in reducing export fees and achieving a higher degree of harmonization with respect to both the definition of the Horizontal Network (HN) and the utilization of regulated cost figures.

During the 10th Florence meeting, which took place on the 8 and the 9th of July 2003, participants stressed the need to move forward in the future towards the implementation of a system of locational signals. A new version of the existing ITC mechanism was presented by ETSO which abolished the export fee.

At the 11th Florence meeting, which took place on the 16 and the 17th of September 2004, parties agreed that the concept of Long Run Average Incremental Cost put forward by the European Commission for the computation of ITCs needed further clarification. ETSO, CEER and the EC committed themselves to continuing working on the definition of a long term mechanism. Application of the provisional mechanism was extended for one more year.

Parties have been working to agree on a more permanent ITC method for a number of years now. January 1, 2007 is the new deadline to design a more efficient mechanism and begin to apply it.

Cross-border tarification and the inter-TSO payment mechanism

Throughout the Florence process parties have agreed to apply a system of local or national transmission tariffs which are independent from short term signals. The purpose of these transmission tariffs is recovering the cost of the transmission service and sending locational signals. Local tariffs grant access to the entire EU system grid and are modified according to a system of compensations among countries. Compensations among countries must cover the cost of building, operating and maintaining transmission lines according to the extent to which these lines are used by external agents as well as the cost of transmission losses associated with cross-border

flows. Computation of ITCs requires the use of real flow network models and standardized costs. Any changes to transmission charges derived from the ITC mechanism should be non-transaction based in order not to interfere with energy charges.

Resulting ITCs must provide correct locational signals at TSO or country level. However, they are not meant to send locational signals at individual agent level. Since the ITCs are applied at TSO level, averaging over many nodes, the final signals must be weak. Strong locational signals must be provided through the congestion management mechanisms already in operation as well as those still to be implemented.

All in all, a system of local L and G charges, modified according to the compensation received by the corresponding TSO, seems to be the only sensible approach in the short and medium term. However, moving towards a system of pan-European tariffs that provide locational signals related to losses and investments in transmission lines should be a goal in the long term.

3.5.1.2 Methods considered for the computation of inter-TSO compensations

Both the doctoral student and the thesis director have been involved now for several years in the comparison of alternative ITC schemes and definition of a suitable long-term method for the computation of inter-TSO compensations. The team has carried out several projects for the European Commission on this subject, see (Pérez-Arriaga, Olmos et al. 2002; Pérez-Arriaga and Olmos 2003).

Methods for computation of Inter-TSO payments can be classified into two major categories. All of them are aimed at allocating the cost of the network based on the electrical usage made of it. These two families of methods, and the most representative methods within each family, are:

1. Methods that take into account all cross-border flows among countries. We can distinguish between methods of two types within this category:
 - o Methods that determine the electrical usage that each agent makes of the grid, i.e. those that result in nodal transmission tariffs. Methods of this type have already been described. They include:
 - Average Participations
 - Marginal Participations
 - o Methods that directly compute the external usage made of the grid of each country by the remaining ones. These include:
 - Simplified Average Participations
 - Modified Average Participations

2. Methods that compute the compensation due to each country from the impact of the transit through the country on the electrical usage made of its grid:
 - o With and Without Transits method
 - o Average Participations applied to Transits method
 - o Provisional or ETSO method
 - o New methods recently proposed by ETSO

After a careful assessment by experts in the subject, see (Consentec / Frontier 2005; FSR 2005), three Inter-TSO Compensation methods were seriously considered for replacement of the currently implemented temporary method as the method to compute inter-TSO payments in Europe. These methods are the Average Participations method, the With and Without Transits method and the Marginal Participations method.

Description of the different methods

Now we discuss each one of the methods that have been mentioned above. Together with the description of each method, we provide a critical analysis of its properties and establish its suitability to be actually applied. Methods that have already been described earlier are only briefly discussed here.

Methods that take into account all cross-border flows among countries

Much of the debate on the election of a long term ITC mechanism has been focused on the issue of whether all cross-border flows or only those considered to be part of the transit through each country should be taken into account when computing ITCs. This subsection presents those methods that do not differentiate between cross-border flows of one type and the other.

Methods that provide nodal tariffs

These methods determine how much each agent is using each one of the lines in the system. The cost of each line is allocated among agents in proportion to the electrical usage of the line by the different agents. The transmission tariff for each agent results from the aggregation of the charges corresponding to individual lines.

Average Participations

It has already been explained in full detail, see section 3.4.1.1. Variants of the method were also presented there.

Marginal Participations

The MP method has already been explained as well, see section 3.4.1.2.

Aumann-Shapley method

This is a method based on the cooperative game theory. The method was developed by Shapley originally. Afterwards, Aumann introduced some modifications to it in order to make it not discriminatory. Agents of the same type (generators or loads) choose in sequence the loads/generators that are consuming/producing the power that the former inject/withdraw into/from the grid. Each agent is held responsible for the flow that the corresponding incremental transaction, taking place between the node where the agent is located and the set of loads/generators it has chosen, makes of the transmission grid. The size of this incremental transaction is X MW, being X the amount of power traded by the agent. Obviously, rational agents will choose the sources/sinks of the power they produce/consume in order to minimize the transmission charge they have to pay. All possible orderings of the existing generators (loads) are considered (each generator is given to opportunity to choose first, second,... and the last) and each agent must pay the average of the transmission charge it would have to pay for each of these orderings. Pereira et al. propose in (PSR 2005) an efficient implementation of this mechanism.

This mechanism has nevertheless not been considered in the comparison among ITC methods provided in section 3.5.1.6 because its application to a regional market like the European one has not yet been explored.

Methods that directly compute the external usage of the grid of a country by others

These are the methods that take into account all cross-border flows but introduce some kind of simplification that is incompatible with the computation of nodal tariffs. They only produce results at country or system level.

Simplified Average Participations

The Simplified Average Participations method (SAP) aims to produce the same results as the AP method using far less input data. SAP should only be considered an option if some of the input data required to apply the full fledged version of the algorithm are not available.

A single snapshot is unable to correctly represent the electrical usage made of the network during any given year. Generally, a minimum number of scenarios are necessary. Therefore, applying a method involves gathering the required set of input data for a group of scenarios that is considered to be representative of the system operation during the whole year.

SAP employs a simplified representation of the IEM network. This method may be regarded as a suitable option, as long as the procedure to gather enough data is not

fully operative, or the data have not been collected. From a methodological perspective, the procedure to be followed is exactly the same as in the AP method. In the aforementioned simplified network model of the IEM each country is represented using a single node and lines between nodes represent the existing electrical interconnections between neighboring countries. The flow over each one of these lines corresponds to the aggregated cross-border flow through the border between the corresponding countries.

Provided that only the net amount of generation or demand is considered for each country, the only information needed to apply the method for a given set of operating conditions is the snapshot of the corresponding aggregated import and export flows at each border between neighboring countries or TSOs. If both the generation and demand within each country are considered, information on the total internal demand or generation in each country k becomes necessary. Disregarding losses, it is straightforward to obtain all the information required to apply the algorithm from the import and export flows aggregated at border level and the total internal demand or generation per country.

Once the AP method is applied to this simplified model of the system, the responsibility of the different countries in the flows on each one of the inter-connectors between systems is known. Then, in order to determine how much each country k is using of the grid of another one j , the amount of flow ‘using’ the grid of country j that country k is responsible for must be compared to the total amount of flow through the grid of country j . The amount of flow through the grid of a country j that another country k is responsible for is computed as the maximum between the aggregate value of import flows to country j caused by country k , on the one side, and the aggregate value of exports from country j that have been allocated to country k on the other.

The major drawback of this simplified implementation of the AP method lies in the fact that it fully ignores the behavior of network flows inside each one of the countries or TSOs. A significant loss of accuracy must therefore be expected. Hence, if this approach were adopted transitorily because of the lack of a more comprehensive set of input data, one should try to replace it as soon as possible.

Variants of the SAP method

Similarly to what happens in AP, considering only the cost of the used part of the grid of each country when computing ITCs is perfectly possible. The cost of the unused fraction of the grid would have to be paid locally. In other words, the compensation due to each country j may be computed as a fraction of the cost of the used part of country j ’s grid. This fraction is computed as the ratio of the flow through country j that is caused by other countries to the total flow through the grid of country j .

Apart from that, one may think of computing the participation of the net generation or demand within each country in the electrical usage made of the regional grid, instead of separately computing the electrical usage of the aggregate generation and the aggregate demand within each country.

Modified Average Participations

Background

Given that the algorithm of Average Participations totally ignores political borders, the generation and demand of a comparatively small, well balanced and heavily transited country may produce flows whose sinks and sources, respectively, are located outside the country. Therefore, when considered as a whole, this country appears to be heavily using the network of other countries. As a result of this, charges for these countries in the inter-TSO payment scheme are larger than what one could initially expect. This apparent difficulty vanishes when we analyze the results provided by AP from a Single System Paradigm perspective.

The Modified Average Participations method (MAP) aims to solve this hypothetical problem. This method may be considered a variant of AP. The MAP method assumes that the internal generation in a country is primarily devoted to serve the internal demand (or that the internal demand is primarily served by the internal generation) and that the flows produced and consumed internally only use the national network. Only the imbalance between the internal generation and demand in a country is assumed to make use of the external network.

Description

As mentioned before, the MAP method is based on the assumption that generators and load in a TSO area primarily use the local grid and only the imbalance between generation and demand within a TSO area affects the flows on the grids of other TSOs.

The MAP method is applied by considering each TSO area in turn. All HN elements in a given area are collapsed into a single node. This will be a generation or load node, depending on whether the TSO area is a net exporter to or net importer from other TSO areas. The representation of the HN in the other TSO areas is left unchanged. The AP method is then applied. If the TSO area is a net exporter the net power flow is traced downstream, while if the TSO area is a net importer the net power flow is traced upstream.

Once the charges that each country j has to pay to all other countries are known, it is straightforward to compute the total compensation due to each country as well as the net balance of charges and compensations for the country.

The MAP method makes extensive use of political borders, which, according to the “single system paradigm”, should not influence transmission charges in a

multinational market. Therefore, MAP does not respect this paradigm. This method is discriminatory, since it applies a specific treatment to each agent depending on the internal balance of generation and load within its country. Note, for instance, that the MAP method may result in very different transmission charges for two identical generators that produce the same output and are located very close one to another but at different sides of a border. This method might only be adequate while the regulatory procedures gradually adapt to a more open model where political borders are ignored for network tarification purposes.

Variants of the MAP method

As it has been explained for other methods, one first election when implementing the MAP method is between using the total cost of the horizontal network when computing compensations for external use and just employing the cost of the used part of the grid of each country.

Additionally, when computing the external usage that agents within a country make of the grid of others, we may consider the net demand or generation at every external node in the system instead of considering generation and demand separately.

Methods that are based on the concept of transit

Methods of this family are based on the idea that only those flows through the grid of a country that are not related to power injections and withdrawals within this country must give rise to a compensation for the usage they make of the grid of the country. First we present the concept of transit and its properties. Then we discuss the different methods within this category.

The concept of transit

The concept of transit is intuitive and, apparently, simple to use. The transit through the grid of a country is normally defined as the part of the flows on the grid of this country that has nothing to do with the power injected or produced by the agents located in the country. In other words, the transit through a country is comprised of those flows that are assumed not to have been injected into the grid or withdrawn from it by the agents inside this country.

Charging those agents responsible for the transit through a country j with the costs born by this country due to the existence of the transit seems to be very reasonable. These costs are related either to the electrical usage by the transit flows of the network of country j or to the losses in that network that are incurred by these transit flows.

However, when we try to translate the intuitive but imprecise definition of transit given above into a clear mathematical expression we soon realize that several definitions of transit are possible. The difficulty stems from our inability to find an indisputable method in order to ascertain which flows at the border of a country go

through it and which ones are caused by the internal agents. We have encountered similar problems when dealing with any method based on the computation of electrical usage.

ETSO has been systematically using a definition of transit that is very simple to compute and has intuitive meaning. Given a country j with multiple interconnections with other countries, ETSO determines the size of the transit flowing through country j as the minimum between the sum of all import flows (IMP_j) and the sum of all export flows (EXP_j) for this country.

$$Transit_j = \min[IMP_j, EXP_j] \quad (3.4)$$

This definition implicitly assumes that any border flow that cannot be justified as either a net export from or an import to country j must be regarded as a transit flow. Once the size of the transit has been determined, one needs to distribute it among the several interconnection lines, i.e. decide which fraction of every cross border flow is part of the transit and which one is not. ETSO has proposed to distribute the transit pro rata of the flows at the interconnections. An example follows that shows the application of this rule.

Let us consider the case in Figure 3.5-1, where a country has four interconnection lines. The flow on two of these lines enters the country. Incoming flows are 2000 MW and 1,000 MW. The flow on the remaining two lines leaves the country. Outgoing flows are 800 MW and 700 MW. According to the definition given above, the transit for this country is the minimum between 3,000 MW and 1500 MW, which is 1500 MW. Since we want to examine the effect that the transit has on the flows within the considered country, we need to assign it to the import and export flows. In this case all the export flows are part of the transit (exports are smaller than imports). The transit is distributed among imports proportionately to their size. Thus, the pro rata rule results in transit flows of 1,000 MW, 500 MW, 800 MW and 700 MW on the lines with actual flows of 2000 MW, 1,000 MW, 800 MW and 700 MW, respectively. Figure 3.5-1 illustrates the process of computation of the size of the transit and the allocation of this transit to cross-border flows.

Problems related to the application of the concept of transit to the allocation of electrical usage among countries

The definition of transit given above obviously ignores the pattern of flows inside country j . Figure 3.5-2 shows how this definition results in the same value for the transit through a country under very different circumstances. Note that in case B there is no electric interconnection between the right hand side and the left hand side parts of the country. Therefore, assuming there is a transit of 4000MW through the grid of the country in case B does not make any sense.

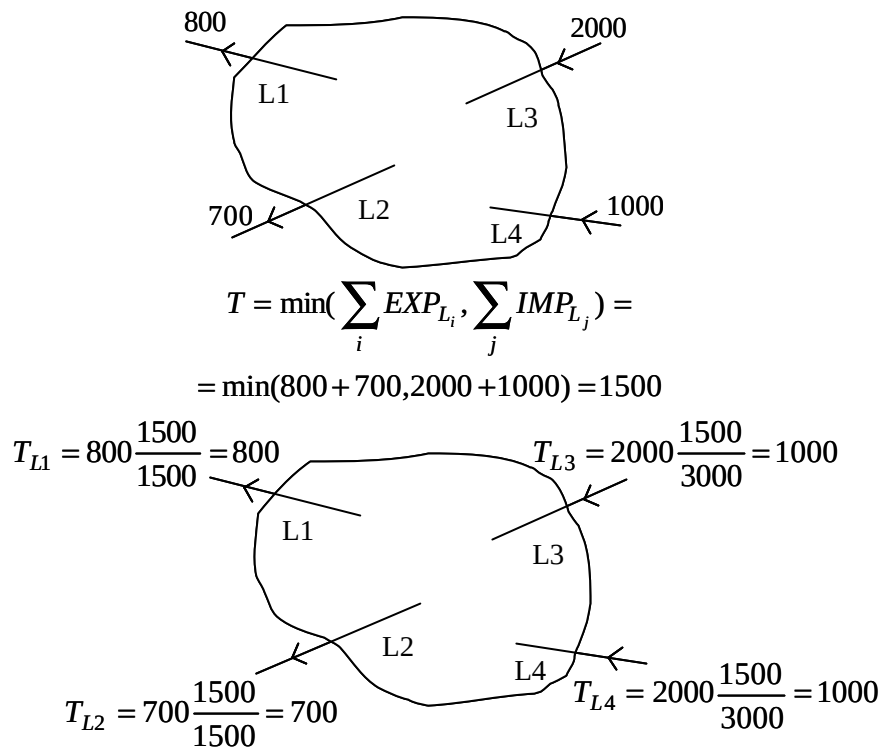


Figure 3.5-1: Definition of the flow pattern of the transit at the borders

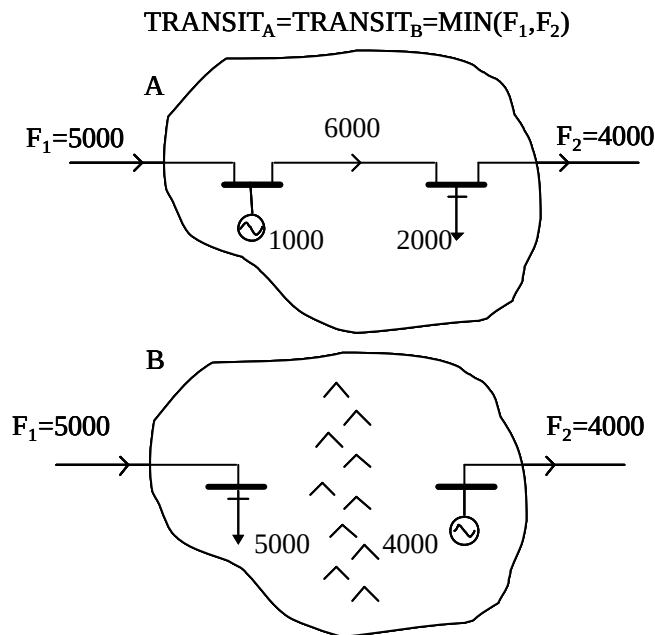


Figure 3.5-2: Transit according to ETSO for two different network topologies

The ambiguity in the definition of the concept of transit translates into several practical problems:

- a) how to specify the flow pattern of the transit when there are more than two interconnection lines between the considered country and others (we have just commented the *ad hoc* pro rata rule of ETSO).
- b) how to assign the responsibility for the transit flows.
- c) how to quantify the electrical usage of the network of the country by each transit flow.

Figure 3.5-3 shows that the economic implications for countries of the application of a method based on transits are clearly dependent on the definition of political borders. According to ETSO's definition of transit, the transit through country A is 3000 MW. This country is therefore entitled to receive an economic compensation based on the electrical usage of the grid of the country by the transit. However, adding a new political border results in new countries A1 and A2 not being crossed by any transit and, therefore, not having the right to receive any compensation.

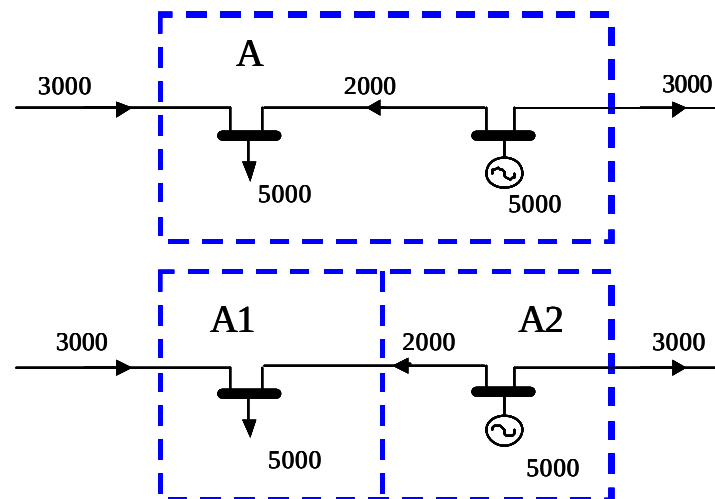


Figure 3.5-3: The influence of political borders on the definition of transits

Despite these difficulties, the concept of transit has found wide acceptance among power system practitioners and it has been frequently used. Two algorithms for the computation of ITCs that are based on the concept of transit are presented in the next subsection. The first one, —the With and Without Transits approach, WWT—, has been presented by ETSO in the Florence Forum. The second one, —the Average Participations applied to Transits method, APT—, is a novel algorithm that uses the same flow tracing rule as AP to offer an alternative to the WWT method. APT overcomes some of the difficulties encountered by WWT.

With and Without Transits method (WWT)

The With and Without Transit algorithm probably is the most simple and intuitive application of the concept of transit to the determination of inter-TSO payments. WWT compares the network flows in the considered country in two different situations. The

real network flows, i.e. those that actually exist, are compared to the flows in a fictitious situation where the transit has been removed.

The fictive “without transit” situation is completely determined once we accept the definition of transit and the allocation rule for this transit provided by ETSO. Computing the network flows in this fictive situation requires running a load flow for the new operating conditions. Care and engineering judgment are essential in the evaluation of the fictive “without transit” situation since the pro rata allocation of flows might result into bizarre local flows that have little to do with the physical reality of the power system. The removal of the transit may result in an augmentation of the power flows and losses in the considered country depending on the “direction” of the transit flows. If transit flows go in one direction they will tend to increase the electrical usage made of the grid as well as transmission losses. If they go in the opposite direction transit flows will provably cause a reduction of power flows and losses. This is just an approximate statement, since flow reversals and other non linear effects may complicate the system behavior.

Once the situations to compare are completely defined, we need to quantify the impact of the transit in the usage made of the network of the considered country. One first option involves determining the volume of flows that are caused by the transit itself in the absence of any other prior network flow. Another possibility is comparing the flows in each line “with and without transit” and considering only those flows that have increased because of the transit. The one we have used to obtain numerical results involves computing a global measure of network usage—the total volume of MWxkm in the entire network—in the two “with and without transit” situations and comparing both amounts.

One of the major drawbacks of the With and Without Transit algorithm lies in the fact that it does not provide any idea on how to allocate the charges corresponding to the compensation due to a country among the remaining ones. ETSO has not provided an indication on how to address this issue, which is critical for the computation of the inter-TSO payments.

By summing all compensation entitlements of all TSOs, we obtain the monetary value of the total compensation fund. Responsibility for contributing to this fund is then allocated to the different TSOs on the basis of their responsibility in generating cross-border/transit flows. This responsibility can be determined in various ways. For example, ETSO proposes using the net import/export from the grid of each TSO as a proxy for the cross-border/transit flows generated by the injections into and withdrawals from that grid. In the WWT method there is no intrinsic consistency in the way that compensation entitlements and responsibilities are computed. Each TSO is assumed to pay every TSO in the region the same fraction of the total compensation due to the latter, irrespective of the geographical or electrical proximity of the grids of the two TSOs and, therefore, of the effect that injections and withdrawals in the grid of one TSO may have on the flows over the other TSO’s grid. Thus, for example,

according to the method proposed for allocating responsibilities in network costs, Portugal would contribute the same fraction of the compensation due to Spain and Finland.

Variants of the With-and-Without Transits method (WWT)

Two main variants on the original WWT algorithm exist. In the first one, only the cost of the used fraction of each line is considered when calculating ITCs. In the second variant the transit through a country is held responsible only for the positive increments in the line flows it causes. ETSO has already obtained results with the latter variant. Variants of WWT present the same drawbacks as the basic approach.

Average Participations applied to transits (APT)

This method aims to solve two major problems of the WWT method while preserving the same underlying philosophy as WWT. APT, like WWT, assesses the electrical usage of the network of a country by the transit through it. However, application of APT does not require defining a fictive “without transit” situation and the method itself determines which countries are responsible for causing the transits and to what extent.

The APT method starts from the same definition of transit employed in the WWT method. Once the transit flow on each one of the interconnections of a country with others have been determined, APT tracks these flows both upstream and downstream, in order to determine: a) how much the transit flows are using the network of country j ; b) where transit flows begin or end in order to allocate responsibilities.

APT uses the same tracking rule as AP. We consider now an interconnection between a country j and another one. The flow over the interconnection is 1,000 MW, of which 300 MW have been determined to be part of the transit. The algorithm tracks the 300 MW transit flow inside country j in order to determine how much it uses the network of this country. Additionally, the 300 MW flow is tracked outside country j in order to assign responsibilities in causing this transit flow.

Quantitative results of the WWT and the APT methods may differ significantly. This comes as no surprise since WWT is an incremental method (when the transit flow is removed, flow reversals and other nonlinear effects may happen) whereas APT only uses the actual flow pattern.

To sum up, some of the criticisms raised against the WWT method are applicable to the APT method as well —such as those derived from the ambiguous concept of transit— but other important ones are properly addressed in APT.

Variants of the Average Participations applied to Transits method (APT)

Variants of the APT method are the same as those of the AP method. Instead of considering the whole cost of the grid of each country when determining the

compensation due to it, we may take into account only the cost of the fraction of each line actually used. Apart from that, we may separately consider the generation and load at each node or just the net demand or generation per node instead. Variants of the APT method complicate it unnecessarily.

The Provisional Method (PM)

This method is presently used to compute inter-TSO compensations. It was proposed by ETSO as a temporary mechanism to be used until a more satisfactory method could be found. Like the WWT and the APT methods PM is based on the concept of transit. However, unlike these two methods, it does not take into account the pattern of flows within each country.

The PM algorithm consists of two steps. First, the compensation due to each country is computed. The size of the total compensation fund results from summing up the compensations due to the different countries. In the second step, the contribution of each country to this fund is determined.

The total share of transit flows in the total flow through the horizontal network of a country, and therefore the share of the HN cost attributable to transits, is defined as the ratio of the volume of transits through that country's grid to the sum of these transits and the load within the country, both on an annual basis. Annual transits equal the annual total of the minimum for each hour of imports into and exports from the country.

Compensations to be paid come from a compensation fund. Each country contributes to this compensation fund according to the sum of the absolute values of the hourly net exports/imports from/into the country throughout the considered year. Additionally, perimeter countries, i.e. those that do not participate in the ITC mechanism, are required to pay 1 € for each MWh exported into the grid of the countries that participate in the IEM.

Variants of the PM method

Contributions to the compensation fund from IEM countries may alternatively be based on other proxies to the amount of cross-border flows created by each country.

Methods recently considered by ETSO

ETSO has been working on a number of alternative methods to WWT as candidates for substituting the Provisional method. One of them is briefly presented in the following paragraphs⁴.

Super-Transactions method (ST)

This method decomposes the existing pattern of flows in the region into a limited number of super-transactions. Specifically, two transactions are considered per country:

- A balanced internal transaction: it is internal to the country, i.e. all the generation and load involved in the transaction is located within the country. Every agent in the country contributes to this super-transaction in proportion to its contracted capacity. The size of the super-transaction equals the minimum between generation and demand in the country. Thus, for example, for a net importing country all the generation would be included. As for the load, the fraction of the power contracted by each consumer considered to be part of the transaction is obtained as:

$$D_{supnac\_i} = D_{total\_i} \frac{G_i}{D_i} \quad (3.5)$$

- An unbalanced regional transaction: a different transaction is considered per country as well. The unbalanced transaction for a net exporting country takes place between the excess of generation over demand in the country and a certain fraction of the excess of demand over generation in net importing countries. This fraction is computed so that generation and demand in this super-transaction are balanced. An analogous reasoning applies to net importing countries. If we focus on the regional super-transaction corresponding to an exporting country j , the excess of generation in j is matched to the demand in net importing countries pro rata of their excess of demand.

Every load and generation within a country participates to the same extent in the corresponding national and regional super-transactions. The ST method aims not to discriminate among the agents of the same type located within a country.

Once a regional super-transaction has been defined for every country, we apply the superposition principle to compute the electrical usage that this transaction makes of

⁴ Once the writing of this thesis has been finalized, still ETSO has proposed a new method called MIMICA, which is being evaluated by ERGEG and also internally by the members of ETSO (March 2006).

the grid of the countries transited by it. This transaction is held responsible for 50% of the incremental flows caused by the super-transaction on the lines of these countries. Note that we are only computing the external usage of the grid of other countries attributable to the generation or demand involved in the transaction.

Therefore, this method makes a distinction between the impact of loop-flows (produced by the national super-transactions) and that of the internal imbalance within each country (produced by the regional super-transaction) on the grid electrical usage of third countries.

3.5.1.3 Further considerations related to the implementation of any method

Application of any method requires addressing some additional issues that are common to all methods. Those are, among others, the definition of the network to be taken into account when computing inter-TSO payments and the selection of the number of scenarios to be considered.

Definition of the Horizontal Network

Identifying the set of lines and other transmission facilities that play a role in cross-border trade is a prerequisite for computing ITCs. These lines comprise what is called the Horizontal Network (HN) of each country. The influence of the definition of the HN in the results obtained for compensations and charges depends on which ITC method we finally choose. We have tried to design a procedure that could be of some help in the definition of the Horizontal Network.

As mentioned before, the Horizontal Network is the set of lines in a country that may be affected to a significant extent by cross-border flows, i.e. those lines whose utilization might correspond, up to a significant degree, to external users. Defining this network a priori - before the responsibility of external users in the line flows has been determined- represents a challenge.

The extent to which the numerical results obtained depend on the definition adopted for the Horizontal Network varies with the method chosen to compute inter-TSO payments. Thus, while the results yielded by the Provisional Method are highly dependent on the size of the Horizontal Network, there is no problem in including as many lines as we want in the Horizontal Network when the WWT method is applied. The WWT method itself is able to determine which lines are affected by the transit and which are not. As for the AP method, it is able to establish which lines within a certain voltage level are used by external agents. However, AP tracks the flows downstream until the final loads are reached. Therefore, if the distribution grid is included in the horizontal network external agents will end up paying for part of it. This does not seem to be right since, assuming there is no distributed generation, the configuration of a distribution grid only depends on the pattern of load in the area. Results provided by APT, MAP and SAP are also dependent on the number and the identity of lines

included in the HN. All in all, unless the WWT method is used to compute inter-TSO payments, the definition of the HN is very relevant. If the WWT method is applied, our only concern should be not to leave out of the HN some lines that are relevant for cross-border trade.

Several possibilities exist to identify which lines belong to the Horizontal Network:

- ETSO has proposed a mechanism that involves checking the flows in an empty network model of the considered country when 100 MW are input at one of the interconnectors on the border of the country and 100 MW are withdrawn from any other interconnector on the border. The process is repeated for all possible pairs of interconnectors. Then, all lines whose flow is above a certain threshold value (e.g. 1 MW), at least in one case, are considered to belong to the Horizontal Network.

This method may be suitable to define the HN to be used by transit flows but not the horizontal network to be used by pure exports or imports. Thus, for example, the HN network defined following this process would not include the transmission lines used by and import flow into Italy that crosses the entire country and is consumed in the South. This method may be applied in combination with transit-based ITC methods, like WWT or the temporary method, but not if the ITC method finally chosen considers all cross-border flows, like AP or MP.

The method proposed by ETSO excludes from the HN any lower voltage lines whose flows are basically not affected by transit flows because higher voltage lines exist that better accommodate these flows. However, it may include in the HN lines that should not belong to it. Consider, for instance, the case of a country with a small interconnector that typically carries a very small flow and is located at a place in the border very far away from the rest of the interconnectors. This is not such an uncommon case: think, for instance, of the weak interconnection of Italy with Greece, which is very distant from the much more important interconnectors of Italy with France, Switzerland, Austria and Slovenia. In this case, the algorithm includes in the HN many lines in the middle of Italy that will certainly not be used by transit flows. One could say that the method is reliable in eliminating lower voltage lines and not excluding any line that should be part of the Horizontal Network to be considered when a transit-based ITC method is applied. It may include, however, too many high voltage lines and it may ignore some, as explained below in the method that is proposed in this thesis.

- An alternative method, which is suitable to define the HN to be considered when applying a transit-based ITC method, involves comparing the line flows in situations that are as similar as possible in the demand pattern but present different flow patterns because of cross-border flows. Only those lines that

show non negligible differences between the flows corresponding to different situations can be considered to belong to the Horizontal Network. Again, this method would not be appropriate to identify those lines that are significantly used by export or import flows, i.e. those that are not transit flows. Applying this method requires having a great amount of snapshots of flows to be compared. These snapshots should have the same generation and demand patterns and grid configuration. This is certainly very difficult. Producing snapshots by means of simulation may be preferable to taking these snapshots out of the actual system behavior.

- We must develop an approach that is able to identify all those lines that are significantly used by cross-border flows of any type in a high enough number of scenarios (lines used by cross-border flows under very infrequent operating conditions should be excluded). Thus, the method we propose is as follows:

The starting point is a representation of the transmission network of the TSO areas included in the analysis, and a number of ‘representative snapshots’, describing the pattern of flows over this network at different times, generally meant to reflect different operating conditions. For each TSO area we define all possible pairs of injection and withdrawal nodes which are relevant for the identification of the HN. These include any pair of nodes at the borders of the TSO area, as well as any pair of a border node and an internal node on the transmission grid. We propose that, for each TSO area, a network element belongs to the HN if a 100 MW increase in injection and withdrawal at the defined pairs of (injection and withdrawal) nodes results in at least 5% of the representative snapshots, and for at least 5% of the node pairs, in a variation of the power flow over the element by more than 1 MW. This approach for identifying the HN may appear cumbersome, but it is not excessively complex, and it would not have to be run every year.

- One final option is making use of statistical correlation methods to identify those lines whose flows show some correlation with the actual flows at the interconnectors or with the magnitude defined as “transit”. This possibility has been explored in a master thesis (Petrucci 2003) jointly supervised by the author of this doctoral thesis and the thesis supervisor.

Selection of the number of scenarios

Parties involved in the design and functioning of the IEM have agreed to compute annual compensations for the electrical usage and losses incurred by each country that are the responsibility of others. Obtaining a solution that accurately reflects the transmission costs that each country should be compensated for would entail taking into account 8,760 scenarios corresponding to each one of the 8,760 hours of the year. The European association of Transmission System Operators (ETSO) has been collecting hourly snapshots of power flows for most IEM countries since 1-1-2003.

Therefore, in principle, 8760 data sets of real network flows are available per year (1 per hour). Despite the availability of these data, application of an ITC method to as many scenarios as hours in a year is not convenient, mainly for political and practical reasons. Collecting and processing 8760 bulky data sets is not trivial. Therefore, determining how many snapshots are necessary to obtain a reasonable estimate of the compensations, charges and net inter-TSO payments within an acceptable margin of error remains an issue of utmost importance. The next subsection presents a method to identify a reduced set of scenarios which is representative enough of the system operation during a given year to compute compensations among countries that are reasonably accurate.

Proposed method

The method explained here is meant to help in the a priori identification of a reduced set of scenarios that could provide an acceptable level of accuracy when computing annual Inter-TSO Compensations. Taking into account the complete set of 8760 snapshots in order to select a small number of them is very unpractical. Therefore, finding a suitable “proxy” of the complete map of network flows in the system becomes necessary. Scenarios to be used in the computation of ITCs would be selected taking only into account their proxies. Hence, the set of system variables used as an approximation to the complete map of flows must verify the following properties:

- a) the complete set of 8760 samples of these variables is easily available.
- b) a high degree of correlation exists between these variables and the feature of interest, namely, a measure of the external network usage of the grid of each country.

The only reasonable proxy that meets the aforementioned conditions is the collection of import/export hourly flows for each country. This information is far more easy to obtain than the map of real network flows. What is more, it seems to capture the essential conditions that influence the external utilization of a network, at least intuitively.

Figure 3.5-4 graphically depicts the method proposed to select those hours that are of interest as representatives of others. In order to pick up representatives we shall classify the 8760 hours of the year into a small number of clusters according to how similar the measures for the different hours are. As said before, we assume that the simple pattern of hourly imports and exports for each country mostly determines the matrix of participations of each country in the use of the grid of others, see Table 3.5-9 to Table 3.5-15 in section 3.5.1.5. This means that the clusters of participation matrices must very approximately coincide with the clusters of patterns of export and import border flows. Therefore, computing clusters of export and import border flows allows us to identify a reduced set of scenarios of real network flows from which to calculate representative participation matrices for network cost allocation purposes.

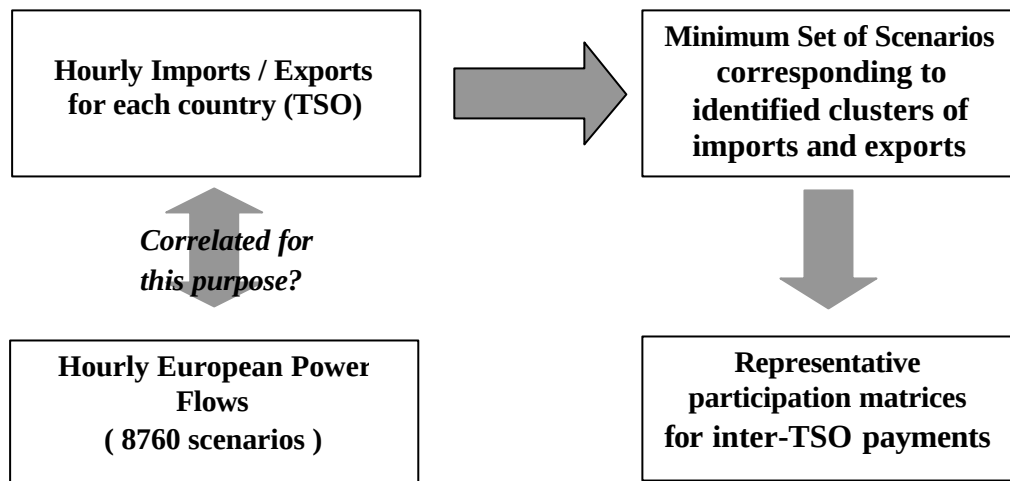


Figure 3.5-4: Selection of the scenarios to be taken into account in the computation of ITCs

Basing the selection of scenarios exclusively on a classification of hourly import and export flows is the only feasible option, taking into account the fact that the amount of hourly data that the author had available was very limited. Section 3.5.1.5 shows numerical results corresponding to the application of the above presented method in order to identify a suitable set of scenarios for the period of time between July 2000 and June 2001. We had available 24 snapshots of the real system operation during the year 2002 for 13 interconnected European countries and 72 scenarios corresponding to the year 2003, but no other variables apart from import and export flows were available for each hour of a whole year.

3.5.1.4 Coherence of the selected method with other aspects of transmission regulation

The method for network cost allocation must be consistent with the overall framework of transmission regulation so that any mutual implications with other aspects of transmission regulation do not create undesirable conflicts, now or in the future. The process of computation of inter-TSO payments is related to the following elements of cross-border transmission regulation:

- Investments in new infrastructures;
- Locational signals for operation and investment and potential pan-European or at least regional transmission tariffs;
- Network tariff harmonization

The relationship that exists between inter-TSO payments and transmission investments provides an example of this need for consistency. The inter-TSO payment mechanism automatically allocates the cost of any new network investment within the HN to the different countries or TSOs, according to a specific procedure. Therefore, the implications of choosing one ITC procedure or another for the expansion of the

regional transmission network have to be examined very carefully. Now we discuss the interaction between the calculation of inter-TSO payments and each one of the three issues listed above. In section 3.5.1.6, we assess the performance of the different methods with respect to these issues.

Allocation of the cost of new network investments

This section aims to explain how the election of an inter-TSO compensation mechanism will affect the allocation of the cost of new network investments among TSOs or countries. Implicit in the inter-TSO payment mechanism, is the fact that any new relevant transmission investment will be included in the Horizontal Network. Hence, its cost will be allocated following the same rules that apply to any other network asset. In conclusion, the cost of any new relevant network investment will be allocated to the different countries via the inter-TSO payment mechanism.

Here we discuss only the three ITC methods that we consider to be most interesting, namely AP, WWT and APT. The following paragraphs discuss the expected implications of using each one of the three considered methods. The MP method, which is presented in (Consentec/Frontier 2005) as the most efficient one conceptually speaking, presents significant flaws that invalidate it for the computation of ITCs in a regional context, see the discussion in section 3.4.2. In section 3.5.1.5 the rules of cost allocation of the three considered methods (AP, WWT and APT) are applied to several case examples of individual lines. These lines are located in different parts of the regional grid and they serve different purposes. Such a quantitative comparison provides a different perspective from which to contemplate the several methods that have been considered for the long-term mechanism of determination of inter-TSO payments.

The AP method

AP renders the consideration of new network investments a trivial matter. Once the line is built and in operation the flows on the line are recorded hourly (for instance), as for all other existing lines. Then, the new line is treated as any existing line. AP assigns the cost of the new line to the different network users individually regardless of the political borders that exist. AP only takes into account the real network flows. Therefore, the resulting inter-TSO compensations and charges automatically include the allocation of the cost of the new line to the different TSOs. The method itself precisely assigns the cost of the new line.

The unit cost of the line is an input to the model. Both standard costs and regulated costs are possible. If the cost of constructing, operating and maintaining a line is assigned by means of a competitive auction, this is the cost figure to be used for this line.

The WWT method

Once the line is built and in operation, the WWT method determines how the flow over the considered line varies when the transit is removed. This information may be used together with some simple pro rata assignment in order to evaluate how much of the total electrical usage of the line corresponds to external agents. Two difficulties appear:

- The WWT method does not provide any guidance on how to allocate the cost corresponding to the external utilization of the line to network users in other countries, either at agent or at country level. One needs to resort to some other allocation method or to an arbitrary assignment rule.
- The change in the flow over the new line attributable to the transit may critically depend on the pro rata assignment rule used by the WWT method to distribute the transit flow through a country among the different interconnections of the country. This becomes quite clear when the new investment is a cross-border line. Then, the WWT method itself determines how much of the transit corresponds to the considered line without performing any technical evaluation.

The APT method

Following the procedure described when presenting this method, APT establishes how much the transit flows use any new transmission line in the system. Given that the method itself computes the end and origin of these flows, we can determine which agents have created these flows and are therefore using the new lines. However, the APT method does not provide any information about the use made of new lines by cross-border flows that are not part of the transit through a country.

Computation of locational signals for operation and investment

Presently, well established and complete EU-wide locational signals for transmission do not exist in the IEM. This situation may remain for some time in the future. Neither energy nodal prices nor pan-European transmission network tariffs, which could potentially have some locational content, are applied at EU level. Only two mechanisms provide some sort of locational signals at regional level:

- Congestion management. Still in its infancy at European level. Some uncoordinated auction mechanisms are already in place. The EU Regulation requires the implementation of market based mechanisms. CEER has worked in the development of practical guidelines to be met by the implemented congestion management scheme. ETSO and other stakeholders are working on the development of coordinated congestion management schemes.
- Inter-TSO payments. Local G and L charges complemented by inter-TSO compensations already provide some locational signals. The locational content

of those signals provided by the ITC mechanism is small since compensations and charges are computed and applied at TSO level.

Inter-TSO payments are primarily meant to compensate economically countries for the costs they bear that are caused by external agents. ITCs are not aimed at sending precise locational signals to the individual agents of the market. However, the inter-TSO payment mechanism adopted in the Florence Forum may also play a role in sending correct locational signals. Inter-TSO payments modify the basic local G and L tariffs resulting in access charges to the entire IEM transmission network.

As explained above, some methods used to compute inter-TSO payments allocate network charges at agent level. These methods result in nodal transmission tariffs, see (Pérez-Arriaga, Olmos et al. 2002). Other methods, only aimed at computing ITCs, are not useful to calculate transmission tariffs.

Transmission tariffs, if computed using an appropriate cost allocation method, are perfect long-term locational signals. No more could be asked for in terms of location and tariff harmonization. Since, for the time being, it has been agreed not to apply a system of pan-European transmission tariffs, nodal tariffs obtained with the methods referred to in of the previous paragraph would be aggregated at TSO level in order to obtain compensations among countries. These compensations must then modify local L and G charges.

Provided an ITC mechanism is in place that results in nodal transmission tariffs, this method itself will provide us with the key to allocate net inter-TSO payments to the internal G and L tariffs in any given country j . An ITC method of this type precisely determines how much the generators in country j are using of the network external to j , as well as the extent to which generators from outside country j use the network of this country. The excess of the former amount over the latter is the extra charge (or credit, if the amount is negative) that generators in country j should pay. Locational signals thus obtained ignore any geographical differentiation within a country since the average value for country j is used instead of the individual value for each generator. The same reasoning applies to consumers.

To conclude, if properly defined and implemented correctly, the inter-TSO payment mechanism contributes adequate long-term locational signals at TSO or country level. These signals complement transmission network tariffs locally implemented by countries or TSOs.

As long as the present lack of harmonization in the computation and design of transmission network tariffs in the IEM persists, the impact of the contribution of inter-TSO payments is expected to be low in general, as (Perán, Pérez-Arriaga et al. 2001) suggests.

Implications in network tariff harmonization

As indicated above, harmonization of network tariffs is essential both to provide a level playing for competition to develop in the IEM and to send adequate locational signals at country or TSO level. First, we outlay the procedure that should ideally be used to harmonize tariffs. Then, a more practical solution is presented for its implementation at European level.

A suggested procedure

Any sound transmission tarification procedure at IEM level should be based on the responsibility of each network user in the utilization or development of each one of the elements of the transmission network in the IEM. This should not be conditioned by political borders. The implementation of such a tarification procedure consists of three steps:

- Computing the responsibility of every network user in the development or the utilization of each one of the facilities of the transmission network, regardless of political borders.
- Translating responsibility in network use into monetary terms. Using standard transmission costs across the IEM plays a key role in this.
- Computing transmission tariffs G and L using a certain structure for these charges (i.e. charges per kWh, per kW or per customer).

Procedures to calculate the remaining transmission charges beyond the G and L charges whose computation is described above should be harmonized as well across member states, if this is possible. For instance, charges (or credits) needed to complete the recovery of transmission costs in a given country or TSO may be assigned to consumers, either uniformly or in such a way that total charges to demand become uniform. Complementary charges may be paid by the demand in the country according to national regulation. The advantages of this harmonization approach are:

- All generators within the IEM receive totally correct long-term locational transmission signals meant to convey correct siting incentives.
- The total remuneration of the transmission activity within each country or TSO is left entirely to subsidiarity. This appears to be reasonable because of the large disparity of values and procedures adopted in the different countries or TSOs. These procedures have their origin in well established commitments of the regulatory authorities in each country with the owners of transmission assets.
- Regulatory authorities in each Member State are responsible for the application to their consumers of the complementary charge necessary to complete the recovery of the cost of the transmission service. The remaining amount to be

allocated to consumers is the difference between the accepted remuneration of the transmission activity in the country and the sum of the amount collected through the application of local G and L tariffs and the net inter-TSO compensation for the country. In general, this will distort long-term locational signals to consumers. However, this is deemed to be an acceptable compromise, since:

1. Incorrect locational transmission signals to consumers are deemed to distort market results to a lesser extent than incorrect signals to generators do. This is particularly true if the extra charge is applied so that the loss of economic efficiency is minimized, as it is the case when Ramsey-like tarification methods are employed.
2. Some Member States have long standing commitments to maintaining uniform tariffs for consumers within their territories.
3. The economic value of the difference between the regulated grid cost and the fraction of this cost paid through the system of IEM wide tariffs has to be assigned somehow and this appears to be the least damaging option.

The procedure to compute IEM-wide transmission tariffs presented here implicitly contains an inter-TSO compensation scheme. Political borders have been ignored when allocating responsibilities in transmission costs. Therefore, the resulting G and L tariffs include the compensations and payments that should take place among countries if agents from each country paid the national transmission grid.

A simplified version

In the absence of a method to compute harmonized nodal transmission charges, the European Commission and the Tarification Task Force of CEER have been working on a simplified approach to tariff harmonization. This approach basically amounts to leave each country to set a value for the G tariff, either uniform or with some geographical differentiation, such that the associated revenues from generation represent a small percentage of the total regulated cost of the Horizontal Network of the country. The remaining cost of the transmission activity within each country, including any net inter-TSO payments, is recovered for each country by the L charges.

Conceptually speaking, transmission charges should be structured as a capacity charge instead of an energy charge in order not to distort the energy dispatch. However, given that G charges obtained according to these harmonization guidelines would be very small, the distortion they would introduce in the short-term operation of the power system, if applied as energy charges, would be negligible. The advisability of using a variable energy charge to recover transmission grid costs largely depends on the relative size of the differences between transmission charges in different areas

(countries) when compared to the size of the corresponding differences in energy prices.

Perán et al. conclude in (Perán, Pérez-Arriaga et al. 2001) that, given the circumstances existing in the different EU countries, harmonizing the way in which transmission tariffs are computed may not be an option. This is due to the fact that an important part of transmission tariffs are related to regulatory payments that have their roots in the history and particular situation of the industry within each country. In any case, some level of harmonization is necessary in order not to distort competition.

3.5.1.5 Results for the European system

This section includes the results of nodal tariffs and inter-TSO payments obtained for the European system. This section also includes the outcome of a study of application of the methodology presented in 3.5.1.3 aimed at selecting a number of scenarios from which to compute annual inter-TSO payments.

Inter-TSO compensations

As for the calculation of inter-TSO compensations, results are divided into two different parts. First, we compare the results obtained with AP, APT, WWT, MAP, SAP and the Provisional method for the year 2001. Only one scenario, corresponding to the winter peak load (17th of January), has been taken into account when computing these results. Then the outcome of a more comprehensive study covering the most promising methods follows. The methods studied are AP, WWT and APT. Numbers in this study correspond to 24 scenarios representative of the real operation of the system during the year 2002.

Compensations for the year 2001

We compare the results of compensations, charges and net payments obtained with AP, APT, WWT, MAP, SAP and the Provisional method for the year 2001.

Description of the data

Due to the scarcity of officially provided data on scenarios of real network flows, the analysis reported here refers to a single case: the UCTE model for January 17th, 2001, at 10:30 am. Values of system variables in this scenario were estimated by the different system operators. A single scenario of a large meshed network comprised of 17 countries is more than enough to assess the methods and identify most of the pros and cons of the different algorithms. Based on this assessment, we identify on a preliminary basis the most suitable methodology for network cost allocation in the European context. Countries considered are those that appear in Table 3.5-1.

Numbers for compensations and losses are not expressed in monetary terms, since the required cost figures were not available at the moment of the study. On the one hand, data on the length of the lines in the model had not been provided and it could not be inferred from line reactances, since more than 5% of the lines of the model did not correspond to physical lines but to fictive equivalent lines. On the other hand, the standard cost per km of 400 kV lines, 220 kV lines, 132 kV lines, as well as for the different types of transformers, were not available either.

The following decisions were adopted in order to address the problem of the lack of data on line lengths and standard network costs:

1. All lines were considered to be of equal length.
2. Based on the available cost information, it was assumed that 1 km of a 220 kV line is equivalent to 0.66 km of a 400 kV line and that 1 km of a 132 kV line is equivalent to 0.44 km of a 400 kV line.
3. Some transformers were considered to be 400 kV lines, others were considered to be 220 kV lines and the remaining ones 132 kV lines.

Once these assumptions are accepted, results can be expressed in terms of numbers of equivalent 400 kV lines.

Table 3.5-1 provides the aggregated import and export flows between the different countries, as well as the aggregated generation and demand, the number of generation and demand nodes, the aggregated net export and the size of the transit for each of the 17 countries. The UCTE model as a whole has 3,655 lines and 3,383 nodes. The total amount of production is 244.00 GW, the total load is 240.86 GW and losses amount to 3.14 GW.

Country	Name	Demand	Generation	#D	#G	Exports	Imports	Net E-I	ETSO-T
E	Spain	29,090	28,289	197	103	554	1,851	-1,298	554
P	Portugal	6,364	6,891	35	20	755	361	395	361
F	France	64,804	74,126	677	184	8,487	193	8,293	193
I	Italy	31,480	26,508	217	112	0	5,321	-5,321	0
CH	Switzerland	5,720	5,877	53	37	2,686	2,604	82	2,604
D	Germany	50,150	49,649	367	93	3,887	5,042	-1,155	3,887
B	Belgium	7,170	6,405	31	12	1,053	1,752	-699	1,053
NL	Netherlands	10,651	7,134	34	14	325	3,923	-3,597	325
SLO	Slovenia	820	929	4	3	814	764	50	764
A	Austria	2,814	3,905	20	20	1,961	577	1,385	577
CZ	Czech Rep.	8,194	9,836	32	16	2,099	791	1,308	791
PL	Poland	15,449	16,032	106	29	843	235	608	235
BIH	Bosnia	316	671	6	7	146	146	0	146
HR	Croatia	995	1,122	7	7	733	684	50	684
H	Hungary	4,146	3,619	24	7	373	725	-352	373
SK	Slovak Rep.	2,616	2,759	24	10	569	567	3	567
UA	Ukraine	84	250	1	1	279	29	250	29

Data in MW

Table 3.5-1: Aggregated data for the considered UCTE scenario

Results for the different methods

Since the results for the year 2001 are based on the analysis of a single scenario, they can only be considered for the purpose of comparing the different network cost allocation algorithms. Under no circumstances are they useful to infer the final results of compensations, charges and inter-TSO payments among countries.

Table 3.5-2, Table 3.5-3 and Table 3.5-4⁵ summarize the results of inter-TSO payments that have been obtained with the different methods. As explained before in the chapter, AP stands for Average Participations, SAP for Simplified Average Participations, MAP for Modified Average Participations, WWT corresponds to the With and Without Transit method, and APT stands for Average Participations applied to Transits. Figures in the tables represent percentages of the total number of equivalent 400 kV lines in the considered UCTE scenario. The number of 400 kV equivalent lines in the considered case is 3,655.

Inter-TSO compensations were not computed for Marginal Participations because this method was discarded from the outset taking into account the following considerations:

1. Results are too much dependent on arbitrary modeling decisions, —such as the choice of reference node—.
2. MP is not able to reflect the actual flow pattern and the physical limits imposed by the grid, such as the limited capacity of the lines.

After giving a quick look at numbers in these tables one can conclude that significant numerical differences exist among the methods. Some persistent patterns of behavior can be easily identified. Heavily importing or exporting countries, such as Italy or France, respectively, receive large compensations (except in the WWT and APT methods) but pay even larger charges, which results in significant negative inter-TSO payments from them, i.e. they have to pay. On the other hand, well balanced and heavily transited countries such as Switzerland receive large compensations and pay small charges (this effect is emphasized in the MAP, WWT and APT methods), thus obtaining large positive inter-TSO payments, i.e. they are net receivers.

Compensations for transited countries that are also net exporters or importers do not follow a clear pattern. Thus, for example, the net inter-TSO payment for Germany or Portugal in the WWT method has the opposite sign to the corresponding result in all

⁵ Note that in the tables E stands for Spain, P for Portugal, F for France, I for Italy, CH for Switzerland, D for Germany, B for Belgium, NL for The Netherlands, SLO for Slovenia, A for Austria, CZ for the Czech Republic, PL for Poland, BIH for Bosnia, HR for Croatia, H for Hungary, SK for the Slovak Republic and UA for Ukraine.

the other methods. Similarly, the net compensation for the Netherlands in AP has the opposite sign to the corresponding result in the other methods.

Additionally, MAP and APT tend to reduce the charges to be paid by well balanced countries, —i.e. countries with approximately the same amount of internal demand and generation—. Net values of inter-TSO payments that are obtained by both methods are quite similar. APT treats unfavorably purely exporting or importing countries.

	E	P	F	I	CH	D	B	NL	SLO	A	CZ	PL	BIH	HR	H	SK	UA
AP	0.35	0.11	0.87	0.73	1.11	1.34	0.22	0.39	0.11	0.44	0.21	0.18	0.08	0.16	0.13	0.22	0.03
SAP	0.24	0.06	1.70	1.01	1.00	1.70	0.31	0.33	0.11	0.40	0.20	0.20	0.00	0.08	0.13	0.20	0.02
MAP	0.21	0.05	0.54	0.48	0.90	0.99	0.21	0.01	0.05	0.14	0.07	0.01	0.00	0.10	0.05	0.16	0.01
WWT	-0.09	0.20	0.00	0.00	1.25	-0.30	0.08	0.09	0.10	0.32	0.07	0.02	0.16	0.25	0.12	0.05	0.00
APT	0.17	0.07	0.03	0.00	1.13	1.16	0.16	0.05	0.12	0.21	0.13	0.10	0.08	0.16	0.10	0.23	0.01

Table 3.5-2: Compensations due to countries because of network usage for the year 2001

	E	P	F	I	CH	D	B	NL	SLO	A	CZ	PL	BIH	HR	H	SK	UA
AP	0.30	0.19	1.48	1.02	0.60	0.94	0.24	0.31	0.10	0.35	0.43	0.16	0.02	0.20	0.13	0.15	0.06
SAP	0.24	0.07	1.73	1.25	0.61	1.01	0.32	0.81	0.11	0.43	0.57	0.13	0.00	0.07	0.18	0.11	0.04
MAP	0.18	0.07	1.49	1.02	0.02	0.10	0.07	0.30	0.01	0.26	0.26	0.08	0.00	0.01	0.05	0.00	0.06
WWT	0.09	0.03	0.81	0.26	0.06	0.27	0.07	0.18	0.04	0.14	0.16	0.07	0.01	0.03	0.03	0.03	0.02
APT	0.09	0.10	0.84	0.64	0.17	0.44	0.03	0.32	0.05	0.30	0.36	0.17	0.02	0.16	0.10	0.07	0.05

Table 3.5-3: Charges to be paid by countries because of network usage for the year 2001

	E	P	F	I	CH	D	B	NL	SLO	A	CZ	PL	BIH	HR	H	SK	UA
AP	0.05	-0.08	-0.61	-0.29	0.51	0.40	-0.02	0.07	0.01	0.09	-0.22	0.02	0.06	-0.04	0.01	0.07	-0.03
SAP	-0.01	-0.01	-0.04	-0.24	0.39	0.69	-0.01	-0.48	0.00	-0.03	-0.37	0.07	0.00	0.01	-0.05	0.09	-0.02
MAP	0.03	-0.02	-0.95	-0.54	0.89	0.89	0.14	-0.29	0.05	-0.11	-0.19	-0.08	0.00	0.09	-0.01	0.16	-0.05
WWT	-0.19	0.17	-0.81	-0.26	1.19	-0.57	0.01	-0.10	0.06	0.19	-0.08	-0.05	0.15	0.22	0.09	0.02	-0.02
APT	0.08	-0.02	-0.81	-0.64	0.96	0.72	0.13	-0.27	0.07	-0.09	-0.23	-0.07	0.06	0.00	-0.01	0.16	-0.04

Table 3.5-4: Inter-TSO payments because of network usage for the year 2001

A first estimate of the economic implications of each method

Using estimated values of the average length and the standard unit cost of lines in the system, we have translated compensations, charges and net inter-TSO payments into monetary terms. In this way we can get a rough idea of the total volume of compensations in each method. Standard costs are mainly based on the current Spanish regulation. These figures are different from the ones used by ETSO in the Provisional Method. They correspond to an estimate of representative values of the system. Results of economic compensations must be taken very cautiously: not only are we using data from a single scenario but we are also employing a crude estimate of the

unit costs and lengths of the lines. However, these results are useful to carry out additional comparisons among the different methods, as well as to guess the order of magnitude of the expected outcome of each one of these methods.

Table 3.5-5, Table 3.5-6 and Table 3.5-7 show the compensations, charges and net payments for the different methods when expressed in monetary terms. Figures in these tables are expressed in million euros. We only provide results for the 9 countries in the study that belonged to the IEM and the UCTE at the time the study was carried out. Every 400 KV equivalent line in the system has been considered to be 50 km long. The total annual cost of one of these lines has been assumed to be 0.653 million euros (0.01306 million euros per km of line). Using these basic data, compensations, charges and inter-TSO payments have been computed. All 8760 hours in the year 2001 have been considered to be equal to the single available hourly snapshot.

	E	P	F	I	CH	D	B	NL	A	TOTAL
AP	8.4	2.6	20.8	17.4	26.5	32.1	5.3	9.2	10.4	132.8
SAP	5.6	1.4	40.2	24.0	23.8	40.4	7.4	7.8	9.5	160.2
MAP	5.0	1.2	12.9	11.5	21.6	23.8	5.0	0.3	3.5	84.7
WWT	-2.3	4.7	-0.1	0.0	29.8	-7.1	2.0	2.1	7.7	55.8
APT	4.0	1.8	0.7	0.0	27.2	27.9	3.9	1.1	5.1	71.6
PM(9)	4.8	2.7	2.1	0.0	27.7	40.9	7.0	1.4	6.4	93.2

Table 3.5-5: Total economic compensations (M€) related to network infrastructure costs for the year 2001

	E	P	F	I	CH	D	B	NL	A	TOTAL
AP	7.2	4.6	35.3	24.4	14.4	22.5	5.7	7.4	8.3	129.9
SAP	5.8	1.7	41.1	29.7	14.5	24.0	7.6	19.1	10.3	153.8
MAP	4.2	1.7	35.7	24.5	0.4	2.4	1.6	7.2	6.2	83.8
WWT	2.2	0.7	19.4	6.1	1.4	6.6	1.7	4.4	3.3	45.9
APT	2.1	2.4	20.2	15.3	4.2	10.5	0.8	7.6	7.2	70.2
PM	4.3	2.1	31.0	10.0	5.0	9.2	3.2	7.3	6.2	78.4

Table 3.5-6: Total economic charges (M€) related to network infrastructure costs for the year 2001

	E	P	F	I	CH	D	B	NL	A	TOTAL
AP	1.2	-2.0	-14.6	-7.0	12.1	9.5	-0.5	1.8	2.2	50.8
SAP	-0.2	-0.3	-0.9	-5.7	9.3	16.5	-0.2	-11.3	-0.7	45.1
MAP	0.8	-0.5	-22.7	-13.0	21.2	21.4	3.3	-6.9	-2.7	92.7
WWT	-4.5	4.0	-19.5	-6.1	28.4	-13.7	0.3	-2.3	4.4	83.1
APT	1.9	-0.6	-19.5	-15.3	23.0	17.4	3.1	-6.5	-2.1	89.4
PM	0.5	0.6	-28.9	-10.0	22.7	31.7	3.8	-5.9	0.3	104.3

Table 3.5-7: Total volume of inter-TSO payments (M€) related to network infrastructure costs for the year 2001

Table 3.5-5, Table 3.5-6 and Table 3.5-7 include numbers for the Provisional Method as well, in order to compare it to the real network flow methods being examined.

As expected, compensations for the Average Participations method are larger than those obtained for other methods, apart from Simplified Average Participations. However, intermediate values of compensations and charges are not significant. Net inter-TSO payments on the contrary represent the amount that each country will receive or pay in the end. They are the true outcome of the inter-TSO payment mechanism. Table 3.5-7 shows the values of the net payments. The leftmost column in the table shows the sum of the absolute values of the figures in the corresponding rows, i.e. the sum of the absolute values of the net payments received or paid by each country. Results show that the total volume of inter-TSO payments is very similar in all the methods. Average Participations yields one of the lowest figures and the Provisional Method yields the highest.

Compensations for the year 2002

Results for the year 2001 together with the qualitative analysis carried out above have led us to conclude that the AP method is the preferred one. The WWT and the APT methods may be suitable alternatives if parties finally opt for a long term method based on the concept of transit. We now move on to further assess these three methods. The Provisional Method has also been included in the study because it is the one presently applied. Results of the study correspond to 24 snapshots of network flows corresponding to the real operation of the system in 2002. The system studied is comprised of 13 IEM countries. Results include bilateral compensations and charges as well as net payments for the 13 countries in the system.

This subsection also explores some less common features of the three selected methods that were not addressed in the study for the year 2001. This completes the comparative evaluation of the three methods.

The methods compared here have some characteristics in common. Each one of them has some degrees of freedom, which result from ambiguities and not entirely justified assumptions in the definition of the methods.

Numbers in the tables presented in this section are expressed in millions of euros unless otherwise stated. An explanation of the meaning of the figures provided in the tables follows the one corresponding to the AP method.

Description of the data

The four selected methods have been applied to 24 scenarios corresponding to the real operation of a system comprising 13 IEM countries during the year 2002. The scenarios have been provided by ETSO. These scenarios were chosen according to the outcome of a statistical analysis of hourly values of power flows on tie lines. Snapshots correspond to 4 different months: March, June, September and December and were recorded at 03:30h, 11:30h and 19:30h of third Wednesdays and preceding Sundays of each of these months. As said before, the data comprise 13 countries: Austria, Belgium, Switzerland, the Czech Republic, Germany, Spain, France, Hungary, Italy, The Netherlands, Portugal, Slovenia and Slovakia. Acronyms used in the tables to refer to the countries are the same as those employed for the year 2001.

Table 3.5-8 summarizes the aggregate data for the average scenario, which aims to represent the 24 scenarios used in the study. Each row provides all the information for a country, which includes the aggregate values of generation, load, exports, imports and transits as well as the absolute values of the net exports. All numbers are expressed in MW. The IEM-13 network has around 3950 nodes and 5700 lines (numbers vary depending on the considered scenario).

Some other data have been used to obtain the results presented here. We have had to estimate the annual cost per km of 400KV line in order to express both compensations among countries and nodal transmission tariffs in monetary terms. First, we computed the number of kilometers of equivalent 400KV lines within the Horizontal Network of 11 European countries. These countries were Austria, Belgium, France, Germany, Italy, Portugal, Spain, Switzerland, the Czech Republic, The Netherlands and Slovenia. The figures obtained were 3896 km for Austria, 3930 for Belgium, 4477 for Switzerland, 5116 for the Czech Republic, 34811 for Germany, 21511 for Spain, 27916 for France, 13019 for Italy, 3737 for the Netherlands, 3627 for Portugal and 696 for Slovenia. Second, we divided the regulated cost of the Horizontal Network of each country (values were those used in the Provisional Method for the year 2002) by the number of kilometers of equivalent 400 KV lines within the network of the country. Thus, a regulated cost per kilometer of equivalent 400KV line was obtained for each country. Finally, the standard unit cost of a 400 KV line used to express all the results in monetary terms resulted from computing the average of the regulated unit costs obtained for the different countries. An annual value of 0.0352 million Euros per km of 400 kV line was obtained. The European Commission issued at the same time of the study an estimate of the replacement costs per km of double circuit 380 KV line for most of the countries considered in this study. Values were in the range of €350,000 to €600,000. This estimate results in annual costs in the same range as the one considered

here, if we assume a depreciation period of 30 years and a cost of capital per annum of 6% real.

	Generation	Load	exports	imports	net E-I	transit
A	2193.2	2134.9	1464.8	1481.6	16.8	1464.8
B	6366.8	7130.2	921.4	1726.1	804.7	921.4
CH	5196.6	4498.7	3488.6	2931.5	557.1	2931.5
CZ	5167.6	3708.8	2250.7	1065.1	1185.6	1065.1
D	44187.6	43411.9	4437.7	4983.5	545.7	4437.7
E	18805.2	18221.1	743.8	1326.7	582.9	743.8
F	53158.3	42804.9	8284.6	91.0	8193.6	91.0
H	2492.6	2943.4	797.0	1275.0	478.0	797.0
I	16565.1	21860.6	13.6	5706.2	5692.7	13.6
NL	7317.5	9317.9	350.3	2455.1	2104.8	350.3
P	3951.9	3957.9	425.3	529.6	104.3	425.3
SLO	1088.9	626.7	986.9	703.4	283.4	703.4
SK	2308.2	1833.8	1059.1	608.2	450.8	608.2

Table 3.5-8: Aggregate data for the average scenario of 2002 in the 13-IEM system

Snapshots also include the capacity of the line as part of the data. Line capacities are necessary to obtain numerical results of inter-TSO payments when only the cost of the used fraction of each line is allocated in the mechanism. However, the capacity of some lines was missing. For these lines, we have used the average capacity of the lines of the same type in the system.

The length of every line has been estimated from its impedance according to some coefficients provided by ETSO. These coefficients are different for each type of line.

Results for the AP method

Table 3.5-9 shows the results obtained for the AP method in its basic form. An explanation of the figures in the table follows. If we take column F in the table, which corresponds to France, numbers in the column must be interpreted in the following way:

- 31.2 in the column means that the fraction of the French grid that Italy must pay is worth 31.2 million euros.
- 1121.9 represents the cost of the fraction of the French grid to be paid by local grid users.
- 1195.2 is the total cost of the French network and it is equal to the sum of all numbers in the column above.
- 1232.0 corresponds to the total electrical usage made by France of the European network, including the French network. This figure is obtained as the sum of all numbers in the row labeled F.

- The difference between the two aggregate numbers: $1195.2 - 1232.0 = -36.8$, is the net inter-TSO payment for France. Since it is negative, France turns out to be a net payer. This means that the electrical usage made by France of the grid of other countries is worth more than the electrical usage that these other countries make of the French network. This is the most important result of the analysis for France. Thus, the row combining the net result for all the countries has been framed.

Considering now the numbers in row F, again for France, we can say that:

- 20.0 represents the cost of the fraction of the Swiss network used by France.

Other numbers of interest in column F are:

- The number 73.3 in column F is the sum of the cost of the fractions of the French network used by others (7.2 by Belgium, 14.5 by Switzerland, 5.8 by Germany, 14.6 by Spain, 31.2 by Italy and 0.1 by The Netherlands). This number does not include the cost of the fraction of the French network allocated to France.
- Similarly, the number 110.1 in column F is the sum of the costs of the fractions of other networks that have been allocated to France. The value of the participations of France in the electrical usage of other networks are shown in row F, (22.1 of Belgium, 20.0 of Switzerland, 29.1 of Germany, 11.1 of Spain, 27.4 of Italy and 0.4 of The Netherlands).
- It should be clear now that the difference between the figures in the two previous bullets is equal to the inter-TSO payment for France.

	A	B	CH	CZ	D	E	F	H	I	NL	P	SLO	SK
A	86.0	0.0	0.0	7.4	6.5	0.0	0.0	1.5	3.5	0.0	0.0	3.4	1.0
B	0.0	123.1	0.0	0.0	0.5	0.0	7.2	0.0	0.0	11.0	0.0	0.0	0.0
CH	0.8	0.0	93.4	0.0	11.6	0.0	14.5	0.0	12.4	0.0	0.0	0.0	0.0
CZ	14.9	0.0	0.0	146.4	16.7	0.0	0.0	1.2	0.2	0.0	0.0	0.3	3.3
D	13.7	0.8	11.3	7.7	1228.8	0.0	5.8	0.0	2.7	26.9	0.0	0.1	0.0
E	0.0	0.0	0.0	0.0	0.0	791.2	14.6	0.0	0.0	0.0	17.2	0.0	0.0
F	0.0	22.1	20.0	0.0	29.1	11.1	1121.9	0.0	27.4	0.4	0.0	0.0	0.0
H	2.0	0.0	0.0	1.3	0.0	0.0	0.0	68.1	0.0	0.0	0.0	0.0	4.0
I	9.1	0.0	30.8	0.1	5.3	0.0	31.2	0.0	463.6	0.0	0.0	4.9	0.0
NL	0.0	2.5	0.0	0.0	9.9	0.0	0.1	0.0	0.0	195.1	0.0	0.0	0.0
P	0.0	0.0	0.0	0.0	0.0	27.2	0.0	0.0	0.0	0.0	133.3	0.0	0.0
SLO	7.3	0.0	0.0	0.2	0.1	0.0	0.0	0.0	1.9	0.0	0.0	14.2	0.0
SK	1.7	0.0	0.0	5.6	0.0	0.0	0.0	7.5	0.0	0.0	0.0	0.0	58.8
	135.4	148.4	155.6	168.7	1308.5	829.6	1195.2	78.3	511.7	233.4	150.6	23.0	67.1
	109.4	141.8	132.6	183.1	1297.7	823.0	1232.0	75.3	545.0	207.5	160.6	23.6	73.6
	26.0	6.6	22.9	-14.4	10.8	6.6	-36.8	3.0	-33.3	25.8	-10.0	-0.7	-6.5
	49.5	25.3	62.2	22.3	79.7	38.4	73.3	10.2	48.1	38.3	17.2	8.8	8.3
	23.5	18.7	39.2	36.7	68.9	31.8	110.1	7.3	81.4	12.4	27.2	9.5	14.8

Table 3.5-9: Annual compensations, charges and net payments because of network usage produced by the AP method for the year 2002

A comparison of the different possible implementation of the AP method follows. Only results for the first scenario of the 24 available have been compared. Results for one

scenario are more than enough to assess which alternative is preferable as far as numerical results are concerned. Furthermore, computing results for only one scenario takes far less time than repeating the process for the 24 scenarios. Once we have obtained numbers for one scenario we know that computing results for more is possible.

Table 3.5-10 shows compensations, charges and net payments produced by the basic AP method for the first scenario of the 24 provided by ETSO corresponding to the year 2002.

	AT1	BE1	CH1	CZ1	DE7	ES1	FR1	HU1	IT1	NL1	PR1	SI1	SK1
AT1	73	0	0	3	5	0	0	0	4	0	0	2	0
BE1	0	168	0	0	2	0	7	0	0	6	0	0	0
CH1	0	0	77	0	10	0	21	0	9	0	0	0	0
CZ1	19	0	0	150	18	0	0	0	1	0	0	1	1
DE7	12	1	15	6	1208	0	7	0	5	23	0	0	0
ES1	0	0	0	0	0	769	18	0	0	0	21	0	0
FR1	0	20	25	0	27	22	1122	0	29	0	0	0	0
HU1	1	0	0	0	0	0	0	64	0	0	0	0	4
IT1	20	0	31	0	9	0	36	0	456	0	0	7	0
NL1	0	1	0	0	6	0	0	0	0	183	0	0	0
PR1	0	0	0	0	0	32	0	0	0	0	128	0	0
SI1	5	0	0	0	0	0	0	0	4	0	0	12	0
SK1	1	0	0	4	0	0	0	7	0	0	0	0	59
	131	190	149	164	1285	823	1211	72	509	212	149	22	64
	88	183	118	190	1277	809	1244	69	561	190	160	22	71
	43	6	31	-26	8	14	-33	3	-52	22	-11	0	-7
	58	22	72	13	77	54	88	8	53	29	21	10	5
	15	15	40	39	69	39	122	5	104	7	32	10	12

Table 3.5-10: Compensations, charges and net payments for the first scenario of the year 2002 obtained using the basic AP method

	A	B	CH	CZ	D	E	F	H	I	NL	P	SLO	SK
A	127	0	0	1	1	0	0	0	1	0	0	1	0
B	0	191	0	0	1	0	1	0	0	1	0	0	0
CH	0	0	135	0	2	0	5	0	3	0	0	0	0
CZ	4	0	0	178	3	0	0	0	0	0	0	0	0
D	2	0	4	2	1291	0	2	0	2	3	0	0	0
E	0	0	0	0	0	816	5	0	0	0	3	0	0
F	0	2	5	0	4	3	1188	0	9	0	0	0	0
H	0	0	0	0	0	0	0	92	0	0	0	0	1
I	5	0	11	0	2	0	12	0	499	0	0	3	0
NL	0	0	0	0	2	0	0	0	0	234	0	0	0
P	0	0	0	0	0	4	0	0	0	0	146	0	0
SLO	1	0	0	0	0	0	0	0	1	0	0	22	0
SK	0	0	0	1	0	0	0	2	0	0	0	0	80
	139.6	192.2	156.0	181.5	1305.4	824.0	1211.8	93.2	514.5	237.9	149.1	25.7	81.3
	130.3	193.5	144.5	185.4	1305.3	823.8	1211.0	92.7	531.7	236.5	150.8	24.3	82.4
	9.3	-1.3	11.5	-3.9	0.1	0.2	0.8	0.5	-17.2	1.4	-1.7	1.3	-1.1
	12.9	1.7	21.0	3.1	14.9	7.8	24.2	1.7	15.7	3.6	2.8	3.8	1.3
	3.6	2.9	9.6	7.0	14.8	7.6	23.4	1.2	32.8	2.2	4.5	2.4	2.5

Table 3.5-11: Compensations, charges and net payments yielded by AP for the first scenario of the year 2002 when only the cost of the used fraction of each line is considered to compute ITCs

Table 3.5-11 presents the results obtained for the first scenario of the year 2002 taking only into account the cost of the used fraction of the grid. As one can easily anticipate, both compensations and charges obtained with this variant of the method are smaller than those obtained when the cost of the entire grid is allocated. The net compensation for some countries changes notably. For instance, France was an important net payer in the basic AP approach whereas now it is a net receiver. This seems to suggest that the French grid is being used to a larger extent than the neighboring grids.

Results for the With-and-Without Transit method (WWT)

Table 3.5-12 provides the results corresponding to the WWT method. Interpretation of the numbers in the table is exactly the same as that of the table for AP.

	A	B	CH	CZ	D	E	F	H	I	NL	P	SLO	SK
A	99.0	0.2	1.6	-0.5	1.9	-0.4	0.0	0.7	0.0	0.3	-0.1	0.1	0.2
B	1.4	141.1	2.2	-0.6	2.7	-0.7	0.0	1.0	0.0	0.5	-0.1	0.1	0.2
CH	2.0	0.6	97.8	-1.0	4.3	-1.3	0.1	1.6	0.0	0.8	-0.2	0.2	0.4
CZ	2.2	0.5	3.6	202.3	4.2	-0.8	0.0	1.5	0.0	0.8	-0.4	0.2	0.5
D	1.8	0.7	3.5	-1.3	1261.8	-1.9	0.0	1.7	0.0	0.6	-0.4	0.2	0.6
E	1.4	0.3	2.2	-0.6	2.4	849.9	0.0	0.9	0.0	0.6	-0.1	0.2	0.2
F	15.3	3.3	24.2	-6.9	28.4	-6.2	1198.3	10.4	0.1	5.4	-1.3	1.5	2.7
H	0.7	0.2	1.2	-0.4	1.5	-0.4	0.0	70.8	0.0	0.2	-0.1	0.1	0.2
I	11.3	2.3	17.6	-5.0	20.6	-4.0	0.4	7.4	516.1	4.1	-0.9	1.1	2.0
NL	3.6	0.9	5.7	-1.8	7.8	-1.6	0.1	2.6	0.0	227.8	-0.5	0.3	0.9
P	0.5	0.1	0.7	-0.2	0.8	-0.2	0.0	0.3	0.0	0.2	154.9	0.0	0.1
SLO	0.5	0.1	0.8	-0.2	1.0	-0.2	0.0	0.3	0.0	0.2	-0.1	22.6	0.1
SK	0.9	0.2	1.4	-0.4	1.6	-0.3	0.0	0.6	0.0	0.3	0.0	0.1	74.7
	140.6	150.5	162.4	183.3	1339.0	831.8	1199.1	99.8	516.4	241.9	150.7	26.7	83.0
	103.1	147.8	105.2	214.7	1267.4	857.4	1275.2	73.8	573.0	245.9	157.2	25.2	79.1
	37.5	2.7	57.2	-31.4	71.6	-25.6	-76.0	25.9	-56.6	-4.0	-6.5	1.5	3.8
	41.6	9.3	64.6	-19.0	77.2	-18.1	0.8	29.0	0.3	14.0	-4.2	4.1	8.2
	4.1	6.6	7.4	12.4	5.6	7.5	76.8	3.0	56.9	18.1	2.3	2.6	4.4

Table 3.5-12: Annual compensations, charges and net payments because of network usage yielded by the WWT method for the year 2002

Results yielded by the variants of the WWT and APT methods are not shown here since they present the same drawbacks as the original method.

Results for the Average Participations applied to Transits method (APT)

Table 3.5-13 includes the results of compensations and charges, corresponding to the average scenario, that are produced by the APT method.

	A	B	CH	CZ	D	E	F	H	I	NL	P	SLO	SK
A	96.2	0.0	0.0	2.8	5.9	0.0	0.0	1.2	0.0	0.0	0.0	3.4	0.6
B	0.0	134.5	0.0	0.0	0.5	0.0	0.1	0.0	0.0	3.3	0.0	0.0	0.0
CH	0.7	0.0	106.8	0.0	8.5	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0
CZ	12.0	0.0	0.0	158.2	14.4	0.0	0.0	0.8	0.0	0.0	0.0	0.3	2.7
D	9.9	0.4	8.9	3.7	1248.8	0.0	0.1	0.0	0.0	4.2	0.0	0.1	0.0
E	0.0	0.0	0.0	0.0	0.0	806.9	0.4	0.0	0.0	0.0	11.9	0.0	0.0
F	0.0	11.8	17.1	0.0	22.8	5.4	1197.3	0.0	0.1	0.0	0.0	0.0	0.0
H	1.7	0.0	0.0	0.4	0.0	0.0	0.0	74.9	0.0	0.0	0.0	0.0	2.4
I	8.4	0.0	24.3	0.0	4.5	0.0	0.3	0.0	516.1	0.0	0.0	3.5	0.0
NL	0.0	2.3	0.0	0.0	8.8	0.0	0.0	0.0	0.0	231.4	0.0	0.0	0.0
P	0.0	0.0	0.0	0.0	0.0	17.8	0.0	0.0	0.0	0.0	138.7	0.0	0.0
SLO	6.1	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	16.2	0.0
SK	1.4	0.0	0.0	4.0	0.0	0.0	0.0	4.7	0.0	0.0	0.0	0.0	62.9
	136.4	149.0	157.1	169.3	1314.3	830.1	1198.8	81.6	516.2	239.0	150.6	23.6	68.5
	110.1	138.4	116.7	188.3	1276.2	819.2	1254.5	79.4	557.1	242.5	156.4	22.5	73.0
	26.3	10.6	40.4	-19.0	38.1	10.8	-55.8	2.2	-40.8	-3.5	-5.8	1.1	-4.5
	40.3	14.5	50.3	11.1	65.5	23.2	1.5	6.7	0.1	7.6	11.9	7.4	5.7
	14.0	3.9	9.9	30.1	27.4	12.3	57.3	4.5	41.0	11.1	17.8	6.3	10.1

Table 3.5-13: Annual compensations, charges and net payments because of network usage yielded by the APT method for the year 2002

Results for the Provisional Method (PM)

Table 3.5-14 includes the results of compensations and charges, corresponding to the average scenario, that are produced by the PM method.

	A	B	CH	CZ	D	E	F	H	I	NL	P	SLO	SK
A	90.5	0.4	1.4	0.9	2.8	0.7	0.1	0.5	0.0	0.2	0.3	0.3	0.5
B	1.7	133.6	1.9	1.3	3.8	1.0	0.1	0.7	0.0	0.3	0.4	0.4	0.7
CH	2.7	0.9	104.9	2.3	6.2	1.7	0.3	1.1	0.0	0.5	0.7	0.7	1.1
CZ	2.6	0.9	3.2	146.4	6.0	1.3	0.1	1.2	0.0	0.6	0.6	0.6	1.1
D	3.1	1.1	3.4	2.5	1228.9	1.8	0.2	1.2	0.0	0.5	0.8	0.7	1.3
E	1.6	0.5	1.9	1.2	3.6	803.9	0.1	0.7	0.0	0.4	0.4	0.4	0.7
F	17.8	6.1	21.3	13.5	40.8	10.1	1196.8	7.8	0.2	3.7	4.5	4.4	7.3
H	0.9	0.3	1.1	0.8	2.1	0.5	0.1	78.1	0.0	0.2	0.2	0.2	0.4
I	12.9	4.3	15.3	9.4	29.5	7.1	0.9	5.7	516.0	2.7	3.1	3.1	5.2
NL	4.4	1.5	5.3	3.5	10.3	2.5	0.3	1.9	0.1	232.3	1.1	1.1	1.9
P	0.5	0.2	0.6	0.4	1.2	0.3	0.0	0.2	0.0	0.1	138.2	0.1	0.2
SLO	0.6	0.2	0.7	0.4	1.4	0.3	0.0	0.3	0.0	0.1	0.1	14.3	0.2
SK	1.0	0.3	1.2	0.7	2.3	0.6	0.1	0.4	0.0	0.2	0.2	0.2	62.3
	140.6	150.5	162.4	183.3	1339.0	831.8	1199.1	99.8	516.4	241.9	150.7	26.7	83.0
	98.7	146.0	123.2	164.7	1245.6	815.4	1334.1	85.1	615.3	266.2	142.1	18.8	69.7
	41.8	4.4	39.1	18.6	93.3	16.4	-135.0	14.7	-98.9	-24.3	8.7	7.9	13.3
	50.0	16.8	57.5	36.9	110.1	27.9	2.3	21.6	0.4	9.6	12.5	12.4	20.6
	8.2	12.4	18.3	18.3	16.8	11.5	137.3	7.0	99.3	33.9	3.9	4.5	7.3

Table 3.5-14: Annual compensations, charges and net payments because of network usage yielded by the PM method for the year 2002

Comments and conclusions

Table 3.5-15 summarizes the compensations, charges and net payments because of network usage that are produced by the four different examined methods. Some of the conclusions that can be drawn from the numerical results in this section have already been outlined when discussing the results for the year 2001. Those methods that base compensations to countries on the usage made by the transit of the grid of these countries tend to favor heavily transited countries, like Germany or Switzerland, at the expense of large importers or exporters, like France, Italy or the Netherlands. The compensation computed by transit based methods for mainly importing or exporting

countries is smaller than the amount these countries would receive if total cross border flows were taken into account. At the same time, according to the WWT method, large importers and exporters are the most important contributors to the compensation fund, i.e. these countries pay the highest charges. A method which uses all cross border flows, and not only part of them, does not favor one type of countries or another.

Quite understandably, results obtained for the Provisional Method resemble most those of the WWT method. After all, both methods base the compensation due to a country on the impact that the transit through the country has on the usage of its grid. Both WWT and PM compute the transit in the same way. Furthermore, both methods follow the same procedure to allocate the compensation due to a country among all the others. When we compare the results for WWT and PM in more detail we realize that negative net payments are concentrated in less countries in the PM method. Thus, for example, according to PM, France has to pay €135 million and Italy €98.9 million, whereas WWT results in France having to pay €76 million and Italy €56.6 million. Consequently, the net result for some countries is positive in the PM method and negative if the WWT method is used. This is the case of Spain, the Czech Republic or Portugal. These differences are caused by the different ways in which the two methods determine the impact that the transit has on the usage made of the grid of a country. The WWT method results in negative compensations when the removal of the transit reduces the usage made of the grid. On the other hand, compensations obtained in the PM method are always positive. In other words, the PM method does not consider the possibility that the transit through a country may actually reduce the utilization of the country's grid instead of increasing it.

Some results look surprising at first sight. For example, WWT yields negative and large net payments for Spain. By contrast, both AP and APT yield positive net payments for Spain. The outcome of the WWT method seems to be at odds with the fact that the transit through Spain is not particularly important. However, results in this method are difficult to predict, since they critically depend on the internal distribution of flows in the country and how removing the transit affects them. In this case, removal of the transit produces an increase in the flow over many Spanish lines. Consequently, the transit reduces the global electrical usage made of the Spanish grid and Spain has to compensate others for the positive effects that the transit through its grid causes. In short, both AP and APT consider the actual pattern of physical flows. Therefore, results of the AP and APT methods are similar among them and different from those produced by WWT, which forces an artificial flow across the grid of the countries.

Results for the Netherlands are unexpected as well. This country is an important net receiver under AP while it is a net payer under both APT and WWT. The Netherlands imports a great amount of power while the transit through its grid is very small. Thus, the compensation due to the Netherlands when transit based methods are used is significantly smaller than when all cross border flows are considered. The same reasoning applies to the cases of Italy and France.

	A	B	CH	CZ	D	E	F	H	I	NL	P	SLO	SK
AP_CO	49.5	25.3	62.2	22.3	79.7	38.4	73.3	10.2	48.1	38.3	17.2	8.8	8.3
AP_CH	23.5	18.7	39.2	36.7	68.9	31.8	110.1	7.3	81.4	12.4	27.2	9.5	14.8
AP_N	26.0	6.6	22.9	-14.4	10.8	6.6	-36.8	3.0	-33.3	25.8	-10.0	-0.7	-6.5
WWT_CO	41.6	9.3	64.6	-19.0	77.2	-18.1	0.8	29.0	0.3	14.0	-4.2	4.1	8.2
WWT_CH	4.1	6.6	7.4	12.4	5.6	7.5	76.8	3.0	56.9	18.1	2.3	2.6	4.4
WWT_N	37.5	2.7	57.2	-31.4	71.6	-25.6	-76.0	25.9	-56.6	-4.0	-6.5	1.5	3.8
APT_CO	40.3	14.5	50.3	11.1	65.5	23.2	1.5	6.7	0.1	7.6	11.9	7.4	5.7
APT_CH	14.0	3.9	9.9	30.1	27.4	12.3	57.3	4.5	41.0	11.1	17.8	6.3	10.1
APT_N	26.3	10.6	40.4	-19.0	38.1	10.8	-55.8	2.2	-40.8	-3.5	-5.8	1.1	-4.5
PM_CO	50.0	16.8	57.5	36.9	110.1	27.9	2.3	21.6	0.4	9.6	12.5	12.4	20.6
PM_CH	8.2	12.4	18.3	18.3	16.8	11.5	137.3	7.0	99.3	33.9	3.9	4.5	7.3
PM_N	41.8	4.4	39.1	18.6	93.3	16.4	-135.0	14.7	-98.9	-24.3	8.7	7.9	13.3

Table 3.5-15: Summary of the results of inter-TSO payments among countries for the AP, WWT, APT and PM methods

The reason why AP results in large net payments for the Netherlands is that, despite the fact that Holland is a heavy importer, AP concludes that Dutch agents are not using much of the grid of neighboring countries (such as Belgium or Germany). This is probably related to the fact that there are important generation centers in Belgium and, specially, in Germany that are very close to the Dutch border. Therefore, Dutch consumers are probably getting their power from external generators located close to them. Dutch imports do not result in a significant usage of the grids of neighboring countries.

Countries crossed by large transits that are well balanced internally are much better off when WWT or APT are used than when AP is the method chosen to compute ITCs. This is, for instance, the case of Switzerland or Germany.

Locational signals produced by AP

The AP method has been used to obtain nodal transmission charges for each one of the 3965 nodes of the IEM-13 network model. Tariffs have been computed for each one of the 24 scenarios corresponding to the year 2002 that have been provided by ETSO. From these nodal tariffs for each agent and scenario average annual tariffs have been obtained for every agent. Annual tariffs are the result of averaging the nodal tariffs corresponding to the different scenarios. The tariff for each agent and scenario is weighted with the amount of power either produced or consumed by the agent. We have neglected the amount of power produced or consumed by an agent in those scenarios where the nodal tariff for the agent is zero. This happens when the agent is not using the grid according to AP. Figures 18, 19 and 20 represent the nodal tariffs computed for the IEM-13 system. Tariffs for agents whose transmission charge is zero in all the scenarios have not been represented⁶.

⁶ The true average transmission charge is obtained when also the agents with zero transmission charges are included in the computation of the average value.

Figure 3.5-5 shows the average nodal L and G tariffs for the 13 countries in the case example. Tariffs are represented using a bar chart. Numbers are expressed in €/MWh. Figure 3.5-6 and Figure 3.5-7 show the nodal transmission tariffs for all the loads and generators, respectively, in the IEM-13 network. Apart from the final average G and L tariffs, figures represent what we call average internal tariffs. Average internal tariffs within each country are the average tariff that generators and consumers in the country would have to pay if no inter-TSO payments took place. That is to say, average internal tariffs represent the average charge that generators and consumers in a country would face if they had to pay the whole cost of the national grid. It is assumed that generation would pay half of the grid and load the remaining half. The average G and L tariff for each country has been represented using a blue line, whereas the average internal tariff for each country has been represented using a solid pink line. The difference between the average tariff and the average internal tariff in each country represents the average impact of the inter-TSO payment mechanism on the transmission tariffs in the country.

Due to lack of information on the geographical location of each agent, it has not been possible to show a map with the spatial distribution of tariffs in the IEM. However, we can appreciate from figures that the distribution of average tariffs across countries is not uniform at all. Thus, average values for G tariffs range between 1.03 €/MWh and 2.89 €/MWh, with an average value for the IEM of 2.02 €/MWh. In the same way, average values for L tariffs range between 1.35 €/MWh and 2.89 €/MWh, with an average value for the IEM of 2.33 €/MWh. The significant differences that exist among the average transmission charges in different countries are not only due to the fact that different patterns of flows exist in different parts of the grid but also to the fact that the amount of transmission assets per MWh of energy traded varies widely from one country to another. Inter-TSO payments among countries are responsible for the difference between the average transmission charge that would exist in each country if there were no cross border exchanges of electricity (pink line in Figure 3.5-6 and Figure 3.5-7) and the actual average transmission charge computed for the country (blue line). Thus, for example, a positive inter-TSO payment of 26 million euros for Austria means that other countries have to pay Austria 26 million euros because of the difference between the cost of the Austrian grid (represented by the average internal tariffs on G and L) and the cost of the grid actually used by Austrian generators and consumers (represented by the actual average G and L tariffs).

The distribution of transmission charges within every country shows a large dispersion with respect to the average value. According to the analysis carried out for the Spanish system in section 3.4.5.1, results should be coherent in most of the cases. In other words, results should, in general, vary more or less gradually with the location of each agent in the grid. Hence, areas of similar tariffs should exist. However, some outliers, or tariffs of extreme values that differ much from the rest of the tariffs in the area, may exist. The existence of outliers may be caused by the non-linear nature of the AP method, which may result in very different charges for agents of the same type that are located close one to another. The unit participation of an agent in the usage made of

the grid may critically depend on the size of generators and loads in the area. This type of situation should seldom occur, however.

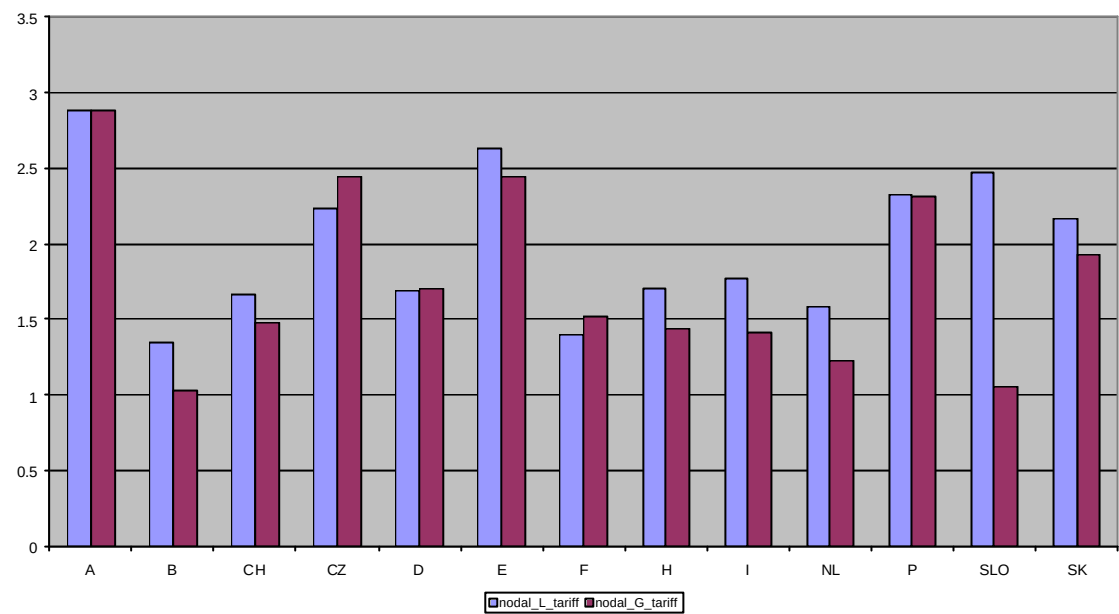


Figure 3.5-5: Average nodal L and G tariffs for the 13 countries in the IEM network

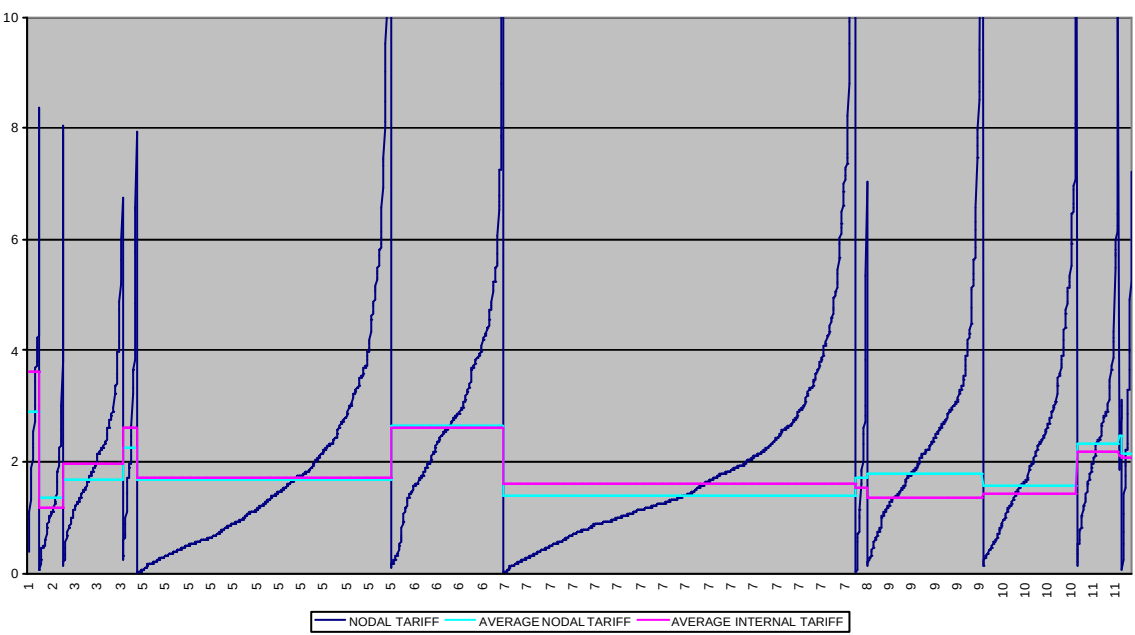


Figure 3.5-6: Nodal tariffs for all the consumers in the IEM-13 network

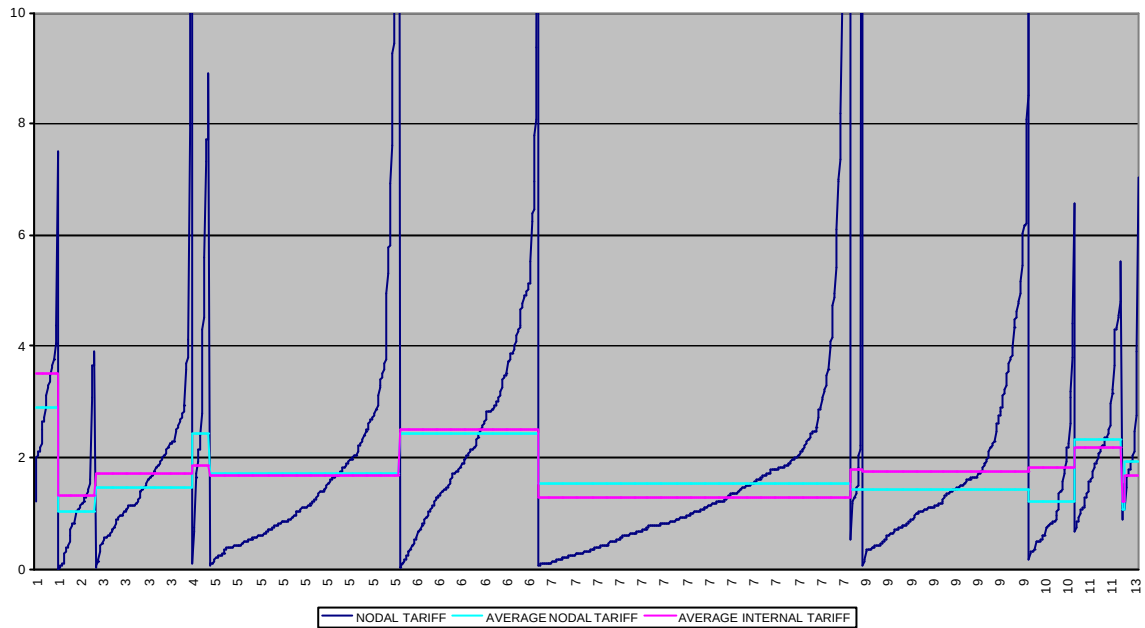


Figure 3.5-7: Nodal tariffs for all the generators in the IEM-13 network.

Focusing on one country like Spain, we can identify areas where nodal tariffs are homogeneous. Thus, for example, L tariffs in the Basque Country region range between 1.3 and 7.3 €/MWh with an average of 2.63 €/MWh and a standard deviation of 1.22 €/MWh. As far as G tariffs are concerned, they range between 0.03 and 2.85 €/MWh with an average of 0.82 €/MWh and a standard deviation of 0.79€/MWh.

A detailed explanation of the existence of outliers has been given in section 3.4.5.1. Outliers above the average may correspond to agents that either produce or consume a small amount of power and are connected to the rest of the grid through a dedicated line or lines. Values well below the average normally correspond to generators in a node where demand is predominant or conversely loads in a node where generation is much more important.

Even within the Basque region, different areas can be distinguished where tariffs are much more similar. One possibility to avoid the dispersion of tariffs is to compute and apply zonal tariffs. Zonal tariffs may be obtained as the weighted average of nodal tariffs over predefined geographical regions. Each nodal tariff in the region would be weighted with the amount of power either consumed or produced by the corresponding agent. Extreme values would have a small influence in the resulting tariff to be applied to every generator or load in the area. Another possibility involves resorting to some method to smooth out individual nodal tariffs by taking into account tariffs for nearby nodes, see the method proposed in section 3.4.5.1.

Smoother tariffs than those presented above may be obtained if the cost of the non-used part of the network is allocated to the demand in the country as a flat charge (variant 2 for AP). Then, a significant fraction of the costs of the network would be socialized.

From these results and those presented in section 3.4.5.1 one can conclude that AP proves to be a very useful aid in computing nodal transmission tariffs with locational content. Average values as well as the volatility of tariffs over the 24 snapshots appear to be normal. An in depth study, once information on the geographical position of each node is available, might help explain the apparently large dispersion of G and L charges within each country. However, section 3.4.5.1 sheds some light over this topic.

Numerical analysis aimed at selecting the number of scenarios to be considered to compute ITCs

Obtaining and processing all the information required to compute ITCs for each hour of the year is not possible at the moment. Therefore, we need a procedure to decide which scenarios to consider when computing ITCs. The next subsection presents the results of application of the methodology proposed in section 3.5.1.3 to select a number of representative scenarios of the operation during a given year. Afterwards, we represent the compensations, charges and net payments obtained for each one of the 24 scenarios provided by ETSO. ETSO provided these scenarios as representative of the operation of the system during the year 2002. We can easily observe that inter-TSO compensations for the different scenarios follow certain patterns.

Selection of a suitable set of scenarios

The application of the method presented in section 3.5.1.3 to select the number and identity of the scenarios to take into account when computing inter-TSO compensations is comprised of two steps: a) Identification of certain patterns in the set of 8760 hourly import/export flows and, b) Checking that the reduced set of scenarios selected according to the information available on import and export flows yield accurate results as far as ITCs are concerned.

Here, we present only the results of the first step of application of the proposed methodology. ETSO has provided for this study a complete set of hourly net export flows for a whole year and nine countries of continental Europe.

A competitive learning algorithm has been employed to identify the underlying patterns that exist in the input data. This process is usually called “clustering” in the specialized literature on data-mining. See, for instance (Hartigan 1975). The competitive learning algorithm applied here is able to “cluster” correlated information in any multi-dimensional space. In this case, the net cross-border flows (exports – imports) for 9 countries define a 9 dimensional space that contains 8760 vectors (one for each hour of the year).

The learning process used here aims to cluster or classify the input data into a set of partitions or clusters such that the intra-cluster variance remains small when compared to the inter-cluster variance. The implemented process is called the PRBFN algorithm, see (Muñoz 1996). This is one out of the several algorithms that might be used. Each algorithm uses different criteria in order to cluster the data. Later in this section, we

apply a second competitive learning algorithm and compare its results to those produced by the PRBFN algorithm.

Discussion of the results

Application of the PRBFN algorithm yields two main results: the reference vector or representative element of each cluster and a label for each vector that identifies the cluster that the vector belongs to. The number of clusters is an input to the model that the user must provide. This means that grouping the data in different numbers of clusters may be necessary before reasonable results are obtained. Figure 3.5-8 and Figure 3.5-9 show the results corresponding to the classification of the set of vectors of cross-border flows into 12 clusters. Now we explain each one of the two figures graphically:

- Figure 3.5-8 shows, for each cluster, the reference vector or representative element of the cluster in blue and each one of the elements of the cluster (samples of net import flows) in yellow. The reference vector does not need to coincide with any of the samples. Since the number of samples is very large, only a yellow shape can be distinguished. In this case, the “x” axis of the 12 graphics represents all the vector components (9 net cross-border flows) and the “y” axis represents the flow in GW. This figure provides a means of easily spotting how homogeneous the clusters are and how well each reference vector “represents” its cluster members.
- Figure 3.5-9 is a color map that shows the cluster each vector belongs to. A color code is used to represent clusters. The “x” axis represents the 24 hours in a day and the “y” axis represents the 365 days of a year. Therefore, each pair (x,y) corresponds to the sample for a certain hour. This color map allows us to identify some patterns by visual inspection. This will become clear later, once the results of the analysis are explained in detail.

The number of vectors within each cluster is indicated above each graph in Figure 3.5-8 (labelled with an M). Clusters 2, 5 and 8 are very similar. The reference vectors of the three clusters follow similar patterns and the dispersion of the samples with respect to these reference vectors is not very large. The same applies to clusters 3 and 6. However, only clusters 4, 7 and 6 are really homogeneous (the yellow area is small). The remaining clusters present a higher dispersion.

We can draw a first conclusion: patterns of net cross-border flows are surprisingly similar. The representative elements of the different clusters have all a similar shape and the dispersion of the samples within each cluster is not very high.

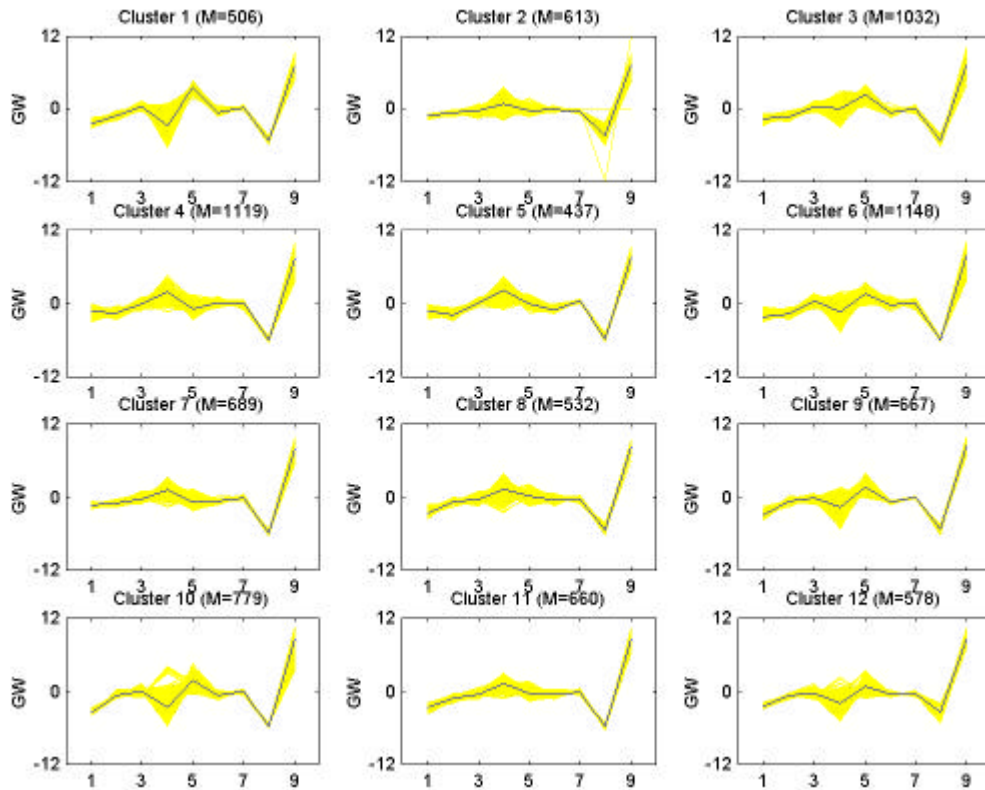


Figure 3.5-8: Vectors corresponding to the representative and all the samples within each of the 12 clusters of net cross-border flows

Figure 3.5-9 shows the possible correlation between the time of the year and the identity of the cluster the corresponding sample belongs to. As mentioned before, the graph shows by means of a colour code the membership of each hourly vector (net border flows of 9 countries).

The obtained results are quite interesting because the colour map confirms that an important correlation exists between the type of flow exchange pattern and the hour of the day, the day of the week and the season when it happens. For example, the first hours of each day (from 00 to 07 hours) present a pattern which is clearly different from that for the rest of the hours. Also, the final hours of each day (around 23 – 24 hours) exhibit a different pattern which is similar in some cases to the pattern of the first hours. Some seasonal patterns can be detected as well. For example, 5 different patterns exist throughout the year for the first hours of the day: July— October , October-January, January-April, April-May and May-July. Slightly different patterns can be identified for the later hours of the day. For example, weekend hours are clearly different from the rest. Defining a different number of clusters changes these results but only slightly.

The cluster analysis presented here provides some evidence that a set of 12 scenarios may be enough to accurately evaluate cost compensations among countries due to energy exchanges. However, results of this analysis will be applicable to the problem

of computing ITCs only if hourly energy exchanges between countries are highly correlated to real network flows.

In the end, this analysis is aimed at defining a minimum set of scenarios representative of the system operation over the whole year. We have selected those scenarios for which the vector of cross-border flows is closest to the corresponding reference vector. Table 3.5-16 presents the 12 scenarios that have been finally selected.

Other selections of scenarios are possible. As Figure 3.5-9 highlights, the identified patterns of cross-border flows are valid for many hours and days. This is due to the fact that net imports and exports happen to be more stable than the generation and demand within each country. For example, import flows through an interconnection may remain at its maximum during most of the time thus not following the evolution of the total load within the importing or exporting country. Using specific patterns of net border flows for weekends is not necessary, according to the obtained results. Cross-border flows (net aggregated value) over the weekend tend to be quite similar to those during the early hours of weekdays.

Cluster ID	Day	Month	Hour
1	4th Tuesday	May	16
2	3rd Thursday	Sep	6
3	2nd Friday	May	24
4	1st Sunday	Feb	14
5	2nd Tuesday	Apr	2
6	2nd Tuesday	Feb	12
7	2nd Monday	Nov	4
8	4th Sunday	Sep	21
9	4th Monday	Jul	14
10	3rd Thursday	Nov	11
11	4th Saturday	Dec	19
12	3rd Saturday	Aug	17

Table 3.5-16: A plausible selection of 12 scenarios representative of the system operation over a whole year

Complementary study regarding the choice of the number of scenarios to be used

Since 1-1-2003, ETSO has been collecting hourly snapshots for most IEM countries. Despite the availability of these data, computing inter-TSO compensations for every hour of the year remains an objective far too ambitious. Therefore, based on a statistical study, ETSO determined that 24 scenarios carefully chosen corresponding to the year 2002 could be enough in order to produce reasonable accurate annual values of compensations, charges and net payments.

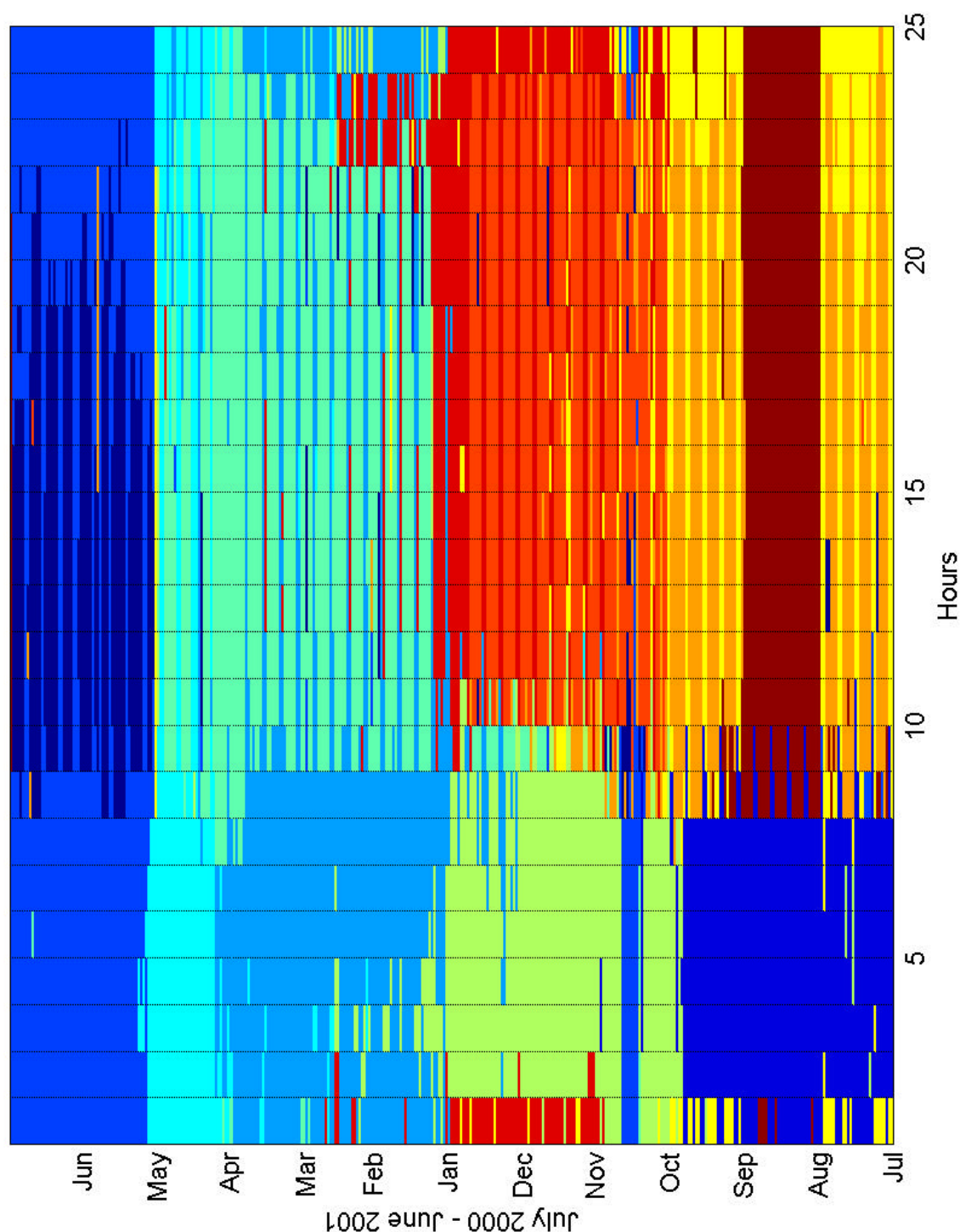


Figure 3.5-9: Representation of the samples of 12 clusters of net cross-border flows among countries according to the time of the year when they occur

Based on the compensations, payments and net ITCs that result from separately applying the AP, WWT and APT methods to each of the 24 snapshots of the 13-IEM system that were provided by ETSO, we have carried out an analysis aimed at drawing some useful conclusions on the number of scenarios required to compute annual inter-TSO payments. In any case, these results are not enough to carry out a statistical evaluation so as to make a reliable prediction of the number of required scenarios.

Evaluation of the results

Each one of the series in Figure 3.5-10, Figure 3.5-11 and Figure 3.5-12 represents the geographical distribution of the net inter-TSO payments for a certain scenario. Figure 3.5-10 corresponds to AP, Figure 3.5-11 to WWT and Figure 3.5-12 to APT. Figures show that net inter-TSO payments for AP or APT are much more stable across the different scenarios than those produced by WWT. For instance, inter-TSO payments for Germany range from –18.4 to 33 million euros in AP, depending on the scenario, whereas values for WWT range between –53.7 and 170.2 million euros. Net compensations in APT range from 4.9 to 63.8 million euros. Similar results hold for other countries.

However, results for different methods and scenarios exhibit broad similarities. Net payments to France and Italy are always negative, regardless of the method used to compute them, whereas those for Switzerland are mostly positive and those for the Czech Republic are mostly negative. Despite varying widely from one method to another, net compensations due to Slovenia, Hungary or Belgium do not change significantly across scenarios. All this seems to suggest, although very preliminarily, that a number of scenarios not far from 24 may be enough to compute accurate values of annual inter-TSO payments.

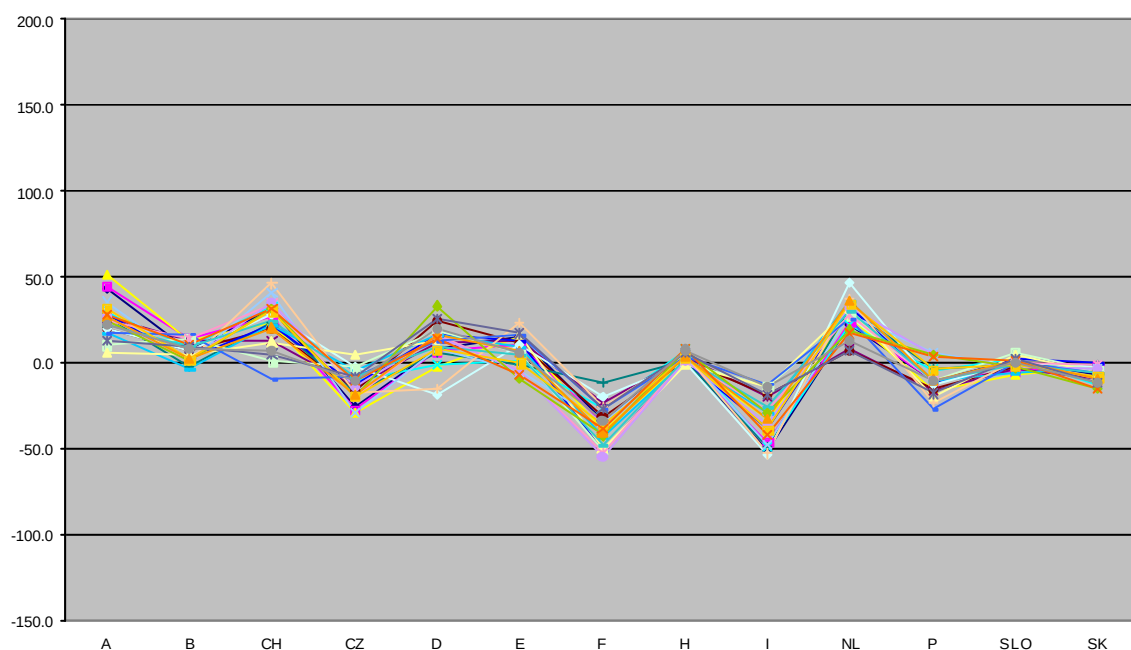


Figure 3.5-10: Geographical distribution of the net inter-TSO payments in 2002 for AP

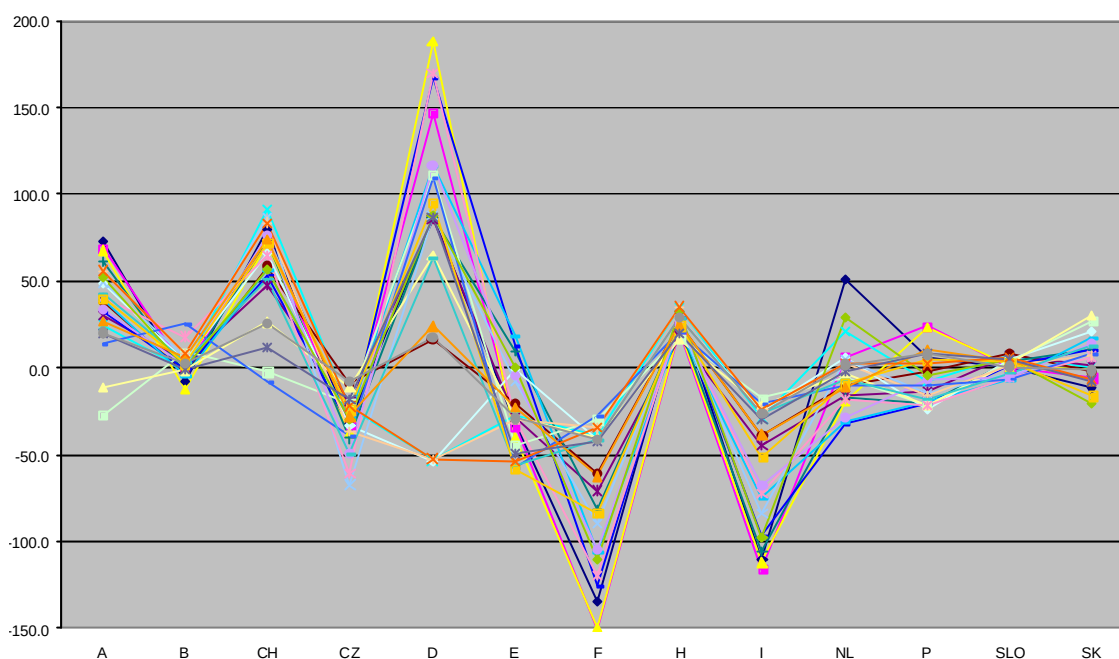


Figure 3.5-11: Geographical distribution of the net inter-TSO payments in 2002 for WWT

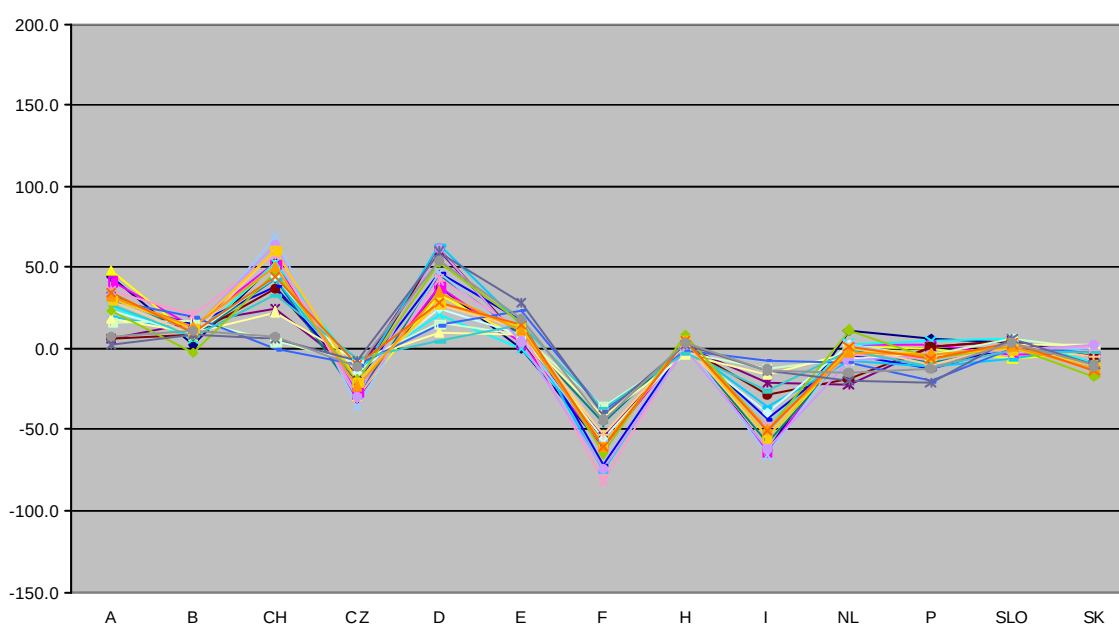


Figure 3.5-12: Geographical distribution of the net inter-TSO payments in 2002 for APT

Figure 3.5-13, Figure 3.5-14 and Figure 3.5-15 show the geographic distribution of the compensations to be received by the different countries. Values in each series correspond to a different scenario. Compensations show more clearly defined patterns than net inter-TSO payments, especially those for AP and APT. This may be related to the fact that net inter-TSO payments result from the difference between compensations and charges.

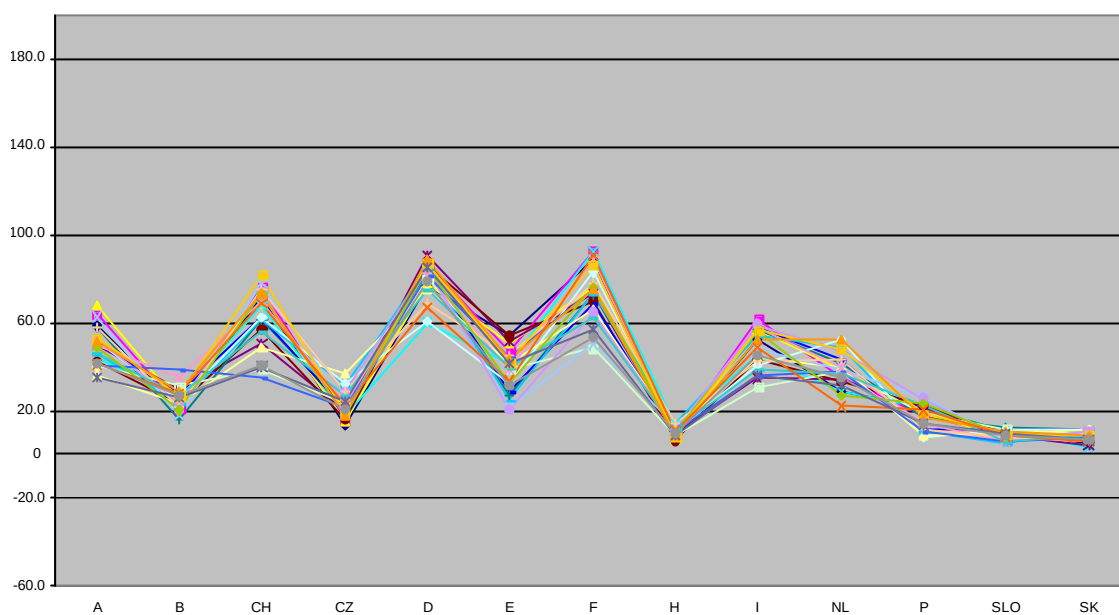


Figure 3.5-13: Geographical distribution of compensations to countries in 2002 for the AP method

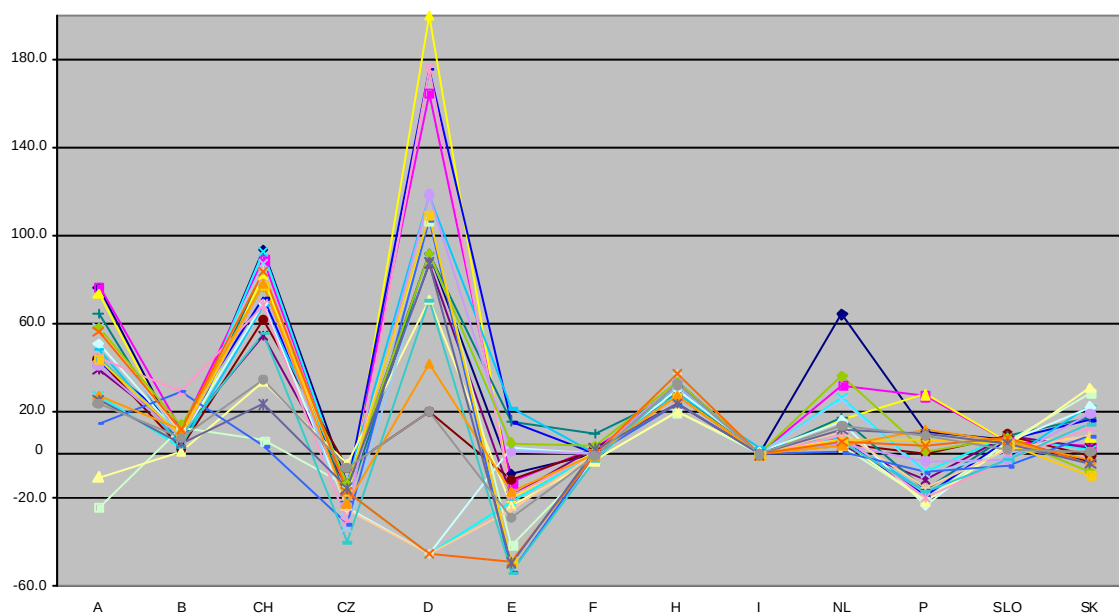


Figure 3.5-14: Geographical distribution of compensations to countries in 2002 for the WWT method

Finally, Figure 3.5-16, Figure 3.5-17 and Figure 3.5-18 represent the geographical distribution of charges. Charges produced by AP show again a clear pattern. Charges paid in the WWT method are quite low for all countries except for France, Italy and the Netherlands. These countries have the largest volumes of net imports or exports. Charges obtained for the APT method resemble those for the AP method though they are smaller.

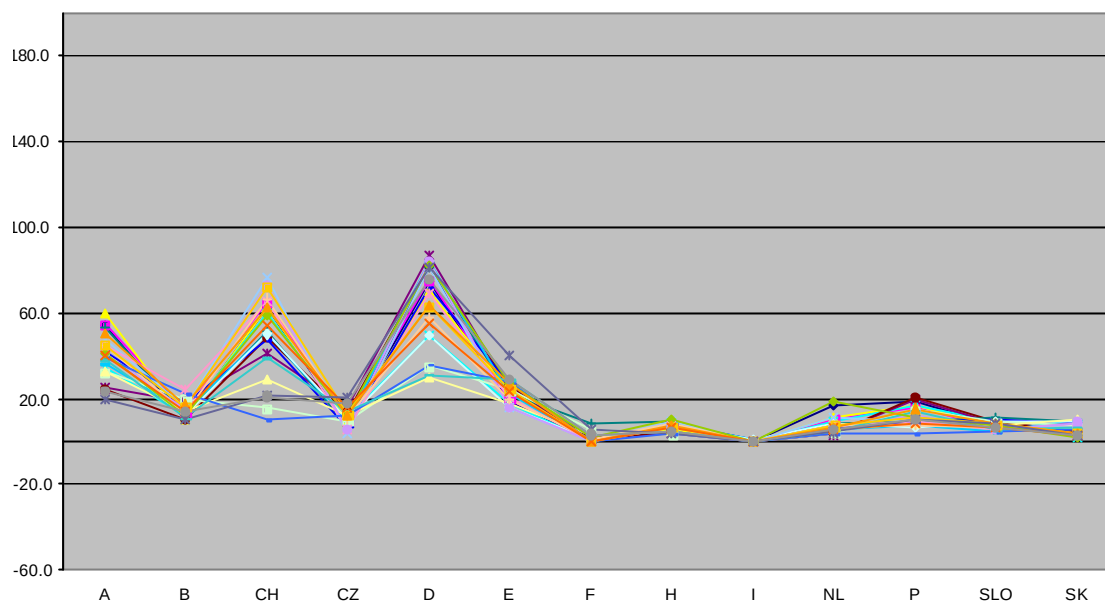


Figure 3.5-15: Geographical distribution of compensations to countries in 2002 for the APT method

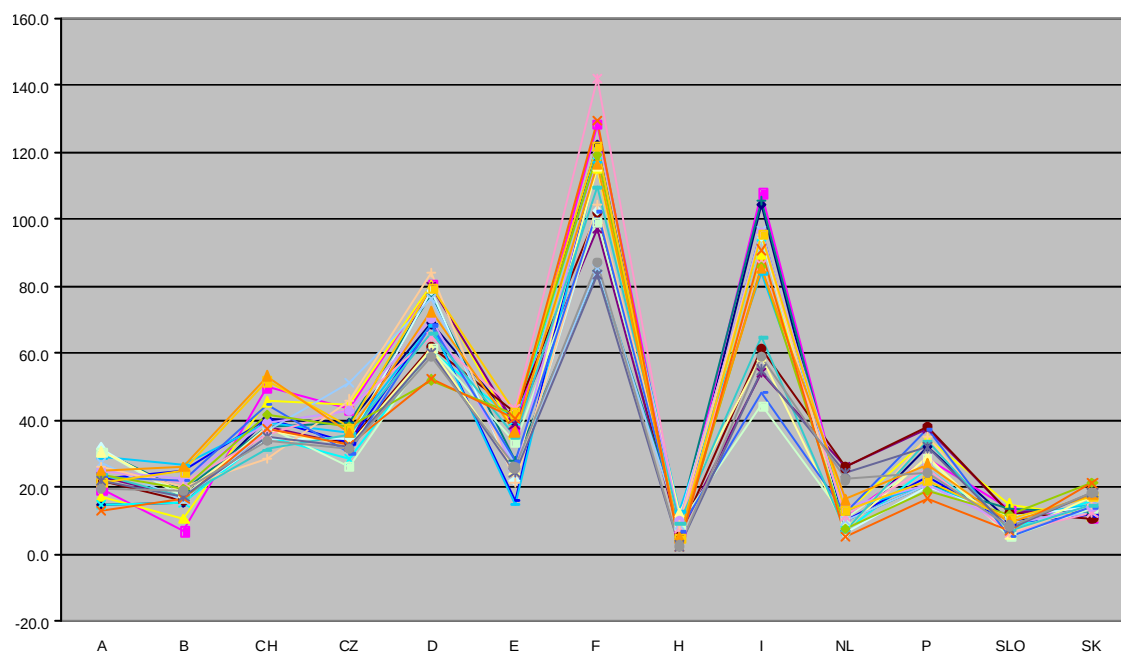


Figure 3.5-16: Geographical distribution of charges to be levied on countries in 2002 for the AP method

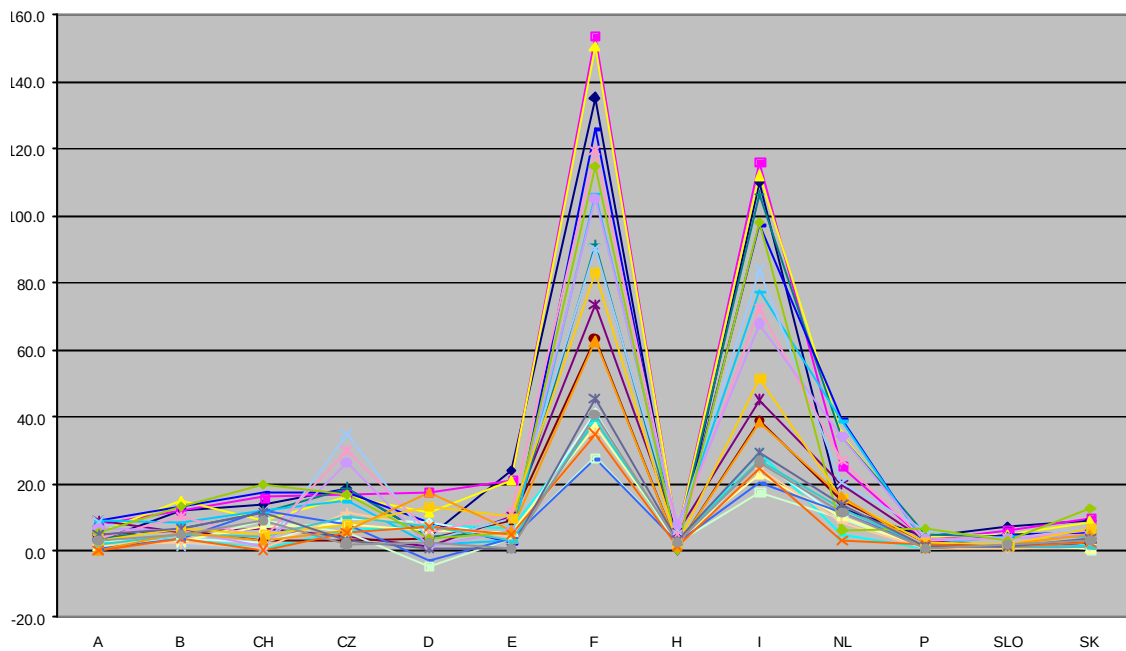


Figure 3.5-17: Geographical distribution of charges to be levied on countries in 2002 for the WWT method

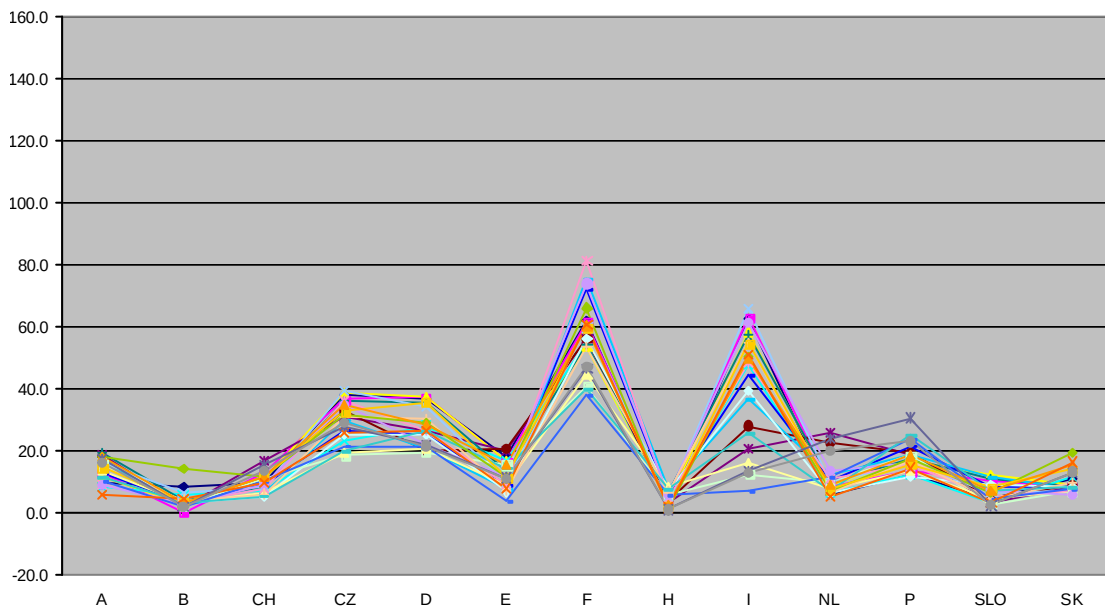


Figure 3.5-18: Geographical distribution of charges to be levied on countries in 2002 for the APT method

3.5.1.6 Comparative analysis of the preferred methods

This section compares the three methods that have been selected in the first instance for computing inter-TSO payments in the European system, namely Average Participations (AP), Average Participations applied to Transits (APT) and the With and

Without Transits method (WWT). Additionally, we will also assess the Marginal Participations method (MP), which has been proposed in (Consentec/Frontier 2005) as the most efficient method to compute ITCs from a conceptual point of view. The four methods make use of the actual physical flows that exist in the transmission network. In the following paragraphs we evaluate these methods according to different criteria and eventually chose one.

The broader context.

Application of this criterion assesses the consistency of a network cost allocation method with the underlying paradigm of the IEM contained in the EC Regulation and the methodological decisions already made within the context of the Florence Forum.

WWT computes the impact of a transaction between two countries on the flows in a third one. Therefore, it follows the “wheeling” philosophy, which is in contraposition to the prevailing principle in multinational markets, the “single system paradigm”. AP and MP compute the usage of the system network by each agent. Political borders are ignored and results obtained for both methods are completely consistent with the “single system paradigm”. Inter-TSO compensations result from the aggregation of nodal charges at country level. Similarly to what happens in the WWT method, APT produces results that are dependent on the existing political borders thus not complying with the paradigm.

Now we examine the coherence of each method with other aspects of transmission regulation, namely investment in new infrastructures, locational signals for operation and investment and computation of regional transmission tariffs.

Investments in new infrastructures

When defining the “without transit” situation, the WWT method forces changes in the flows over many lines within each country. Many times, these flows are not much related to import and export flows in the country. The fact that the WWT method may result in negative compensations for the external usage of the grid of a country implies that a country may have to compensate others for building a line in its territory, which seems very difficult to be accepted by the country. Additionally, after having determined the fraction of the cost of a new line within a country to be paid by external agents, WWT resorts to an arbitrary rule to allocate this cost among the remaining countries.

The AP and MP methods directly allocate the cost of any line to grid users both external to the country and internal. These methods automatically determine which countries are using a new line and to what extent. However, MP may yield both positive and negative participations in the use of new lines whereas those produced by AP are always positive. Therefore, AP is, in principle, well suited to the task of assigning the cost of new lines in a regional market, whereas the allocation of the cost

of new lines resulting from MP may be contested by those countries that have to pay them.

APT solves two major problems of WWT. Thus it always results in positive participations in the recovery of the cost of new lines and the method itself determines how much other countries must contribute to the cost of new lines. APT provides results for the allocation of the cost of individual network investments that are similar to those obtained for AP, though the external usage of new lines typically is smaller than under AP.

Locational signals for operation and investment

WWT is not capable of sending locational signals to agents since it only computes compensations among countries. As for AP and MP, they assign the cost of the entire network of the IEM to all market agents. AP and MP produce network charges with locational content for agents. APT only allocates the responsibility for causing transit flows to individual generators and loads. It does not track all flows in the network and, therefore, it does not allocate the cost of the whole network. All in all, APT does not result in nodal tariffs or any other kind of locational signal related to the entire transmission network.

Computation of regional transmission tariffs

Only AP and MP result in regional transmission tariffs and can therefore be of some help in the process of tariff harmonization.

Justification of the method

Economic and Technical soundness

All the methods rely on the assumption that electrical usage is a suitable proxy to economic benefit. However, the underlying assumptions behind each of the methods differ from one to another. Thus:

- **WWT** is based on the idea that the external use of the grid of a country can be measured as the incremental impact of the transit through the country on the grid flows. This requires to define “transit” & to allocate the “transit” among the several cross-border lines, therefore creating a new “without transit” situation to be compared to the actual one. The responsibility of each country k in the external use of the network of any other country m is assigned pro rata of the net flow into or out of the country k , regardless of geographical proximity or any other consideration.
- Accepting the technical impossibility to assign line flows unambiguously to sinks or sources, **AP** assumes that this responsibility can be simply obtained by

tracking each flow upstream & downstream following the same paths & branching proportions that exist in the actual flows provided by ETSO. By construction, the use of each line is assigned 50/50 to demand & generation, although this proportion may be easily modified afterwards if the user of the algorithm has a reason to do so.

- In **MP**, it is assumed that any power inflow or outflow in any node of the system has as source or sink the same (slack) node, which has to be chosen somehow before the method can be applied. Therefore, the responsibility of a given inflow (outflow) in the use of each line of the network is obtained by injecting (retrieving) power at the considered node & retrieving (injecting) it at the slack node, & examining which is the impact on each line of the entire network. Obviously the result critically depends on the choice of the slack node. The choice of the slack node also determines the split $X / 100-X$ between the total charges to all demands & to all generations.
- **APT** is based on the idea that the external usage of the grid of a country, C, is only related to the use that the transit through the country makes of this grid. Thus, using the same rule as AP, transit flows are tracked inside the grid of C in order to compute the usage that they are making of this grid. Transit flows are also tracked outside the country in order to assign responsibilities for the external use of country C's grid.

The WWT method is unable to compute nodal tariffs. Regarding AP, provided the rationality of the tracking rule it employs is accepted, this method has all the good economic properties, including the additivity of charges. AP and MP produce locational signals at agent level. However, the underlying assumption behind MP does not seem to be acceptable. Finally, APT shares some basic assumptions with both WWT and AP, which make it even more arbitrary than AP or WWT. Additionally, APT does not produce locational signals.

WWT relies on the assumption that only those cross-border flows that are neither injected nor withdrawn by the internal agents of a country should give raise to compensations for the external usage of the grid of the country. However, contrary to what the method states, it may be argued that exporting countries are using the grid of the importing ones to deliver the power they produce to the nodes where it is consumed. Analogously, importing countries would be using the grid of exporting countries to obtain the power they consume.

Once we have determined the volume of the transit through a country, we need to allocate this transit among the border lines of the country. Deciding which cross-border flows are part of the transit and which ones are not is always going to be arbitrary. The “without transit” situation artificially created by the method may be illogical or it may lack any physical sense. What is more, results will always depend on the political borders considered. This is by no means a desirable property. Several ways exist of

computing the impact of the transit on the usage made of the grid of a country. All these alternative methods produce different results. Lastly, an arbitrary rule is employed to allocate the compensation due to a country among the remaining ones. The method itself does not provide any clue on how to do this.

The AP method relies on the branching rule (or proportionality rule) presented above to track the flows in the grid. Application of this branching rule yields the true average use that each agent makes of the grid, according to the physics laws that rule the functioning of power systems, see section 3.4.3.2. Provided this branching rule is accepted, AP addresses most of the issues involved in the regulation of transmission in a coherent and very sensible manner. These issues are inter-TSO payments, tariff harmonization or locational signals, among others. The method assigns responsibilities in network cost based on the average electrical usage of the grid by each agent, which may be the most appropriate way to allocate sunk costs, i.e. costs that have already been incurred. On the other hand, marginal charges may be preferable to allocate the cost of new lines since they are stronger incremental signals. Results obtained from AP under very different circumstances seem to be meaningful and very robust. Similarly to what happens in WWT, AP has several degrees of freedom in its implementation and the results obtained depend on the way the method is applied.

Quite logically, APT shares some weak points both with AP and with WWT. Criticisms of the concept of transit are applicable here as well. APT arbitrarily allocates the transit among the border lines. This method resorts to the same rule as AP to track those cross-border flows deemed to be part of the transit. However, unlike WWT, APT does not produce an artificial “without transit” situation. In addition, this method determines which country is responsible for each of the cross-border flows that comprise the transit. Results obtained are in accordance with the rationale of the method and they make engineering and economic sense.

If the marginal generator in the dispatch were the same for the whole region, applying MP with a slack bus located in the same node as this marginal generator would yield the true marginal use by each agent of the transmission grid in the region, see 3.4.3.2. However, in a regional market like the European one, the marginal generator varies from one area to another. This is due to the fact that the inter-connectors between these areas are frequently congested. Additionally, transmission losses incurred when transporting power between different areas tend to be very large. Thus, the marginal generators in different areas are different and applying MP does not make sense. If the assumption behind the MP method were accepted, this method could be used to compute locational signals to generators and loads, allocate the cost of new lines or harmonize transmission tariffs, among other things.

Legal soundness

WWT does not comply with the existing regulation (as we interpret it) because it does not take into account all cross-border flows when computing inter-TSO compensations

but only those flows that are considered to be part of the transit. Additionally, it does not track the flows in the grid in order to decide which agent (country) is responsible for the external usage of the grid of a country. Instead, WWT computes the contribution of each country to the global compensation fund based on the size of the net export from or import into the country. Compensations to all countries are drawn from this global compensation fund.

AP is coherent with the regulation. Apparently, this method does not consider the benefits that a country may obtain from the existence of a transit through the country's grid. Transit flows may reduce the total electrical usage of the grid of the transited country. In this case, applying AP results in small participations of other countries in the usage of the grid of the considered one. Note, however, that the strict application of the "benefits" that are caused by external flows results in bizarre payments, such as having to compensate other countries when an internal line is built.

Unlike the WWT method, APT does search for the origin and end of transit flows thus avoiding inconsistency with the EU regulation in this regard. However, APT only tracks transit flows instead of all cross-border flows (as the AP method does). APT cannot directly show the potential benefits to the transited country brought about by transits because in APT flows only appear to create positive losses and positive usage of infrastructures. This is the same criticism raised against AP (which seems rather an advantage of AP than a criticism). Both AP and APT deal with this problem although not in an explicit way as WWT does.

The MP method is coherent with the European regulation on the subject. It takes into account all cross-border flows, results at agent level do not depend on the location of political borders, is able to provide locational signals and, according to the rationale of the method, determines the benefits that a network incurs as a result of hosting cross-border flows.

Implementation issues

Issues such as the amount and nature of data required to apply each method, the complexity of the method, the assumptions in its implementation and the experience that exists of its utilization may prove to be critical when trying to get a method accepted. A discussion of all these topics follows.

Data requirements

AP and APT need the same input data, namely the flows over the lines of the system and the amount of generation and demand at each node for every scenario that is used to compute annual compensations and charges. Additionally, information on the cost of the lines is necessary to express compensations in monetary terms. Compensations to countries can be computed separately for each one of the countries in the system. On the other hand, determining how much a country must contribute to the compensation

that is due to others requires having available input data corresponding to all the countries or, at least, part of them.

Implementing the WWT method requires having available all the information required to run a load flow for each country. Each country may compute on a separate basis how much each of the remaining ones must pay it. In order to determine the cross-border compensations that other countries owe it, each country only needs to know the aggregate value of net imports and exports of all the remaining countries in the system. Knowing how much the country needs to compensate others for requires being able to run a load flow for each one of the remaining countries, however.

Application of the MP method requires running a load flow at regional level. Therefore, in order to apply this method to compute inter-TSO compensations within the IEM, we need all the information that is relevant to run a load flow comprising all the countries within this market. Obviously, this includes an electricity model of the transmission grid of the region. Compensations for countries cannot be computed separately. Therefore, we need to integrate the information corresponding to all the different countries into a single model. As a consequence, the MP method is the one whose data requirements are largest.

Availability and complexity of procedures or commonly agreed definitions required to apply the method. Experience with its utilization

AP makes use of a very simple rule in order to track the flows in the system. Experience of its utilization exists since it has been applied in several countries (New Zealand and Central America and proposed in Poland, among others). It has been used, together with other methods, to compute the total regional transmission charge to be paid by each country in the Central American Regional Electricity Market. Results are very sensitive to the definition of the Horizontal Network, so developing a satisfactory method to identify the HN is of great importance.

APT follows the same rules as AP to track the flows in the grid. This method has never been used in reality since it was recently developed by the research team the author belongs to. As in the AP method, applying APT involves having available a reasonable method to define the HN. Furthermore, it relies on the concept of transit. Therefore, APT needs of a precise definition of what is the transit through a country and how to compute it.

Applying WWT requires defining the transit through the grid of each country. The method employs a load flow model to compute the electrical usage of each system grid in the “without transit” situation. Load flow models are commonly used in the electricity industry but most people involved in the decision making process leading to the adoption of a long term ITC mechanism are not familiar with them. Provided the HN defined is large enough, WWT is not as sensitive to the precise definition of the HN as AP or APT.

Application of the MP method requires running a large number of load flows (as many as nodes where at least one agent is located). Load flows are well known by power system engineers. However, they are more difficult to understand by people that do not have a technical profile. Results produced by the method could be very sensible to the definition of the HN, depending on the identity of the node taken as the slack bus. This method has been applied in several countries (Argentina, Chile,..) but has never been used to compute inter-TSO compensations.

Difficulty involved in understanding the method. Rationality of the results that it produces

The concept behind the application of the WWT method is very simple and intuitive. We only need to compare the system flows in two different situations: the situation that really occurs and the one that results from removing the transit through the country. However, despite being a common practice among power engineers, running a load flow is quite complex. Additionally, the “without transit” situation may have no physical meaning or it may be impossible. Compensations obtained in the WWT method are larger than those produced by other methods and they are more volatile as well. Results are difficult to accept when, according to the method, the transit is deemed to reduce the electrical usage of the grid of a country. This country should compensate others for the external usage of its grid. Negative participations in electrical usage are especially difficult to accept when allocating the cost of a new line that is built within the country.

The AP method is easy to explain and understand. However, the tracing rule is difficult to justify from a theoretical point of view. Some degrees of freedom exist in the application of the method. The method itself is known to very few people (only those working on this specific topic). Results for AP make sense and they are quite stable. They are also less volatile and smaller than those for the WWT method. The method is well suited to allocating the cost of new investments.

Understanding APT requires understanding both AP and WWT, since it combines some of the features of both methods. In order to apply the method, we have to define the transit through the grid of each country, though running a load flow is not necessary. Unlike the WWT method, APT does not require changing the flows in the system. It relies on the same arbitrary, albeit reasonable, tracing rules as AP.

The basic idea behind the MP method is easy to understand. However, as mentioned above, computing results with this method requires running a load for the whole region. Running load-flows is common practice among power-system engineers but people that do not have a technical background are unable to understand the equations of a load-flow. Results produced by the method are highly dependent on the location of the slack bus. When applied to allocate the cost of new lines, some TSOs may oppose paying others for the use the latter make of these lines (TSOs that are deemed to make a negative use of transmission lines must be compensated for this use).

Application of AP, MP and WWT to a stylized representation of the EU system

Here, we want to assess the properties of the numerical results provided by the AP, MP and WWT methods in the context of their application to a system that is a stylized representation of the IEM. This makes sense, because drawing conclusions from a simple case example is much easier than doing this from the results obtained when applying the methods to a real scale system.

Results obtained from the application of the AP and MP methods to this same system were shown in section 3.4.2.7. Now, we only focus on the inter-TSO compensations produced by the three methods and on some specific features of the WWT method.

The considered system is the one represented in Figure 3.5-19. When applying the WWT method, we have assumed that only 50% of a cross-border line is located within each country. Thus, when computing the compensation to be received by a country for the external use of a cross-border line, only 50% of the flow on the line has been allocated. Compensations produced by the WWT method when applied to this system are shown in Table 3.5-17. The meaning of numbers in this table is the same as those of Table 3.4-3, Table 3.4-6 and Table 3.4-9 in section 3.4.2.7. The rows labeled with the name of a country show the use that the corresponding country makes of the grid of all the countries in the system. The row labeled 'Use by others' provides the use of the grid of each country by others. The row labeled 'Use of others' provides the use that each country makes of others' grids. Finally, the row labeled 'Net use' provides, for each country, the difference between the two previous results. Numbers in the table are expressed in '000 MWkm. Due to the fact that only P, E and D are crossed by transits, the remaining countries do not receive compensations.

If country P is split into two new systems P1 and P2, then transit through systems P1 and P2 is zero and they are not entitled to any compensation. In addition, total net flow produced by P1 and P2 is much higher than that produced by P. Consequently, contribution factors change as well. As a result of this, compensations among countries change substantially, as Table 3.5-18 shows.

From the numbers in Table 3.5-17 and Table 3.5-18, one can draw the following conclusions: 1) the electrical use made of the grid of a country by the remaining ones changes when this country is split into 2 areas. The total use made of the regional grid by consumers and generators in this country changes as well; 2) some countries (like E and I in the example) must compensate others for the use the latter make of their grids and 3) countries that are not balanced pay compensations to all the transited countries in the region (no matter how far they are). These results suggest that the WWT method is not well suited to compute compensation among countries within the IEM.

Table 3.5-19 and Table 3.5-20 show again the compensations among countries previously obtained for the AP and MP methods, respectively. One can immediately see that results obtained for different methods are significantly different, as Table 3.5-21 confirms. Thus, choosing one method or another is by no means irrelevant.

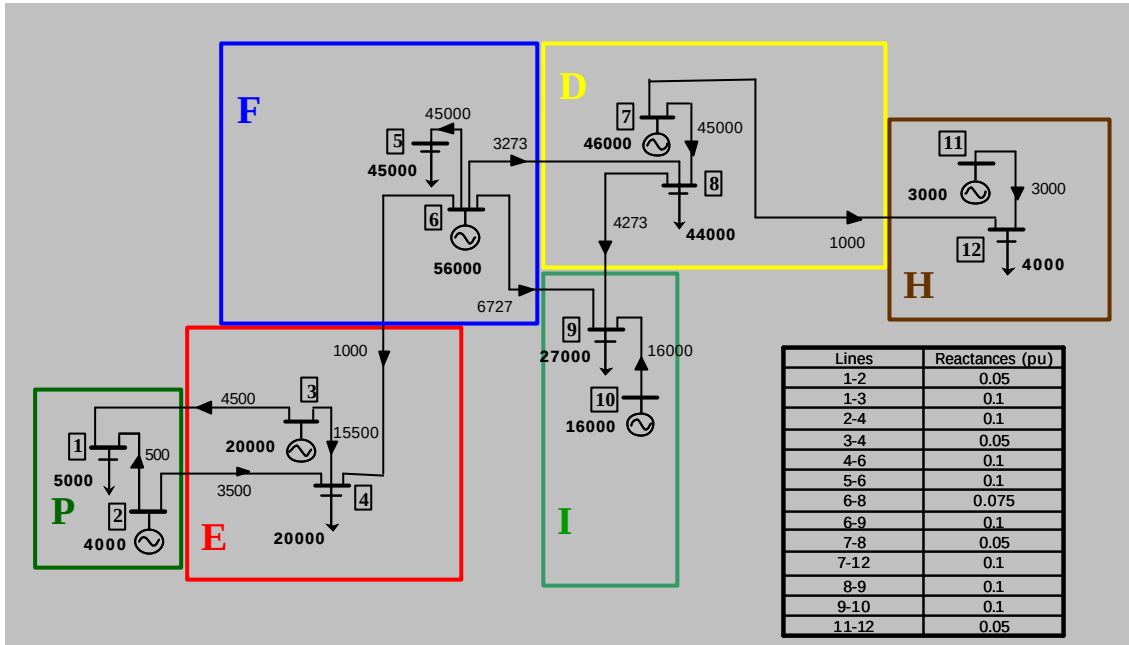


Figure 3.5-19: stylized representation of the IEM used in the case example.

	P	E	F	D	I	H
P	1285	45	0	51	0	0
E	0	5000	0	0	0	0
F	368	473	23364	536	0	0
D	70	90	0	12007	0	0
I	368	473	0	536	10750	0
H	35	45	0	51	0	1000
Use by others	840	1125	0	1175	0	0
Use of others	96	0	1376	160	1376	131
Net use	744	1125	-1376	1015	-1376	-131

Table 3.5-17: inter-TSO compensations produced by the WWT method when applied to the stylized representation of the IEM system.

	<i>P1</i>	<i>P2</i>	<i>E</i>	<i>F</i>	<i>D</i>	<i>I</i>	<i>H</i>
<i>P1</i>	1188	0	169	0	192	0	0
<i>P2</i>	0	938	135	0	153	0	0
<i>E</i>	0	0	5000	0	0	0	0
<i>F</i>	0	0	360	23364	409	0	0
<i>D</i>	0	0	68	0	11982	0	0
<i>I</i>	0	0	360	0	409	10750	0
<i>H</i>	0	0	34	0	38	0	1000
<i>Use by others</i>	0	0	1125	0	1200	0	0
<i>Use of others</i>	360	288	0	769	68	769	72
<i>Net use</i>	-360	-288	1125	-769	1133	-769	-72

Table 3.5-18: inter-TSO compensations produced by the WWT method when applied to the stylized representation of the IEM system and P is split into two new systems P1 and P2.

	<i>P</i>	<i>E</i>	<i>F</i>	<i>D</i>	<i>I</i>	<i>H</i>
<i>P</i>	1125	1000	0	0	0	0
<i>E</i>	1000	5000	125	0	0	0
<i>F</i>	0	125	23773	343	877	0
<i>D</i>	0	0	280	11655	498	125
<i>I</i>	0	0	868	1059	9375	0
<i>H</i>	0	0	0	125	0	875
<i>Use by others</i>	1000	1125	1273	1527	1375	125
<i>Use of others</i>	1000	1125	1345	903	1927	125
<i>Net use</i>	0	0	-72	625	-552	0

Table 3.5-19: Compensations among countries produced by the AP method corresponding to the usage that external agents make of the grid in each country

	<i>P</i>	<i>E</i>	<i>F</i>	<i>D</i>	<i>I</i>	<i>H</i>
<i>P</i>	1292	292	45	-68	0	0
<i>E</i>	833	5833	0	0	0	0
<i>F</i>	0	0	24750	750	0	0
<i>D</i>	0	0	0	11500	0	0
<i>I</i>	0	0	250	1000	10750	0
<i>H</i>	0	0	0	0	0	1000
	833	292	295	1682	0	0
	269	833	750	0	1250	0
	564	-542	-455	1682	-1250	0

Table 3.5-20: Compensations among countries produced by the MP method corresponding to the usage that external agents make of the grid in each country. We take node 8 as the slack bus

	<i>P</i>	<i>E</i>	<i>F</i>	<i>D</i>	<i>I</i>	<i>H</i>
MP1	564	-542	-455	1682	-1250	0
MP2	458	-542	712	1895	-2417	-107
AP	0	0	-72	625	-552	0
WWT	744	1125	-1376	1015	-1376	-131

Table 3.5-21: comparison of the net compensation to be received by countries in the stylized example according to MP, considering two different choices of the slack bus, AP and WWT

Taking into account the numbers shown in the previous tables, as well as the conclusions drawn both in section 3.4.2.7 and here about the properties of the numerical results produced by the different methods, one can immediately realize that the AP method is the only one capable of producing, under any circumstances, reasonable results that cannot be easily challenged. Results for the MP method are conditioned by the assumption, made in this method, that the same node responds to every power injection or withdrawal in the region. Consequently, these results depend on the slack bus that is chosen, and locational differences between transmission tariffs in different nodes do not depend on the existence of congestion and losses in the grid. Results produced by the WWT method do not make sense either, as explained above.

3.5.1.7 Final recommendation of a procedure

Based on the analysis carried out in the previous section we are now in the position to conclude that the Average Participations method (AP) is clearly preferable to the With-and-Without Transits method (WWT). The Average Participations applied to Transits method (APT) may be regarded as an alternative to AP if parties agree to apply a method that relies on the concept of transit. The MP method may provide reasonable results when applied in a national market but is not well suited for regional markets.

Only AP satisfies the “broader context” set of criteria. Neither WWT nor APT comply with the single system paradigm. They do not provide a satisfactory solution to other aspects of transmission regulation apart from the computation of ITCs. MP complies with the single system paradigm but the results it provides for the allocation of the cost of new investments may be politically unacceptable. Locational signals sent by transmission tariffs resulting from the MP method do not take into account the existence of congestion and losses in the system.

Unlike WWT and APT, AP can be justified on the grounds of economic theory. WWT, on the other hand, requires a lesser amount of data than the rest of the methods in order to be applied. Additionally, more experience of its utilization to compute inter-TSO payments exists. MP is the method with largest data requirements.

The lack of coherence of WWT and APT with some aspects of the EU Regulation, the higher quality of the numerical results yielded by AP, the inability of WWT and APT to provide any kind of locational signals and the poor performance of WWT and MP in the implicit allocation of the cost of new network investments definitely tilt the balance

towards AP when selecting a method for the long-term mechanism of inter-TSO payments in the Internal Electricity Market.

3.5.2 The regional electricity market in Central America

This section discusses the situation that exists in the MER in Central America regarding the allocation of the cost of the regional grid and proposes the best regulation possible in this regard. First we recall the relevant regulation adopted in this region. Afterwards we present our own proposal.

3.5.2.1 Description of the adopted regulation

Given the situation in the region, authorities have decided to develop a regional market that achieves coordination among the operation (dispatch) and planning processes of the different national systems without interfering, to the extent possible, in the internal functioning of these systems.

Thus, parties in the region have agreed to establish a system of nodal prices at regional level. Transactions taking place in the context of the regional market are valued according to these prices. Each country faces a net revenue or payment resulting from the regional transactions in which the agents of the country take part. Afterwards, this net amount is allocated to the internal agents.

Applying nodal prices results in a revenue for each line of the regional grid that is considered in the dispatch. This revenue, which is named IVT (for *Ingreso Variable de Transporte*, in Spanish), comes from the congestion rents plus transmission losses that correspond to the line in the dispatch. As it is explained in chapter 4, these rents coincide with the difference between the energy prices at both ends of each line times the flow over the line. These funds may be devoted to either of two uses:

- Congestion rents corresponding to a line could be devoted to pay for part of the cost of the line.
- Alternatively, they could be used to recover a fraction of the cost of the national transmission system where the line belongs to.

Regulators in the region have opted for leaving this subject to the discretion of each country. Some countries make use of the first criterion, others have opted for the second one.

The decision on how to spend the congestion rents of a regulated line corresponds to the owner of the line. This may be either the country where the line is located or the region as a whole (in the case of lines owned by the EPR, which owns an important transmission line crossing the whole region). It is well known that revenues from the

application of nodal prices typically only recover a small part of the grid costs. It is therefore necessary to resort to additional grid charges.

Transmission charges aimed at recovering the cost of the used fraction of each line are different from those applied to recover the cost of the unused fraction. Therefore, we describe separately both types of charges.

Recovery of the cost of the used fraction of each line

Only regulated lines earn transmission charges aimed at completing the recovery of the cost of the used fraction of each line. Merchant lines and investments at risk only earn the congestion rents that correspond to them. Alternatively, private promoters may negotiate the collection of use-of-the-system payments in return for losing part of the transmission rights they hold over the capacity of the line. Charges applied to recover the cost of the used fraction of each line are named Use-of-the-System charges (or *Peajes* in Spanish). Use-of-the-System charges must recover the difference between the cost of the used fraction of each line and the congestion rents corresponding to the line. As its name suggests, Use-of-the-System charges are related to the use that each agent, or the agents within each country, make of a line.

The cost allocation method used to compute Use-of-the-System charges is named MFD (for *Método del Flujo Dominante*, in Spanish) and has been developed in the context of this regional market. A description of this method follows. The MFD method boils down to separately applying the Average Participations method to each one of the several super-transactions defined in the region, see section 3.3.3.4. As explained in this section, national super-transactions are comprised of all the power injections and withdrawals taking place within each national market while the regional super-transaction is comprised of all the power exchanges happening in the regional market. The flows attributed to each super-transaction result from the application of the superposition principle.

Different rules apply to each super-transaction depending on whether it is a national super-transaction or the regional one. When computing the line flows attributable to national super-transactions one must only consider the fraction of the RTR located within the corresponding country. When determining the flows corresponding to the regional super-transaction the whole RTR must be considered. Only those super-transactions whose contribution to the flow over a line goes in the same direction as the net flow pay for part of the cost of the line. The cost of the used fraction of each line is allocated among those super-transactions with a positive contribution to the line flow in proportion to the size of these contributions.

Recovery of the cost of the unused fraction of each line

Different transmission charges are applied to recover the cost of the non-used part of each line. These are the Complementary Charges (CC). Only those regulated lines whose construction has been centrally planned by the EOR have the right to earn a complementary charge, which completes the recovery of the whole cost of these lines.

The cost of the unused fraction of the aforementioned lines is allocated among the network users using a variant of the MFD method. In this variant of the method both the super-transactions that have a positive contribution to the flow over a line and those that have a negative contribution must pay part of the cost of the line. The cost of the line is allocated pro rata of the absolute values of the contributions of the different super-transactions.

3.5.2.2 Description of the proposed solution

The general scheme adopted for the computation of transmission charges in the region must be preserved. Thus, rents resulting from the application of nodal prices should contribute to the recovery of the cost of the grid. Additional transmission charges should be applied to complete the recovery of the cost of regulated lines. On the other hand, investments by private entrepreneurs should only earn the corresponding congestion rents.

There are certain aspects of the regulation that should be modified. Instead of applying different methods to allocate the cost of the used and unused fractions of each line, we recommend implementing one of the following two options:

- Apply the same usage-based method to complete the recovery of both the used and the non-used fraction of each regulated line.
- Apply a usage-based method to allocate the cost of the used-fraction of regulated lines and socialize the cost of the unused fraction of each line among the agents in the country where the line is located (preferably to demand).

Both options are suitable, though socializing to the demand the unused fraction of each line seems preferable. Actually, excess capacity in a line is not used. Additionally, SOs built transmission lines that are larger than needed in order to deal with possible contingencies and accommodate transactions resulting from the future installation of new generation and demand in the system. Neither the reliability benefits resulting from the existence of excess capacity, nor the estimated future use of the transmission grid, can be estimated based on the present electrical use of lines by agents. Thus, socializing the cost of the excess capacity of lines appears to be more sensible.

But we must also determine whether the usage-based method that is applied to compute Use-of-the-System charges (the MFD method) is conceptually sound. Thus,

we are going to compare it with a variant of the Average Participations method that is described now.

Variant of the AP method to be considered for the MER

In order to be in line with the transaction based philosophy adopted in the region, national and regional power injections and withdrawals at each node are dealt with separately. This is contrary to how the method is applied normally. Here we distinguish between the injections at a certain node that are part of a local or national super-transaction and those that are part of an international or regional transaction. When tracking the power injected or withdrawn by each agent only the net flow over each line is considered. Thus, this implementation of the method cannot be considered transaction based. Results produced by the method may differ from those obtained if only the net amount of power produced or consumed by each agent is deemed to use the transmission grid.

Both the MFD method and the variant of the AP method just described are compared now. We discuss the arguments in favor and against choosing each one of the two methods:

- Average Participations: the method proposed is very similar to the original AP method. Only in the unlikely case that an agent were scheduled to consume (respectively produce) power in the national market and produce (respectively consume) power in the regional one, could the results obtained for both variants of the AP method differ. Thus, an agent selling 100 MW in the national market and then buying 30 MW in the regional market is deemed to use the system grid separately for each of these two transactions. Whereas the AP method in its basic form would only consider a generator producing 70 MW, the modified version considered here charges the agent both for the grid usage corresponding to a generator producing 100 MW and for that corresponding to a load consuming 30 MW.

Formally speaking, the method no longer complies with the single system paradigm, since it is applied separately to power injections and withdrawals within national super-transactions and to those within the regional super-transaction. This method takes only into account the actual flows occurring in the system. Both the internal flows within each country and the cross-border flows influence the computation of grid charges. However, these charges could depend on other factors apart from the location and the consumption or generation profile of each agent. The transmission charge paid by an agent may vary depending on where the agent trades its energy. Hence, this variant of the AP method does not produce locational signals independent of the commercial transactions taking place. Participations in the use made of the lines are less dispersed than under the MFD method. The allocation method may be

formulated as the result of a cooperative game, see section 3.4.3. In this regard it has interesting properties from the point of view of equity and efficiency.

No distinction is made between the flows produced by national transactions and those produced by regional transactions. However, power injections and withdrawals corresponding to national commercial transactions are dealt with separately from those belonging to regional transactions. Distinguishing between power injections of one type and the other helps compare the results produced by AP and MFD. Line flows considered when determining the path followed by any injection are always the same regardless of the nature of the injection (national/regional). Therefore, the method does not discriminate between local and regional injections.

- MFD method: participations in the electrical usage of the grid by agents buying or selling their energy locally are restricted to the grid of the corresponding country. Participations of power injections and withdrawals within the regional super-transaction are not restricted in any sense. By principle, this method discriminates between the agents that trade their energy locally and those that trade it in the regional market since a different pattern of flows is considered in each case (those flows created by the corresponding national super-transaction versus those produced by the regional one). Additionally, the grid considered when computing the transmission charges levied on national injections and withdrawals is not the same as that considered for injections and withdrawals taking place in the regional market. MFD does not provide locational signals since charges depend on the transactions realized.

From the assessment previously made, one can easily conclude that the variant of the AP method proposed here is clearly preferable to the MFD method proposed by the regional institutions. In most cases, results obtained with the AP method will coincide with those obtained from the original version of this same algorithm. Besides, contrary to what happens in the MFD method, the AP method does not discriminate between national and regional power injections.

Allocation of the net payment to be made by each country to the individual agents within the country

Here, we discuss some general aspects of the problem of designing efficient transmission charges within each country. The following paragraphs discuss the structure of transmission charges as well as the split of these charges between generation and demand.

Tariff structure

The total regional transmission charge for a country results from the aggregation of charges to all power injections and withdrawals within the country that are part of the

regional super-transaction. This charge should be levied on all the injections and/or withdrawals in the country and not just on those which participate in the regional market. Otherwise, we would be discriminating again among agents according to the transactions they get involved in.

Conceptually speaking, transmission charges should be independent from the operational decisions by agents. However, structuring transmission tariffs as energy charges may be preferable to other options, as long as these charges are small compared to differences in energy prices among nodes in the system and/or times of the year. If this is the case, then small changes in the variable costs of generators and consumers are unlikely to significantly affect the dispatch in the region.

Split of transmission charges between generation and demand

Once determined the fraction of the cost of each line to be paid by each super-transaction, computing the individual participations of market agents in the MFD method requires applying the AP method to the flows produced by each super-transaction. The allocation of the charge corresponding to national super-transactions to the agents within a country is left in the hands of the corresponding country. However, the charge corresponding to the regional super-transaction can only be distributed among the different countries using the AP method to obtain the individual charges for the different agents within each country.

One may compute how much external agents are using of the grid of each country by applying the AP method to the injections, withdrawals and line flows associated with the regional super-transaction. Even more, this allows us to determine how much the generation within each country is using the grid of other countries and analogously for demand.

Mixing the total charges to be levied on generation and demand would result in a complete degradation of locational signals. Consequently, generators within a country as a whole should pay for the fraction of the cost of the grid in other countries they are responsible for. The same applies to demand. Again, we should never differentiate between agents within each country according to whether they trade their energy locally or regionally. In this way, G and L locational signals and charges will be meaningful and reasonably fair, at least at country level.

Given that the decision on how to split the cost of each line between generation and demand affects the net payment that each country must face, applying a 50%/50% split seems unavoidable. The AP method implicitly assigns 50% of the use of each line to generators and 50% to loads, since the flow on each line is tracked both upstream and downstream until it reaches the generators that inject it into the system and the loads that withdraw it.

3.5.2.3 Numerical results

This section presents and discusses the numerical results obtained from the application of the two proposed methods to a case study corresponding to the Central American system. The MFD and the AP methods have been used to allocate the cost of the regulated lines that belong to the regional transmission grid (RTR).

Description of the case study

Both methods have been applied to 48 scenarios for the year 2002. Four different load levels were considered per month: peak load, base load and 2 intermediate scenarios. The model of the system is comprised of 179 nodes, belonging to 6 countries, and 80 lines. Each one of the generating units within each power plant has been explicitly represented using a different node. Figure 3.5-20 shows the net power balance over the whole year for each one of the 6 countries in the region.

Results obtained for the case study

Table 3.5-22 shows the annual payments and revenues obtained with the MFD method, whereas Table 3.5-23 shows the corresponding results for the AP method. As mentioned above, power sales and purchases in the national and regional markets have been considered separately in both methods. As mentioned above, the cost of the regulated lines belonging to the RTR has been allocated to the countries in the region. Regulated regional lines can be of any of two types: they may belong to the EPR or they may be part of the transmission system of a country. The total cost of the lines that belong to the EPR has been allocated according to the electrical usage that is made of them. However, the electrical usage made of each line that belongs to the national grid of a country only determines how to allocate the cost of the used fraction of the line. Each country pays the remaining part of the cost of its national lines. Despite not being the most efficient and fairest solution conceptually speaking, we have opted for using each method to allocate only the cost of the used fraction of the regulated lines owned by countries. After all, this is what the regulation states at the moment. Therefore, we can compare the results produced by the MFD method as it is applied now to the ones obtained with the AP method in the same conditions.

Results have been aggregated at country level since, at least in the MFD method, the net payment to be faced by the generation and demand within each country should only be used to determine the average G and L access charges in the country.

Numbers in both tables are expressed in thousands of dollars. Note that ‘PA’ stands for Panama, ‘CR’ stands for Costa Rica, ‘NI’ stands for Nicaragua, ‘HO’ stands for Honduras, ‘ES’ stands for El Salvador and ‘GU’ stands for Guatemala. ‘Pay_EPR_REGT’ means the part of the cost of the network owned by the EPR that the corresponding country must pay because of the usage that regional power injections and withdrawals within this country are making of this grid. Analogously,

‘Pay_EPR_NATT’ means the part of the cost of the network owned by the EPR that each country must pay because of the usage that local injections and withdrawals within this country are making of this grid. ‘Pay_NAN_REGT’ refers to the part of the cost of the national networks of other countries that the corresponding country must pay because of the usage that regional injections and withdrawals in the country are making of these grids. ‘Pay_NAN_NATT’ is the charge associated with the usage of the national grids of other countries by the local injections and withdrawals in the corresponding country. ‘Rev_NAN_REGT’ is the revenue of the corresponding country in return for the usage that regional power injections and withdrawals in other countries are making of its grid. Finally, ‘Rev_NAN_NATT’ is the revenue of the corresponding country in return for the usage that local power injections and withdrawals in other countries are making of its grid.

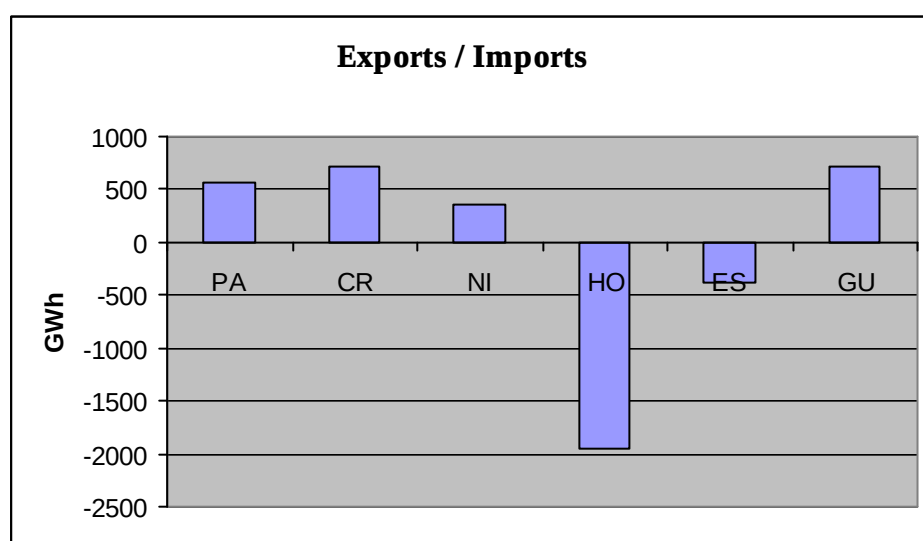


Figure 3.5-20: Net power balance for each country in Central America over the year 2002

The last four lines in each table correspond to the part of the regional grid that each country must pay, the part that each country must pay of other countries’ grids, the compensation due to each country for the usage others are making of its grid and the total net payment faced by each country, respectively.

	PA	CR	NI	HO	ES	GU
Pay_EPR_REGT	4887.9	4789.7	1944.1	6699.9	3662.3	4108.9
Pay_EPR_NATT	455.7	3019.9	392.2	2668.5	2158.7	811.9
Pay_NAN_REGT	1452.8	2309.9	619.7	2611.8	1496.5	860.7
Pay_NAN_NATT	0.0	0.0	0.0	0.0	0.0	0.0
Rev_NAN_REGT	443.3	1495.2	3613.9	1437.3	867.6	1494.1
Rev_NAN_NATT	0.0	0.0	0.0	0.0	0.0	0.0
Pay_EPR	5343.6	7809.6	2336.3	9368.4	5821.1	4920.9
Pay_NAN	1452.8	2309.9	619.7	2611.8	1496.5	860.7
Rev_NAN	443.3	1495.2	3613.9	1437.3	867.6	1494.1
Tot_pay	6353.1	8624.3	-657.9	10542.9	6450.0	4287.4

Table 3.5-22: results provided by the MFD method for Central America in 2002

Table 3.5-22 and Table 3.5-23 show that participations in the usage of the grid located within each country are more dispersed in the MFD method than in the AP method. The difference between the highest total compensation to be paid by a country and the lowest is 11200 m\$/year in the MFD method whereas that for the AP method is 5238 m\$/year, i.e. differences among compensations in the MFD method are much larger than those in the AP method. Compensations in the AP method are always positive and much more similar among them than those produced by the MFD method.

	PA	CR	NI	HO	ES	GU
Pay_EPR_REGT	801.8	1608.2	1002.7	1916.8	366.1	1025.7
Pay_EPR_NATT	2990	7219.4	4056.5	6211.9	3937.3	4463.5
Pay_NAN_REGT	281.1	520.2	396.9	943.1	223.9	296.8
Pay_NAN_NATT	825.8	2222.6	2286.8	1994.4	1945.6	1204.2
Rev_NAN_REGT	185.9	587.7	914.3	352.5	289.8	331.8
Rev_NAN_NATT	684.7	2614.4	2977.2	1624.1	1052.8	1526.2
Pay_EPR	3791.8	8827.6	5059.2	8128.7	4303.4	5489.2
Pay_NAN	1106.9	2742.8	2683.7	2937.5	2169.5	1501
Rev_NAN	870.6	3202.1	3891.5	1976.6	1342.6	1858
Tot_pay	4028.1	8368.3	3851.4	9089.6	5130.3	5132.2

Table 3.5-23: results provided by the AP method for Central America in 2002

This is so despite the fact that national super-transactions cannot enter the grid of other countries in the MFD method, since they are confined to the borders of the corresponding country. Thus, rows 'Pay_NAN_NATT' and 'Rev_NAN_NATT' in Table 3.5-22 are zero for every country. Due to the fact that the MFD method decomposes the flows in the system into those produced by each one of the super-transactions, the usage that regional power injections and withdrawals within each country make of the grid of other countries and the regional grid is much higher in the MFD method than in the AP method. The total volume of net compensations and payments is a bit higher in the MFD method than in the AP method. The difference, however, is small.

Highest exporters tend to pay the highest compensations in AP. This trend is somewhat less clear in the MFD method. Thus, the absolute value of the net export is approximately the same for Nicaragua and El Salvador. However, the total net payment to be faced by Nicaragua in the MFD method is negative, i.e. it is entitled to a compensation, whereas El Salvador must pay a large compensation (6450m\$).

There are significant similarities between the results obtained from both methods. Thus, Honduras and Costa Rica are the two largest payers both in AP and in MFD. Nicaragua is the only net receiver in MFD and it is the country that has to pay least in AP.

Final Recommendation

Taking into account the properties of both methods, one can draw the conclusion that the AP method is clearly preferable to the method already proposed for the region (MFD). AP has much better properties than the MFD method and the results it produces seem to be more reasonable and less volatile than those for the MFD method. The following list summarizes the main advantages of one method over the other:

- The AP method does not discriminate between local and regional power injections and withdrawals when computing the electrical use they make of the grid. This is due to the fact that the set of line flows considered when determining the path followed by each power injection or withdrawal is always the same, regardless of the market where this power injection has taken place. On the other hand, MFD clearly discriminates between local and regional transactions since the set of line flows considered in each case, as well as the model of the grid, is different.
- The AP method provides an answer to other aspects of transmission regulation apart from the allocation of grid costs at country level. Thus, for example, it may provide long term locational signals or nodal transmission tariffs if so desired.
- The results for compensations among countries produced by the AP method seem to be more in line with the actual external usage that each country is making of the grid of others. Furthermore, results for AP are less volatile and the total volume of compensations is smaller.
- More experience exists of the utilization of the AP method. This would be the first time the MFD method is applied.

This is in line with the decision previously made in section 3.5.3.1, which was based on the arguments laid out when qualitatively comparing both methods. Some of these arguments have been recalled here.

3.5.3 Other regional markets

Presently, many regional markets are in the process of formation. After having discussed the IEM and the Central American one in depth, we move on to analyze two other regional market models: the market that has been developing in South America (MERCOSUR) for a number of years already, and the Standard Market Design (SMD) proposed for the US. MERCOSUR and the SMD are compared to the Regional Electricity Market in Central America and the IEM in Europe, respectively.

3.5.3.1 The MERCOSUR regional electricity market

The general design of the regional market in South America was described in section 1.3.4 of chapter 1. Now we focus on the solution that has been proposed in the region for the computation of transmission tariffs.

Solution proposed by the World Bank for the remuneration of the transmission service

As far as we now, no transmission pricing scheme has been adopted in this region yet. Here, we shall describe the scheme proposed for the region by the World Bank in (WB 2000). The structure of grid charges is almost identical to that of transmission tariffs in Central America. As in the MER, tariffs proposed for the market in South America include a component associated with the application of nodal energy prices to the regional transactions. Thus, power injections and withdrawals in the regional market are priced at their corresponding nodal energy prices.

Given that revenues from the application of nodal prices can only pay a small fraction of the regulated revenue of the transmission activity, market agents must pay additional grid charges. These charges may be Use-of-the-System charges, resulting from the application of some usage-based method, or they may be computed as the long run marginal costs corresponding to the nodes of the regional transmission grid. Grid charges include a charge that allows the recovery of the cost of the unused fraction of a line.

Several uses of transmission charges are possible:

- Charges may only be used to recover the cost of the regulated lines in the system, i.e. those lines that belong to a country and result from the centralized expansion of the national grids. Additionally, financial transmission rights would be awarded to private promoters of cross-border lines.
- Alternatively, owners of efficient interconnection lines, regardless of whether these lines were promoted by the private initiative or they were the result of central planning, may be guaranteed the recovery of their investments through the application of transmission charges.

Lastly, owners of interconnection lines between countries would be entitled to a revenue resulting from the arbitrage of quality of service they carry out between both sides of the interconnection. In other words, they would earn as much as the difference between the cost of the non served energy avoided in the system where the quality of service is lower and the additional costs born by the system where the quality of service is higher. See the discussion on the topic included in (WB 2000).

The solution we propose for the remuneration of this service

Parties in the region have agreed to implement a system of nodal energy prices. Applying nodal prices, which is the optimum solution to manage the energy dispatch, results in part of the grid cost being automatically recovered through the collection of congestion rents. Congestion rents should be allocated among lines in the system according to the marginal value of each line in the dispatch, as chapter 4 explains. Allocation of congestion rents is irrelevant when we are talking about regulated lines but it is an issue of utmost importance when some of the lines are owned by private investors, who expect to make a profit out of the exploitation of the line.

However, congestion rents do not suffice to pay the whole cost of a normal and well developed grid. Then, additional grid charges are necessary to recover the remaining part of the cost of the regulated lines. In principle, transmission tariffs should be set according to the economic benefit that each agent obtains as a consequence of the existence of each transmission line or, alternatively, according to the responsibility of each agent in the grid costs incurred by the system.

Theoretically speaking, the responsibility of a generator or load in the network costs can be computed as the Long Run Marginal Cost (LRMC) corresponding to the node where the agent is located and the time when the generator or load is installed. However, the calculation of LRMCs is plagued with difficulties which make this alternative highly impractical, see the discussion on this method in section 3.3.2 of this chapter.

Regarding the allocation of grid costs to beneficiaries, we can immediately conclude that only part of the benefits produced by a line are related to the electrical usage that agents make of this line. Some of these benefits are related to the increase in the quality of service provided to consumers. Therefore, allocating charges based also on the quality transfers taking place seems reasonable. However, estimating the reliability benefits and others that market agents obtain from each line turns out to be very difficult. In practice, we have not been able to come up with a criterion other than electrical usage to measure how much each agent benefits from each grid, except for new lines. Whenever reliability benefits, and other benefits not related to electrical usage, are significant, allocating the whole cost of each regulated line based on the electrical usage that each agent makes of it does not seem to be justified. Hence, we should allocate the cost of the used fraction of each line according to a usage based method and socialize the remaining cost of the line. Obviously, this is not applicable to merchant lines, which may not recover the total cost of the investment.

AP is probably the usage-based method that has best properties among the ones that we have studied. However, AP provides locational signals corresponding to the average use of the grid by each agent. If we want to send locational signals corresponding to the cost that the system would incur if an additional generator or load were installed, we must apply different transmission charges from those

computed with AP. In this case, transmission charges produced by AP could be modified in order to strengthen the causal relation between the grid costs that the system incurs because of the installation (or retirement) of an agent and the transmission charges this agent has to pay. See our proposal in this regard in section 2.5.1.4 of chapter 2.

As for merchant lines, they should only earn the congestion rents they are able to raise or, alternatively, revenues from the sale of the corresponding transmission rights. Congestion rents should not be limited. However, a private promoter would not be guaranteed the recovery of the initial investment. Applying transmission charges that allow private grid promoters that make efficient investments to recover the cost of these investments may be an interesting approach in some cases. In general, private entrepreneurs should decide whether to earn the congestion rents corresponding to the lines they own or earn certain agreed regulated revenues.

3.5.3.2 The Standard Market design in the US

Several different regional markets exist in the USA. The size of these regional markets is comparable to that of large national markets in Europe. Here, we use the term “regional” to talk about each one of the regional markets that are forming throughout the US, i.e. we are not referring to the country as a whole. Given that the regulation adopted in each regional market is different, we shall only describe the approach proposed by FERC in order to compute transmission tariffs. Afterwards, we shall outline our own proposal.

Remuneration of the transmission service proposed by FERC

The transmission grid within each state in the US has traditionally played a similar role to that of national grids in the EU. In both cases the transmission grid was developed for the generation belonging to local vertically integrated utilities to supply the local load. Some power transfers between states, or vertically integrated utilities, were envisaged as a support mechanism in case any of them ran into security or stability problems. Therefore, some interconnections existed though, typically, they were very limited in number and capacity. Both these interconnections and the internal grid of the states are not developed enough to accommodate the level of interstate power trade presently demanded. Important reinforcements to the grid are still necessary since very small progress has been made in this respect. This situation strongly resembles that between the grids of some member countries in the EU.

FERC is the entity in charge of setting tariffs for the transmission of electricity between states. This organization has provided some guidelines in this respect through various orders and notices recently issued, including the notice of proposed rulemaking where it outlined the main features of what could be the Standard Market Design to be implemented in the different regional markets in formation, see (FERC 2002; FERC

2003). Now, we summarize the content of these guidelines and compare it to the agreements that have been reached within the IEM in Europe:

- Pricing mechanisms applied in the short and the long run must foster the efficient usage of existing grid and generation facilities as well as optimal investment in new transmission and generation. Unless retailing takes place independently from generation and transmission, FERC will not be able to decide which pricing mechanisms to apply. It will have to leave these decisions to the will of state authorities. Therefore, unbundling of activities is necessary.
- Transmission tariffs proposed by FERC bear some resemblance to the tariff setting process in Europe. Similarly to what happens in Europe, FERC distinguishes between short term and long term signals. However, whereas short term signals in the US relate both to congestion and losses, decision makers in the EU have agreed to consider losses as part of the long term signals to be sent to market agents through the mechanism of inter-TSO payments.
- FERC proposes a system of nodal prices to be applied within each region in the short run. Nodal pricing manages energy and transmission capacity in an integrated manner. On the other hand, congestion signals in Europe are presently dealt with separately from energy in most cases, though implicit auction schemes are also being considered for some borders or areas. The systems proposed both for Europe and the US make use of market mechanisms to solve congestion in the grid.
- Long term signals proposed by FERC play a two part role. On the one hand, they allow grid owners to recover their investment. On the other hand, they send locational signals to agents. FERC has proposed two alternative systems of access charges: postage stamp type transmission tariffs and license plate rate tariffs. In any case, only consumers within each regional market, or retailers buying power on their behalf, would have to pay access charges. Postage stamp tariffs are the same for every consumer within each regional market. In order to apply a system of license plate rate tariffs, a series of price areas must be defined a priori. License Plate Rate tariffs are exit transmission tariffs that vary from one area to another but are uniform within each one of these areas. Consumers located in different areas would have to pay different unit transmission tariffs. Therefore, access tariffs of the second type send locational signals to consumers. Grid locational signals for generators in the US would be limited to connection charges. Both deep and shallow connection charges could exist, however. Payment of the transmission access tariff would grant access to the entire regional grid (the grid within each regional market).

This system of charges would be separately applied to each regional market in the US. At the same time, compensations between different regional markets

would take place. Compensations among regions for the use that consumers in each one of them make of the grid of other regions would depend on the net aggregated imports by each region from each of the remaining ones. Thus, consumers located in region 'Y' would only pay part of the cost of the grid of region 'X' if region 'Y' imports a significant net amount of power from 'X'. Otherwise, grid charges would apply independently for agents within each region and there would be no compensation among the different regions.

In principle, Europe has opted for a system of entry / exit access tariffs, i.e. both consumers and generators would have to pay access charges. However, access charges for generators will normally be much smaller than those for consumers. Transmission tariffs are the result of modifying local grid charges according to the net inter-TSO payment faced by each country. The allocation of the net inter-TSO compensation to be received by a country to the loads and generators in the country is carried out by the competent local authority. Agents gain access to the entire European transmission grid by paying the corresponding access tariffs. Given that the electrical size of many national markets in Europe is more or less equivalent to that of some RTOs in the US, the inter-TSO compensation mechanism in Europe can be considered to be equivalent to the payment by external agents of part of the cost of the grid of each regional market.

The pricing scheme proposed by FERC has been successfully implemented in some of the regional markets already created in the country (like PJM).

Our proposed solution

We shall outline our proposal by criticizing the approach proposed by FERC. In any case, some recommendations made by FERC are very reasonable. Thus, short-term price signals in the SMD are perfectly efficient. Nodal prices represent the optimal solution to the problem of placing a value on the energy produced or consumed at each node and time. Additionally, making transmission rights available is very reasonable in order to help market agents hedge against the differences between energy prices at different nodes.

On the other hand, long term price signals can be significantly improved. Both generators and loads (instead of only loads) should contribute to grid costs in proportion to their responsibility in these costs. Connection charges would be employed to recover the cost of those lines that are necessary to physically connect a new generator or load. Apart from them, both generators and loads should pay access charges. Access charges should be computed based on the electrical usage that each agent makes of the grid in the region. The method of Average Participations should be applied for this. If applying nodal access charges is not possible, zonal G and L unit charges might be computed by averaging the charges obtained for all the agents of the same type (generators or loads with the same profile) located within the same area. Obviously, several price areas should be defined a priori. If the regional grid were

meshed enough and no major grid investments were necessary, postage stamp tariffs might be the best option and sending locational signals would not be a priority. However, this is not the situation of regional markets in the US.

Apart from that, transmission tariffs should reflect the fact that agents in a regional market benefit from the grid in other regions. Therefore, a system of compensations for the usage that agents in each regional market make of the grids of other markets is necessary. This inter-regional compensation mechanism would resemble that implemented in the IEM and it should take into account the pattern of cross-border flows between the existing regional markets. Ideally, a system of nodal or zonal transmission tariffs, based on network usage as a proxy to economic benefits, should apply at country level.

3.6 Conclusions

The remuneration of the transmission activity is discussed in chapter 2 of this thesis. Here, we have dealt with the problem of collecting from network users the regulated revenues of transmission owners. This chapter has analyzed and compared several approaches to allocate the cost of transmission lines. Nodal prices are optimal short term energy signals. They also send long term signals to agents and help recover part of the cost of the transmission grid. Transmission rights are used by agents to hedge against the volatility in the difference of energy prices between different parts or nodes of the grid. The owner of a transmission right receives the difference between the corresponding nodal energy prices in the dispatch. Hence, revenues of transmission right owners are part of the total revenue from the application of nodal prices in the dispatch. Therefore, revenues for the system from the initial sale of transmission rights are based on the expected revenue from the application of nodal energy prices at the nodes that these rights are defined between.

Revenues obtained by the system from the application of nodal energy prices, or from the initial sale of transmission rights, must be used to partly pay the cost of the existing transmission grid or to finance new grid investments.

Typically, revenues from the application of nodal prices (or any other dispatch model) only collect a small fraction of the cost of the grid. Therefore, additional transmission charges are necessary. These charges must promote efficiency in investments in the long run and allocative efficiency in the short run. Conceptually, grid tariffs should be based on the benefits that agents obtain from the network (or the cost they cause the system to incur). In practice, this is only possible for connection lines (dedicated to the physical connection of an agent to the grid), for new lines and for specific network investments whose beneficiaries are not very much dispersed. When building a new line, we may only consider the benefits that network users will obtain from it in the years immediately following the time when the line goes into operation. This facilitates

the identification of the beneficiaries of this line. *The cost of the lines that are strictly necessary to connect network users to the grid must be paid by these agents. Additionally, new lines whose beneficiaries can be clearly identified should be paid by them.*

If identifying the beneficiaries of a line or set of lines is not possible, and sending long-term locational signals is necessary (due to the fact that the decision by agents on where to install new generation or load may cause the system to build new transmission lines), tariffs should be based on the usage that each agent is making of the grid. An in depth discussion of the technical, economic and legal aspects of the problem of allocation of sunk grid costs has led us to conclude that the Average Participations method is better suited than other usage-based methods to compute transmission tariffs. Properties of AP include the following:

1. AP yields very reasonable results under any circumstances. Results seem to be quite stable. They do not vary excessively from one scenario to another within the same year, nor from one year to another.
2. The AP method is very easy to implement. Applying AP only requires having available the power injections and withdrawals in the system as well as the flows over the lines for a number of representative snapshots of the system operation during a given year.
3. The AP algorithm agrees with the basic economic and engineering principles that must rule the allocation of sunk grid costs. According to the physical laws governing the functioning of power systems, it computes the true average use that generators and loads make of the transmission grid. Additionally, the proportionality branching rule applied in the method can be justified from the point of view of the economics theory. In fact, the solution provided by the AP method coincides with that of a cooperative game where the inflows into each node of the network are distributed among the outflows.
4. This method provides an answer to other aspects of transmission regulation, including the allocation of the cost of new network investments or the computation of locational signals.

The AP method computes the average use that each agent makes of the grid. Thus, locational signals obtained from the AP method may not reflect the incremental cost incurred by the system as a result of the installation or retirement of a generator or load. However, charges provided by AP may be modified in order to strengthen the relation between the cost caused by an agent and the transmission charges it faces. Usage-based methods based on the Cooperative Game Theory may provide efficient transmission charges. However, most of them have not been applied yet in large power systems and they are still in the development phase.

If identifying the beneficiaries of regulated lines is not possible, and sending locational signals is a priority, usage-based methods must be applied to compute transmission tariffs. The AP

method is best suited to this task, among the usage-based methods considered in this thesis. Usage-based methods based on cooperative game theory are a promising alternative.

Transmission charges based on the responsibility of each agent in the grid costs, or on the network usage made by each agent, are likely to influence the short-term decisions by agents. Instead, we may first compute the total transmission charge to be paid by each group of generators or loads, and then allocate this charge to the agents within this group. In this case, the structure of network charges becomes an issue of utmost importance. The relative importance of grid charges when compared to energy prices determines which is the best tariff structure to use. If transmission charges are of the same order of magnitude as energy prices, structuring grid tariffs as capacity charges is preferable. If transmission tariffs are small compared to energy prices, making these tariffs part of the variable cost faced by agents is acceptable from the point of view of efficiency and it is easier to implement.

In order not to distort operational decisions by network users, a set of price areas may be defined. Price areas defined for the purpose of computing transmission tariffs should comprise groups of agents that make a similar use of the network per unit of traded power. Agents with the same profile that are located within the same price area should pay the same unit transmission charges. If transmission tariffs are low compared to energy price differences, the former could be structured as energy charges. Otherwise, they should be structured as capacity ones.

The internal split of the total amount of charges between generation and demand in a country is also very relevant to our discussion. It is widely accepted that generators should pay a smaller part of the cost of the grid than loads. However, harmonizing the way in which transmission tariffs (to generators mainly) are computed may be central to achieving a leveled playing field where competition among agents in different countries can develop. *The fraction of the cost of the transmission grid within each state or country that is paid by generators should be approximately the same in all the countries included in a regional market.*

Regarding the allocation of grid costs in the context of regional markets, in many of them transmission charges are computed by the local authorities in each country or state (the IEM, the market in Central America, some regional markets in the US, etc.). Local charges are modified according to the net compensation faced by each country resulting from the application of some allocation method at regional level. This allocation method must be efficient. However, if, despite using an efficient method at regional level, local charges within each country are not computed according to a sound criterion (for example, if uniform unit transmission charges are applied), most of the locational content of transmission tariffs will be lost. *Therefore, as long as sending locational signals is a priority, harmonizing some basic features of transmission tariffs computed within each country or state is essential in order for transmission tariffs to encourage agents to make efficient investment decisions. Regulated transmission tariffs computed at regional level must pay the whole cost of transmission lines, both existing and new, that are relevant to cross-border trade. Additionally, they must pay the cost of the used fraction of those*

lines whose construction has been unilaterally decided by countries or states and are not of special interest to market agents trading their energy regionally. In principle, regulated transmission investments should not be discriminated based on the party that has proposed them (regional institutions versus countries).

Transmission tariffs influence and are influenced by other aspects of transmission regulation. Now, we assess the mutual influence between transmission tariffs and congestion on the one hand, and between tariffs and grid investments on the other:

1. Network access and congestion management: the allocation of sunk grid cost must not affect the economic dispatch. This consideration should influence the election of the structure of transmission charges, as discussed in the previous paragraph. Additionally, the way transmission rights are initially allocated may have an impact on the variable production costs incurred by market agents, and hence on the dispatch. This implies that transmission rights cannot be initially allocated according to the transmission charge that has been historically (or is presently) paid by each agent, which should be a function of the electrical usage made of the grid lines by this agent.
2. Investments in new lines: the allocation of the cost of new lines may affect investment decisions by market agents and countries. It is therefore important that grid charges corresponding to these lines are determined on the basis of how much each agent or group of agents benefits from their existence. Otherwise, efficient lines may not be constructed, due to the opposition of agents or countries that feel their interests are being harmed. Apart from that, transmission tariffs are locational signals that may affect the decision by agents on where to install, thus also influencing the optimal expansion of the grid.

As a result of the mutual influences existing among different aspects of the regulation of transmission, when designing transmission tariffs we must take into account the effect that these transmission tariffs will have on the management of congestion in the grid (which should be none) and on the development of this grid.

Bibliography

- ACCC Australian Competiton and Consumer Commission (ACCC). Web page:
<http://www.accc.gov.au/content/index.phtml/itemId/142>.
- ACCC (1999). Regulatory Test for New Interconnectors and Network Augmentations, Australian Competiton and Consumer Commission, <http://www.accc.gov.au/content/index.phtml/itemId/54367/fromItemId/54368>.
- ACCC (2004). Review of the Regulatory Test for Network Augmentations, Australian Competition and Consumer Commission (ACCC). Web page:
<http://www.accc.gov.au/content/index.phtml/itemId/532002/fromItemId/54368>.
- Agurto, R. (1992). Cargos por uso de la transmisión en Chile: Análisis del concepto de áreas de influencia. 6º Congreso de América Latina y el Caribe sobre Economía de la energía eléctrica, Mendoza, Argentina.
- Anderson, K. P. and A. McCarthy (1999). "Transmission pricing and expansion methodology: lessons from Argentina." Utilities Policy **Vol. 8**(No. 4): 199-211.
- Bakirtzis, A., P. Biskas, et al. (2001). "Comparison of two methods for long-run marginal cost-based transmission use-of-system pricing." IEE Proceedings on Generation, Transmission and Distribution **148**(4): 477-481.
- Baran, M. E., V. Banunarayanan, et al. (1999). "A Transaction Assessment Method for Allocation of Transmission Services." IEEE Transactions on Power Systems **Vol. 14**(No. 3): 920-927.
- Baughman, M. L. (1997). "Pricing of open-access transmission services in Texas." Utilities Policy **6**(3): 195-201.
- Bialek, J. (1996). "Tracing the flow of electricity." IEE Proceedings on Generation, Transmission and Distribution **143**(4): 313-320.
- Bialek, J. (1997). "Topological Generation and Load Distribution Factors for Supplement Charge Allocation in Transmissin Open Access." IEEE Transactions on Power Systems **Vol. 12**(No. 3): 1185-1190.
- Bialek, J. W. and P. A. Kattuman (2004). "Proportional sharing assumption in tracing methodology." IEE Proceedings on Generation, Transmission and Distribution **Vol. 151**(No. 4): 526-532.

- Bialek, J. W. and S. Ziemianek (2003). Tracing based transmission pricing of cross-border trades: fundamentals and circular flows. IEEE Power Tech Conference, Bologna, Italy.
- Bjorndal, E., G. C. Stamtsis, et al. (2005). "Finding core solutions for power system fixed cost allocation." IEE Proceedings on Generation, Transmission and Distribution **Vol. 152**(No. 2): 173-179.
- Bushnell, J. and S. Oren (1997). "Transmission pricing in California's proposed electricity market." Utilities Policy **6**(3): 237-244.
- CER (2004). Electricity Tariff Structure Review: International Comparisons, Commission for Energy Regulation (CER). Web page: <http://www.cer.ie/cerdocs/cer04101.pdf>.
- Chang, Y.-C. and C.-N. Lu (2001). "An electricity tracing method with application to power loss allocation." International Journal of Electrical Power and Energy Systems **Vol. 23**(No. 1): 13-17.
- Chicco, G., G. Gross, et al. (2002). "Allocation of the Reactive Power Support Requirements in Multitransaction Networks." IEEE Transactions on Power Systems **Vol. 17**(No. 4).
- Consentec/Frontier (2005). Study on the further issues relating to the inter-TSO compensation mechanism, European Commission. Directorate-General Energy and Transport. Web page: http://europa.eu.int/comm/energy/index_en.html. Report prepared by Consentec and Frontier Economics.: 72.
- Consultants, C. E. and H. B. Putnam (1995). Transmission Services Costing Framework. Volume 2: Framework Description and Application, EPRI. Ref. EPRI TR-105121-V2.
- Contreras, J. and F. F. Wu (1999). "Coalition Formation in Transmission Expansion Planning." IEEE Transactions on Power Systems **14**(3): 1144-1152.
- Contreras, J. and F. F. Wu (2000). "A Kernel-Oriented Algorithm for Transmission Expansion Planning." IEEE Transactions on Power Systems **15**(4): 1434-1440.
- CRIE (2002). Reglamento Transitorio del Mercado Eléctrico Regional, Comisión Regional de la Interconexión Eléctrica, Web page: <http://www.omca.net/documentos/Rtmer01.pdf>.
- Daniel, J. S., R. S. Salgado, et al. (2005). "Transmission loss allocation through a modified Ybus." IEE Proceedings on Generation, Transmission and Distribution **Vol. 152**(No. 2): 208-214.

- Danitz, F., H. Rudnick, et al. (2002). Use Based Allocation Methods for Payment of Electricity Transmission Systems. PowerCon 2002 IEEE-PES/CSEE International Conference on Power System Technology, Kunming, China.
- EC (2004). Medium Term Vision for the Internal Electricity Market. DG Energy and Transport Working Paper. Strategy paper. Brussels, DGTREN. European Commission.
- EC (2005). Annual Report on the Implementation of the Gas and Electricity Internal Market. Brussels, European Commission (EC): 10.
- EOR (2005). Reglamento del Mercado Eléctrico Regional (RMER). Libro III del RMER. De la Transmisión., Ente Operador Regional (EOR).
- Evans, F., J. M. Zolezzi, et al. (2003). "Cost Assignment Model for Electrical Transmission System Expansion: An Approach Through the Kernel Theory." IEEE Transactions on Power Systems. **Vol. 18**(No. 2): 625-632.
- Fang, W. L. and H. W. Ngan (2002). "Succinct method for allocation of network losses." IEE Proceedings on Generation, Transmission and Distribution **Vol. 149**(No. 2): 171-174.
- Farmer, E. D., B. J. Cory, et al. (1995). "Optimal Pricing of Transmission and Distribution Services in Electricity Supply." IEE Proceedings on Generation, Transmission and Distribution **142**(1): 1-8.
- Farmer, E. D., B. L. P. P. Perera, et al. (1995). "Optimal Pricing of Transmission Services: Application to Large Power Systems." IEE Proceedings on Generation, Transmission and Distribution **142**(3): 263-268.
- FERC (2002). Notice of Proposed Rulemaking, Remedying Undue Discrimination through Open Access Transmission service and Standard Electricity Market Design. Washington, Federal Energy Regulatory Commission. Web page: <http://www.ferc.gov/industries/electric/indus-act/smd/nopr.asp>.
- FERC (2003). White Paper: Wholesale Power Market Platform, Federal Energy Regulatory Commission (FERC). Web page: <http://elibrary.ferc.gov/idmws/nvcommon/NVViewer.asp?Doc=9686957:0:32>.
- Fernández, J. L. and M. Remacha (2005). Comparison on transmission pricing in Europe: Synthesis 2004, Tariffs Task Force, European Association of Transmission System Operators (ETSO): 10.
- Fradi, A., S. Brignone, et al. (2001). "Calculation of Energy Transaction Allocation Factors." IEEE Transactions on power Systems **Vol. 16**(No. 2): 266-72.

- FSR (2005). A study on the inter-TSO compensation mechanism, Florence School of Regulation. Web page: <http://www.iue.it/RSCAS/ProfessionalDevelopment/FSR/>: 96.
- Galiana, F. D., A. J. Conejo, et al. (2003). "Transmission Network Cost Allocation Based on Equivalent Bilateral Exchanges." IEEE Transactions on power Systems **Vol. 18**(No. 4): 1425-1431.
- Green, R. (1997). "Electricity transmission pricing: an international comparison." Utilities Policy **6**(3): 177-184.
- Green, R. (1997). "Transmission Pricing in England and Wales." Utilities Policy **6**(3): 185-193.
- Happ, H. H. (1994). "Cost of wheeling methodologies." IEEE Transactions on Power Systems **Vol. 9**(No. 1): 147-156.
- Hartigan, J. (1975). Clustering Algorithms. New York, John Wiley & Sons.
- Hur, D., C.-I. Yoo, et al. (2004). "Transmission embedded cost allocation methodology with consideration of system reliability." IEE Proceedings on Generation, Transmission and Distribution **Vol. 151**(No. 4): 427-432.
- Jong, J. d. (2004). The "regional approach" in establishing the internal EU electricity market, Clingendael International Energy Programme. Netherlands Institute of International Relations. Web page: http://www.clingendael.nl/publications/2004/20041200_ciep_paper_dejong.pdf.
- Joskow, P. L. (2005). "Transmission Policy in the United States." Utilities Policy **Vol. 13**: 95-115.
- Kattuman, P. A., R. J. Green, et al. (2001). A Tracing Method for Pricing Inter-Area Electricity Trades, Department of Applied Economics, University of Cambridge, Cambridge Working Papers in Economics with number 0107. Web page: <http://ideas.repec.org/p/cam/camdae/0107.html>. **June**.
- Kim, B. H. and M. L. Baughman (1997). "The Economic Efficiency Impacts of Alternatives for Revenue Reconciliation." IEEE Transactions on Power Systems **Vol. 12**(No. 3): 1129-1135.
- Kirschen, D., R. Allan, et al. (1997). "Contributions of individual generators to loads and flows." IEEE Transactions on Power Systems **12**(1): 52-60.
- Kirschen, D. and G. Strbac (1999). "Tracing Active and Reactive Power between Generators and Loads Using Real and Imaginary Currents." IEEE Transactions on Power Systems **Vol. 14**(No. 4): 1312-1318.

- Kristiansen, T. (2004). "Congestion management, transmission pricing and area price hedging in the Nordic region." International Journal of Electrical Power & Energy Systems **Vol. 26**(No. 9): 685-695.
- Lima, J. W. M., M. V. F. Pereira, et al. (1995). "An Integrated Framework for Cost Allocation in a Multi-owned Transmission System." IEEE Transactions on power Systems **Vol. 10**(No. 2): 971-977.
- Llarens, D. G. (2004). Un nuevo modelo mercantil para la determinacion de cargos de transmisión en mercados eléctricos competitivos con acceso abierto. Buenos Aires, Argentina, Mercados Energéticos.
- Lo, K. L. and S. P. Zhu (1993). "Wheeling and marginal wheeling rates: theory and case study results." Electric Power Systems Research **Vol. 27**(No. 1): 11-26.
- Marangon, J. W. and E. J. d. Oliveira (1998). "The Long Term Impact of Transmission Pricing." IEEE Transactions on Power Systems **Vol. 13**(No. 4): 1514-1520.
- Muñoz, A. (1996). Aplicación de técnicas de redes neuronales artificiales al diagnóstico de procesos industriales. Departamento de electrónica y automática. School of engineering. Madrid, Universidad Pontificia Comillas.
- Mutale, J., G. Strbac, et al. (2000). "Allocation of losses in distribution systems with embedded generation." IEE Proceedings on Generation, Transmission and Distribution **Vol. 147**(No. 1): 7-14.
- Pan, J., Y. Teklu, et al. (2000). "Review of Usage-Based Transmission cost Allocation Methods under open Access." IEEE Transactions on Power Systems **Vol. 15**(No. 4): 1218-1224.
- Pantos, M. and F. Gubina (2004). "Ex-Ante Transmission-Service Pricing Based on Load-Flow Patterns." IEEE Transactions on Power Systems **19**(2): 796-801.
- Park, Y. M., J.-B. Park, et al. (1998). "An Analytical Approach for Transaction Costs Allocation in Transmission System." IEEE Transactions on Power Systems **Vol. 13**(No. 4): 1407-1412.
- Perán, F. L., I. J. Pérez-Arriaga, et al. (2001). "Benchmark of electricity transmission tariffs in the countries of the internal electricity market of the European Union." Utilities Policy **Vol. 10**: 47-56.
- Pérez-Arriaga, J. I. (2002). Methodology for cross-border tarification in the Internal Electricity Market of the European Union. Power Systems Computing Conference (PSCC), Seville.
- Pérez-Arriaga, J. I., F. L. P. Montero, et al. (2002). Benchmark of Electricity Transmission Tariffs, Directorate-General for Energy and Transport / European

- Commission. Web page:
http://europa.eu.int/comm/energy/electricity/publications/doc/bench_trans_tarif_en.pdf.
- Pérez-Arriaga, J. I. and L. Olmos (2003). Extension of the project on cost components of cross border exchanges of electricity, Directorate-General for Energy and Transport / European Commission. Web page:
http://europa.eu.int/comm/energy/electricity/publications/index_en.htm.
- Pérez-Arriaga, J. I., L. Olmos, et al. (2002). Cost components of cross border exchanges of electricity, Directorate-General for Energy and Transport / European Commission. Web page:
http://europa.eu.int/comm/energy/electricity/publications/index_en.htm.
- Pérez-Arriaga, J. I. and F. J. Rubio (1995). "Marginal pricing of transmission services: An analysis of cost recovery." IEEE Transactions on Power Systems **Vol. 10**(No. 1): 546-553.
- Pérez-Arriaga, J. I. and Y. Smeers (2003). Guidelines on tariff setting. Transport Pricing of Electricity Networks. F. Levesque, Kluwer Academic Publishers.
- Petrucelli, M. (2003). Estimación del uso del sistema de transporte por los agentes del mercado de electricidad. Madrid, Comillas University.
- PHB-HAGLER/SYNEX (2000). Diseño General del Mercado Eléctrico Regional. Buenos Aires, Informe elaborado por Mercados Energéticos S.A., PHB-HAGLER BAILLY y SYNEX Ingenieros Consultores para Unidad Ejecutora del Proyecto SIEPAC.
- Poletti, S. (2005). "Optimal Time-of-use Access Charges." Journal of Regulatory Economics **Vol. 27**(No. 2): 203-224.
- PSR (2002). Alocacao de Custos de Transmissao com Base no Uso da Rede, Komite de Revitalizacao/MME: 23.
- PSR (2005). Allocation of Transmission Service Costs. Methods and Tools for Transmission Costing, Power Systems Research PSR. Web page:
<http://www.psr-inc.com.br/>: 25.
- Ramsey, F. (1927). "A contribution to the theory of taxation." Economic Journal **32**: 47-61.
- Rau, N. S. (1989). "Certain Considerations in the Pricing of Transmission Service." IEEE Transactions on Power Systems **Vol. 4**(No. 3): 1133-9.

- Reta, R. and A. Vargas (2001). "Electricity tracing and loss allocation methods based on electric concepts." IEE Proceedings on Generation, Transmission and Distribution **Vol. 148**(No. 6): 518-522.
- Riechmann, C. (2000). "Strategic pricing of grid access under partial price-caps-electricity distribution in England and Wales." Energy Economics **Vol. 22**(No. 2): 187-207.
- Rivier, M. and J. I. Pérez-Arriaga (1993). Computation and decomposition of spot prices for transmission pricing. 11th PSCC Conference, Avignon, France.
- Rubio, F. J. (1999). Metodología de Asignación de Costes de la Red de Transporte en un Contexto de Regulación Abierta a la Competencia. Doctoral Thesis. Escuela Técnica Superior de Ingeniería. Madrid (in Spanish), Universidad Pontificia Comillas.
- Rubio, F. J. and J. I. Pérez-Arriaga (2000). "Marginal Pricing of Transmission Services: A Comparative Analysis of Network Cost Allocation Methods." IEEE Transactions on Power Systems **VOL. 15**(NO. 1): 448-454.
- Rudnick, H., R. Palma, et al. (1995). "Marginal Pricing and Supplement Cost Allocation in Transmission Open Access." IEEE Transactions on Power Systems **Vol. 10**(No. 2): 1125-1142.
- Rudnick, H., R. Varela, et al. (1996). Evaluation of alternatives for power system coordination and pooling in a competitive environment. IEEE/PES Winter Meeting, Baltimore.
- Shirmohammadi, D., P. R. Gribick, et al. (1989). "Evaluation of transmission network capacity use for wheeling transactions." IEEE Transactions on Power Systems **4**(4): 1405-1413.
- Shirmohammadi, D., C. Rajagopalan, et al. (1991). "Cost of transmission transactions: an introduction." IEEE Transactions on Power Systems **Vol. 6**(No. 4): 1546-1560.
- Shirmohammadi, D., X. Vieira, et al. (1996). "Some Fundamental Technical Concepts about Cost Based Transmission Pricing." IEEE Transactions on Power Systems **Vol. 11**(No. 2): 1002-1008.
- Silva, E. L., S. E. C. Mesa, et al. (1998). "Transmission Access Pricing to Wheeling Transactions: a Reliability Based Approach." IEEE Transactions on Power Systems **Vol. 13**(No. 4): 1481-1486.
- Skokljev, I. and V. Maksimovic (2004). "Transaction allocation symbolic analysis method." European Transactions on Electrical Power(No. 14): 261-275.

- Stamtsis, G. C. and I. Erlich (2004). "Use of cooperative game theory in power system fixed-cost allocation." IEE Proceedings Generation, Transmission and Distribution **Vol. 151**(No. 3): 401-406.
- Strbac, G., D. Kirchen, et al. (1998). "Allocating transmission system usage on the basis of traceable contributions of generators and loads to flows." IEEE Transactions on Power Systems **13**(2): 527-534.
- Tan, X. and T. T. Lie (2002). "Application of the Shapley Value on transmission cost allocation in the competitive power market environment." IEE Proceedings on Generation, Transmission and Distribution **Vol. 149**(No. 1): 15-20.
- TP-NZ (1990). Transmission Pricing. Calculation of Network Charges, Trans Power NZ Limited.
- Turvey, R. (2000). "Infrastructure access pricing and lumpy investments." Utilities Policy **Vol. 9**(No. 4): 207-218.
- Usaola, J. (2005). "A transaction-based method for allocation of transmission grid cost and losses." Electric Power Systems Research **76**: 395-403.
- Vazquez, C., L. Olmos, et al. (2002). On the choice of slack bus in mechanisms for transmission network cost allocation that are based on network utilization. Power Systems Computation Conference, PSCC, Seville.
- Vlachogiannis, J. G. (2000). "Accurate model for contribution of generation to transmission system effect on charges and congestion management." IEE Proceedings Generation, Transmission and Distribution **Vol. 147**(No. 6): 342-8.
- WB (2000). Interconexiones Regionales de Mercados Eléctricos en Sudamerica. Propuesta para Facilitar el Incremento de los Intercambios Energéticos en la Región. Buenos Aires, The World Bank, ESMAP, Department of Energy US, CIER.: 151.
- Wei, P., Y. Ni, et al. (2002). "Load flow tracing in power systems with circulating power." International Journal of Electrical Power and Energy Systems **Vol. 24**(No. 10): 799-885.
- Wu, F. F., Y. Ni, et al. (2000). "Power Transfer Allocation for Open Access Using Graph Theory—Fundamentals and Applications in Systems Without Loopflow." IEEE Transactions on Power Systems **Vol. 15**(No. 3): 923-9.
- Yu, C. W. (2002). "Transmission embedded cost allocation using sensitivity analysis." IEE Proceedings Generation, Transmission and Distribution **Vol. 149**(No. 5): 627-632.

- Yu, C. W. and A. K. David (1997). "Pricing Transmission Services in the Context of Industry Deregulation." IEEE Transactions on Power Systems **Vol. 12**(No. 1): 503-510.
- Zolezzi, J. M. and H. Rudnick (2002). "Transmission Cost Allocation by Cooperative Games and Coalition Formation." IEEE TRANSACTIONS ON POWER SYSTEMS **VOL. 17**(NO. 4): 1008-15.

Annex 1: Mathematical properties of the participations of agents in the use of lines according to MP

The total participation of a generic agent i in the flow through line j is:

$$f_{i,j} = A_{i,j} \cdot (g_i - d_i) \quad (3.6)$$

where $A_{i,j}$ is the sensitivity of the flow through line j with respect to the power injection at node i and $g_i - d_i$ is the net injection at this node. If T_i represents the per-MW transmission charge for a generic agent i , and the cost C_j of each line j is allocated to all grid users in proportion to their total participation in the flow over that line, then net-work charges for agent i can be obtained as:

$$T_i \cdot (g_i - d_i) = \sum_j \left(C_j \cdot \frac{A_{i,j} \cdot (g_i - d_i)}{\sum_z A_{z,j} \cdot (g_z - d_z)} \right) \quad (3.7)$$

where z is a dummy index representing the different agents in the system.

As it is shown in (Rivier and Pérez-Arriaga 1993), changing the slack node modifies the sensitivity factors for line j in a fixed additional term, which is the same for all nodes. If X_j represents this fixed term, the total participation of agent i in the flow over line j can be expressed as:

$$f'_{i,j} = (A_{i,j} + X_j) \cdot (g_i - d_i) \quad (3.8)$$

And the total network charge for agent i turns out to be:

$$T'_i \cdot (g_i - d_i) = \sum_j \left(C_j \cdot \frac{(A_{i,j} + X_j) \cdot (g_i - d_i)}{\sum_z (A_{z,j} + X_j) \cdot (g_z - d_z)} \right) \quad (3.9)$$

Ignoring losses, and assuming that the system is in balance, implies that the sum of net injections by all agents is zero. Hence $\sum_z X_j \cdot (g_z - d_z) = 0$.

Then, the total payment faced by agent i can be rewritten as:

$$\sum_j \left(C_j \cdot \frac{A_{i,j} \cdot (g_i - d_i)}{\sum_z A_{z,j} \cdot (g_z - d_z)} \right) + (g_i - d_i) \cdot \underbrace{\sum_j \left(\frac{X_j \cdot C_j}{\sum_z A_{z,j} \cdot (g_z - d_z)} \right)}_K \quad (3.10)$$

The first term in equation (3.10) equals the total network charge obtained when using the original slack bus. The second term is the same for all agents i , except for the $(g_i - d_i)$ factor, which is common to both terms. This implies that, for every network user, a change in the slack node results in an additional term K in the per-MW transmission charge. This additional term is the same for all of the network users.

$$T'_i(g_i - d_i) = (T_i + K)(g_i - d_i) \quad (3.11)$$

Equation (3.11) is only true if negative participations are fully taken into account. If we consider only positive contributions, ignoring the negative ones, the property presented above does not hold.

The fixed term K that changes the unitary transmission price is multiplied by the net injection at each bus in order to compute new transmission network charges. Hence, whenever it is additive for any generator, it is additive for all of the generators in the system and subtractive for all of the demands, and vice versa. Thus, a slack bus located near the major load centre tends to increase total network charges paid by generators as well as reduce total transmission charges born by consumers. On the other hand, a slack bus close to a generation area will normally increase the share of total grid payments faced by the demand and will reduce the charges paid by generators. Therefore, choosing the slack bus is just a way of determining the global percentage of the network costs to be paid by producers and hence by consumers as well.

4 Access to the grid

4.1 Introduction

Access to the transmission grid is essential for generators to be able to supply power to loads. Traditionally, transmission networks have been owned and operated by vertically integrated utilities. These vertically integrated utilities could deny access to their transmission networks to third parties. In the new deregulated environment, open access to the transmission grid is a prerequisite for the development of competition among generators to supply power to loads. In other words, fair competition on the market place can only occur if market agents are not discriminated when accessing the grid.

Open access to the transmission grid implies that every agent should, in principle, have the right to access the transmission grid in the same conditions. In order to obtain the right to use the grid agents must pay the corresponding transmission tariff, which should grant them access to the entire network. However, the transmission grid may not be able to accommodate all the transactions resulting from the market. In this case, some network constraints become binding. Then, determining which market agents (or transactions) are allowed to undertake the transactions they want to engage in becomes an issue of utmost importance.

Network restrictions in a regional market may be of any of two types:

- System-wide congestion in the transmission grid resulting from the lack of enough transmission capacity in certain lines or corridors to accommodate all the commercially desirable transactions among the agents in the region.
- Local restrictions associated with the inability to inject or withdraw all the power produced or consumed in a certain local area (one node or a small set of neighboring nodes).

The aforementioned two types of network restrictions are dealt with separately since different mechanisms must be implemented for each one of them. Local restrictions within each area must be solved by the corresponding SO. Rules applied by the SO to solve this kind of restrictions must only affect the system locally. Local restrictions may exist both in those systems where the transmission grid is well meshed and in systems where the grid is underdeveloped.

On the other hand, system-wide congestion in the grid may be infrequent within the boundaries of national transmission systems and control areas, where the network is well meshed most of the times. However, the development of regional markets, where agents from different countries or areas trade their energy, may place significant stress on transmission systems. As said before, transmission systems have not been developed to support an important volume of international commerce. Thus, reinforcing the interconnection capacity among systems becomes necessary. However,

efficient management of the existing available capacity is the only option in the short term. What is more, some lines are likely to remain congested, since completely solving the congestion would probably imply carrying out investments that are not justified from an economic point of view (based on a cost/benefit analysis): network investment is an expensive means of solving concentration and market power problems that have a regulatory solution. Contrary to what happens with local grid restrictions, the solutions that must be implemented to deal with system-wide congestion either at national or regional level (though it is normally more significant in regional markets) tend to affect the dispatch of energy all around the system.

The chapter is divided into the following sections: section 4.2 describes the solutions that can be implemented to deal with local grid restrictions, section 4.3 discusses the management of the scarce capacity of congested lines and section 4.4 addresses some specific issues related to the management of congestion. Finally, section 4.5 concludes.

4.2 Treatment of local grid restrictions

The connection of unlimited generation or load in any point of the grid is not compatible with the safe operation of the system. Systematic local network constraints may limit the amount of power that can be injected into or withdrawn from the grid in a certain area. This section focuses on the mechanisms that are employed to deal with these restrictions. Sporadic local restrictions may be addressed using ad-hoc rules. Local restrictions should only have a local impact. In other words, in principle, they should affect neither the development nor the operation of the system in other areas.

This section is structured in the following subsections: Section 4.2.1 presents the connection policy to be applied for conventional, non-renewable, generation. Section 4.2.2 discusses the connection of renewable energy producers to the grid.

4.2.1 Regulation of the connection of conventional, non-renewable, generation capacity

Here we discuss methods to decide which agents should be granted access to the transmission grid in those areas where the interconnection capacity with other areas is scarce. Local restrictions resulting from the accumulation of generation in an area beyond the interconnection capacity of this area with others are still infrequent in most well developed systems. However, several facts may prompt the occurrence of such constraints. Generation from renewable resources and distributed generation have an increasing importance. Additionally, some deficiencies in regulation may undermine the coordination between the different parties that are involved in the process of investing in new transmission and generation facilities.

In most cases, this should be a temporary problem. The fact that the amount of generation capacity in an area is larger than the amount of power that can be effectively shifted out of this area should encourage the System Operator to install new lines in order to solve this transmission constraint. The case where part of the installed generation capacity is highly inefficient and should be substituted with more efficient generation may be an exception to this rule.

However, the System Operator may want to limit the amount of power that is produced at a given node because of reliability reasons. Otherwise, the system may run into stability problems if the generation in this node is lost.

This is an issue of growing concern for the System Operator in several countries like, for instance, Spain. Two alternatives exist when dealing with this problem:

1. One may limit the amount of generation capacity that can be installed in a node or area to the interconnection capacity of this node or area with the remaining ones. Alternatively, one can give priority of access to this interconnection capacity to the generators that were installed first. This is what is commonly called as 'capacity reservation'.
2. Alternatively, one may think of not restricting the installation of new generation, thus leaving any new and existing generators to compete locally.

As we have just said, implementing some mechanism of capacity reservation involves restricting the installation of new generation or giving priority to old generators over the new ones when being dispatched. Thus, the development of new more efficient generation in a certain area could be hampered by the existence of old inefficient generators. Advocates of applying a capacity reserve mechanism argue that new generators always have the possibility of installing in other areas or other nodes within the same area.

On the other hand, allowing new generation to compete on equal terms with existing generators located in the same area means that generators in areas or nodes with scarce interconnection capacity would permanently run the risk of being replaced in the dispatch by new generation installing there.

Having studied the pros and cons of both alternatives, we propose discarding the adoption of a capacity reservation mechanism. However, some measures may be necessary to soften the strict implementation of open access regulation.

First of all, the System Operator and power producers should exchange all the information relevant to know how the system will evolve in the future. An adequate level of coordination among them is essential in order to optimize the development of the system.

Thus, power producers must inform the System Operator about their plans to install new generation facilities. A system of financial guarantees should be in place whereby those market agents that apply for obtaining the permit to install a new power plant must pay the cost of the studies conducted by the System Operator in order to assess the feasibility of this connection. These guarantees should be reimbursed once the promoter has obtained the license to begin to operate the new generation facility. The unitary price of the financial guarantees an agent is required to pay should be higher the higher the number of applications for connection this agent has made. Otherwise, agents may apply for connection at several locations while they have no real intention of building new generation facilities in most of these locations. Unless the System Operator has a good estimate of the future installation of new generation capacity in the region, it will not be able to develop an expansion plan for the grid so that it adapts in time to the system requirements.

At the same time, the System Operator must inform agents of its future plans to reinforce the grid as well as of the interconnection capacity between any problematic area or node and the rest of the system. This information will be used by the agents to make their investment decisions.

These measures, albeit necessary, will probably not be enough to avoid local systematic grid restrictions. In other words, in some nodes or areas the installed generation capacity may be larger than the amount of power that the grid can absorb. Therefore, in addition to the preventive actions previously mentioned other corrective actions must be taken if necessary.

To begin with, whenever capacity payments are in place in an area, they must be limited to the amount of power that can effectively be transported to other areas or nodes. Capacity payments are employed to guarantee that enough generation capacity is installed at system level. As far as the security of supply is concerned, generators should not be discriminated on the basis of their production costs. This is due to the fact that capacity payments serve the purpose of avoiding load shedding and every generator must be in operation under emergency conditions. Thus, if the installed firm generation capacity in one node is 4000MW and the interconnection capacity with the rest of the system is only 2500MW, a generator with 400 MW of firm capacity at this node should only receive capacity payments for $400 \times 2500 / 4000$ MW of firm capacity.

Another possible measure, which is complementary of the previous ones, involves selling congestion rights over the scarce transmission interconnection capacity between an area or node and the rest of the system. Congestion rights or transmission capacity rights may be physical or financial. Physical transmission rights would allow its owner to inject a certain amount of power into the system at any time. The rights may be defined over a flowgate (flowgate rights) or among several points in the system (point-to-point rights), see section 4.3 in this chapter or (Pérez-Arriaga and Olmos 2005). Physical rights may influence the energy dispatch since they grant priority in the effective use of the physical transmission capacity. Hence, they may be the source of

inefficiencies in this dispatch if the rights holder does not behave in an “economically rational” manner. Financial rights, on the other hand, do not affect the physical dispatch, whose efficiency depends on the mechanism adopted to solve congestion in the grid.

Regardless of whether physical or financial rights are used, agents would buy them in order to be hedged against the risk of being constrained off by other generation. The capacity product bought and sold would depend on the nature of these rights. If rights are physical, generators buying them would be entitled to use the physical transmission interconnection capacity between the corresponding generation pocket and other zones. If rights are financial, generators buying them would be entitled to receive the corresponding congestion rents in the dispatch. As it will be explained in section 4.3, both products are very similar as long as enough liquidity exists in secondary markets of transmission capacity.

Implementing congestion rights requires including constraints in the energy dispatch that limit the amount of power produced in generation pockets. Hence, energy prices computed for these generation pockets may be different from those for the rest of the system. If congestion occurs on an interconnection, the energy price would be different on both sides of the interconnection and congestion rents would be different from zero. Thus, implementing congestion rights, even if only used for local systematic congestions, would affect the formulation of the energy dispatch in the region.

4.2.2 The connection of renewable energy producers to the grid

Some conventional energy sources that have played a central role in the last decades are close to be exhausted. This represents a major threat to the security of supply of power systems in the long term. At the same time, the impact of the production of electricity from fossil fuels on the environment is more and more evident with the passing of time. Nuclear power has its own serious problems. This has prompted authorities in many systems to promote the installation of power generation from renewable energy sources (RES) through the development of new regulation. However, increasing the production from some types of RES (those that are termed “non-dispatchable” such as wind, solar and mini hydros) must be made compatible with the security in the operation of the system. System Operators, local and regional governments and regulators and promoters of generation from RES must exchange information and coordinate.

Regulation of connection to the grid is central to encourage the installation of new renewable generation and autoproducers. Regulation should be as clear and stable as possible. It should give producers from non-dispatchable RES priority in the energy dispatch as well as in the access to the transmission grid, as long as the installation of new RES generation is compatible with the system security. The establishment of caps to the amount of RES capacity to be installed represents an obstacle to the increase in

the share of electricity produced from RES, which has been adopted in most systems as one of the major objectives in the medium and long terms.

Depending on the existing market rules, when systematic local grid restrictions exist and a generation pocket forms, renewable generation may have priority over conventional generation in the dispatch. In any case, RES should not be discriminated against. An order of precedence must be established for the different types of renewable generation, just in case the amount of RES generation is larger than the interconnection capacity of the generation pocket.

As a general rule, RES promoters should not install in an area more effective generation capacity with zero production costs than the interconnection capacity between the area and the rest of the system.

In the case of RES generators whose utilization factor is low, as it is the case of wind generators, more generation nominal capacity may be allowed to connect to the system in an area than the amount that can be effectively exported from this area. However, the System Operator should be provided with the means to regulate the output of these power plants.

As far as the allocation of connection costs is concerned, RES producers should be responsible for the cost of the transmission facilities that are totally devoted to connect them to the system. Objective, transparent and non-discriminatory criteria should determine the recovery of the cost of other grid facilities. The benefits obtained by market agents should be taken into account in this regard.

In order to facilitate the installation of new RES generation, tariffs should consider benefits or reductions in costs that the system may get from the installation of these generators.

4.3 Congestion management mechanisms in the short term and the long term

In this section, we investigate the allocation of scarce transmission capacity on lines where systematic system-wide network constraints exist that affect the dispatch at system level.

Not all patterns of commercial transactions can be allowed, since many of them result in unacceptable operating conditions because of a diversity of security reasons. The set of commercial transactions that is arranged by market agents, either bilaterally or in organized auctions, frequently result in line flows that exceed the transmission capacity of some lines. There is a limit to the amount of power that a line can dissipate in the form of heat without deteriorating. Consequently, the flow over the line is also

limited. Apart from this, there are cases where increasing the power flowing over a line, or a collection of them, beyond a certain limit could result in some unstable situation for the power system as a whole. Whenever the flow over a line cannot increase further without violating one or several restrictions, we say that this line is “congested”. However, transmission networks may impose other types of limitations, such as upper or lower voltage limits, which typically require that some generators that are located at specific points in the network have to operate. Maintaining the system in stable conditions may exclude particular generation patterns.

Although improperly, the term “congestion management” is frequently used to describe the ensemble of actions that are needed to meet the security criteria imposed by the System Operator. “Network constraint management”, or even better, “System constraint management” –since both generation and the network are involved in most constraints- is a more adequate expression. The situation of interest is the inability of the grid, because of the limited capacity of its lines, to accommodate all power transactions resulting from the unconstrained economic dispatch.

The analysis of the competitive case (i.e. when no market power exists or is exercised) is the central part of our study since it is by far the easiest to deal with. Hence, we will initially assume that no market power exists. In practice, market power is always present in power markets in some degree. Sometimes the level of horizontal concentration in the industry is very high. Other times the companies have interests in different activities and can take advantage of this fact to increase their profits by discriminating competitors. Both problems may occur at the same time. The effect that the existence of market power has on several specific aspects of congestion management will be briefly discussed as well.

We shall begin by presenting and comparing the different mechanisms that exist for the management of congestion in the short term, see section 4.3.1. First, organizational aspects are considered. Then, the influence of market power on the outcome of each one of these congestion management methods is briefly commented. Based on this analysis, we make specific recommendations about the method to implement in each one of several regional markets that presently are in the process of formation or have already started to function.

Under ideal conditions, many of these mechanisms are equivalent to carrying out a security constrained economic energy dispatch. In reality their outcome is different because existing conditions are far from being ideal. Additionally, different methods may yield different locational signals for generators and loads.

Most of the mechanisms that we shall discuss for the dispatch of energy in real time (or just ahead of real time) are compatible with the trading of capacity and energy products in the long and medium terms. Trading capacity rights as well as negotiation of contracts must take place some time in advance of the energy dispatch.

We shall focus on the definition of products that relate to the capacity of lines and corridors in the transmission grid, see section 4.3.2. This section covers several issues. Subsection 4.3.2.1 introduces the concept of capacity rights. Afterwards, subsection 4.3.2.2 presents and compares the different formats that congestion rights may adopt. A recommendation on the use of a specific type of rights is made. Subsection 4.3.2.3 explains the initial allocation of rights over the transmission capacity in the long term and the subsequent trading of these products in secondary markets. Both the design of the sequence of auctions leading to the energy dispatch and the determination of the amount of transmission capacity to be put up for sale in the different auctions draw our attention. Finally, subsection 4.3.2.4 briefly discusses the effect that the existence and allocation of congestion rights may have on the exercise of market power.

There are other aspects very relevant to the problem of allocating the scarce transmission capacity of the transmission grid. These are, among others:

- the allocation of congestion rents resulting from the energy dispatch or the sale of congestion rights,
- the definition of single price areas, i.e. areas where power injections and withdrawals may reasonably be priced at the same level without incurring into significant efficiency losses,
- the treatment of special grid restrictions and
- the definition of a balancing point for transmission contracts.

All these issues are introduced in section 4.3 but are analyzed more in depth in section 4.4, which complements section 4.3 focusing on very specific aspects of the management of congestion.

4.3.1 Discussion of different algorithms for the short term energy dispatch

Congestion management in transmission networks is one of the areas where the power system academic community has been more prolific recently, see (Chao and Peck 1996; Stoft 1998; Ruff 1999; Chao and Peck 2000; Tabors and Caramanis 2000; Boucher and Smeers 2001; ETSO 2001; Henney 2002; Hogan 2002; O'Neill, Helman et al. 2002; ETSO 2004; ETSO/EuroPEX 2004) as a sample of recent relevant literature on the subject. Despite this, the subject cannot yet be considered closed, particularly in the context of large multi-national markets, such as the Internal Electricity Market (IEM) of the European Union or the Standard Market design (SMD) in the US. Any plausible approach must comply with sound engineering and economic requirements but it must also be politically acceptable. This imposes further constraints and requirements, such as the need for it to be compatible with the diversity of regulatory regimes and

market structures that already exist in the region, or the need to be easy to understand by decision makers and to implement by System Operators.

This section describes and critically analyses realistic alternatives for congestion management in regional markets. The objective is to compare and contrast all the models. Afterwards, we provide specific recommendations for the effective implementation of a short-term congestion management algorithm at regional level. The method finally chosen may not be the fanciest and more accurate model since this would not be feasible in some cases. In addition to implementing an efficient and practicable scheme to allocate the scarce transmission capacity, SOs should make use of any tool at their disposal to improve the management of congestion. In this sense, authors in (Deladreue, Brouaye et al. 2003) establish statistical relationships between the flows in a system. Knowing these relationships may help the SO anticipate how flows in the system will change when they modify the dispatch to alleviate congestion on some lines.

The considered set of alternatives to allocate capacity includes a full fledged centralized nodal pricing scheme encompassing the entire region. Nodal prices of electricity were developed by Schweppe et al., see (Schweppe, Caramanis et al. 1988). Afterwards, Rivier and Pérez-Arriaga propose a meaningful decomposition of spot prices in their generator and network components, see (Rivier and Pérez-Arriaga 1993). Afterwards, others have decomposed nodal prices in more general components, see (Chen, Suzuki et al. 2002) among others. The generalized use of system-wide nodal prices for the short-term energy market is very far from being acceptable in some cases like that of the European regional electricity market, where the current regulatory and political contexts require the adoption of less ambitious design options. Therefore, we need to explore some other approaches that are not the most efficient ones but are at least viable. In any case, full nodal pricing remains as the ideal reference.

This pragmatic approach resembles efforts to find an acceptable scheme for a Standard Market Design in the USA, see (FERC 2002). One has to keep in mind that a Regional Transmission Organization (RTO) in the USA may be of the same electrical size as the average European country. The regulatory diversity is very large among the IEM countries as well as among different regions in the USA. Implementing a more harmonized and compact regulatory scheme of the transmission activity is easier in smaller multinational markets, such as the Central American Electricity market, see (KEMA 2002).

In those cases where market agents need to decide where to address their energy bids, the publication of information by the System Operators and other institutions in the different areas or countries is essential so that agents are aware of the existing trading opportunities. Traditionally, SOs have been in charge of the computation and publishing of Transfer Capacities (TCs) between areas. These are a set of numbers meaning the amount of power that can be transacted between each pair of areas without violating any network restriction. Of course, the transfer capacity between two

areas A and B depends on the power exchanges taking place between other areas in the system. Thus, some standard conditions must be established for their computation. The academic literature contains many examples of methods to compute TCs. Thus, (Hamoud 2000) takes into account the economic merit order when simulating an increase in the power exchange between two areas or countries in order to compute the corresponding ATC. Authors in (Ou and Singh 2002) describe several methods for the assessment of ATCs that internalize reliability margins. ATCs may be dependent on a number of aspects like the pattern of generation and load within each area. Congestion management methods that rely on the allocation of the ATC between areas are less accurate but very simple to design and apply. Some algorithms take into account reactive power flows that also contribute to the congestion in the grid, see (Grijalba, Sauer et al. 2003). Others use novel techniques such as those based on the fuzzy logic, see (Khairuddin, Ahmed et al. 2004).

4.3.1.1 Terminology and approach classification

First, we present some very simple methods presently used for congestion management. Contrary to the more developed ones, these are not market based methods, in the sense that transmission capacity is not allocated based on the willingness of market agents to pay for it. The strong point of these methods lies in the fact that they are easy to understand and apply.

Some methods are based on priority lists. That is to say, when allocating scarce transmission capacity, some agents are given priority over others that are based on some criteria other than their economic offer. Probably, the most representative example of this family of methods is the capacity allocation carried out on a first come first served basis. Those agents that first bid for capacity are the ones that get it. Capacity is assigned until no more is left. Despite being very simple, this mechanism is conceptually flawed since it does not promote the efficient use of transmission capacity. This method is used to manage congestion on the borders between France and Spain, France and Germany, France and Belgium and France and Switzerland. It is also used internally in Austria as well as between Austria and Germany or Switzerland, see (Consentec/Frontier 2004).

In other methods, the volume of capacity that agents ask for determines how much capacity they obtain. This methods are normally referred to using the term “pro-rata assignment”. Capacity is allocated proportionately to the amount of it that each agent requests. These are again inefficient methods, since capacity may end up in the hands of agents that will not engage in the most valuable transactions. This method is employed to manage congestion between France and Spain, France and Italy, France and Germany, France and Belgium and France and Switzerland. It is also used on the borders between Italy and Greece, Italy and Slovenia and Italy and Switzerland, see (Consentec/Frontier 2004).

Despite being very simple and having been actually implemented in many cases, we must discard now these methods because they do not result in an efficient use of the

transmission capacity. What is more, these methods can easily be gamed by market agents willing to withhold transmission capacity in order to take advantage of the market power they have in an area or country.

We discuss now market based methods. These are more refined methods that aim to optimize the use made of the transmission capacity. In a first attempt, methods of this type can be classified according to whether they are coordinated or not, i.e. whether the allocation of capacity in all congested lines takes place at the same time or not. In principle, coordinated methods are superior from an efficiency point of view, though they require a higher degree of cooperation among TSOs or areas.

Alternatively, methods can be classified based on the interface between short term energy and capacity markets. Figure 4.3-1 represents a system comprising 10 nodes and 3 different control areas that has been made up in order to explain each one of the approaches that are described in the following paragraphs.

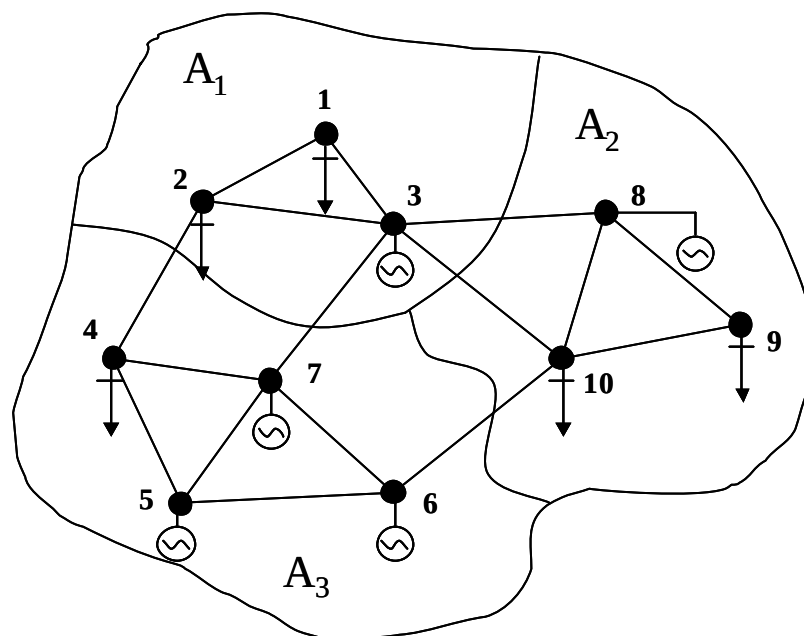


Figure 4.3-1: 10 node system comprising 3 different control areas: A1, A2 and A3.

Implicit auctions

Implicit auctions is a general term that applies to any congestion management method where energy and capacity are traded simultaneously. The effect of network constraints and losses is internalized in the prices of energy. Nodal pricing, see (Schweppe, Caramanis et al. 1988; Hogan 1992), also termed Locational Marginal Pricing (LMP), is the ideal implementation of this principle. In LMP energy prices differ from one node to another at a given moment in time. Nodal prices happen to be perfect short-term economic signals. According to this approach, every node of the system in Figure 4.3-1, from node '1' to node '10', would have a different energy price.

Nodal pricing may be seen as the outcome of a joint competitive auction of energy and rights to use the transmission capacity. See (O'Neill, Helman et al. 2002) for an example of implementation of an contingency constrained auction for both energy and transmission rights where the authors consider both options and obligations. Auctions proposed in (O'Neill, Helman et al. 2002) are different from other designs of implicit auctions in the sense that authors propose using them both in the short and the long term.

The academic community has come up with several designs to run implicit auctions in a decentralized manner. Thus, (Aguado, Quintana et al. 2004) decomposes the original problem into several simpler ones. The optimal outcome at regional level is found through an iterative process.

Apart from that, Stoft provides a design of the implicit auctions where only the energy price at a central hub and the capacity prices of several flow-gates are considered, see (Stoft 1998). This would allow bilateral trading of rights over the transmission capacity of these corridors. At the same time, prices to be taken into account may probably be fewer.

Hogan concludes in (Hogan 2002) that only the security constrained economic dispatch provides efficient prices that give agents more flexibility to engage in contracts.

A simplification of nodal pricing, only legitimate when the pattern of network utilization meets certain conditions, is zonal pricing, also termed market splitting. In this regard there are different opinions as to what market splitting refers to. The implementation of market splitting, according to the interpretation of this term adopted here, implies the definition of internally well-meshed areas which can be considered as super nodes for congestion management purposes. Market splitting is applied within the Nordic region and in Italy. If necessary, TSOs in the NORDEL system may move power among certain areas, defined a priori, in order to solve congestions arising among these areas. This implementation of market splitting agrees with that of many others in the academic literature and the industry (see (ETSO 1999) and (1999; Newbery, Damme et al. 2003)).

Nodal energy prices within any of the predefined areas may probably be very similar if losses are ignored and serious congestion is limited to the interconnections between areas. In this case, areas can be regarded as "Single Price Areas" (SPAs) as far as congestion management is concerned (Christie and Wangenstein 1998; Stoft 1998; Chao and Peck 2000).

Thus, if in the system of Figure 4.3-1 we define Single Price Areas that are coincident with existing control areas (A1, A2 and A3), then all the nodes within each area (nodes 1, 2 and 3 corresponding to area A1, for example) would have the same energy price.

Balanced transactions within a SPA should not significantly affect the flow over the inter-area links. In other words, any bilateral transaction within a SPA should not create loop flows outside this area which may significantly contribute to congestions inter single price areas. If we refer to the system in Figure 4.3-1, a transaction taking place between nodes 1 and 3 in single price area A1 must not modify significantly the flows over structurally congested lines. The definition of single price areas, whenever applicable, is not a trivial matter without practical consequences, see (Boucher and Smeers 2001). In fact, it will affect the validity of the results of congestion management. Intra single price area transactions need not be subject to coordinated congestion management procedures. Hence, these transactions may be thought as having priority in the dispatch over intra-area transactions.

The geographical pattern of Single Price Areas may not respect the frontiers between control areas. Each control area is normally under the jurisdiction of a different Transmission System Operator (TSO). Talking about multinational markets, like the IEM in the EU or the Central American Market, the footprint of each control area often coincides with the political borders of a Member State. For the sake of simplicity, we will assume that each Member State has one or more control areas, each one under a different TSO. Each control area may comprise several single price areas, which are defined just for the purpose of congestion management. This can always be achieved by conveniently subdividing any Single Price Area if necessary. However, assuming SPAs that are the same as existing control areas or countries may result in an inefficient dispatch or, even worse, in one that is far from being feasible. Thus, the proposal by ETSO and EuroPEX for an implicit auction between Power Exchanges in Europe combined with the allocation of transmission capacity to bilateral transactions should be reconsidered, see (ETSO/EuroPEX 2004). In addition, both institutions claim that a simple iterative process would suffice to achieve coordination between PXs (areas). Hence, running a centralized auction at European level would not be necessary according to (ETSO/EuroPEX 2004).

Redispatch and counter-trading

Running the energy market as if no transmission constraints existed may be acceptable in power systems where network congestions are not serious or frequent. Once the outcome of the energy market is known, one can check whether the pattern of commercial transactions violates any network constraint. Only when a constraint is violated does the System Operator need to re-dispatch some generation. Several implementations of re-dispatch are possible. According to some of them, the cost of the re-dispatch carried out to solve any violation of the network constraints should be minimum, see (1999; Rau 2000; Tao and Gross 2002). In these cases, market-based mechanisms must be used to modify the pattern of generation in the system. In other words, changes to the dispatch must be based on the bids sent by market agents indicating how much they ask for in order to change their market positions. Other re-dispatch algorithms aim to minimize the number and size of the adjustments to the

original dispatch, see (Galiana and Ilic 1998; Alomoush and Shahidehpour 2000). Authors in (Fang and David 1999) describe other possible schemes.

Algorithms proposed in (Alomoush and Shahidehpour 2000; Biskas and Bakirtzis 2002) are aimed at re-dispatching generation and load in the context of regional markets. These algorithms must achieve coordination between the different zones. Thus, (Biskas and Bakirtzis 2002) decomposes the original problem using Lagrangian relaxation techniques. The coordination variables are the prices of the power exchanges between zones.

Counter-trading is a specific implementation of the method of re-dispatch. In the counter-trading method the System Operator nominates pairs of generators that modify their outputs to create a power flow that goes in the opposite direction to the one causing the congestion. If the line between nodes 6 and 10 is congested in the direction from 6 to 10, the SO may schedule an increase in the power production at node 8 and an identical reduction at node 6 so that the flow resulting from the bilateral transaction between both nodes eliminates the congestion.

Obviously, one could generalize and say that re-dispatch is nothing but counter-trade, since any increase in the output of a generator has to be matched by a corresponding and identical (except for losses) reduction in the output of another generator. Typically, the extra cost of re-dispatch or counter-trade is socialized to all consumers. In this case, any economic signals resulting from the management of congestion, which could have been used to emulate nodal pricing, are lost. Conceptually speaking, assigning the cost of re-dispatch to those market agents that “create” the network constraint is possible. Economic signals would thus not be completely lost. This is a technically complex task, nevertheless, see (Rivier and Pérez-Arriaga 1993) and others. Authors in (Tao and Gross 2002) allocate the cost of re-dispatch taking into account the participations of agents (injections and withdrawals considered separately) in the flow over the congested lines. In order to do this, they express the flow over the congested lines as a function of power injections and withdrawals. Similarly, (Baran, Babunarayanan et al. 2000) determines the participation of each transaction in the flow over congested lines. Afterwards, the total cost of re-dispatch is allocated among congested lines taking into account both the marginal cost of the restriction on the flow through each congested line and the incremental cost of the re-dispatch necessary to avoid violating this restriction.

Experience with counter-trade in California shows that those schemes based on re-dispatch may be subject to gaming by market agents who artificially create congestion in the grid in order to be paid afterwards to remove it. In any case, nodal pricing or implicit auctions seem to be superior to congestion management mechanisms based on re-dispatch. In (Singh, Hao et al. 1998) Singh et al. compare nodal pricing to a mechanism based on decentralized bilateral trade among market agents, followed by the minimum cost re-dispatch necessary to solve infeasibilities. They conclude that price signals resulting from nodal pricing are more efficient, unless the cost of re-

dispatch is efficiently allocated to the agents responsible for congestion in the grid. However, as we have explained before, efficiently allocating the cost of re-dispatch is not an easy task nor is there an indisputable way of doing it.

Explicit auctions

The third approach considered here assigns energy and transmission capacity in inverse order of re-dispatch, namely, network capacity first and energy afterwards. Prior to running the energy market, market agents must obtain the right to use the transmission capacity they need to carry out the commercial transactions they want to get involved in. Agents may buy this capacity (the right to use it) in a centralized explicit auction. In this case, every transaction that might be using transmission lines likely to get congested has to participate in this capacity auction. The right to use the transmission network is allocated to those agents who value it most. Transmission rights to be acquired by market agents may relate to the capacity of flow-gates or corridors. These rights must be initially allocated to market agents in some way. Afterwards, trade in these rights may take place in a decentralized manner, i.e. agents may negotiate bilaterally the acquisition of these rights. In this case, running a centralized auction for transmission capacity would not be necessary.

Chao and Peck were the first ones to propose the utilization of flow-gate rights that might be acquired through bilateral trading. Once agents had acquired the flow-gate rights required to engage in commercial transactions, an unconstrained energy dispatch would take place where unbalanced bids to inject or withdraw power would be matched. Only those agents that had acquired flow-gate rights over the transmission capacity they use to market their power would be allowed to enter into commercial transactions with others, either through bilateral transactions or through a centralized energy auction, see (Chao and Peck 1996). In this article, authors demonstrate theoretically that this system would converge towards an efficient equilibrium.

Similarly, Oren and Ross propose in (Oren and Ross 2002) an auction for flow-gate rights prior to the energy dispatch. Again, only agents having acquired the necessary rights could take place in the energy auction. Authors propose a system where SOs responsible for the energy dispatch in the different control areas would coordinate to manage the flow on the congested lines that is the responsibility of transactions taking place within different areas. There are other works on the use of flow-gate rights in combination with unconstrained energy markets, see (Tabors and Caramanis 2000) for an example.

Once transmission capacity rights have been assigned in one way or the other, the energy auction takes place. Only those transactions that have acquired capacity rights to access the congested transmission they use can participate in the energy market. Therefore, if the generator at node 8 wanted to sell its energy on a power exchange whose balancing node is located at node 4, it would first have to participate in the

explicit capacity auction to acquire the transmission rights needed to access the aforementioned PX or buy the corresponding flow-gate rights directly to their owner. On the other hand, a bilateral transaction between nodes 1 and 3 would not have to acquire any transmission rights due to the fact that both nodes are in area A1. Auctioning transmission capacity at regional level requires some centralized coordination, see (ETSO 2001). If flow patterns due to the different transactions were not considered jointly they might result in unexpected violations of network constraints unless significant security margins were applied. But employing security margins would most likely result in an underutilization of the transmission grid.

Similarly to what happens in implicit auctions, the definition of SPAs becomes a matter of utmost importance when dealing with capacity and energy separately. If we were not able to define SPAs, separating the allocation of energy and capacity would not be possible. The final unconstrained energy dispatch would have to replicate exactly the outcome of the previous capacity auction (or that of the bilateral trading process taking place among agents to buy and sell capacity rights). Given that any transaction would significantly affect the flow through the congested lines, it would have to be taken into account when allocating transmission capacity.

In those systems where explicit auctions are used, local authorities are in charge of the dispatch of energy within their corresponding areas. Thus, areas or countries enjoy a high level of independence. For this reason, capacity auctions have been widely applied in real life power systems. This is the method used to manage congestion on the borders between Austria and the Czech Republic, Belgium and the Netherlands, Denmark and Germany or France and the United Kingdom, among others, see (Consentec/Frontier 2004).

Both implicit and explicit auctions may take place locally, i.e. affecting only one country or interconnection, or at system level. A coordinated implicit auction involves all the countries or control areas in the region. The dispatch in any area depends on that in the neighboring ones. Correspondingly, transmission capacity in a coordinated explicit auction is assigned on all the inter-connections at the same time. Due to the fact that Kirchhoff laws rule the distribution of flows over the grid, any transaction between two nodes has an impact on the line flows in the rest of the system. Strictly speaking, every transaction in the system makes use of the capacity of congested lines. Therefore, all transactions should be managed in a jointly manner. Only if the definition of SPAs makes sense may it be acceptable to deal with transactions internal to any of these areas independently from those transactions taking place outside the area. In any case, transactions between different areas must be dispatched all at the same time. If the error committed when defining SPAs is not acceptable, nodal pricing would be the only efficient alternative left. Thus, computing SPAs becomes an issue of utmost importance when systematic congestion exists as it is the case in most regional markets in the process of creation. Section 4.4.3 discusses the creation of SPAs.

The academic community has proposed several mechanisms to carry out a decentralized dispatch in the presence of congestion in the grid. As mentioned above, the most popular of them is based on the bilateral trade of physical transmission rights over the capacity of flow-gates, as proposed in (Chao and Peck 1996; Tabors and Caramanis 2000). However, other mechanisms have been proposed as well, whereby agents negotiate bilaterally power exchanges between them. The joint feasibility of the resulting transactions is achieved through the exchange of information between these agents and the System Operator. See (Wei, Ni et al. 2002; Ilic, Hsieh et al. 2003) among others.

As explained above, network constraints that only have a very local effect are not considered here. System Operators tend to resort to local measures to solve them. Applying market-based mechanisms to solve local constraints is counterproductive in general. Local voltage problems, whose solution typically involves starting up a particular generator or one among a very small number of them (often belonging to the same company), are a clear example of this. Agents who solve a local constraint should be given a previously agreed administrative payment. TSOs need not coordinate with others to address limits imposed by the grid at local level. The study presented here takes for granted that actions have been taken to deal with confined network problems.

The following subsections present, discuss in detail and compare different congestion management schemes. Losses have been ignored in all of them since this is not the focus of this chapter. In any case, we should bear in mind that nodal prices are the ideal reference model for short-term signals. Therefore, the combination of all the economic signals from the system operation (energy, congestion and losses) should be roughly equivalent to the use of full nodal prices. Every agent should pay the marginal losses caused by a unit transaction between the node where it is located and that where this agent buys or sales its energy. Hence, the reference node used to compute marginal losses must be the same as that used for congestion management purposes. For a study of the decomposition of nodal prices into their energy, losses and congestion terms, see (Rivier and Pérez-Arriaga 1993).

All the implementations of a congestion management scheme discussed here consider the existence of both bids for the capacity that is needed to execute a bilateral contract and unilateral bids to buy energy from or to sell energy to a power exchange. Each market agent has the possibility of choosing the type of transactions that better fits its needs. Thus, the resulting market is more flexible than those that only allow for one type of transactions or the other (Singh, Hao et al. 1998).

After analyzing different congestion management schemes, we provide specific recommendations over the adoption of an scheme in several regional markets. Section 4.3.1.2 presents the mathematical formulation of the implicit auction scheme. Despite the fact that nodal or zonal implicit auctions (if the definition of SPAs is possible) are the best design option conceptually speaking, implementing an auction of this kind may not be viable in all the regional markets under consideration. Taking the implicit

auction formulation as a reference, section 4.3.1.3 discusses the use of a coordinated explicit auction of transmission capacity followed by local energy auctions where grid effects are not considered. Section 4.3.1.4 analyses the interaction between the use of the previous two types of short-term auctions and the exercise of market power by agents in the region. Finally, section 4.3.1.5 discusses the implementation of a short-term congestion management scheme in different regional markets.

4.3.1.2 Formulation of the implicit auction scheme

A coordinated implicit auction where transmission capacity allocation is fully integrated with the energy short-term market seems to be the best solution from an efficiency point of view. In addition, according to some publications, this mechanism may be preferable to others in order to mitigate the exercise of market power, see (Singh, Hao et al. 1998; Neuhoff 2002) among others. We shall refer later to some practical difficulties that may arise when implementing this method in the context of some regional markets like the IEM.

The central part of this scheme is an IEM-wide auction that may take place one day ahead of real-time operation. In this auction both capacity and energy bids are accepted for any given hour:

- Bids in €/kWh from individual agents to buy or sell energy. Agents do not specify on which power exchange they want to trade their energy. These are “unbalanced” or “one-sided” bids.
- Bids in €/kW to acquire the transmission right to buy energy from a certain geographical zone and sell it in another specific zone during the considered hour. These are “balanced” or “bilateral” bids.

In principle, all market transactions have to participate in the auction since all of them have the potential to create network constraints. However, if existing conditions make the definition of SPAs possible, bilateral intra-SPA transactions do not need to be considered in the formulation of the model. In the latter case, we assume that the corresponding TSOs redispatch power transactions in order to remove any congestion internal to a SPA. SOs are also in charge of solving any minor inter-SPA congestion caused by the assumption that intra-SPA transactions have no effect outside the corresponding SPA. Using an accurate model of the system grid is necessary to compute the impact that each possible CBT transaction may have on the transmission network flows. The accuracy of the results may improve if the TSOs provide each other with an estimation of the internal patterns of generation and demand within each SPA¹.

¹ One argument frequently employed in the IEM context, despite being widely criticized in many instances, is that the use of Single Price Areas is preferred to the use of all network nodes because of confidentiality reasons. Agents request that nodal injections and withdrawals should

Based on the information provided by market agents and System Operators, an optimization algorithm determines the set of bids compatible with the network constraints that maximizes global welfare. Only these bids are accepted, whereas the remaining ones are discarded. This certainly constitutes a level playing field for capacity and energy bids.

The following subsection presents a concise mathematical formulation for this implicit auction. Despite being efficient, this model, which will be termed Model I from now on, may be difficult to implement in the context of some regional and national markets. Thus, we develop Model II in another subsection. In some regional markets Model II may attain more practicability than Model I at the expense of making some efficiency concessions. However, it may still not be viable in some markets. Lastly, we describe other alternative implementations of the implicit auction that seek to combine efficiency and additional practicability.

Mathematical formulation of Model I

The formulation basically corresponds to that of an optimal power flow, see (Bartholomew, Siddiqui et al. 2003). For the sake of simplicity, we will only consider here congestion on transmission lines. Power Transfer Distribution Factors (PTDFs) quantify how much a per unit transaction i between two nodes or areas (or a transaction corresponding to a one-sided bid of an agent j at a power exchange) changes the flow on every network line k . Baldick demonstrates in (Baldick 2003) that Power Transfer Distribution Factors are invariable with the system operating conditions. According to this work, PTDFs only depend on the topology of the network. We assume that a reasonable number N of Single Price Areas exists in the region. In this case, the potential number of sets of PTDFs to be computed is $N(N-1)/2$, where N is much smaller than the total number of nodes. Each set of PTDFs corresponds to a different pair of Single Price Areas and consists of one PTDF for each line or corridor of interest for congestion management. If the number of SPAs were very high a nodal formulation of the problem would be preferable to the formulation based on PTDFs. In this case we may adopt a schematic representation of the system where intra-area networks could be somewhat simplified². Confidentiality

be aggregated into SPAs so that individual transactions could not be easily identified when collecting the data that are needed for congestion management. Experience exists, however, of the utilization of detailed network models where every node is represented (in PJM for instance). Agents in PJM do not seem to be concerned about the fact that submitting their offers at node level reveals more information on their commercial transactions.

² A formulation that is based on PTDFs has been used throughout this document. The main ideas, however, also apply to a formulation that is based on a complete load flow model for the entire network (see the corresponding subsection below). In this case, and for confidentiality reasons, individual injections and withdrawals taking place within a SPA could be somehow distributed pro-rata among the different nodes of the SPA. Strictly speaking, this formulation

reasons may lead to the modification of the actual load and generation patterns within each SPA. This should not cause further problems as long as intra-area transactions do not significantly affect the flow over congested links. However, if SPAs of enough size do not exist, preserving the confidentiality of commercial transactions may be incompatible with an efficient dispatch. Computer models for large-scale systems are available to perform these tasks, see for instance.

Market agents bid with the objective of maximizing their profits. Assuming a competitive market, agents' offers reflect their marginal production costs in the case of producers and the marginal benefit they obtain from electricity in the case of consumers.

Evaluating the impact on the network flows of each individual unbalanced transaction requires defining a slack or reference node as the counterpart in the transaction. In the formulation below a single slack node responds to all the unbalanced transactions since all of them are cleared at a single central clearinghouse. The implications of the choice of the slack bus will be discussed later.

Model I

$$\max_{p_{bi}, p_{uj}} \sum_i B_{bi} p_{bi} + \sum_j B_{uj} p_{uj}$$

s.t.

$$-\bar{F}_k^- \leq \sum_i PTDF_{kbi} p_{bi} + \sum_j PTDF_{kuj} p_{uj} \leq \bar{F}_k^+ \quad , g_k \quad \forall k = 1, 2, \dots, \text{number of corridors} \quad (4.1)$$

$$\sum_i PTDF_{kbi} p_{bi} + \sum_j PTDF_{kuj} p_{uj} \leq \bar{F}_k^+ \quad , m_k \quad \forall k = 1, 2, \dots, \text{number of corridors} \quad (4.2)$$

$$\sum_j p_{uj} = 0 \quad \text{global balance equation} \quad (4.3)$$

$$p_{bi} \leq \bar{P}_{bi} \quad , b_{bi} \quad \forall i = 1, 2, \dots, \text{number of balanced transactions} \quad (4.4)$$

$$p_{uj} \leq \bar{P}_{uj} \quad , b_{uj} \quad \forall j = 1, 2, \dots, \text{number of unbalanced transactions} \quad (4.5)$$

would require a full network representation but a simplified grid model may be adopted. The size of this system does not depend on the number of SPAs or the number of potentially congested lines. Besides, it can accommodate the formulation of complex system network constraints more easily, see section 4.4.

$$p_{bi} \geq 0, a_{bi}, a_{uj} \quad \forall i, j \text{ balanced and unbalanced transactions} \quad (4.6)$$

where:

i : set of balanced transactions (contracts) between agents located within different SPAs.

j : set of unbalanced transactions.

k : set of congested lines or corridors.

B_{bi} : bid price (€/kW) per kW of capacity in the considered hour offered by bilateral transaction i .

B_{uj} : bid price (€/kWh) offered by agent j per kWh of energy bought or sold at the power exchange in the considered hour. It is positive if j is a load and negative if it is a generator.

p_{bi} : scheduled power for the balanced transaction i .

p_{uj} : scheduled power purchased (respectively sold) from a power exchange by agent j (j is involved in an unbalanced transaction). This variable is positive if j is a load (purchases electricity) and negative if it is a generator.

$PTDF_{kbi}$: power transfer distribution factor of balanced transaction i with respect to line k .

$PTDF_{kuj}$: power transfer distribution factor of unbalanced transaction j with respect to line k when the single slack node previously chosen for the entire system acts as the counterpart.

$\overline{F}_k^+$: maximum flow allowed over line k in the ‘positive’ direction. Its value must internalize any security criteria that might be considered by the system operator (as a result of a contingency analysis, for instance)³.

$\overline{F}_k^-$: maximum flow allowed over line k in the ‘negative’ direction.

$\overline{P}_{uj}$: maximum quantity that the agent j is willing to generate or consume.

³ This may result in conservative, and therefore inefficient, estimates. Alternatively, additional constraints corresponding to contingency situations may be included. The alternative formulation that is commented in a previous footnote may help here, see (Hogan 2002).

$\bar{P}_{bi}$: contracted power for balanced transaction i . In other words, this is the maximum amount of physical point-to-point rights that balanced transaction i is willing to acquire.

Discussion of the model

The objective function to be maximized in Model I guarantees that those individual agents or transactions who value most the utilization of the network have access to it, regardless of whether they try to execute a bilateral contract or to bid on a PX. Welfare maximization is therefore achieved.

The set of equations (4.1) and (4.2) limit the flow on every line of the system (only those lines that might become congested need to be considered). All inter-SPA transactions have been expressed in a bilateral format. The slack bus associated with the central auction house acts as the counterpart in one-sided transactions. Equation (4.3) keeps the balance between bids to buy and sell energy.

Equations (4.4) and (4.5) limit the amount of power that each transaction is assigned. Finally, (4.6) forces scheduled quantities of balanced transactions to be non-negative. Losses have not been taken into account in this streamlined formulation.

The amount paid by each transaction to access the transmission network depends on the amount of power allocated to this transaction, the PTDFs of the flows on all congested lines with respect to this transaction and the dual variables g_k and m_k of the constraints limiting these flows (equations (4.1) and (4.2)). The mathematical expression of the congestion charge that a bilateral transaction i has to pay may be decomposed into several terms each of them corresponding to a congested line k :

$$C_{ki} = (g_k + m_k) PTDF_{kbi} p_{bi} \quad (4.7)$$

Equation (4.7) implies that the selection of a slack node as a counterpart for unbalanced transactions does not influence which bids are accepted in the auction. Accepted bids may be paired (arbitrarily) as bilateral transactions where the slack node does not have any function to play. However, defining a balancing node for unbalanced transactions is necessary due to the fact that Model I makes use of PTDFs, which can only be computed for balanced transactions. Hence, some node or nodes must play the role of responding node in unbalanced transactions. The location of the slack node determines the PTDFs of the flow on each congested line k with respect to each transaction j . Therefore, it affects the corresponding congestion charge, whose expression is:

$$C_{kj} = (g_k + m_k) PTDF_{kuj} p_{uj} \quad (4.8)$$

One may feel uneasy about the fact that the arbitrary choice of the reference node determines the congestion charge of unbalanced transactions. The balancing node may be placed at the physical location of the central clearinghouse although many other

locations are also possible. However, this effect is counteracted by the difference among energy prices in different SPAs resulting from congestion. For instance, assume that a generator g presents an unbalanced bid j to the central auction house. Assume this bid is accepted and paid the corresponding marginal price, which is the price of the SPA where the slack node is located. Generator g also has to pay the capacity charge required to access that SPA. The overall result is equivalent to the generator selling its output locally and being paid the energy price of the SPA where it is physically connected. Energy prices in the different SPAs (i.e. the energy price at the reference node plus the corresponding transmission capacity charge) are equivalent to those obtained under a zonal pricing scheme. Agents involved in bilateral transactions must estimate zonal prices when shipping power from one area to another. As a consequence of all this, the choice of the reference node does not have any practical implications for the physical dispatch of generation and the net remuneration of each generator.

Revenues of the TSOs and those of the owners of regulated lines must not depend on the net outcome of the management of congestion in the grid. Article 6 of the European regulation for the IEM of the EU, approved by the EU Commission and Parliament, bans this use of congestion rents, see (EP 2003). Due to the monopolistic nature of the transmission activity, both System Operators and transmission owners (with some rare exceptions, as it is the case of merchant lines) are regulated entities subject to price control. Under ideal conditions (if economies of scale were not present in the planning of the grid and investments were continuous) revenues resulting from the application of nodal prices would suffice to pay the construction and maintenance of the grid. In real-life systems congestion rents only cover a small part of these costs (around 20% according to some estimates in (Rubio and Pérez-Arriaga 2000)). Congestion rents should be devoted to either of two uses: they may reduce the authorized revenue of transmission owners to be recovered via network tariffs. Alternatively, these rents could finance new grid investments aimed at relieving congestion.

In the first case, congestion rents may be allocated among lines in the system in one of two different ways. We may use these rents to reduce network access tariffs line by line thus devoting the difference in prices between both ends of a line times its flow to reduce transmission charges paid by agents that otherwise would have to afford the whole cost of the line. Otherwise, they can reduce network access tariffs in the system as a whole. Section 4.4 discusses the choice between these two alternative uses of congestion rents.

Note that the use of Model I for congestion management is not in conflict with the existence of shorter-term energy markets, such as intra-daily markets, regulation markets or markets for operating reserves. Market agents may use these markets to better adjust their positions to their actual interests or possibilities in real time or to participate in the provision of ancillary services.

Model is simple and it ensures that congestion is managed with maximum efficiency (acknowledging the approximations involved in the definition of Single Price Areas). However, this model requires a level of centralization, or at least coordination among the institutions in charge of the market and the system operation, that may be impossible to achieve under the present regulatory conditions existing in some regional markets, like the IEM. Note that in Model I no PX other than the central auction house exists.

The next subsection presents an alternative formulation of Model I based on a full nodal representation of the grid. This formulation may be of special interest when the number of Single Price Areas is so high that the number of PTDFs to be computed becomes unreasonably high.

Nodal Formulation of Model I

This section provides an alternative formulation of the reference auction model presented above. This formulation corresponds to that of a classical optimal load flow where some modifications have been introduced to include bilateral (multilateral) contracts.

$$\max \sum_j B_j p_j$$

s.t

$$\sum_w \frac{q_y - q_w}{x_{yw}} + \sum_i z_{iy} p_i = 0, h_y \quad \forall y = 1 \dots \text{number of nodes} \quad (4.9)$$

$$\frac{q_y - q_w}{x_{yw}} \leq \overline{F}_{yw}^+, g_{yw} \quad \forall y, w \in k = 1 \dots \text{number of nodes} \quad (4.10)$$

$$-\overline{F}_{yw}^- \leq \frac{q_y - q_w}{x_{yw}}, m_{yw} \quad \forall y, w \in k = 1 \dots \text{number of nodes} \quad (4.11)$$

$$q_{ref} = 0 \quad (4.12)$$

$$p_j \geq 0 \quad \forall j = 1 \dots \text{number of agents} \quad (4.13)$$

$$p_j \leq P_j \quad \forall j = 1 \dots \text{number of agents} \quad (4.14)$$

$$\sum_j u_{ij} p_j = 0 \quad \forall i = 1 \dots \text{number of balanced transactions} \quad (4.15)$$

The parameters and variables that are common to this model and Model I have the same meaning in both cases. However, there are some new parameters and variables that need further explanation: j is the set of market agents, y is the set of nodes, w is also the set of nodes (it is an alias of y), B_j is the price bid (€/kWh) per kWh of energy offered by agent j , p_j is the scheduled power corresponding to the bid by agent j , q_y is the phase angle for node y , P_j is the maximum power that agent j is willing to transact, $x_{y,w}$ is the reactance of the line between nodes y and w , z_{jy} is a Boolean parameter which is '1' if agent j is located at node y and '0' otherwise. Finally, u_{ji} is also a Boolean parameter whose value is '1' if the bid from agent j is part of balanced transaction i and '0' otherwise.

Discussion of the nodal model

The objective function to be maximized here is the same as that in Model I. Balanced transactions have been decomposed into two or more different unbalanced transactions, each one of them corresponding to either a withdrawal or an injection of power into the grid. Of course, total generation scheduled as part of a balanced transaction must equal total demand.

Equation (4.9) represents the set of balance equations for all the nodes in the system. The dual variable of each of these constraints is the nodal price at the corresponding node. Equations (4.10) and (4.11) limit the flow over each line in both directions. Equation (4.12) sets the reference phase angle. Equations (4.13) and (4.14) ensure that the amount of power granted to each unbalanced transaction in the auction is neither negative nor greater than the maximum amount of power that each agent is willing to transact. Finally, equation (4.15) guarantees that total generation and demand for each balanced transaction are equal to each other.

As said earlier, employing the nodal formulation of Model I makes sense when the number of SPAs is very high. The size of the optimization problem based on the computation of PTDFs grows proportionally to the square of the number of SPAs, whereas that of the nodal model remains unchanged. Thus, the nodal formulation of the implicit auction may be the only reasonable option if there are more than a handful of SPAs.

In the next subsection we propose a new model for the implicit auction, Model II, which tries to preserve as much as possible the current functions and independence of the individual PXs. This model may be easier to implement than Model I in those systems where achieving a high level of centralization is not possible.

Mathematical formulation of Model II

In Model II every unbalanced transaction has to specify the PX 'e' to which it will address its offer to buy or sell power. In the first place, an energy and capacity auction

goes on at regional level. Each PX is assigned a different reference node in this auction. Once the coordinated auction has taken place, PXs may conduct individual auctions either at control area level or, at least, within the SPA where the corresponding reference node is located. These subsequent auctions allow market agents to change their positions, therefore giving PXs a more significant role.

The core of Model II is a coordinated regional energy and capacity auction. Every market agent that engages in a bilateral transaction or bids on a PX must choose where to trade its energy. Provided that the defined SPAs can be reasonably represented using a single node, our model is able to accurately assess the impact on the network flows of each individual transaction between two different SPAs. The academic literature discusses some similar designs to the one proposed here, see (Stoft 1998) for instance. The selection of the reference node for each PX does not affect the price that each agent trading his energy at this PX is paid or charged, see (Rivier and Pérez-Arriaga 1993). Given a set of bids by agents, the algorithm maximizes the global value of the accepted bids subject to the specified network constraints. The identity of the set of accepted bids does not depend on the reference node used at each PX. The energy and network components of the final price of energy faced by each agent are coherent and always refer to the node where the agent is located. Thus, agents bidding on a PX would have to pay for the capacity they use to access the PX where they trade their energy. Accordingly, they would then buy or sell their energy at this PX's price. Bilateral contracts would have to pay for the capacity used to carry out the transaction between the injection and withdrawal points of the contract.

However, the decision by agents on whether to bid at one PX or another affects the final outcome of the optimization process. Unless perfect arbitrage exists, the set of accepted bids will depend on where agents address them. We cannot talk about nodal prices because an agent located at a specific node may get different prices depending on where it trades its energy. If agents are not perfectly informed they will not take advantage of every possible trade opportunity in the market. As explained below, the liquidity of the market at each PX may be a matter of concern in this regard. The mathematical formulation of the coordinated energy and capacity auction is presented now.

Model II

$$\max_{p_{bi}, p_{uje}} \sum_i B_{bi} p_{bi} + \sum_j \sum_e B_{uje} p_{uje}$$

s.t

$$-\bar{F}_k^- \leq \sum_i PTDF_{kbi} p_{bi} + \sum_j \sum_e PTDF_{kuj} p_{uje} \leq g_k \quad \forall k = 1, 2, \dots, \text{number of corridors} \quad (4.16)$$

$$\sum_i PTDF_{kbi} p_{bi} + \sum_j \sum_e PTDF_{kuj} p_{uje} \leq \bar{F}_k^+ \quad m_k \quad \forall k = 1, 2, \dots, \text{number of corridors} \quad (4.17)$$

$$\sum_j p_{uje} = 0 \quad \text{set of balance equations} \quad , t_e \quad \forall e = 1, 2, \dots \text{number of power exchanges} \quad (4.18)$$

$$p_{bi} \leq \bar{P}_{bi} \quad , b_{bi} \quad \forall i = 1, 2, \dots \text{number of balanced transactions} \quad (4.19)$$

$$p_{uje} \leq \bar{P}_{uje} \quad , b_{uje} \quad \forall je = 1, 2, \dots \text{number of unbalanced transactions} \quad (4.20)$$

$$p_{bi} \geq 0 \quad , a_{bi} \quad \forall i \text{ balanced transactions} \quad (4.21)$$

where:

je : set of unbalanced transactions that bid at PX e .

B_{uje} : bid price (€/kWh) per kWh of energy traded by agent j in the considered hour when buying or selling power at PX e . It is positive if j is a load and negative if it is a generator.

p_{uje} : amount of power scheduled corresponding to a transaction addressed by agent j to PX e . It is positive when agent j purchases power and negative when it sells power.

$PTDF_{kue}$: power transfer distribution factor of the flow over line k with respect to the power traded under unbalanced transaction j when a prescribed slack node representing PX e is used as the counterpart.

$\bar{P}_{uje}$: maximum quantity that agent j , who offers its energy at PX e , is willing to generate or consume.

Discussion of the model

Most of the constraints and the objective function to be maximized in Model II are common to Models I and II. Apart from them, additional constraints exist in Model II whose impact on the global economic efficiency of the dispatch needs to be evaluated. As said before, unbalanced transactions in Model I are not forced to bid at a specific PX. On the other hand, agents willing to buy or sell their energy in Model II must decide to which PX they will address their bids.

Again, all transactions are expressed in a bilateral format. The counterpart for one-sided transactions addressed at each power exchange e is the slack bus associated with this PX. The set of restrictions represented by equation (4.18) forces a global match

of all the scheduled unbalanced transactions addressed to each PX e . The congestion charge paid by unbalanced transactions depends on the identity of the PX they are addressed to since the value of the PTDFs for a transaction depends on the origin and end of this transaction.

Employing a different slack node for each PX may have some implications. In order for Models I and II to be equivalent in practice, market agents that send unbalanced bids under Model II should be able to accurately predict the outcome of the market, i.e. the energy prices in the different PXs. This in turn requires that the amount of transactions addressed to each PX is high enough (market at the different PXs are liquid enough). Model II does not jointly consider all the unbalanced transactions as Model I does. Bids are separately matched at the PX where they have been sent. Errors committed by market agents that make unbalanced bids when estimating the energy price at the different PXs may result in the dispatch outcome of Model II not being the same as that of Model I. For instance, a bid to a PX e may be rightly rejected but would have been accepted at a different PX f . Hence, agents in Model II may not make an optimal use of the grid. Some efficient generators, which would be dispatched in Model I, may be replaced in Model II by others whose energy is more expensive.

If market agents were perfectly informed and enough liquidity existed, traders would arbitrage any trade opportunity and energy prices would be the same as those in the centralized auction. The resulting pattern of generation and load in the system and the use made of the grid would be the same as well. Similarly to what happens in Model I, the congestion rent to be paid by each transaction in Model II would depend on the PX to which it is addressed. However, the price this transaction would charge or pay for its energy would be that of the corresponding PX. In the end, the net payment for the energy traded (either positive or negative) would be independent of the identity of the PX where the unbalanced bid is sent. But again, this would only happen if agents have access to perfect information.

Note that, properly speaking, the concept of price areas does not exist in Model II. Two different PXs could theoretically be located quite near each other and within the same SPA. However, if agents are not capable of arbitraging prices efficiently, the price of energy in both markets could differ. At the same time, agents connected to the same node of the grid could see different final energy prices depending on where they submit their offers.

The rigidity of the unbalanced bids in Model II, which have to be addressed to specific PXs, could impair the achievement of full economic efficiency. The interest of Model II resides in the fact that it may be applicable in those cases where a centralized implicit auction is not an option. If Model II were implemented, individual PXs would play a significant role since trade would take place at each of these PXs independently. However, PXs would not be able to manage cross-border transactions addressed to them in the coordinated auction. These transactions would be handled by the central clearing house.

Model II relies on arbitrage among market agents in order to achieve an economically efficient dispatch of energy and capacity. Given that the outcome of Model II will probably be sub-optimal, unless there are liquid enough secondary markets where agents can adjust their commercial positions, perfect arbitrage will never take place and some efficiency losses will occur. Thus, maintaining the individual functionality of PXs (at least in a certain way) causes the system to incur some efficiency losses. In other words, Model II increases flexibility and decentralization at the expense of efficiency.

The following is a description of what could be a pragmatic blueprint for implementation of Model II:

1. Define a number N of single price areas (SPAs) that are accurate enough just for the purpose of the bid-matching algorithm. If we refer back to the system in Figure 4.3-1 we may consider 3 SPAs if taking SPAs that are the same as the existing control areas makes sense in this case. Afterwards, we must define the balancing node for each PX. The balancing node of a PX may be a single physical node or a distributed bus comprising several nodes. In the latter case, the nodes that are part of the balancing bus may be located within the same SPA or they may be located in different SPAs. However, considering a distributed balancing node within a representative (large in electrical size) SPA seems superior. Local auctions following the coordinated one may include all transactions that happen within the same SPA. Hence, if the balancing bus of a PX e is located in a large single SPA a large number of transactions addressed to e by agents in this area may take part in the corresponding local auction, see (Rivier and Pérez-Arriaga 1993). Again, in the system of Figure 4.3-1, nodes 1 and 7 may be the balancing nodes corresponding to two different PXs 1 and 2. A balanced transaction between nodes 6 and 10 has to participate in the coordinated auction but one between nodes 6 and 7 does not need to take part. Unbalanced transactions that are addressed to PXs 1 and 2 by agents located in areas $A1$ and $A2$, respectively, do not need to participate either.
2. Despite the existence of several PXs, a central auction must run the joint energy and transmission capacity auction of Model II. All transactions that are scheduled by Model II are firm and they have to be physically executed, including those bilateral contracts that have obtained the prescribed transmission rights. Power injections and withdrawals scheduled at any PX (those accepted at PX 1, on the one hand, and those accepted at PX 2, on the other) must be balanced.
3. The central auction house publishes the results of the coordinated auction. At this point, each PX may run its individual auction. Local auctions may adopt one of two formats:

- a. Each local auction only includes the generation and demand within the SPA where the reference node is located. This auction does not modify the scheduled transactions involving agents located outside this SPA. In the system of Figure 4.3-1, only agents at nodes 1, 2 or 3 could participate in the auction held at PX 1 whereas agents at nodes 4, 5, 6 and 7 may participate in the auction organized by PX 2.
- b. Alternatively, PXs might be deemed to cover certain geographical areas. Then, each local auction may include transactions taking place within this area. If PX 1 extends over SPAs 1 and 2 and PX 2 extends over SPA 3 agents at nodes from 1 to 3 and from 8 to 9 could participate in the auction run by PX 1 while agents at nodes from 4 to 7 could take part in the one organized by PX 2. This design option is more appealing in those cases where each PX is somehow associated with a control area, such as OMEL in Spain or APX in the Netherlands. In this case, each local auction would apply to the several SPAs within the corresponding control area therefore increasing liquidity and making the balancing requirement that applies to one-sided bids less restrictive. The result of this auction may well be unfeasible, since part of the involved transactions make use of the capacity of congested lines (they take place between nodes in different SPAs). A balancing market could then be run at control area level to solve these infeasibilities. However, given that transactions scheduled in the balancing market must take place between different SPAs within the control area, the outcome of this market may create new infeasibilities affecting lines outside the control area. This design option makes more sense when applied to “regional markets”, such as Scandinavia, UK plus Ireland or the Iberian Peninsula. It is specially indicated for those cases where transactions between SPAs within a control area only cause negligible flows outside the control area because of the topology of the network.

A regional balancing market taking into account all the SPAs at the same time would certainly solve all infeasibilities as well as imbalances occurring in real time as a consequence of outages or unplanned operational conditions. This regional balancing market should be limited to solving infeasibilities and imbalances. Otherwise it would completely alter the final dispatch and render the previous auctions useless. All the scheduled transactions are settled on the basis of prices resulting from the final dispatch once auctions in steps II and III have taken place.

Model II, like Model I, requires PXs to transfer part of their current responsibilities to regional institutions. Again, Model II requires a high level of centralization of the energy market, a level that countries (or control areas, states, etc) in some regions will probably not be willing to accept.

Therefore, although a centralized regional auction house for the short-term markets of energy and transmission capacity conceptually looks like an appealing option, such a scheme seems unlikely to be adopted in some regions. For this reason, an alternative approach which interferes less with the functioning of presently existing PXs is proposed in the next section.

Other alternatives to the reference formulation

Other alternatives to congestion management approaches in Model I and Model II are possible. Thus, PXs would play a more significant role if separate energy only auctions ignoring network constraints were organized at PX level. Every agent could decide on which PX to offer its energy. Given that the result of this set of auctions would probably not be feasible, a regional security constrained implicit auction would take place afterwards, where infeasibilities would be solved. If transactions initially accepted by each PX were left unchanged no guarantee would exist that congestion created in PX's auctions could be fixed afterwards in the security constrained regional one. Alternatively, transactions accepted by PXs in the first place might be rejected afterwards if they contribute to congestion in the network. In this case the outcome of the last auction will be the same as that of Model I. In other words, individual PXs would become irrelevant.

Apart from this, EuroPEX, the European Association of PXs, and ETSO, the European association of Transmission System Operators, have recently proposed a scheme that is named Decentralized Flow Based Market Coupling or DFBMC. The algorithm comprises two stages. In the first one Power Exchanges clear (there is one per area) and a net demand curve is computed for each of them. Local generation and demand within each PX are supposed to be located at the same node. In the second stage, net imports to each PX are computed together with the use of interconnection transmission capacity (transmission capacity between areas) by bilateral transactions. Several implementations of this scheme are possible. Subsection 4.3.1.5 provides a description and a critical assessment of the DFBMC method within the discussion corresponding to the management of congestion in the Internal Electricity Market of the EU.

One last option, only possible if areas covered by PXs coincide with Single Price Areas, involves holding separate only energy auctions at each PX. Only bids from agents located within the corresponding PX could participate in these auctions. Each PX would then create an offer curve according to the bids it has received in this auction. They would submit these curves to a central clearing house. Then, the central auctioneer would buy energy in some PXs and sell it in others so as to maximize the value of the resulting commercial transactions. The auctioneer would compute a different energy price for each PX. However, this scheme, if applicable, would in practice be totally equivalent to the reference model. Therefore, it would not leave any room for PXs to carry out any relevant function.

4.3.1.3 Coordinated explicit auction of transmission capacity followed by an energy auction where the grid is not considered

Under ideal conditions, the contribution of a regional auction for transmission capacity and subsequent implicit auctions held within each area is equivalent to a regional implicit auction. This is shown in Annex 1. Some conditions must apply in order for this to be true: the market must be competitive, agents must be perfectly informed and the system conditions must not change between the time the explicit auction is held and the time local implicit auctions take place. Unfortunately, these conditions never apply. Thus, dealing separately with energy and transmission capacity results in some loss of efficiency.

Despite being the best option from an efficiency point of view, holding an integrated auction of energy and capacity at regional level may not be possible because of the implementation difficulties discussed in the previous sections. Then, a coordinated explicit auction followed by independent energy auctions within each Single Price Area (PX if possible) may strike a reasonable balance between efficiency and practicability.

According to this scheme, the interconnection capacity among SPAs would be auctioned first. Then, local energy auctions would take place. This design aims to replicate the outcome of Model I while preserving the independence that PXs and control areas enjoy at the moment. The European Association of Transmission System Operators (ETSO) initially proposed this mechanism for the European market, see (ETSO 2001).

The formulation of the explicit auction adopts a breakdown of the system into SPAs as Models I and II do. Unbalanced bids at each PX are referred to a common balancing point. This allows us to compute the impact on the network flows of these transactions.

The core of this scheme is a region-wide explicit auction of transmission capacity that takes place at a central auction house before the opening time of short-term (day-ahead, typically) local energy markets. In this auction, agents buy the right to use the transmission capacity. The format of these bids is the same both for bilateral contracts (balanced transactions) and for unbalanced transactions addressed at PXs. In all cases, the bid price is expressed in €/kW of capacity for the considered hour. Agents bid for the transmission capacity they need to market their energy at a specific PX (assigned to a specific geographical area) or to exercise a bilateral contract between two areas.

Capacity products offered in this auction may include obligations and options. If options are offered, the auctioneer needs to check that all possible realizations of the set of sold capacity options are feasible. In other words, executing any subset of the sold options must be compatible with the constraints imposed by the grid. Normally, this results in an underutilization of the transmission capacity between areas, since netting of the flows produced by the execution of sold options is not possible.

Given that the coordinated explicit option of transmission capacity takes place just before the short term markets for energy (the day-ahead market in each area or PX), market agents may probably be quite certain about the physical transactions they want to get involved in. Therefore, there is no real need for options to be offered in the short term. However, options over transmission capacity may be more interesting as long term hedging products because uncertainty is greater the further ahead we are of the real time operation of the system. Thus, long and medium term auctions for transmission rights should probably include both options and obligations. Section 4.4.1.1 briefly discusses the implementation of a dispatch including contingencies and options.

The auctioned product is the right to establish a transaction between two single price areas. Therefore, market participants do not have to acquire separate transmission rights for individual lines that could become congested (flowgate rights), but a point-to-point right (Single-Price Area to Single-Price Area, if defining SPAs is an acceptable approximation), which implicitly includes any necessary transmission rights on the relevant lines. This is the reason why the formulation below is expressed in terms of transactions instead of employing transmission rights over the capacity of individual lines.

Some members of the academic community consider that selling individual rights on individual lines may be more convenient for market agents, since they would be able to trade these rights in a decentralized manner afterwards (see (Chao and Peck 1996; Chao and Peck 2000)). In our view, trading rights on individual lines is a nuisance when compared to an algorithm that allows agents to deal only with the useful product –the right to transact power among several nodes of the grid. There is no reason for an auction of point-to-point rights to run into liquidity problems since the algorithm would implicitly manage individual flow-gate rights.

Mathematical formulation of the explicit auction, Model III

In contrast to the implicit auction in Model I, now only bids for capacity are accepted. Agents who are granted transmission capacity in the auction are the ones who value it most, subject to the condition that the set of accepted bids be feasible.

The formulation proposed here is very similar to that of the implicit auction in Model I. In this formulation only obligations have been considered.

Model III

$$\max_{p_i, p_j} \sum_i B_{bi} p_{bi} + \sum_j B_{uj} p_{uj}$$

s.t.

$$-\bar{F}_k^- \leq \sum_i PTDF_{kbi} p_{bi} + \sum_j PTDF_{kuj} p_{uj} , g_k \quad \forall k = 1, 2, \dots, \text{number of corridors} \quad (4.22)$$

$$\sum_i PTDF_{kbi} p_{bi} + \sum_j PTDF_{kuj} p_{uj} \leq \bar{F}_k^+, m_k \quad \forall k = 1, 2, \dots, \text{number of corridors} \quad (4.23)$$

$$p_{bi} \leq \bar{P}_{bi}, b_{bi} \quad \forall i = 1, 2, \dots, \text{number of balanced transactions} \quad (4.24)$$

$$p_{uj} \leq \bar{P}_{uj}, b_{uj} \quad \forall j = 1, 2, \dots, \text{number of unbalanced transactions} \quad (4.25)$$

$$p_{bi}, p_{uj} \geq 0, a_{bi}, a_{uj} \quad \forall i, j \text{ balanced and unbalanced transactions} \quad (4.26)$$

where B_{uj} is the bid price in (€/kW) offered by agent j when bidding for the transmission capacity it needs to access the PX where it plans to trade its energy. Agent j offers an amount B_{uj} per kWh of energy to be produced or consumed in the considered hour. Note that this meaning is different from the definition of B_{uj} in Model I.

The notation used here is completely analogous to that of Model I with the sole exception that agents in Model III are bidding only for capacity. As said before, Model III does not incorporate energy bids.

Due to the fact that only bids for capacity are considered, we do not need to include the set of restrictions associated with the balance of generation and load at each PX. Unbalanced bids at a PX are referred to the reference node of this PX in order to determine how much capacity the agent needs to acquire on each link so as to market its energy.

Transmission rights resulting from the auction have to be physical obligations, i.e. they have to materialize in the energy market (see the comments on the relation with the short-term energy market below). This auction determines the physical use that is made of the capacity of cross-border lines. However, similar auctions could be held in the long and medium terms in order to provide agents with the possibility of hedging against the volatility of differences among energy prices in different areas of the system. Transmission rights sold in these auctions could be either physical or financial. If we accept that the TSO should sell both types of rights without running any risks, then all the transmission rights issued should be simultaneously feasible. Rights sold should comply with the revenue adequacy requirement (see (Hogan 2002) among others). The owner of a right has the possibility to sell it in a subsequent auction.

Similarly to what happens in Model I, TSOs determine the limits expressed in constraints limiting the flows on the lines on the basis of the characteristics of the network, the expected pattern of generation and load and any potential contingencies

that may occur in the system. The *a priori* declaration of all the transactions that can affect the flows on inter-SPA congested lines may reduce the uncertainty in the definition of limits to the line flows. The ability to accurately estimate line flow limits also depends on the number of SPAs defined: the higher the number of areas, the higher the accuracy.

Discussion of the model

This subsection discusses several issues of much relevance in the implementation of Model III.

Coordinated versus individual auctions

Due to the lack in many cases of an agreed region-wide method for congestion management, explicit auctions for transmission capacity are presently being held, or have been proposed, for several frequently congested lines or flow-gates, see (IAEW 2001). This lack of coordination may probably cause a loss of efficiency. Infeasible flow patterns may result. This is due to the fact that avoiding the mutual interdependence among flows in different parts of the network is simply not possible. A coordinated allocation algorithm, such as the one in Model III, is the adequate response to this problem.

Assumptions behind the definition of SPAs

As with Models I and II, the successful implementation of Model III critically depends on whether we are able to find a set of SPAs that satisfactorily meet the basic conditions expressed in section 4.3.1.1. Otherwise, implementing a nodal scheme would be necessary.

Relation with the short-term energy market

The coordinated explicit auction for transmission capacity determines which transactions have the right to participate in the different short-term energy markets in the region in order to attain the most efficient use of transmission capacity. This requires a careful design of the coupling between the explicit auction of capacity and the subsequent short-term energy markets held at PX level. Short-term energy markets typically consist of a day-ahead energy market followed by even shorter-term markets.

Unlike Model I, the combination of an explicit capacity auction and a short-term energy market, where only inter-SPA transactions that were scheduled in the former auction are allowed to bid, is not a fully complete market for energy in the sense indicated by (Boucher and Smeers 2001). The fact that separate auctions are employed to optimize the allocation of transmission capacity and energy may entail some loss of efficiency and even the risk of unfeasibility. Organizing an explicit auction followed by independent energy dispatches for each country or area is less efficient than Model I. In Model III energy and capacity are not allocated at the same time. First, agents

willing to transact power outside their SPA need to acquire transmission rights in the explicit capacity auction. Only those that get the required rights can schedule a bilateral transaction or bid to buy or sell power.

Therefore, any bilateral/ multilateral transaction among nodes located in two or more areas in the system of Figure 4.3-1 must participate in the coordinated explicit auction. Additionally, any offer to a PX, except for those from agents at nodes 1, 2 or 3 when addressed to PX 1, or those from agents at nodes 4, 5, 6 and 7 when addressed to PX 2, also has to acquire the corresponding transmission rights. Afterwards, energy auctions take place at each individual PX. Only those agents external to the area where a PX is located that have acquired the transmission rights necessary for them to access this PX can participate in the corresponding energy auction. The use made of the interconnection capacity between areas (the amount of power flowing among areas A1, A2 and A3 in Figure 4.3-1) is fixed at the time local energy auctions take place. Hence, power transfers among PXs cannot adapt to the conditions and offers made by agents in each energy auction. The fact that the use made by agents of the interconnection capacity between areas is fixed a priori represents an additional constraint in the energy dispatch at each PX. These additional constraints are necessary in order to ensure that the resulting final dispatch is feasible. The next section expands on this idea.

The commitment of an agent to physically execute the right it has acquired in the capacity auction varies depending on the type of rights traded. Section 4.3.2.2 depicts two possibilities in this regard: options and obligations.

4.3.1.4 Dealing with Market Power in the short term. Implications of the use of different types of short-term auctions for energy and capacity in the exercise of market power

The previous two sections have discussed the implementation of coordinated implicit and explicit auctions under conditions of perfect competition. However, perfect competition among agents never exists. Traditionally, vertically integrated utilities in control areas or countries supplied the local load. With the passing of time, competition has been introduced in many former monopolies. However, vertical integration is still present in some areas. Additionally, the fact that the activity of generation of electricity is very intensive in capital together with the existence of entry barriers for new competitors result in high levels of horizontal concentration in most systems.

Market power (MP) is defined as the ability of an agent to unilaterally affect the market price of a product in order to maximize its profits. The amount of market power an agent enjoys depends on the slope of the net demand curve faced by the agent. The larger the slope of this curve, the greater the amount of market power it has.

Assessing the existing level of market Power in electricity markets is particularly difficult. Thus, simple indexes such as the Herfindahl Hirschman Index (HHI) do not provide a representative measure of the extent to which market agents can affect prices

in the market. The HHI is calculated as the sum over all market participants of the square of their market shares. Theoretically speaking, the best method to assess the level of static market power in an industry is the computation of the market equilibrium, see (Cournot 1838; Bertrand 1883; Klemperer and Meyer 1989). Authors in (Neuhoff, Barquin et al. 2002) compare different methods to compute the market equilibrium taking into account the limits imposed by the grid. However, simpler measures of MP also exist such as those proposed in (Wang, Xiao et al. 2004) where authors determine the percentage of the load within an area or node that must be served by a certain generator. Similarly, (Gan and Bourcier 2002) considers the percentage of the generation capacity of each agent that is necessary in the energy dispatch. These simple indexes can internalize the fact that congestion in the grid limits the size of the relevant market where each generator is competing to sell power.

Analogously, measures to curb market power include some very simple ones such as bid caps used in California. These are cost based caps on bids whenever prices are high, see (Rothkopf 2002). This article proposes to prevent pivotal suppliers with the incentive to exercise market power from participating in setting the market price. This is an alternative to price caps that may also prove effective in some cases. In addition to all this, analyzing the effect that the different types of auctions for energy and capacity have in the exercise of market power is relevant to the decision on which model to implement among the ones discussed above. Consequently, despite the fact that the analysis of market power is not the focus of this thesis work, we must, at least, briefly comment on the impact of the use of different types of auctions on the exercise of market power.

Regarding the mechanism used to compute the clearing price in an auction, it is widely accepted that marginal auctions are superior to the pay-as-bid ones. Thus, comparing both designs from a MP perspective, (Rassenti, Smith et al. 2003) concludes that, although pay-as-bid auctions may achieve a reduction in the price volatility, they result in price levels that are clearly above those yielded by the marginal price ones. Now we move on to discuss the impact that the interface between the energy and capacity markets has also in this regard.

If we analyze the short term energy dispatch as an static play, it is immediate to conclude that explicit auctions, where the use made of interconnection capacity is determined before the dispatch of energy within each area takes place, are more prone to the exercise of market power than implicit auctions. In effect, explicit short term auctions break the system into several areas, separated by congested lines or corridors, within which the short term energy dispatch is independent from that in the rest of the system. Therefore, generation in neighboring areas cannot react to the decision by a market agent in an area to withhold power. Each generator in an area only faces competition from the remaining local generators. On the other hand, implicit auctions dispatch energy and capacity at the same time. Hence, the use made of the interconnection capacity adapts to the offers by generators and consumers within the different areas.

However, market agents in explicit auctions can react in the next auction period to the exercise of market power by a certain agent. Thus, they bid for transmission capacity according to their estimation of the value that this capacity will reach in the subsequent energy dispatch. This estimation depends, to a great extent, on the value reached by transmission capacity in recent history. Hence, if the value of transmission capacity increases because an agent exercises its market power they will react increasing the amount of money they bid to obtain this transmission capacity. The dispatch of energy and capacity in Model III can be model as a two-stage equilibrium problem.

Some research conducted at the University of Cambridge concludes that the combination of an explicit auction and several subsequent only energy auctions results in firms exercising their market power to a higher degree than under implicit auctions, see (Neuhoff 2002).

4.3.1.5 Managing congestion in the short term, the case of three different regional markets

This subsection discusses how to best deal with congestion in some of the most representative regional electricity markets that exist today. This analysis is carried out in the light of the conclusions drawn from the comparison of different methods in the previous point.

Nodal pricing, a marginal implicit auction where energy and capacity are allocated simultaneously, is the more advanced design conceptually speaking. However, it requires an important level of integration, or at least coordination, of the markets existing in the different countries or areas of a region. Thus, implementing it in the first stage of the process of building a market may not be possible in some regions. Other methods that interfere less with the functioning of local markets, like the combination of a coordinated explicit (capacity) auction and independent local energy auctions, may be a suitable alternative in these cases.

Authors in (Bompard, Correia et al. 2003) provide a useful comparison of the congestion management mechanisms in place in some of the most representative electricity markets in the world. Specifically, they discuss the mechanisms in PJM, California, Norway, Sweden and England and Wales. Optimal price signals to market agents are only provided by LMP implemented in the PJM market. Zonal pricing followed by re-dispatch (Norway and Sweden) results in close to the same social surplus as LMP. However, some of the locational content of prices is lost. Last, arrangements in E&W and California result in lower social surpluses due to the additional constraints imposed on the final dispatch (balance constraints within each Scheduling coordinator in California, no modification of the unconstrained load pattern in E&W). Incentives to alleviate congestion may depend greatly on the use made of congestion rents.

The following paragraphs review the situation in several regional markets and provide recommendations on the market based congestion management mechanism to use in each case.

The European Internal Electricity Market (IEM)

The situation in Europe, the regulatory decisions already made and the mechanisms already adopted are presented jointly with, what we think is, the most efficient short-term congestion management scheme possible, taking into account that the solution chosen must be compatible with the existing regulatory and institutional settings in the region.

According to European regulation, market based mechanisms, whereby agents put a value on the transmission capacity required to engage in transactions at regional level, must be employed to manage congestion in the European grid (EP 2003). Therefore, regulation implicitly rules out simple non market-based mechanisms like the allocation of capacity on a first come first served or on a pro-rata basis. These and other basic mechanisms are presently used to assign the interconnection capacity between some pairs of countries in Europe, see section 4.3.1.1. ETSO provides in (ETSO 2004) a comprehensive discussion of the congestion management methods presently applied in Europe to allocate the cross-border transmission capacity.

An IEM system-wide nodal pricing scheme, or an IEM-wide energy and capacity auction run by a single PX as in Model I, are far from what appears to be achievable in the European Union in the short and medium terms. Regulations and institutional frameworks present in the countries of the EU are very different amongst themselves. Existing Power Exchanges and System Operators are unlikely to agree to organize a single regional auction for energy and capacity, although significant progress has been achieved recently.

In the short term, there seems to be no room for implementing an implicit auction at European level. However, this type of auctions may be used to manage the allocation of capacity and energy within each one of several pre-defined sub-regions. Sub-regions should be comprised of countries that have achieved a higher degree of integration amongst themselves than with other areas in Europe. The following is a tentative list of these sub-regions: a) The Iberian peninsula (where an Iberian market is scheduled to begin to operate in the coming months), b) the Nordic region (where the NORDEL market exists since 1998), c) Britain and Ireland, 4) Italy and 5) the central plate (comprising France, Germany, Switzerland, Austria, Belgium and the Netherlands). Countries from east Europe that have attained a higher degree of development of their internal markets may be included in the latter (Poland being an example of this). Others like Bulgaria, Romania or most of the former Yugoslavian republics should create their own sub-regional market with the intention to join other sub-regional markets once they are prepared, see (ERGEG 2005).

The allocation of interconnection capacity among sub-regions could take place in a coordinated capacity auction like the one described in Model III. ETSO proposed this scheme for Europe in 2001, see (ETSO 2001). Similar proposals had been made before. Therefore, the global model devised for Europe would correspond to that described in section 4.3.1.3. That is, a coordinated explicit auction at European level would be followed by implicit auctions within the different sub-regions.

At the same time, market organizations and arrangements within each of these sub-regions must gradually adapt to a common design in order to achieve the eventual integration of all sub-systems into a single system where a global implicit auction determines the dispatch of energy. This integration will take place in the medium to long term. This general scheme is in line with the one proposed by Consentec and Frontier Economics in their report on Congestion Management methods for the IEM, see (Consentec / Frontier 2004).

Some initiatives exist to foster the creation in the short term of an integrated auction for energy and capacity at European level. Thus, ETSO, the European organization of Transmission System Operators, jointly with EuroPex, the European Association of Power Exchanges, have devised a system called Decentralized Flow Based Market Coupling (DFBMC). This method aims to replicate the outcome of a regional implicit auction, see (Consentec / Frontier 2004). The DFBMC method considers both power exchanges between neighboring national markets and reservations of transmission capacity for bilateral contracts. It consists of two stages. First, Power Exchanges cleared and a net demand curve is obtained for each of them (located in different areas). Then, the use of the interconnection capacity between areas, both by imports to and exports from PXs and by bilateral contracts, is optimized. The DFBMC method is explained and assessed below.

The DFBMC mechanism proposed for Europe resembles the congestion management mechanism implemented in the Nordic region where the interconnection capacity between several areas is allocated in an implicit auction that results in the regional market dispatch. Price areas may or may not be the same as countries. Thus, Sweden is one of the price areas whereas Norway is comprised of several (Christie and Wangensteen 1998). In principle, a single short-term energy price is computed for the whole region. Energy prices in the different areas only differ when congestion occurs. See (Kristiansen 2004) for a complete description of the regulation of transmission in the Scandinavian market. Presently, the forward market is several times larger than the short term one. The existing long-term bilateral contracts are financial. Therefore, they do not affect the short term energy dispatch, at least under ideal conditions (perfect competition and enough liquidity).

The Nordic market is considered by many an example of the successful development of a regional market. However, there are some controversial aspects in it⁴. Additionally, implementing market splitting in a meshed network like the European one would probably require centralizing the dispatch of energy in the region contrary to what is stated by the proponents of the market coupling mechanism.

Decentralized Flow Based Market Coupling

The DFBMC mechanism is based on previous congestion management mechanisms proposed separately by EuroPEX and ETSO. EuroPEX proposed a decentralized market for exchanges of energy among Power Exchanges (EuroPEX 2003). This proposal did not consider bilateral contracts. ETSO had previously proposed a centralized market for energy and capacity (ETSO 2002).

The DFBMC is an implicit auction scheme, i.e. the short term markets for energy and transmission capacity clear at the same time. This method considers the region divided into a number of areas and aims to replicate the outcome of the corresponding optimal zonal dispatch. The operation of the market and the system within each area is left in the hands of the corresponding local (national) authorities. Several possibilities exist to determine the use to be made of the interconnection capacity between areas, which are explained below.

The transmission grid within each area is assumed to be well meshed. Therefore, the method only deals with congestion taking place on the interconnection between areas. The considered areas may coincide with the countries existing in the region though this is not necessarily efficient from an economic point of view. In principle, there could be areas where no PX is operating but there should not be more than one PX in the same area, as it is explained below.

The method comprises two stages: a first one where the preliminary energy dispatch within each of the existing PXs is determined and a second one where the use made of the interconnection capacity in the region is optimized. In the first stage, PXs within the different areas compute the preliminary energy dispatch for the PX and the corresponding net demand curve is obtained. Agents within each area may choose between bidding on the local PX and engaging in a bilateral contract with another agent. Bilateral contracts are not considered in this stage of the algorithm.

⁴ Glachant and Pignon consider that allowing the SO to retain part of the congestion rents resulting from the dispatch sets an incentive for them to artificially increase congestion [Glachant, J.-M. and V. Pignon (2002). Nordic electricity congestion's arrangements as a model for Europe: Physical constraints or operators' opportunism? ISNIE International Conference, MIT, Boston.

In the second stage of the algorithm, the use of the interconnection (regional) transmission capacity is determined. The interconnection capacity may be used either by power transfers taking place among PXs or by bilateral contracts arranged among market agents. The objective here is maximizing the social welfare in the region taking into account the system restrictions, mainly the limits of the transmission capacity of interconnectors. This is equivalent to maximizing the total value that the set of arranged transactions place on the capacity of the corridors among areas. As we have just mentioned, the set of arranged transactions must be compatible with the limits imposed by the transmission grid.

At least, two different methods exist to allocate the transmission capacity of corridors. Capacity may be centrally allocated taking into account the net demand curves for all the PXs in the region, the capacity bids of bilateral transactions and the topology of the network. Then, it all boils down to solving an optimization problem, which yields the net imports by PXs and the size of bilateral transactions (according to the amount of transmission capacity they are allocated). PXs may react to these results (the net amount of power that each PX will buy from others) by submitting new net demand curves. If new demand curves are submitted, the use to be made of the regional transmission capacity would be computed again. Thus, we would obtain new power transfers among PXs and new volumes of bilateral transactions. This iterative process would continue until PXs accept the resulting net power imports into their markets or until the maximum number of iterations allowed is reached.

In a second method for computing the use of the interconnection capacity between areas, as many day-ahead bilateral implicit auctions as corridors take place separately. Each of these auctions yields the amount of transmission capacity of a corridor to be used by power transfers among PXs and the amount to be used by each bilateral transaction. These sequence of elementary bilateral auctions must be repeated until results obtained for all of them are consistent (mainly the imports into and exports from PXs), see (EuroPEX 2003).

Focusing on one of these bilateral implicit auctions, incremental offers issued by the different PXs to import/export power are represented by two net demand curves. Each curve is associated with one of the two areas linked by the corridor that is considered in the auction. Power transfers between PXs will result in certain imports/exports taking place between an area and the neighboring ones. In order to compute the equivalent net demand curve corresponding to an area in an auction, we must consider the imports/exports of this kind taking place between this area and those areas that are not explicitly taken into account in the auction. From the net demand curve of the PX located in each of the two areas connected by an interconnector (if there is a PX in the area), and the aforementioned imports/exports into or from the area, we compute the net demand curve of the area in the corresponding auction. Finally, using the net demand curves for the two areas linked by an interconnector, and the offers by bilateral transactions to buy capacity on the interconnector, we determine the use that each party would make of this interconnection capacity. This problem may be solved

as an optimization one where we aim to maximize the value placed by agents (power transfers between PXs and bilateral transactions between agents) on the transmission capacity auctioned. The detailed mathematical formulation of this second algorithm remains to be put forward.

Critical assessment of the method

Similarly to what happens in other zonal methods, in order to successfully apply DFBMC we must obtain a sensible division of the system into areas. As explained above, dividing a region into accurate SPAs is not always possible. See section 4.4.3 for a method to compute SPAs. If the division into areas is the same as that employed in a zonal implicit auction, both methods should yield very similar results, though this depends on the reaction of PXs to the net imports computed by DFBMC.

If the use made of interconnectors is computed centrally, the algorithm proposed here is almost identical to that of an implicit auction, except that in DFBMC PXs are allowed to react to the results of application of the method by changing their net demand curves. If this variant of the method is used, some central institution will have to carry out the dispatch of energy and capacity at regional level, which may not be desirable from the point of view of control areas (countries) that want to preserve all their independence.

Regarding the variant of the method whereby power transfers among areas are computed using a sequence of bilateral auctions, the practical viability of this algorithm remains to be proved. In addition, efficiency losses may result from the fact that bilateral transactions must submit separate bids for the capacity on all the interconnectors they need to use. On the other hand, control areas would in this case preserve all their independence.

In any case, the effort required to coordinate the system operation and functioning of the different markets is a major obstacle in the way to implement the DFBMC method.

The Development of Regional Markets in the USA

The Federal Energy Regulatory Commission, with authority over the inter-state trade of energy in the US, devised an entire reference framework for the development of markets in the different regions of the country. This general framework is called the Standard Market design (SMD). According to the SMD, congestion within each region should be managed using nodal pricing (Locational Marginal Pricing (LMP) in the specialized jargon used in America). Therefore, federal authorities envisaged the adoption of implicit auctions from the start of the process. In the long term, congestion rights would complement implicit auctions thus allowing agents to shield against the risk of volatility in the price of transmission capacity, see (FERC 2002).

Some systems, like that covered by the PJM regional market, have been successful in the implementation of this congestion management mechanism. In fact, the FERC took

the PJM market as a model for the development of the SMD. For a description of the functioning of daily and intra-daily markets in PJM see (Ott 2003).

However, implementation of the SMD in other regions where markets are less mature has found strong opposition from state regulators and utilities. Local generation belonging to strong, vertically integrated utilities use the transmission system owned by these companies to serve the local load. This is in line with the wheeling philosophy that prevailed for a long time in the United States, see (Tobin 1990). According to this philosophy, only the spare capacity that is left once the local load has been served should be devoted to interstate trading. Using implicit auctions in these regions seems to be too ambitious in the short term. Due to generalized fierce opposition, the FERC had to water down the SMD design and transform it into a set of recommendations for the development of regional markets, see (FERC 2003). Meanwhile, the “PJM model” continues its expansion in neighboring systems.

Rau identifies in (Rau 2003) many technical problems to be dealt with before implementing a Standard design at national level. However, some aspects should receive the same treatment across states as soon as possible. These include the use of Locational Marginal Prices in the short-term, the allocation of scarce transmission capacity in the long term through congestion rights (discussed in section 4.3.2.6) and an adequate treatment of losses.

Measures proposed by Rau would certainly contribute to solving the seams problem, i.e. they would make it easier for market agents to engage in transactions with parties located in other regions or to ship power between regions. Obviously, if every region in the country adopted a common market design, like the one proposed by FERC, seams would not exist within each region and seams issues between regions would be very much eased. However, the FERC proposal seems less likely to be implemented, at least in the short-term. In addition to achieving a harmonization of those market rules in the different regions that are relevant to inter-regional trade, some mechanism must be put forward in order to manage the interconnection capacity among these regions.

This thesis' proposal

From what has been said, it can be concluded that allowing regions with less developed electricity industries to manage their internal congestions, while encouraging those regions where the industry is mature enough to implement a system based on implicit auctions, may be a sensible compromise. Ideally, market authorities in a region should choose between applying the SMD scheme, backed by the FERC, and making use of a less centralized, less intrusive market based solution, like zonal pricing or bilateral trading. Energy trade taking place between different regions might be managed using a coordinated explicit auction of a national scope, whereby the inter-regional transmission capacity would be auctioned. This auction could be centrally organized by an entity subject to the oversight of FERC. This set of rules are likely to solve most, if not all, seam problems faced by market agents.

The utilization of long-term financial transmission rights (FTRs) is far less problematic than that of nodal prices, since the former do not condition the short-term physical energy dispatch and they are compatible with the existence of any short-term congestion management mechanism. Hence, FTRs should become a standard across the country to manage congestion in the long term. The utilization of transmission rights is discussed in more depth in section 4.3.2.6.

The regional market in Central America

Due to the small size of electricity systems in this region and the fact that existing regional institutions are not as deeply rooted as those in the EU (EU Commission, Council and Parliament) or the USA (FERC, Federal Government), national systems may attain a higher level of integration more easily. In fact, some of the regional institutions that exist nowadays in Central America were created within the context of the implementation of a market for electricity in the region. Due to these circumstances, authorities in the region were able to develop and implement a congestion management scheme that is very much in line with the most efficient possible solution. Thus, now we only describe the solution that has already been adopted, together with a few critical comments aimed at improving the treatment given to some specific aspects of the problem, i.e. we do not outline a whole alternative design.

The market design proposed for the region, which has already been partly implemented, involves the creation of a system of nodal prices used to value power injections and withdrawals corresponding to regional transactions. Transactions among agents in the regional market are the result of an integrated economic dispatch of energy and capacity. Consequently, the transmission capacity of the regional grid is efficiently priced in the short term.

The regional market coexists with 6 national markets. The process of computing the final dispatch in the region is divided into two stages. At the first stage of the process, national markets clear. Based on the preliminary dispatch obtained for national markets, net demand curves are obtained for each node of the RTR. These net demand curves can be thought of as incremental bids submitted to the regional market. In the second stage of the process, the regional economic dispatch is computed using the aforementioned incremental bids. In the regional dispatch, incremental power injections and withdrawals, located in the nodes of the RTR, are scheduled. These injections and withdrawals are valued at the corresponding nodal prices. The final dispatch is the result of combining the transactions resulting from national dispatches with the incremental transactions resulting from the regional dispatch.

Market procedures and institutions have no tradition in many of the countries in the region. Hence, using locally the same market mechanism implemented at regional level is not possible in some of these countries. Creating the conditions to implement national markets that are consistent with the regional one should be the long term

objective. In the meantime, countries where the level of development of the market is lower, like Costa Rica, may resort to a regulated system whereby power plants are dispatched based on their audited production costs.

The total net payment to a country in the regional dispatch, which results from pricing incremental power injections and withdrawals that are part of regional transactions at their corresponding nodal prices, must be assigned somewhat to the generators and consumers operating in the country. When doing this, power injections and withdrawals in a country should not be discriminated based on the market (national vs. regional) where they have been scheduled. Thus, if locational differences in energy prices are accepted in a country, the nodal energy prices that are computed in the regional dispatch for the nodes of the RTR located within this country should be taken as a reference to price power injections and withdrawals. If locational energy prices are not an option, generators and consumers trading their energy locally or regionally should all see the same energy price. In any case, countries are responsible for deciding on the way energy prices within each of them are computed.

Regarding the use of transmission rights, two types of rights exist: financial transmission rights, which are aimed at securing access to the grid by paying an amount fixed at the time of buying the rights, and physical rights, which are needed to sign firm supply contracts. Long term firm supply contracts have priority over other transactions in case some transactions must be curtailed for security reasons. In addition to long term firm contracts, market agents may engage in long term non-firm contracts, which do not have priority over others. Transmission rights in the MER are discussed in section 4.3.2.6.

4.3.2 Definition and subsequent trading of capacity contracts previous to the short term energy dispatch

Efficient congestion management systems discussed in section 4.3.1 result in energy prices that exhibit spatial and temporal differentiation, i.e. energy prices in different places and at different times are different.

Fuel prices, the production of hydro plants as well as that from renewable energy sources, the amount of available generation capacity, the regulatory framework and many other factors vary over time causing zonal energy prices and differences among them to vary as well. Market agents perceive a risk associated with the volatility of energy prices and feel the need to shield against it. Risk hedging products were created to address this problem. Markets for these products have developed. In fact, these markets have already reached their maturity in some regions like the Nordic countries.

Contrary to what some authors state, congestion is likely to arise in electricity markets, see (Stoft 1999). Therefore, shielding against the volatility in the value of transmission makes sense. Two main different risk hedging products exist: contracts signed with a

counterpart that fix the price an agent will pay or receive for its energy in the future and capacity contracts or transmission rights that fix the price paid to use some amount of transmission capacity between an injection and a withdrawal point(-s). In fact, depending on the terms of the agreement, at least one of the signatories of an energy contract may find it advisable to sign a capacity contract in order to know in advance the price it will have to pay to access the area where its counterpart is located.

Energy contracts may be negotiated bilaterally between agents or unilaterally between an agent and a central mediation house. On the other hand, capacity contracts need in general to be mediated through the competent TSO or market operator, who manages the use of the transmission capacity available in the system. The analysis of energy contracts, such as the contracts for differences, is beyond the scope of this thesis work. Here, we shall focus on the definition and exchange mechanisms for network capacity products. These products allow agents to hedge against the risk of volatility in the differences of prices among nodes or areas in the system.

Network capacity products are exchanged in long and medium term capacity auctions that take place before the dispatch of energy is computed. In these capacity auctions, agents involved in bilateral contracts, as well as those planning to trade their energy on a certain power exchange, can obtain the transmission rights they need to secure the price they must pay to access the grid. Transmission rights are a hedging tool for market agents to protect themselves against the risk associated with the volatility in the short term value of transmission capacity.

This section is organized as follows. Subsection 4.3.2.1 provides an introduction to the topic. Subsection 4.3.2.2 explains the different types of transmission rights that exist. Typically, several capacity auctions take place on different time horizons to allow market agents to adjust their positions over time. Section 4.3.2.3 deals with the design of the sequence of capacity auctions leading to the short term energy and capacity dispatch. This section also discusses how much of the total transmission capacity of lines must be put up for sale in each auction. Section 4.3.2.4 briefly comments on the influence that the allocation of congestion rights may have on the exercise of market power. Finally, subsection 4.3.2.5 presents the formulation of a long term capacity auction. This auction includes both options and obligations over the transmission capacity between areas defined in the context of a regional market.

4.3.2.1 Introduction to the use of transmission rights

As their name suggests, transmission rights allow their owner to use some amount of transmission capacity or, alternatively, receive congestion rents corresponding to the value of this capacity in the short term dispatch. Congestion rents are the net result of applying nodal prices to power injections and withdrawals in the system.

Transmission rights are issued by the Transmission System Operator in each area. It is unlikely that any other institution or firm will be in the position to issue them. In principle, they must be initially allocated to the owner of the corresponding

transmission capacity. Afterwards, rights can be traded in capacity auctions. Payments to the owners of transmission rights in the short term dispatch must come from congestion rents. The TSO must not issue more transmission rights than the amount of available capacity. Otherwise, it would run the risk of not being able to afford the payments that the owners of these rights are entitled to. This property is normally referred to as simultaneous feasibility of the existing rights or revenue adequacy.

System Operators need to use detailed enough models of the system transmission grid to check the simultaneous feasibility of transmission rights. What is more, the allocation of rights may need to take place in a coordinated manner depending on the format of these rights. Thus, a single institution may have to allocate or, at least, coordinate the allocation of congestion rights in all the areas that a meshed network is comprised of. Unless accurate enough SPAs can be defined, transmission rights must refer to nodes instead of areas. In this case, the grid model used by the TSO should be nodal as well.

Non strategic market agents, i.e. agents that are not able to unilaterally modify energy prices, put a value on transmission rights that is equal to their best estimate of the difference in energy prices among the nodes or areas that the right is defined in between. Some studies carried out recently conclude that the price paid for transmission rights in capacity auctions deviates substantially from the true value of these rights. This may well have to do with the fact that the market for transmission rights defined between nodes very far from one another or belonging to different control areas is not liquid enough, see (Hogan 1996). However, prices of transmission rights may deviate from the corresponding real time differences in energy prices because of other reasons. In some markets, owners of certain transmission rights may obtain additional benefits. For example, as explained when discussing the congestion management mechanism implemented in Central America, ownership of the physical rights required to sign a firm contract entitle a generator to earn the capacity payment existing in the area where the load of the contract is located. Additionally, transmission rights may provide their owner with access to markets for green certificates, etc. Subsection 4.3.2.4 deals with the case where market power exists. This subsection discusses the valuation of transmission rights by strategic generators.

4.3.2.2 Different types of transmission rights

Different types of congestion rights exist. Congestion rights may be physical or financial. Physical rights entitle their owner to actually use the transmission capacity whereas financial rights entitle him to the corresponding congestion rents. Rights can be defined between any two points of the system (point-to-point rights) or relate to a particular line or corridor (flowgate rights). A right that relates to a line or corridor is not equivalent to a point-to-point right defined between the two extremes of this line, as we shall see below. Finally, congestion rights may be characterized either as the option or as the obligation to use the corresponding transmission capacity or receive the value of this capacity in the dispatch.

Options (“use-it” rule) versus obligations (“use-it-or-lose-it” rule)

An option transmission right provides its owner with the possibility to use the corresponding transmission capacity or receive the corresponding congestion rents. However, owning an option does not entail having the obligation to exercise it.

Agents that buy obligations commit to using the transmission capacity that the obligation relates to or receiving an amount equivalent to the value of this capacity in the dispatch (which may be negative, thus involving a payment).

When options are used, the transmission capacity of a line or corridor in the two directions must be auctioned separately. In other words, we cannot take into account the amount of options over the capacity of a line that has been sold in one direction in order to increase the amount of transmission capacity that can be auctioned in the other direction. Otherwise the revenue adequacy criterion might not be fulfilled. Deducting the amount of capacity of a line used in one direction from that used in the other, in order to compute the amount of capacity on a line that is still available for sale, is commonly known as netting. If netting is applied when options are used, the resulting power flows may not be feasible (Ruff 2000; Hogan 2002). Then, an additional centralized mechanism, such as a very liquid balancing market, should be put in place in order to solve infeasibilities. This centralized auction may not leave PXs any significant role to play.

On the other hand, if the auctioneer only issues obligations, netting the flows that transmission rights relate to does not mean that we are compromising the simultaneous feasibility of the set of rights that are sold at the auction. Therefore, preserving the independence of PXs through the use of explicit short-term auctions does not prevent the auctioneer from maximizing the use made of the transmission capacity. Hogan states in (Hogan 2002) that the feasibility of implementing obligations has been demonstrated, whereas options are yet to be physically implemented in any real scale system. Contrary to this, (Chao and Peck 2000; Méndez and Rudnick 2004) recommend the implementation of options.

Options give agents more flexibility to adjust their transmission rights to the commercial positions they hold. All other things being equal, the more flexible a market is the more liquidity it has and the more efficient it is. We recommend the use of obligations in short term explicit auctions until an integrated energy and capacity market becomes a viable alternative. If obligations are adopted some measures may need to be taken to attain more flexibility. These measures may include defining large SPAs or allowing agents from different SPAs to participate in the short-term energy market taking place on each PX. A balancing market may follow short-term energy markets. This market would aim to solve infeasibilities resulting from the application of options or the flexibility mechanisms just mentioned.

Flowgate versus point-to-point (or area to area) rights

There is abundant academic literature both in favor and against the use of each of these two types of transmission rights (Chao and Peck 2000; Ruff 2000; Tabors and Caramanis 2000; Hogan 2002). Point-to-point rights relate to the capacity required to carry out a transaction among two or more nodes or areas. Flow-gate rights relate to the capacity on a specific line or corridor which may get congested (a flow-gate).

Point-to-point rights allow their owner to execute a transaction between two nodes or areas or capture the congestion rents corresponding to a hypothetical transaction between these two nodes. Congestion rents in this case are equal to the volume of the transaction times the difference in energy prices between the injection and withdrawal points. Carrying out this transaction (or earning the corresponding congestion rents) would be possible for the owner of a point-to-point right regardless of the network configuration, the set of active constraints or the contingencies that may happen. However, all these changes may affect the amount of capacity on the different flow-gates used by a given transaction.

Flowgate rights allow their owner to use a specific line or corridor or capture the corresponding congestion rents. As suggested above, owning X MW of a right over flow-gate K is not equivalent to owning the same amount of point to point rights between the two extremes of flow-gate K . In the first case, the owner of the right is entitled to use X MW of the transmission capacity of flow-gate K whereas in the second one the owner can use as much transmission capacity over flow-gate K as needed to transport X MWs of power from one end node of corridor K to the other. Congestion rents to be earned by the owner of the financial flow-gate right are:

$$CR_{FG} = X \cdot g_K \quad (4.27)$$

where g_K is the shadow price of the constraint limiting the flow over flow-gate K in the corresponding direction. On the other hand, assuming flow-gate K is the only congested line used by a transaction taking place between both ends of this flow-gate, the owner of a financial point-to-point right would earn:

$$CR_{FG} = X \cdot PTDF_{tranK} \cdot g_K \quad (4.28)$$

If a transaction between the two extremes of flow-gate K uses other congested corridors, the owner of the point-to-point right would be entitled to the corresponding congestion rents as in (4.28).

Therefore, a collection of flow-gate rights does not automatically provide their owner with the right to carry out a transaction between any two nodes or areas. In other words, the owner of these rights may not be fully hedged against the risk associated with the energy price difference between two nodes or Single Price Areas. But owning flow-gate rights may have some advantages. For instance, contrary to what happens

with point-to-point rights, secondary trade of flow-gate rights might not need to be mediated through a centralized institution.

Many have expressed the opinion that auctions for point-to-point rights may have liquidity problems. However, implementing the short term explicit auction of Model III for agents to buy and sell point-to-point capacity rights does not entail any loss of liquidity since these algorithms are *implicitly* exchanging flow-gate rights.

Chao and Peck recommended in (Chao and Peck 1996) the creation of a capacity market for flow-gate rights for the first time. Once the transmission capacity is allocated, an energy auction would take place where only those market agents that have reserved the capacity they plan to use would be able to take part. According to this scheme, trading of these rights could be bilateral, i.e. it would not need to be mediated through a centralized institution. Tabors and Oren also propose the utilization of flow-gate rights in (Tabors and Caramanis 2000; Oren and Ross 2002). On the other hand, Ruff and Hogan in (Hogan 1992; Ruff 2001), respectively, support the use of point-to-point rights, since only them can provide agents with a perfect hedge against the risk associated with the volatility of the transmission price due to the existence of multiple commercially sensitive flow-gates related both to the base-case and to the contingency scenarios. Méndez concludes in (Méndez and Rudnick 2004) that point-to-point rights are better adapted to centralized trading whereas flow-gate rights are superior in a decentralized environment. Hence, the election of the type of rights to implement should be made conditional on the trading model previously chosen.

Physical versus financial rights

Financial transmission rights allow their owner to collect the corresponding congestion rents. In the case of a point-to-point contract from node j to node k , congestion rents amount to the contracted capacity (MW) times the difference between energy prices at nodes (or single price areas) j and k . This difference may be negative, thus involving a payment instead of a credit if rights are defined as obligations.

Physical transmission rights entitle their owner to use the network transfer capacity between two points or areas in the case of point-to-point rights or the capacity of a line or corridor in the case of flow-gate rights. Therefore, owners of point-to-point physical rights can inject electricity at one node and retrieve it at the other (Chao and Peck 1996; Ruff 2000) at no extra cost. Owners of point-to-point financial rights might as well purchase power at one node and sell it at the other with the congestion rent providing a perfect hedge against the corresponding price difference. Therefore, in principle, physical rights do not give their owner any additional economic advantage with respect to financial rights.

Hogan recommends in (Hogan 1992) the use of financial rights because, earning the corresponding congestion rents, the owner of a financial capacity right is indifferent between carrying out the physical transaction and receiving the corresponding

payments. He also argues that financial rights do not affect the physical dispatch of energy in the short-term, contrary to what may happen if physical rights are implemented. Chao and Peck put forward the utilization of physical rights necessary to carry out any physical energy transaction in the short term. Similarly, (Tabors and Caramanis 2000; Oren and Ross 2002) propose the use of physical rights that are necessary to get involved in physical transactions in the short term dispatch. However, in (Chao and Peck 2000) Chao and Peck argue that these rights would give priority to use the network only if their owner commits to using them. Otherwise, they would revert back into the system in order for this capacity to be used by others. Finally, Kench states that both physical and financial rights can result in an efficient congestion management system. However, physical rights are superior for market agents to contract power at efficient prices, see (Kench 2004).

There are two instances, however, where transmission rights must be physical. Thus, if a explicit capacity auction precedes the short term energy market (as in model III) transmission rights acquired in this auction must be physical to guarantee that the corresponding power transactions will take place. Thus, if Model III is implemented then financial rights acquired by agents in the long term must either be sold or exchanged for the corresponding physical rights in the last capacity auction before the short term energy market.

Apart from this, transmission rights associated with firm energy supply contracts, which provide long-term security of supply, need to be physical. An agent who earns the capacity payment that exists in an area must be able to access this area when there is not enough local generation capacity to serve the local load. Physical rights provide this agent with instant access to the corresponding physical transmission capacity since this capacity is actually reserved (until some time before the operation of the system) for the owner of these rights to use it, see (EOR 2005).

Financial rights may be preferable in the long and medium term, since the allocation of these rights to market agents does not entail any physical use of the underlying transmission capacity. Hence, it could be argued that, all other aspects being equal, financial rights are less prone to affect the physical energy dispatch than physical rights.

4.3.2.3 Sequence of capacity markets leading to the short term energy dispatch

Regardless of the congestion management scheme finally adopted, most market participants may wish to hedge against the uncertainty that the existence of network constraints introduces in the congestion charges they have to pay. Agents can protect themselves from the risk associated with the volatility in the price of transmission by acquiring transmission rights in explicit long and medium term capacity auctions.

These auctions would take place at longer time intervals than short term, typically daily, auctions. The more time in advance of the system operation long-term capacity auctions take place the longer the time interval between two consecutive auctions

should be. The first auction for rights over (part of) the network transmission capacity might take place some years before the operation of the system (3, but even up to 10, for example). Afterwards, annual capacity auctions would follow where agents would buy and sell transmission rights to better adjust their transmission rights portfolios to their commercial positions. In addition, the TSO would put up for sale in these auctions part of the transmission capacity not yet sold. As we draw closer to the short term dispatch the time period between capacity auctions could decrease from one year to several months initially, from several months to one month afterwards and from one month to one week finally. Thus, during the last one or two months before the operation of the system, weekly auctions might be held. In order to promote the liquidity of this market, transmission rights traded in long and medium term capacity auctions may be financial instead of physical.

Division of the capacity of lines into slices to be put up for sale in each auction

Another issue relevant to the design of long term capacity auctions is the split of the transmission capacity on each line or corridor into the fractions to be sold at the different auctions. We must determine how much of the total capacity of each line must be up for sale at each auction. Market power considerations will be left out of the discussion that follows. The interaction between the allocation of congestion rights and market power is briefly analyzed in subsection 4.3.2.4.

We must first bear in mind that, as long as the transmission capacity market is liquid enough, any significant difference between transmission prices in different auctions will be arbitrated. Consequently, regardless of the amount of new transmission capacity offered in each auction, and provided the general situation in the system does not change, the price of transmission capacity should not vary significantly from one auction to another⁵.

In addition to this, assuming that the market is competitive and liquid enough, the amount of transmission capacity auctioned in the long run will not affect the outcome of the energy dispatch even if transmission rights issued in the long term are physical. In effect, issuing physical rights in the long term does not mean that these rights will have to materialize. Under ideal conditions, they are completely equivalent to financial

⁵ When bidding for a transmission right, arbitrageurs base their offers on their best prediction of the difference between the short-term energy prices in the injection and the withdrawal points of the right. On the other hand, consumers and producers that will take part in the short-term capacity auction may have an extra incentive to acquire transmission rights, in order to be able to market their energy. In the absence of market power, the latter agents will probably bid according to their marginal production cost (generators) or the marginal value they place on the energy they consume (load). However, this does not imply that the value that agents of one type or another place on transmission capacity must vary over time (from one auction to another).

rights. Only physical transmission rights acquired in the short-term capacity auction of Model III require their owner to physically use the corresponding transmission capacity. What is more, the fact that an uneconomical transaction takes place in the energy market does not necessarily involve any efficiency loss. This transaction would enable market agents to get involved in other more profitable ones, since it would increase the amount of valuable transmission capacity in the opposite direction that is available. Efficiency losses only occur if the generator taking part in an inefficient transaction is dispatched instead of other more efficient generators located in the same area. But then the former generator would have an incentive not to produce and buy the energy it has to deliver from the latter.

Therefore, under ideal conditions, constraining more some parts of the network than others when slicing the total available transmission capacity will not have a dramatic impact on capacity prices nor will it significantly alter the physical use made of the grid. In any case, TSOs should slice transmission capacity so that its price in the long run is as similar as possible to the price realized in the short run, i.e. TSOs must aim to minimize differences between long and short term capacity prices. Allocating transmission capacity among long and medium term auctions in proportion to the volume of capacity negotiated in each one of these auctions may be a way of achieving the convergence of long and short term capacity prices. Auctions should be liquid enough so that the only driver of capacity prices is the expected real time value of this capacity.

In conclusion, auctioning slices of transmission capacity both in the long and medium terms will help market agents hedge against locational differences in energy prices and establish long term contracts. The decision by the SO on the amount of transmission capacity to be put up for sale in each auction is unlikely to have significant consequences, neither in the price paid for transmission capacity, nor in the system dispatch. However, dividing the transmission capacity of lines in slices that are proportional to the volume of capacity negotiated in each auction may be a sensible option. In any case, transmission capacity must be progressively offered for sale so that everything is available to the market just before short-term energy auctions take place. A security capacity margin should always exist in case the system operating conditions change unexpectedly.

4.3.2.4 Effect that the existence and allocation of congestion rights may have on the exercise of market power

This section briefly discusses the mutual implications that exist between the exercise of market power in real life systems and the allocation and subsequent trading of congestion rights. Here we just present a succinct review of the state of the art on the subject. At the same time, the discussion serves as an introduction to section 4.4.4, which describes an algorithm to identify a balancing node to be used in the definition of transmission contracts.

It can be easily checked that transmission rights may affect the exercise of market power by generation companies, see (Cardell, Hitt et al. 1997; Borenstein, Bushnell et al. 2000). In effect, when an agent holding transmission contracts unilaterally modifies the price of energy in the node or area where it is located (and probably in others as well) this agent is also modifying the value of its transmission right portfolio. Hence, when bidding on the energy market, agents holding market power take into account the impact that their decision to increase or decrease production has on the value of its transmission rights.

However, at the same time, the extent to which an agent is able to modify the price of energy in the market affects the value that this agent places on transmission rights. This is a consequence of the fact that, once a generator has bought some transmission rights, its decision about the amount to produce, and hence its profits from the sale of power, will depend on the quantity and type of the transmission rights it owns.

Agents may find in transmission rights an extra incentive to exercise the market power they hold. We may take two different types of measures to avoid this situation from happening. One first option is designing methods to allocate the transmission capacity whereby market agents do not find it profitable to acquire congestion rights that enhance their market power. Banning the acquisition by market agents of rights that encourage them to abuse market power is the other.

Solutions of the first type are preferable to those of the latter since they interfere less with the market. Instead of limiting the ability of agents to contract any product, measures of the first type encourage agents not to acquire certain type of rights. Authors in (Joskow and Tirole 2000; Gilbert, Neuhoﬀ et al. 2004) conclude that if transmission rights are allocated in a uniform price auction and market agents are perfectly informed, generators will not buy market power enhancing contracts. Arbitrageurs systematically beat generators that hold market power when bidding for these contracts since the former place a higher value than the latter on them. However, this result only applies to the case of perfectly arbitrated markets where no uncertainty exists and agents have access to complete information. Auctions in reality are far from taking place under these stylized conditions. The academic community has not come up yet with an allocation mechanism capable of dissuading agents from buying contracts that encourage them to manipulate energy prices under real life conditions.

Consequently, regulators may have to resort to more aggressive measures to tackle this issue. Preventing agents explicitly from buying some transmission rights is one possibility. In any case, the ability of market agents to hedge against the risk associated with the transmission price volatility should never be compromised.

Rothkop et al. state in (Rothkop 2001) that, in order for transmission rights not to aggravate the problem of the exercise of market power, agents should only be allowed to acquire transmission rights of the kind that provides a hedge for the energy they buy or sell in the market. In other words, generators must only be allowed to buy the

kind of transmission rights that hedges their sales of power in other areas. Analogously, consumers must only be allowed to buy the kind of transmission rights that hedges their purchases of power in other areas. Similarly, authors in (Bautista and Quintana 2005) propose implementing a capacity auction that encourages agents to buy those transmission rights that perfectly match their commercial positions. As an alternative, they propose selling transmission rights to consumers and traders instead of generators. This may not be viable nevertheless.

This simple rule may be effective in most cases. However, under certain circumstances a generator decreasing its output may be able to increase the energy price of a node other than that where this generator is located to a higher degree than the price at its own node, see (Gilbert, Neuhoff et al. 2004). In this case, holding a transmission right in the exporting direction between these two nodes would encourage the generator to exercise its market power. Thus, defining the set of transmission rights that each generator should be allowed to buy is not as easy as it may seem.

One may think of defining transmission rights with respect to a specific node whose energy price is as independent as possible from the commercial decisions made by individual agents. This node may be termed balancing node since it acts as the counterpart of any agent willing to buy a transmission contract. As mentioned above, section 4.4.4 addresses the definition of a suitable balancing node for transmission rights.

Market agents who have bought transmission contracts between their node and a balancing node R are implicitly buying or selling their power at the price of node R . If this price P_R is stable and evolves smoothly with market conditions, these agents have implicitly contracted their output at a more or less fixed price. Consequently, they find it unprofitable to increase market prices by withholding power. This is an argument in favor of offering in the long term as much transmission capacity as possible. The more capacity is available in the long-term, the more transmission contracts agents having market power may engage in. This would reduce the exercise of market power.

The fact that market agents holding certain types of transmission rights exercise their market power to a lesser extent may cause these agents to place a lower value on these rights than arbitrageurs. This is because their future revenue stream from the sale of energy in the market may well be negatively affected by the fact that they own these rights. Hadn't them bought these contracts they would have increased energy prices to a larger extent thus increasing their revenues in the short term market. If they own these contracts, increasing their revenues in the short term energy market by withholding power causes the value of their transmission rights portfolio to decrease. Thus, they must achieve a balance between revenues from the sale of energy and those from the sale of transmission rights or use of the corresponding capacity.

Agents internalize the adverse effect that owning certain rights has on their revenues from energy sales when they make an offer for these rights. Consequently, they may

bid less than arbitrageurs or other agents, thus getting beaten by them. It is not clear at all how many rights defined towards a balancing point agents holding market power will buy, since these agents have both incentives to buy them (as a risk hedging tool) and incentives not to buy them (since these rights will cause their profits from energy sales to decrease). As a consequence of all this, determining a priori the amount of transmission capacity to be offered at each auction is not a trivial issue. Probably, splitting transmission capacity on each line according to the historical demand for it on each auction is the most sensible option.

4.3.2.5 Formulation of a long term capacity auction including both options and obligations

After having discussed several issues related to the use of transmission rights, now we present the formulation of a typical long term capacity auction in the context of a regional market. This auction includes both options and obligations over the transmission capacity between Single Price Areas. All transmission rights offered are point-to-point rights defined between two (or more) SPAs.

As mentioned above, agents may be uncertain about the value of transmission capacity in the long term. They may also find it difficult to decide, long time before the system operation, whether they are going to carry out a certain transaction. Therefore, offering options in addition to obligations makes sense in this case.

$$\begin{aligned}
 & \max_{p_{pobi}, p_{popj}, p_{sobk}, p_{sopm}} \sum_i B_{pobi} p_{pobi} + \sum_j B_{popj} p_{popj} - \sum_l B_{sobl} p_{sobl} - \sum_m B_{sopm} p_{sopm} \\
 & \text{s.t.} \\
 & -\bar{F}_k^- \leq \sum_i PTDF_{kpobi} p_{pobi} + \sum_j PTDF_{kpopj} p_{popj} S_{kpopj}^- - \sum_l PTDF_{ksobl} p_{sobl} - \\
 & \quad - \sum_m PTDF_{ksopm} p_{sopm} S_{ksopm}^-, g_k \quad \forall k = 1, 2, \dots, \text{number of corridors} \quad (4.29)
 \end{aligned}$$

$$\begin{aligned}
 & \sum_i PTDF_{kpobi} p_{pobi} + \sum_j PTDF_{kpopj} p_{popj} S_{kpopj}^+ - \sum_l PTDF_{ksobl} p_{sobl} - \\
 & - \sum_m PTDF_{ksopm} p_{sopm} S_{ksopm}^+ \leq \bar{F}_k^+, m_k \quad \forall k = 1, 2, \dots, \text{number of corridors} \quad (4.30)
 \end{aligned}$$

$$p_{pobi} \leq \bar{p}_{pobi}, b_{pobi} \quad \forall i = 1, 2, \dots, \text{number of bids for obligations} \quad (4.31)$$

$$p_{popj} \leq \bar{p}_{popj}, b_{popj} \quad \forall j = 1, 2, \dots, \text{number of bids for options} \quad (4.32)$$

$$p_{sobl} \leq p_{sobl}^{\text{owned}}, b_{sobl} \quad \forall l = 1, 2, \dots, \text{number of offers of obligations} \quad (4.33)$$

$$p_{sopm} \leq p_{sopm}^{owned}, b_{sopm} \quad \forall m=1,2,..number\ of\ offers\ of\ options \quad (4.34)$$

$$p_{pobi}, p_{popj}, p_{sobl}, p_{sopm} \geq 0, a_{pobi}, a_{popj}, a_{sobl}, a_{sopm} \quad \forall i,j,l,m\ bids\ and\ offers\ in\ the\ auction \quad (4.35)$$

where B_{pobi} is the bid price (€/KW) corresponding to an offer i to buy obligations, B_{popj} is the bid price (€/KW) corresponding to an offer j to buy options, B_{sobl} is the bid price (€/KW) corresponding to an offer l to sell obligations, B_{sopm} is the bid price (€/KW) corresponding to an offer m to sell options, p_{pobi} is the amount of obligations (MW) sold to bid i , p_{popj} is the amount of options (MW) sold to bid j , p_{sobl} is the amount of obligations (MW) sold by the agent submitting bid l , p_{sopm} is the amount of options (MW) sold by the agent submitting bid m , $PTDF_{kpobi}$ is the power transfer distribution factor with respect to line k of the point to point rights allocated to bid i to buy obligations, $PTDF_{kpopj}$ is the power transfer distribution factor with respect to line k of the point to point rights allocated to bid j to buy options, $PTDF_{ksobl}$ is the power transfer distribution factor with respect to line k of the point to point rights corresponding to offer l to sell obligations, $PTDF_{ksopm}$ is the power transfer distribution factor with respect to line k of the point to point rights corresponding to offer m to sell options. S_{kpopj}^- is a binary parameter that is 1 if $PTDF_{kpopj}$ is negative and is 0 otherwise, S_{ksopm}^- is a binary parameter that is 1 if $PTDF_{ksopm}$ is negative and is 0 otherwise, S_{kpopj}^+ is a binary parameter that is 1 if $PTDF_{kpopj}$ is positive and is 0 otherwise and S_{ksopm}^+ is a binary parameter that is 1 if $PTDF_{ksopm}$ is positive and is 0 otherwise. $\bar{F}_k^-$ and $\bar{F}_k^+$ are the amount of transmission capacity on line k (MW) to be put up for sale in the ‘negative’ and ‘positive’ directions, respectively. $\bar{P}_{pobi}$ is the maximum amount of obligations that the agent issuing bid i is willing to buy and $\bar{P}_{popj}$ is the maximum amount of options that the agent issuing bid j is willing to buy. Finally, p_{sobl}^{owned} is the maximum amount of obligations of this type that the agent who issues offer l can sell and p_{sopm}^{owned} is the maximum amount of options of this type that the agent who issues offer m can sell.

Discussion of the formulation

Long term capacity auctions try to maximize the total value placed by agents on the transmission capacity they own. Restrictions are similar to those in models I, II and III. Differences between the formulation of the long term auction and the formulation of the short term ones mainly relate to the fact that options, unlike obligations, may be exercised or not. Different realizations of the set of options in the hands of market

agents result in different flows on the grid. The set of transmission rights issued (options and obligations) must be simultaneously feasible for any realization of the set of options. Specifically, equations (4.29) and (4.30) represent the two extreme cases. Thus, restrictions under equation (4.30) correspond to the case where all the options that produce a positive flow on each line k are exercised to their full extent whereas those that produce a flow in the opposite direction are not exercised. On the other hand, restrictions under equation (4.29) correspond to the case where all the options that produce a flow over a line k in the negative direction are exercised whereas those that produce a positive flow are not. The flow over a line must always be within limits. Equations (4.31) and (4.32) guarantee that offers to buy options and obligations are not allocated more rights than the maximum amount considered in the offer. Equations (4.33) and (4.34) guarantee that agents do not sell more rights than the amount they own (or the maximum amount they are willing to sell). Last, equation (4.35) restricts the amount of rights of each type that are bought and sold by each agent to positive values.

Options provide market agents with more flexibility to adapt their transmission rights portfolio to their commercial positions. On the other hand, the maximum amount of options that can be issued is smaller than the amount of obligations that are simultaneously feasible. Flows created by options cannot be netted since the realization of an option is uncertain.

Options also differ from obligations in the value that market agents place on them. However, the valuation of a portfolio of transmission rights is beyond the scope of this thesis work.

4.3.2.6 Regulatory recommendations on the implementation of a system of long-term transmission contracts in a regional market

This section discusses the implementation of a system of long-term transmission rights in a regional context and provides recommendations on this issue. Given that the use of transmission rights is compatible with any short-term congestion management mechanism implemented in a region, we shall not refer specifically to each one of the different regional markets in this regard. However, some comments will be made on the system of transmission rights used in the MER, since it differs substantially from any other.

Broadly speaking, congestion rents are long-term instruments aimed at protecting agents from the volatility of transmission capacity prices. In addition, they may help private promoters to undertake some network investments since, by selling transmission rights in advance of the construction of a line, they may secure some revenues that will reduce the risk associated with the investment that they perceive. Congestion in regional markets is likely to be much more significant than in local markets. Hence, the use of transmission rights in regional markets is even more useful, if possible.

Recommendations on the format of transmission rights

Transmission rights should be financial since, leaving aside market power considerations, transmission rights of this type cannot influence the energy dispatch. On the contrary, physical rights may affect the dispatch of energy in the region if the transmission capacity market is not liquid enough. However, physical rights may be necessary if transmission rights perform any other function apart from the common ones, see the discussion below on the use of rights in the MER.

Both point to point rights and flow-gate rights should be offered in long term capacity auctions. Thus, market agents would have the opportunity to buy the transmission capacity products that best suit them. Analogously, both options and obligations could be offered in the long term. In spite of this, provided transmission and energy are allocated separately in the short-term, only obligations can exist just before the energy dispatch. This is the case of the coordinated explicit auction proposed for the allocation of transmission capacity among regions in the US, for example.

Recommendations on the split of capacity among auctions

Transmission capacity should be split among auctions in proportion to the volume of trade in each one of these auctions. In other words, the amount of capacity of a corridor that is put up for sale in each auction should be proportional to the volume of capacity over this line that has traditionally been bought in the auction. This rule should be separately applied to split the capacity of each one of the corridors in the region into slices to be offered in each auction. The SO or SOs in a region should try to eliminate any persistent difference between the price of the capacity of a line in different auctions. In any case, provided enough liquidity exists, neither the price of transmission capacity, nor the short-term energy dispatch, should be significantly affected by the fact that the ability of agents to contract transmission capacity is limited to higher degree in an auction than in others.

Recommendations on the design of transmission rights in order not to aggravate the problem of the exercise of Market Power

Market power existing in local markets may not be reduced by increasing the size of the market. The relevant market size may be restricted by network restrictions or market rules that hinder the development of trade between different areas of the region. In addition, the exercise of Market Power in a regional market may have more serious consequences than in local ones, since the resulting increase in prices may affect a larger amount of generation and load. Consequently, this topic must be devoted enough attention when designing regional markets.

Transmission rights owned by strategic market agents may affect the degree to which these agents abuse their market power. Consequently, implementing some measures so that the set of transmission rights issued do not aggravate this problem is necessary. Measures aimed at encouraging market agents not to buy those transmission rights

that would provide them with an extra incentive to exercise the market power they hold are preferable to other more intrusive measures, like banning agents from acquiring these rights.

When assessing the impact that different types of auctions have in the exercise of Market Power by agents, we conclude that marginal, single price, capacity auctions are superior to pay-as-bid auctions. Assuming ideal conditions, strategic generators bidding for Market Power enhancing rights in a marginal, single price, auction would be beaten by traders. This may not be the case in pay-as-bid auctions. Therefore, auctions for transmission capacity in a region should be marginal and a single price should be computed for the right to move power between each two areas. Unfortunately, this result may not hold true under realistic conditions.

Consequently, we must try to find other solutions even when they are more intrusive. Thus, we may only allow market agents to buy transmission rights defined between the point where they are located and a swing bus, which would be the same for the whole region. The energy price at this swing bus must be as independent as possible from unilateral output decisions by strategic generators. Section 4.4.4 presents a methodology to select a balancing point in a regional market, though other possibilities exist, as it is explained there.

Specificities of the system of transmission rights implemented in the Central American market (MER)

The use of transmission rights will be briefly commented on now since the use made of some rights and its properties is specific of this regional market. Transmission rights are used as long term instruments that help market agents shield against the transmission price volatility. In addition, generators signing firm transmission contracts need to acquire physical transmission rights over the capacity they will use in order to access the area where the load in the contract is located. Firm contracts allow a generator to earn the capacity payment that is paid in this area. Consequently, two types of transmission rights exist: financial transmission rights, which shield agents from changes in the price of transmission capacity but do not involve any physical use of this capacity, and physical rights, which are needed to sign firm contracts. Given that physical and financial rights are different products that have different intrinsic values, we must distinguish between rights of one type and the other when auctioning them. In other words, bids to buy or sell physical rights must be different from bids to buy or sell financial rights (we cannot issue an offer to buy indistinctly rights of either of the two types). However, rights of both types can be traded jointly in the same auction.

Regarding the format of transmission rights, as we have mentioned, both physical and financial rights may be bought or sold. Transmission rights should preferably be defined between two nodes of the RTR (point-to-point rights) instead of being associated with a particular line (flow-gate rights). Both options and obligations are

possible though options suit best long term trading of rights, whereas obligations suit best short term trading.

4.4 Treatment of some specific issues related to the management of congestion

The previous section provides an overview of the different allocation mechanisms and products that exist to manage congestion in the long and the short terms. Now we move on to some more specific issues related to the management of congestion.

Section 4.4.1 presents the formal deduction of the mathematical expression for the marginal value of a line in the dispatch. Congestion rents must be allocated to lines in proportion to their marginal value. This is linked to the topic that is covered immediately afterwards: the treatment of certain restrictions not associated with the limited transmission capacity of lines. These restrictions have a market value equal to their marginal cost in the dispatch. Market agents responsible for the restriction to be binding must afford this cost. Section 4.4.2 discusses how to distribute the corresponding revenue.

The definition of SPAs in the system is central to the implementation of any zonal congestion management scheme, as well as to the utilization of transmission capacity products that relate to zones or areas instead of nodes. Section 4.4.3 deals with this issue. Finally, section 4.4.4 proposes a methodology to choose a balancing node to be used in the definition of transmission rights.

4.4.1 Allocation of congestion rents to the lines in the system. Marginal value of a line in the dispatch

The per unit marginal value of a line in the dispatch may be defined as the decrease in the system operation costs resulting from a unit increase in the capacity of this line. Thus, it represents the marginal value of a unit of capacity of this line. Then, the total marginal value of a line can be defined as the unit marginal value of the line times the capacity of the line. This section contains the formal deduction of the expression for the marginal value of a line in the dispatch. According to marginal theory, transmission lines should be allocated an amount of congestion rents proportional to the marginal value of each line in the dispatch. Congestion rents corresponding to a line must be earned by the owners of the line or those parties owning transmission rights over the capacity of the line.

We use a DC model of the transmission grid and neglect transmission losses. This section proves that, under certain conditions, the marginal value of a line in the

dispatch is equal to the difference between the nodal energy prices at both ends of the line times the flow over the line.

The notation used in this mathematical proof follows:

i : set of agents

j, m : set of nodes

T_{jm} : transmission capacity of the line between nodes j and m

x_{jm} : reactance of the line between nodes j and m

q_j : phase angle of node j

u_{jm} : binary variable which signals whether there is a line linking nodes j and m

v_{ij} : binary variable that specifies whether agent i is located in node j

p_i : amount of power that agent i is dispatched

C_i : unit bid by agent i

f_{jm} : power flow over the line between nodes j and m

p_j, a_i, x_{jm}, b_{jm} : dual variables corresponding to restrictions in the dispatch optimization problem

Consider the following formulation of the dispatch problem:

$$\begin{aligned} \max \quad & \sum_i C_i p_i \\ \text{s.t.} \quad & \sum_i v_{ij} p_i + \sum_m u_{jm} f_{jm} = 0 \quad \forall j \quad p_j \end{aligned} \quad (4.36)$$

$$p_i \leq P_i \quad a_i \quad (4.37)$$

$$\frac{q_j - q_m}{x_{jm}} u_{jm} = f_{jm} \quad \forall j, m \quad x_{jm} \quad (4.38)$$

$$f_{jm} \leq T_{jm} \quad \forall j, m \quad b_{jm} \quad (4.39)$$

$$q_{ref} = 0 \quad (4.40)$$

where equation (4.36) represents the set of power balance equations for every node in the system, equation (4.37) limits the amount of power that each agent is dispatched to the maximum it wants to sell or produce, equation (4.38) relates the flow over each line to the system state variables (phase angle at both ends of the line) using the line impedance (linear model of the system). Finally, equation (4.39) limits the flow over each line.

Assuming a linear relationship between the admittance of a line and the thermal capacity of the line, we have:

$$y_{jm} = \frac{1}{x_{jm}} = kT_{jm} \quad (4.41)$$

Then, equation (4.38) may be expressed as:

$$(q_j - q_m)u_{jm}kT_{jm} - f_{jm} = 0 \quad (4.42)$$

The expression of the Lagrangian of this optimization problem is:

$$\begin{aligned} L = & \sum_i C_i p_i + \sum_j p_j \left(\sum_i v_{i,j} p_i + \sum_m u_{j,m} f_{j,m} \right) + \sum_i a_i (p_i - P_i) + \\ & + \sum_{j,m} x_{jm} \cdot \left[(q_j - q_m)u_{j,m}K \cdot T_{j,m} - f_{j,m} \right] + \sum_{j,m} b_{j,m} (f_{j,m} - T_{j,m}) \end{aligned} \quad (4.43)$$

Hence, the value of a line in the dispatch, computed as the product between the unit marginal value of the line and its capacity, results:

$$\begin{aligned} Val_{jm} &= T_{jm} \frac{\partial L}{\partial T_{jm}} = T_{jm} \left((q_j - q_m)u_{jm}kx_{jm} - b_{jm} \right) = T_{jm} (q_j - q_m)u_{jm}kx_{jm} - T_{jm} b_{jm} = \\ &= f_{jm}x_{jm} - T_{jm} b_{jm} \end{aligned} \quad (4.44)$$

First order conditions of the dispatch problem are, among others:

$$\frac{\partial L}{\partial p_i} = C_i + \underbrace{p_j}_{v_{ij}=1} + a_i = 0 \Rightarrow p_j = -C_i - a_i \quad (4.45)$$

$$\frac{\partial L}{\partial f_{jm}} = p_j - p_m - x_{jm} + b_{jm} = 0 \Rightarrow p_j - p_m = x_{jm} - b_{jm} \quad (4.46)$$

$$\frac{\partial L}{\partial q_j} = \sum_m u_{jm} kT_{jm} x_{jm} = 0 \quad (4.47)$$

From the slack complementarity conditions we know that $b_{jm} \neq 0$ if and only if $f_{jm} = T_{jm}$. Two possibilities exist: either the line is congested or it is not. If the line is congested, we have that $b_{jm} \neq 0$ and $f_{jm} = T_{jm}$. Hence

$$\begin{aligned} Val_{jm} &= f_{jm} x_{jm} - T_{jm} b_{jm} = \\ &= f_{jm} x_{jm} - f_{jm} b_{jm} = f_{jm} (x_{jm} - b_{jm}) = \\ &= f_{jm} (p_j - p_m) \end{aligned} \quad (4.48)$$

If the line is not congested, i.e. if there is some spare capacity, then $b_{jm} = 0$. Hence:

$$\begin{aligned} p_j - p_m &= x_{jm} - b_{jm} = x_{jm} \\ Val_{jm} &= f_{jm} x_{jm} - T_{jm} b_{jm} = f_{jm} x_{jm} = f_{jm} (p_j - p_m) \end{aligned} \quad (4.49)$$

Thus, regardless of whether a line is congested or not, congestion rents earned by its owner(s) should be equal to the flow over the line times the difference between the energy prices at both ends of the line. Throughout this proof, the admittance of a line has been assumed to be proportional to the line capacity. The resulting expression for the marginal value of a line is only true when the aforementioned relationship between the admittance and the capacity of a line holds. If we assume a different relationship between both parameters of a line, the mathematical deduction of the marginal value of a line is analogous –but not identical– to the one presented here. We would only have to replace the expression for y_{jm} in equation (4.41) with the appropriate one.

4.4.1.1 Extension of the analysis to the case where contingencies are considered in the dispatch. Application to an auction for transmission rights that includes options

If we consider contingencies in the operation of the system the formulation of the dispatch changes. It follows a possible formulation of the optimization problem to be solved:

$$\begin{aligned} \max \quad & \sum_i C_i p_i \\ \text{s.t.} \quad & \sum_i v_{ij} p_i + \sum_m u_{jme} f_{jme} = 0 \quad \forall j, e \quad p_{je} \end{aligned} \quad (4.50)$$

$$p_i \leq P_i \quad a_i \quad (4.51)$$

$$\frac{q_{je} - q_{me}}{x_{jm}} u_{jme} = f_{jme} \quad \forall j, m, e \quad x_{jme} \quad (4.52)$$

$$f_{jme} \leq T_{jm} \quad \forall j, m, e \quad b_{jme} \quad (4.53)$$

$$q_{ref} = 0 \quad (4.54)$$

where the sub-index e refers to the considered contingency case. The following relationship is fulfilled for each contingency state e as it is in the original dispatch formulation:

$$(q_{je} - q_{me})u_{jme}kT_{jm} - f_{jme} = 0 \quad (4.55)$$

The value of a line in the dispatch is:

$$Val_{jm} = T_{jm} \frac{\partial L}{\partial T_{jm}} = T_{jm} \sum_e ((q_{je} - q_{me})u_{jme}kx_{jme} - b_{jme}) = \sum_e (f_{jme}x_{jme} - T_{jm}b_{jme}) \quad (4.56)$$

Apart from that we have:

$$\frac{\partial L}{\partial f_{jme}} = p_{je} - p_{me} - x_{jme} + b_{jme} = 0 \Rightarrow p_{je} - p_{me} = x_{jme} - b_{jme} \quad (4.57)$$

Considering a line between nodes j and m and a contingency state e we have that $b_{jme} \neq 0$ only if $f_{jme} = T_{jm}$. Hence:

$$\begin{aligned} - \text{ If } b_{jme} \neq 0 \Rightarrow Val_{jme} &= f_{jme}x_{jme} - T_{jm}b_{jme} = f_{jme}x_{jme} - f_{jme}b_{jme} = f_{jme}(x_{jme} - b_{jme}) \\ \Rightarrow Val_{jm} &= \sum_e Val_{jme} = \sum_e f_{jme}(x_{jme} - b_{jme}) = \sum_e f_{jme}(p_{je} - p_{me}) \end{aligned} \quad (4.58)$$

$$\begin{aligned} - \text{ If } b_{jme} = 0 \Rightarrow \begin{cases} Val_{jm} = \sum_e Val_{jme} = \sum_e f_{jme}x_{jme} \\ x_{jme} = p_{je} - p_{me} \end{cases} \Rightarrow Val_{jm} = \sum_e f_{jme}(p_{je} - p_{me}) \end{aligned} \quad (4.59)$$

This result is in line with that obtained for the case where no contingencies are considered. Equation (4.59) is the sum of a set of expressions analogous to that in equation (4.49). Each one of these expressions corresponds to a possible contingency state. As a result of the problem represented by constraints from equation (4.50) to equation (4.54), a different set of power flows and nodal prices may be computed for each scenario e . The marginal value of a line jointly depends on all of them.

Capacity auctions where both options and obligations are issued must be dealt with in the same way as in the case with contingencies. Each realization of the set of options results in a different set of power flows as well as a different set of node phase angles. Therefore, the formulation of the dispatch including options is completely analogous to the one presented in the optimization problem with restrictions from (4.50) to (4.54)⁶.

⁶ An alternative formulation that makes use of Power Transfer Distribution Factors (PTDFs) is perfectly possible as well.

Were options used, the value of a line might be obtained as the sum, over all the possible realizations of the set of options sold, of the differences between the shadow prices of the power balance equations for both end nodes of the line times the flow over it. The number of possible realizations of the set of options is very high since it grows exponentially with the number of bids for options accepted. However, not every realization needs to be considered in practice. Only those that result in some lines or corridors being congested must be explored. As a consequence of this, the number of realizations of options to be considered is not higher than twice the number of active grid restrictions.

If transmission losses are considered in the dispatch, the mathematical expression computed previously for the value of a line no longer holds. However, the process to be followed remains the same. The dispatch problem to be solved in this case is the following one:

$$\begin{aligned} \max \quad & \sum_i C_i p_i \\ \text{s.t.} \quad & \sum_i v_{ij} p_i + \sum_m u_{jme} \left(f_{jme} + \frac{f_{jme}^2}{2r_{jm}} \right) = 0 \quad \forall j, e \quad p_{je} \end{aligned} \quad (59)$$

$$p_i \leq P_i \quad a_i \quad (60)$$

$$\frac{q_{je} - q_{me}}{x_{jm}} u_{jme} = f_{jme} \quad \forall j, m, e \quad x_{jme} \quad (61)$$

$$f_{jme} \leq T_{jm} \quad \forall j, m, e \quad b_{jme} \quad (62)$$

$$q_{ref} = 0 \quad (63)$$

$$\begin{aligned} \max \quad & \sum_i C_i p_i \\ \text{s.t.} \quad & \sum_i v_{ij} p_i + \sum_m u_{jme} \left(f_{jme} + \frac{f_{jme}^2}{2r_{jm}} \right) = 0 \quad \forall j, e \quad p_{je} \end{aligned} \quad (4.60)$$

$$p_i \leq P_i \quad a_i \quad (4.61)$$

$$\frac{q_{je} - q_{me}}{x_{jm}} u_{jme} = f_{jme} \quad \forall j, m, e \quad x_{jme} \quad (4.62)$$

$$f_{jme} \leq T_{jm} \quad \forall j, m, e \quad b_{jme} \quad (4.63)$$

$$q_{ref} = 0 \quad (4.64)$$

Operating the first order conditions for this problem, we obtain:

$$\frac{\partial L}{\partial f_{jme}} = p_{je} - p_{me} - \frac{f_{jme}}{r_{jm}} p_{je} - \frac{f_{jme}}{r_{jm}} p_{me} - x_{jme} + b_{jme} = 0 \rightarrow$$

$$x_{jme} - b_{jme} = p_{je} \left(1 - \frac{f_{jme}}{r_{jm}} \right) - p_{me} \left(1 + \frac{f_{jme}}{r_{jm}} \right) \quad (4.65)$$

The value of a line can be easily obtained by substituting the difference between nodal prices in the mathematical deduction from (4.55) to (4.59) for the expression in the right hand side of equation (4.65). In any case, using a linear expression for losses is essential to implement a capacity auction where they are taken into account.

4.4.1.2 Possible applications of this formula in the regulation of transmission systems

This section provides an expression for the congestion rents to be earned by the owner of a line. Obviously, revenues of transmission owners in the dispatch must be limited to the amount of capacity of each line that they have not previously sold (in the form of rights). Thus, the revenues of the owner of line k in the dispatch should be equal to the unit marginal value of the line times the amount of capacity of line k that has not been sold in long or medium term explicit auctions.

Owners of transmission rights should receive the remaining part of the revenues corresponding to line k . The unit value of a point-to-point obligation is equal to the difference between the energy prices in the injection and withdrawal points of the right. Thus, the revenues of the owner of some rights of this type must be equal to the unit value of the rights times the amount of rights he has got.

When deciding on the construction of a grid reinforcement, the entity in charge of the expansion of the grid should take into account the total impact that the construction of this reinforcement may have on the system operation costs. The expression obtained here for the marginal value of a line may be of great help to estimate the saving, in the operation costs for any snapshot, that the reinforcement of this line would bring about. In fact, if the marginal value of a line, located between two points A and B , did not change when we increase the capacity of the line, the stand-alone effect that the construction of a line of X MW between the same two points would have on the operation costs would be $Saving = \text{Unit marginal value of the line} \cdot X$. Additionally, merchant investors may use this expression to estimate their future revenues in the form of congestion rents (or revenues from the sale of rights). In any case, both merchant investors and the central planner must take into account the fact that the value of a line (and therefore revenues from it) may change dramatically if other lines are built afterwards.

4.4.2 Treatment of special restrictions regarding the allocation of the corresponding revenues

Annex 2 demonstrates that, in the absence of any non-conventional restriction (one not associated to congestion), the net result of application of nodal prices to all power injections and withdrawals in the system is equal to the total cost of restrictions on the flow over the transmission lines (the total payment to be made by users of the transmission grid). Annex 2 also demonstrates that, under the same conditions, revenues from congestion are equal to the sum of the marginal values of all the lines in the system (the aggregate amount to be paid to the owners of the lines). We investigate now the allocation of the revenues associated with a non-conventional restriction. We will assume that this restriction is related to the system reliability and name it *relia*. As an example, restriction *relia* may impose a limit on the total output of the generation located within an area. Allowing more production in the area may represent a risk for the stability of the system if, as a result of a contingency, this area becomes isolated from the rest of the grid. The expression obtained here for the allocation of rents corresponding to restriction *relia* is by no means universal. However, from the process followed one may learn how to proceed in similar cases.

The formulation of the capacity auction is analogous to that in Model III. Now, a reliability restriction exists. This restriction limits the use of a scarce resource R . There are $\bar{R}$ units of this resource available. Each party i within a certain set contributes an amount $\bar{R}_i$ of resource R .

$$\begin{aligned} \max \quad & \sum_j B_j p_j \\ \text{s.t.} \quad & -\bar{F}_k \leq \sum_j PTDF_{j,k} p_j, g_k \quad \forall k=1,2,\dots \text{ number of corridors} \end{aligned} \quad (4.66)$$

$$\sum_j PTDF_{j,k} p_j \leq \bar{F}_k, m_k \quad \forall k=1,2,\dots \text{ number of corridors} \quad (4.67)$$

$$p_j \leq \bar{P}_j, b_j \quad \forall j=1,2,\dots \text{ number of transactions} \quad (4.68)$$

$$p_j \geq 0, a_j \quad \forall j=1,2,\dots \text{ number of transactions} \quad (4.69)$$

$$relia(p, \bar{R}_i) \leq 0, x \quad \text{reliability restriction} \quad (4.70)$$

We proceed to allocate the revenues corresponding to reliability restriction *relia*. The Lagrangian of this optimization problem is:

$$\begin{aligned} L = & -\sum_j B_j p_j + \sum_k g_k \left(\sum_j -PTDF_{j,k} p_j - \bar{F}_k \right) + \sum_k m_k \left(\sum_j PTDF_{j,k} p_j - \bar{F}_k \right) + \\ & + \sum_j b_j (p_j - \bar{P}_j) + \sum_j -a_j p_j + x relia(p, \bar{R}_i) \end{aligned} \quad (4.71)$$

According to the Kashrus-Kuhn-Tucker optimality conditions of this problem, we have:

$$\frac{\partial L}{\partial p_j} = -B_j - \sum_k g_k PTDF_{k,j} + \sum_k m_k PTDF_{k,j} + b_j - a_j + x \frac{\partial \text{relia}(p, \bar{R}_i)}{\partial p_j} = 0 \quad (4.72)$$

Therefore, the total amount offered by those agents whose bids are accepted can be expressed as:

$$\begin{aligned} \text{total amount offered} &= \sum_j p_j B_j = \\ &= \sum_j p_j \left(\underbrace{-\sum_k g_k PTDF_{k,j} + \sum_k m_k PTDF_{k,j}}_{\text{associated with congestion in the grid}} + \underbrace{b_j - a_j}_{\text{associated with restrictions on the amount of power transacted by agent } j} + x \underbrace{\frac{\partial \text{relia}(p, \bar{R}_i)}{\partial p_j}}_{\text{associated with the reliability restriction}} \right) \end{aligned} \quad (4.73)$$

An alternative expression exists for those terms that are associated with restrictions limiting the flow on lines:

$$\sum_j p_j \sum_k -g_k PTDF_{k,j} = -\sum_j \sum_k p_j g_k PTDF_{k,j} = -\sum_k g_k \underbrace{\sum_j p_j PTDF_{k,j}}_{-\bar{F}_k \text{ if } g_k \neq 0} = \sum_k g_k \bar{F}_k \quad (4.74)$$

$$\sum_j p_j \sum_k m_k PTDF_{k,j} = \sum_k m_k \underbrace{\sum_j p_j PTDF_{k,j}}_{\substack{\bar{F}_k \text{ if } m_k \neq 0 \\ \text{(the line is congested} \\ \text{in the negative direction)}}} = \sum_k m_k \bar{F}_k \quad (4.75)$$

Thus, the aggregate value of all these terms is:

$$\text{congestion value} = \sum_k (g_k + m_k) \bar{F}_k \quad (4.76)$$

We also know that $a_j p_j = 0$ for any of two possible options:

- If equation (4.69) is binding then $p_j = 0$
- If equation (4.69) is not binding then $a_j = 0$

Shadow price b_j corresponds to the value for the system of an additional unit of transmission rights bought by agent j . It represents the increase in the objective function that would result from selling an extra unit of rights to agent j . Hence, the

term $b_j p_j$ relates to the value for the system of purchases by this agent, i.e. it does not correspond to a revenue to be perceived by any other party.

Finally, the term $\sum_j p_j x \frac{\partial \text{relia}(p, \bar{R}_i)}{\partial p_j}$ relates to the cost for the system of reliability restriction *relia* where p is the vector of transmission rights amounts bought by the different agents.

Then, total payments by all agents are:

$$\text{total payments} = \sum_k (g_k + m_k) \bar{F}_k + \sum_j p_j x \frac{\partial \text{relia}(p, \bar{R}_i)}{\partial p_j} \quad (4.77)$$

According to the revenue adequacy criterion, revenues from the capacity auction must be enough to pay the amounts owed to the different parties. The revenue adequacy criterion must hold in order for the financial position of the TSO not to be compromised. In this case, creditors are the owners of the lines, who must receive an amount proportional to the value in the dispatch of the line or lines they own, and the providers of the scarce resource R . Revenues to be received by the owners of the lines should, in principle, be:

$$\text{total revenues line owners} = \sum_k \bar{F}_k \left(-\frac{\partial L}{\partial \bar{F}_k} \right) = \sum_k \bar{F}_k (g_k + m_k) \quad (4.78)$$

This amount is equal to the congestion cost of lines. Revenues to be perceived by the providers of scarce resource R should, in principle, be:

$$\text{total reliability revenues} = \sum_i \bar{R}_i \left(\frac{\partial L}{\partial \bar{R}_i} \right) = x \sum_i \bar{R}_i \left(\frac{\partial \text{relia}}{\partial \bar{R}_i} \right) \quad (4.79)$$

Therefore, we have:

$$\begin{aligned} \overbrace{\sum_k (g_k + m_k) \bar{F}_k + \sum_j p_j x \frac{\partial \text{relia}(p, \bar{R}_i)}{\partial p_j}}^{\text{payments by agents}} &= \\ &= \overbrace{\text{Coef}_{\text{cong}} \sum_k \bar{F}_k (g_k + m_k) + \text{Coef}_{\text{relia}} x \sum_i \bar{R}_i \left(\frac{\partial \text{relia}}{\partial \bar{R}_i} \right)}^{\text{revenues}} \end{aligned} \quad (4.80)$$

Factors $\text{Coef}_{\text{cong}}$ and $\text{Coef}_{\text{relia}}$ may be necessary to ensure that revenues are not higher than payments. The amount distributed among creditors must not be larger than the amount collected. Assuming that revenues from the allocation of a scarce resource

must be devoted to paying the providers of this resource, we conclude that $Coef_{cong} = 1$. Hence, the following expression tells us how to allocate the revenues associated with reliability restriction *relia*.

$$x \sum_j p_j \frac{\partial relia}{\partial p_j} = Coef_{relia} x \sum_i \bar{R}_i \left(\frac{\partial relia}{\partial \bar{R}_i} \right) \quad (4.81)$$

According to equation (4.79), revenues from the existence of a reliability restriction should be given to those parties that help keeping the system under safe operating conditions by providing the required service. Using the notation proposed here, party *i* contributes to the reliability service providing an amount $\bar{R}_i$ of the scarce resource *R*. Hence, according to marginal theory, the contribution of party *i* to the reliability service is, in economic terms, $x \bar{R}_i \frac{\partial relia}{\partial \bar{R}_i}$. The same applies to the rest of parties.

Coefficient $Coef_{relia}$ is necessary to guarantee that payments to parties are not higher than revenues in the auction corresponding to the reliability restriction. Then, the final payment due to each party is $Coef_{relia} x \bar{R}_i \frac{\partial relia}{\partial \bar{R}_i}$.

If the party or parties contributing to the reliable operation of the system could not be identified, revenues from restriction *relia* should be socialized among agents. Thus, revenues from this restriction might be allocated in proportion to the transmission tariff each agent has to pay.

In the example provided above, where the total electricity generation in an area was limited to a value, generators producing in the area should pay the cost of this restriction in proportion to the output of each of them. On the other hand, owners of the rest of generation capacity in the system should earn the corresponding revenues. The limit on the generation output in the area is related to the generation capacity reserve in the region. In fact, if more generation capacity existed, more generators would be able to contribute to automatically restoring the balance between generation and load in the system, whenever the electricity output in the considered area or node is lost. Consequently, the limit on the production in this area could be higher. Thus, the total cost of the restriction ($\bar{R}_i x$ according to the formulation above) would have to be allocated to generators in other areas in proportion to their effective generation capacity.

4.4.3 Definition of Single Price Areas

A Single Price Area (SPA) is a set of nodes that is considered as a single one for congestion management purposes. Differences in energy prices among the nodes within a SPA are only related to losses. Defining Single Price Areas is central to the

implementation of any zonal congestion management scheme. Nodal pricing is the only reasonable scheme in those systems where congestion plays an important role and accurate enough SPAs do not exist.

Local markets can manage the dispatch of energy within a SPA independently from the rest of the system. Thus, the identity of agents involved in commercial transactions taking place within each area, as well as the amount of power transacted and the price paid for this power may be kept confidential. Lastly, managing congestion taking place on the border between SPAs is certainly far less complex than considering the whole system grid.

A distinction must be made between sporadic and systematic congestions. Sporadic congestions occur rarely and are not relevant for the outcome of the market. Systematic congestions occur more frequently and affect significantly the economic dispatch in a country or region. Strictly speaking, all congestions in the system should be managed using efficient coordinated market mechanisms. In practice, leaving authorities within each area in charge of solving sporadic congestions may be acceptable.

The remainder of this section is structured as follows. Subsection 4.4.3.1 analyzes from a conceptual point of view the problem of defining SPAs. It includes a description of the identification method that has been developed and applied during the realization of this thesis. Afterwards, 4.4.3.2 presents some numerical results corresponding to a preliminary division into SPAs, both of the Iberian peninsula and of the system comprised of France, Spain and Portugal. Finally, section 4.4.3.3 draws some conclusions.

4.4.3.1 Conceptual analysis of the problem

This section begins with a review of different techniques for the identification of SPAs. A description of the method that we have investigated follows. Finally, some alternatives to this method are described as well.

Introduction. Brief discussion of the different approaches to the identification of SPAs

In the absence of losses, all nodes in a Single Price Area should have the same energy price. A transaction taking place between nodes belonging to the same SPA should not affect the flow on the congested lines or corridors. Different patterns of use of the network may result in different congestion patterns. Thus, boundaries of SPAs may vary from one scenario to another. Any correctly defined SPAs should be valid for any possible network configuration and under any possible system conditions.

Using a linear model to represent the way power flows over the network, one may conclude that all the power transactions between two SPAs must have the same impact on the flows of congested lines. Therefore, all the nodes in a SPA may be represented using only one. Two transactions of the same size starting and ending in the same two areas are then totally equivalent.

Two nodes belong to the same SPA if, and only if, two transactions taking place between each one of these two nodes and any other one produce the same flows on the congested lines. This must be true under any set of operating conditions.

Thus, provided SPAs actually exist, nodes belonging to each SPA may be identified as those such that all elementary transactions taking place between them and a common reference node produce the same flows on the congested lines. This may be assessed using the Power Transfer Distribution Factors of the flow through the congested lines with respect to these elementary transactions. This is the approach that we have followed to define SPAs. Others in the academic community have proposed similar algorithms to identify these areas, Thus, authors in (Kumar, Srivastava et al. 2004) divide a system into several areas according to how similar the PTDFs obtained for the different nodes are. Afterwards, they re-dispatch generation within each area to manage congestion in the grid. Generators picked up to participate in re-dispatch are those whose PTDFs are large and different from those obtained for any other node, i.e. those that are close to the congested lines. Kavicky and Shahidehpour identify price areas using the PTDFs of the flow through the interconnection lines between control areas with respect to transactions, see (Kavicky and Shahidehpour 1997). Unfortunately, defining clear-cut SPAs in meshed systems may not be possible. Using a simplified representation of the system grid may require accepting some margin of error in our calculations. Therefore, defining SPAs may probably result in some efficiency losses in the management of congestion.

If the resulting error is significant, we may have to use different criteria to classify nodes into areas. Specifically, when several congested paths exist between two nodes (the usual situation in meshed networks), the difference in prices between them depends on the following two sets of variables:

1. The flows on the congested lines that are caused by a unit transaction between the two nodes.
2. The marginal value of the restrictions limiting the flow on congested lines.

As a consequence of this, nodes may be classified according to their nodal prices. Basically, this would amount to classifying nodes into areas using nodal prices as the classification variable, instead of the flows on congested lines. However, other methods to identify SPAs exist that also make use of economic information. Thus, (Björndal and Jörnsten 2001) proposes a method aimed at minimizing the welfare loss caused by the consideration of SPAs. However, the way authors implement this principle is questionable. Thus, the authors assume that the marginal energy price computed for all the nodes within the same SPA must be the same. Taking into account this constraint, the paper computes the division of a system into SPAs so that the social welfare resulting from the corresponding zonal dispatch is maximum. Authors argue that simply grouping nodes based on their nodal prices in the optimal dispatch is unlikely to result in the best possible division into areas. They provide examples where

application of their algorithm yields SPAs that include congested lines within them (instead of connecting different SPAs). They also show that the optimal division into areas may result in negative congestion rents. See at the end of this subsection the description of some alternatives to the method we propose.

If the best possible division into areas is not accurate enough, one may think of dealing independently with those transactions for which the error committed is highest. The error committed for a transaction may be obtained as the difference between the power flows on the congested lines produced by this transaction and the flows produced by a typical transaction between the same two areas. Transactions that are not typical may be called special. Both transactions internal to an area and transactions between different areas may qualify as special. Special transactions must receive specific treatment in the congestion management mechanism. They must be dealt with as transactions taking place between two specific nodes instead of two areas (or instead of transactions internal to an area that can be ignored).

Identifying those transactions that deserve specific treatment involves setting a threshold for the error committed when considering a transaction as a typical inter-area or intra-area one. If, once a system is divided into SPAs, the difference between the set of flows produced by a transaction T on the congested lines and those characteristic of a transaction between the same two SPAs is above this threshold, transaction T must receive a specific treatment. Specific PTDFs of the flows on the congested lines with respect to transaction T must be computed. Obviously, this is at odds with the application of a zonal congestion management system. Hence, dealing individually with certain transactions is only possible if there is a small number of these transactions. We must compute the optimum trade-off between the error committed when dividing the system into areas, which is higher the higher the threshold is, and the number of special transactions, which is higher the lower the threshold is.

The next paragraphs present the method that has been applied in our study. This method classifies the nodes in a system according to the flows that elementary transactions taking place between them and a reference node cause on the congested lines. As discussed above, this method is valid only when the fraction of the scheduled intra-area transactions that flows through congested lines is not significant. This section also includes an example of application of the described method.

Description of the method that is used in this thesis

We have used clustering techniques to classify nodes into zones. The methodology proposed here makes use of the topology of the grid. We must minimize the use that intra-area transactions make of the lines that are likely to be congested. In addition, nodes in an area must be electrically connected among themselves.

We have computed Power Transfer Distribution Factors of the flows over the potentially-congested lines with respect to elementary transactions between each node

in the system and the reference node. Once this set of PTDFs has been computed, estimating the impact of a transaction between any two nodes in the system on any potentially congested line is immediate. We only need to compute the difference between the PTDFs corresponding to the injection and withdrawal nodes in the transaction. Nodes belonging to the same area have the same elementary PTDFs since a transaction between them does not affect the flows on the congested lines.

Clustering algorithms produce groups of samples such that the distance (difference) among members of the same group is minimum, while the distance between elements belonging to different groups is maximum. Two different algorithms have been used: *KSOM* and *kmeans*, see (Hartigan 1975; Kohonen 1995). The number of groups, or clusters, is an input to the model. We do not know in advance which is the ideal number of groups (areas). Therefore, several classifications corresponding to different numbers of clusters must be obtained.

Due to the fact that most clustering algorithms are iterative processes, an initial solution (classification) must be provided. The initial solution is used by the algorithm as a starting point in the iterative process. The algorithm rarely arrives at the true optimum and the final solution obtained depends on the starting point considered. For this reason, several experiments have been conducted for any given number of clusters or areas. Thus, we can assess the stability of results with respect to the choice of the starting point.

All nodes within any of the areas must be electrically connected among themselves. Prices in nodes that are not electrically close within a congested system cannot be thought to evolve jointly with each other. Thus, nodes in an area that are not electrically connected to the remaining nodes have been assigned to another area.

Afterwards, we computed the overall error for each division into areas. Intra-area transactions should not modify the flow on congested lines. Therefore, the error corresponding to an intra-area transaction T may be defined as the amount of power produced by transaction T that flows through congested lines. The overall error committed for a given classification has been characterized using the maximum and average transaction errors over all the possible transactions. In addition, we have taken into account the maximum and average errors over all the transactions within each area.

After computing the error corresponding to a classification, we assessed the performance of this classification according to several criteria. The list of considered criteria follows:

- The maximum transaction error committed for the whole classification as well as within each zone must be as small as possible.
- The average transaction error committed for the whole classification as well as for each zone must also be as small as possible.

- The number of nodes within each area must be high enough.
- The division into areas obtained must be compatible with the existing political and organizational structures.

Finally, the classification that ranks highest, taking into account at the same time the above mentioned criteria, must be implemented.

Clustering techniques applied

Two different clustering algorithms have been used in the study: the Autoregressive *Kohonen* Maps (KSOM) and the *kmeans* algorithm. These two algorithms are of two different types. When dividing a set of samples into several groups or clusters using the *KSOM* method, we may define a higher number of cells (elementary groups of samples) than the number of clusters we want to create. Thus, we say that the *KSOM* algorithm is a flexible method. On the other hand, the number of cells employed in the *kmeans* algorithm must be the same as the number of clusters we want to define. Thus, we say the *kmeans* algorithm is a rigid one.

The *KSOM* algorithm represents on a two-dimensional cell map the samples to be classified into areas. The number of cells in the map is an input to the model. A color code is used to specify the distance (difference) that exists between any two samples. When several classification variables exist (as it is the case here, where each variable corresponds to the flow on a different congested line) separate maps can be drawn for each of them. In these maps, colors represent the value that the corresponding variable adopts in the different samples.

From these maps it is possible to identify groups of samples that are similar among themselves but are quite different from the rest of the samples. The number of cells in the map may coincide with the number of groups of similar samples identified or the number of clusters we want to obtain. In this case, each cell corresponds to a different cluster (area). Otherwise, we have to group the cells into clusters.

We have applied the *KSOM* algorithm to compute areas in the Iberian Peninsula. The resulting maps contain 2, 12 and 25 cells. Several maps have been produced for each number of cells in order to assess whether the results are robust. Maps in Figure 4.4-1 correspond to the winter peak bad of the Iberian system in 2004. Labels in the map relate to different regions in the system. The label of each cell represents the region in the system where samples in this cell are most abundant. Classification variables in this case are the flows on three congested lines. The map on the top left corner represents the samples (nodes) in the three dimensional space defined by the three classification variables (flows on the lines). Each one of the other three maps corresponds to a different classification variable (the meaning of the maps has been explained before in this section).

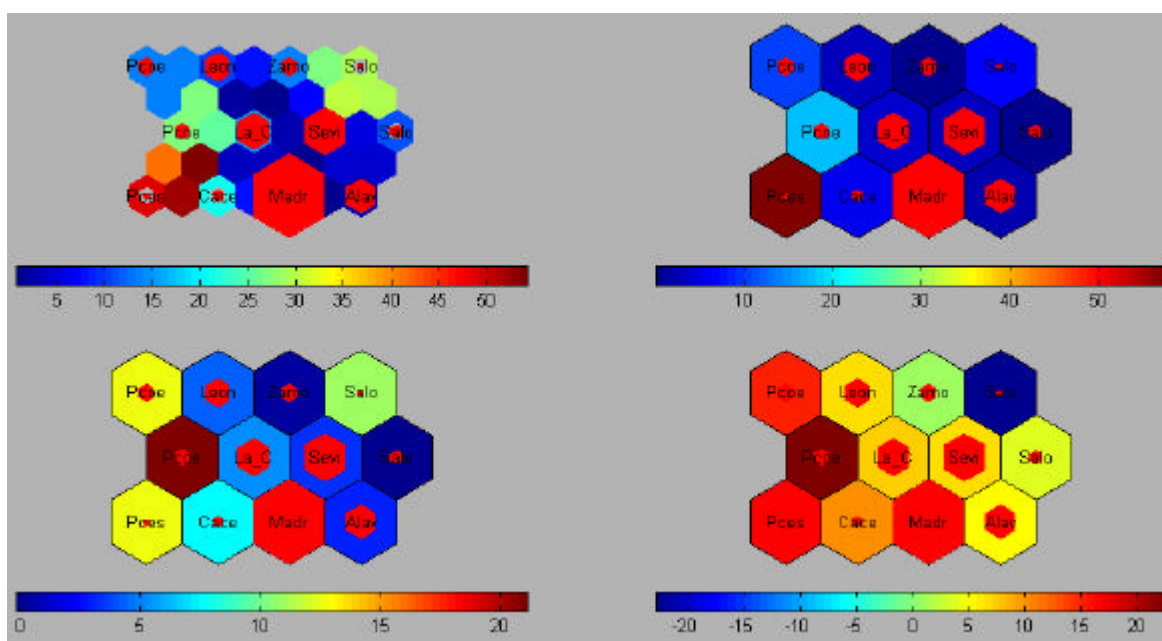


Figure 4.4-1: multivariable Kohonen map and maps corresponding to each of the classification variables for the division into areas of the Iberian peninsula when 12 cells are used

Figure 4.4-2 shows the content of each one of the cells of the map in Figure 4.4-1. Using both figures it is possible to define two SPAs in the Iberian peninsula and locate each one of these areas on a map. Samples belonging to the smaller area are shown in bold in Figure 4.4-2.

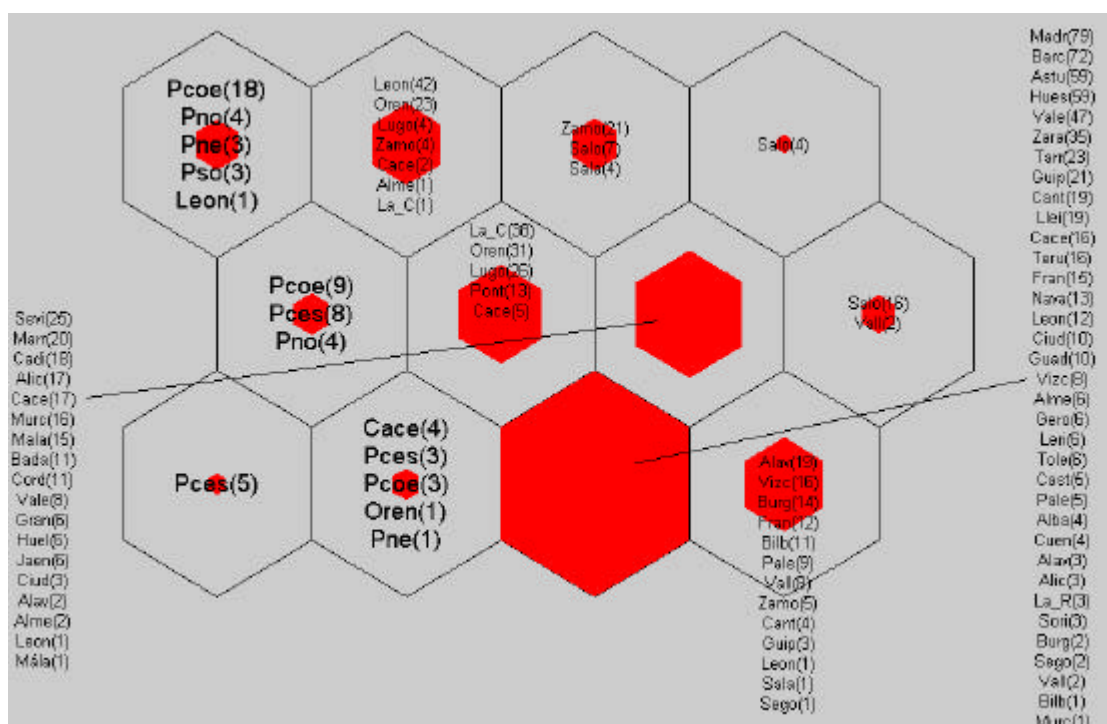


Figure 4.4-2: void map showing the identity of the samples belonging to each of the cells in the previous kohonen map. Two clusters (SPAs) have been identified. We show in bold the identity of the samples within the smaller area.

The algorithm of *kmeans* does not make a distinction between cells and clusters. Hence, we must use as many cells as SPAs we want to identify in the system. Like the *KSOM* algorithm, the *kmeans* algorithm may produce cell maps. However, the distance between two cells on the map bears no relation to the difference between the values adopted by the set of classification variables in the two corresponding clusters. Hence, representing samples on a map does not make sense.

The clustering techniques used here are just a tool for the computation of results but not the focus of our research. Hence, an in depth description of these techniques is beyond the scope of this thesis work. A description of other, more sophisticated, methods for the identification of SPAs follows.

Alternative methods to compute SPAs

This subsection describes two other methods to define SPAs in a meshed system. Conceptually speaking, both methods are superior to the one applied here. Their main drawback lies in the difficulties we may face when implementing them. These are mainly related to the fact that both make use of economic information that is not easily accessible. Part of this information would certainly be considered confidential by market agents in many systems. However, if market agents and institutions are ready to submit information on their costs or bids in the future, we should reconsider their application.

Classification of nodes into SPAs using nodal prices

More accurate methods exist to compute SPAs. One possible modification with respect to the procedure followed here has already been commented in the introduction of this section. Identifying clear-cut areas is not possible in meshed networks. Thus, strictly speaking, nodes should be grouped according to their energy price in each one of the scenarios where a line gets congested. Classifying nodes according to the flows on the congested lines that can be attributed to each one of them is only a second best option. Consequently, the process of identification of SPAs should ideally be structured in the following stages:

1. Simulation of an optimal power flow for all those scenarios where a line gets congested.
2. Classification of nodes into areas for each one of these scenarios according to the corresponding nodal prices.
3. Identification of SPAs as the sets of nodes that belong to the same area in every scenario.

An alternative to stage 3 of this procedure exists. If SPAs resulting from this process were too small, or we could not identify any SPA at all, nodes could be grouped into areas according to the vector of nodal prices obtained for each node and the set of

scenarios considered as a whole. The vector of nodal prices corresponding to each node would be the classification variable to be used to form clusters (SPAs).

However, application of this mechanism faces many difficulties. First, running an optimal load flow for a regional system like Europe, where many different markets exist, and agents can contract power both in centralized markets and bilaterally, is unthinkable nowadays. This is partly due to the fact that the economic information required to compute nodal prices in Europe is unlikely to be available in the short and medium term.

Second, there will always be transactions whose impact on binding constraints is significantly different from the impact of a typical transaction between any two areas in the system (including those transaction internal to an area, which are supposed not to create any flow on the congested lines). Therefore, even if we classify nodes according to their nodal prices, we will also have to identify problematic transactions and manage them separately from any other transaction in the dispatch. In order to identify problematic transactions, we would also have to set a threshold for the error committed in each transaction. Computing the impact of elementary transactions on the flows of congested lines would be necessary.

Thus, classifying nodes into areas according to their nodal prices has some important drawbacks. In any case, computing the optimal system dispatch for the set of possible operating conditions, if feasible, would always be useful. This would allow us to identify those lines that would actually get congested if the system were operated optimally. Note that, presently, congestion is not managed efficiently in most regional systems. However, SPAs must be defined assuming generation and demand are dispatched optimally.

Obviously, computing SPAs in radial systems according to the flows produced by elementary transactions on the congested lines is equivalent to identifying them using nodal prices. However, when the division in areas is not completely accurate both methods may well result in different sets of SPAs.

Method based on the minimization of the welfare losses

Any division of a system into Single Price Areas must be assessed with respect to two main features:

- How simple to implement it is.
- How efficient the operation of the corresponding zonal system is.

The level of compliance with the first objective can be measured in terms of the number of SPAs defined and their identity (whether they can be assimilated to existing political or organizational structures). The level of compliance with the second one may be measured in terms of the welfare losses incurred when implementing the proposed

division into areas with respect to the most efficient solution possible (the nodal pricing scheme).

Provided the defined SPAs are of enough size, we could say that, for any number of SPAs, the best division is the one that most closely approximates the outcome of a nodal pricing scheme. Thus, considering a certain snapshot of a system, we may assess how efficient a division into areas is following the process described now:

1. We compute the optimal dispatch for the system. This needs to be done only once for each snapshot.
2. We then compute the outcome of the zonal dispatch assuming certain areas and the corresponding set of PTDFs for the flows on the likely-to-be-congested lines with respect to the transactions between areas.
3. Finally, we compute the efficiency losses incurred for this division into areas. Losses may be of two types:
 - a. The difference between the total system welfare in the reference case and that corresponding to the zonal dispatch obtained.
 - b. The economic cost associated with the infeasibilities resulting from the zonal dispatch. This may be approximately measured as:

$$ECI = \sum_l f_{excl} g_l$$

where ECI is the economic cost of the infeasibilities, l is the set of congested lines, f_{excl} is the flow on line l in excess of the capacity of the line and g_l is the dual variable of the binding constraint that limits the flow on line l .

Total welfare losses may be obtained as the sum of losses of both types.

We now have a method to assess the efficiency of a classification. A heuristic process similar to the one described in (Latorre-Bayona and Pérez-Arriaga 1994) could be employed to search for the most efficient classification of the system nodes into a certain number of areas. In this process, nodes would be transferred from one area to another, one at a time, and efficiency losses would be evaluated for both the original and the final classifications. In this way, we could assess whether a certain change in the division of the system into SPAs results in a more efficient dispatch or not. A certain initial solution (classification of the nodes into areas) must be computed somehow. This method is similar to the one described in (Björndal and Jörnsten 2001).

Once we have computed the best possible division of the system into different numbers of areas we must determine how many areas to define. In principle, we

should choose the division into SPAs that achieves the optimum trade-off between the simplicity of implementation and the efficiency of the dispatch.

Now, we move on to present the results obtained when applying the method proposed above to identify SPAs in those cases where no economic information is available. This method has been used to identify SPAs both in the Iberian market and in the system comprised of Spain, France and Portugal.

4.4.3.2 Identification of Single Price Areas in the Iberian Peninsula and the system Spain-France-Portugal

We have applied the approach that was proposed in the previous section to define SPAs within two actual systems: the Iberian peninsula and the sub-region including France, Spain and Portugal. The classification variables used in our analysis are the flows produced by elementary transactions on a set of lines identified by the Spanish TSO as belonging to a corridor that is likely-to-be-congested. We were not able to use other, more sophisticated, classification variables, like nodal energy prices, since we did not have available the required economic information for the two systems.

When computing SPAs, we have not dealt separately with those transactions for which the error committed is highest. However, the error for some transactions may be excessive. Specifically, 50% of the power transacted between some pairs of nodes located within the same area flows through a congested line. Probably, these transactions and others should be considered as node-to-node transactions instead of area-to-area transactions. In other words, they should be dealt with separately from typical inter or intra-area transactions. Considering some transactions separately from the rest may result in SPAs that are slightly different from the ones obtained when all transactions are taken into account. At the end of this section we carry out a preliminary analysis to determine the level of the threshold to be used in the identification of special transactions. The identification of special transactions may be part of future research work.

The piece of research presented here is just a first attempt to define SPAs. Further work is needed to obtain satisfactory results. At least, we show that the methodology proposed here is viable and that it provides a tentative division into areas of the two explored systems.

Results for the identification of areas in the Iberian peninsula

According to some first estimates by the Spanish System Operator, three lines may be systematically (frequently) congested in the Iberian system presently. All of them are cross-border lines, i.e. they cross the border between the Spanish and the Portuguese systems. A list of these lines follows:

- Line between the 220KV nodes named ‘Aldeadávila’ and ‘Bemposta’. The former node is located in the Spanish system whereas the latter is located in Portugal.

- Line between the 220KV nodes named 'Aldeadávila' and 'Pocinho'. The latter is located in Portugal.
- Line between the 220KV nodes named 'Saucelle' and 'Pocinho'. The former node is located in the Spanish system.

All the nodes are part of one of the main electrical corridors between the Spanish and the Portuguese systems. The three lines have a nominal capacity of 380MW.

We have applied the proposed methodology to a snapshot corresponding to the winter peak load of the year 2004. The snapshot is comprised of 1200 nodes. Each node has been assigned a label representative of the Spanish province or the region in Portugal where it is located.

Results corresponding to the division of the system into two SPAs

We have classified the nodes of the Iberian peninsula into two groups of them or SPAs. Several classifications have been obtained. As mentioned above, two different clustering algorithms have been applied: *KSOM* and *kmeans*. Maps of 2, 12 and 25 cells have been produced using the *KSOM* method. If the number of cells is higher than the number of specified areas, then cells must be grouped into areas. Several experiments have been conducted for each combination of a clustering algorithm and a number of cells to be explored.

Figure 4.4-3 shows the maximum transaction errors committed for each of the experiments that were carried out. Labels used in the horizontal axis relate, in this order, to the clustering algorithm that was used, the number of considered cells (only for *KSOM*), the experiment number and the way cells have been grouped to form areas (only for *KSOM*). Thus, 'kmeans1' relates to the first experiment carried out using the *kmeans* algorithm, while 'ksom 12.2.1' relates to the second experiment where the *KSOM* algorithm has been used to produce maps of 12 cells. We found three different reasonable ways of grouping the 12 cells resulting from this experiment into areas. The classification labeled 'ksom12.2.1' corresponds to the first one. Finally, the first classification in the horizontal axis, labeled 'area=country', relates to the case where areas are coincident with countries.

The maximum error committed for each classification is represented in the vertical axis. Errors are expressed in MW and result from simulating the impact of 100 MW elementary transactions on the flows through the congested lines. Therefore, the maximum error for a classification represents the percentage of the volume of an intra-area transaction that flows through the congested lines in the worst case.

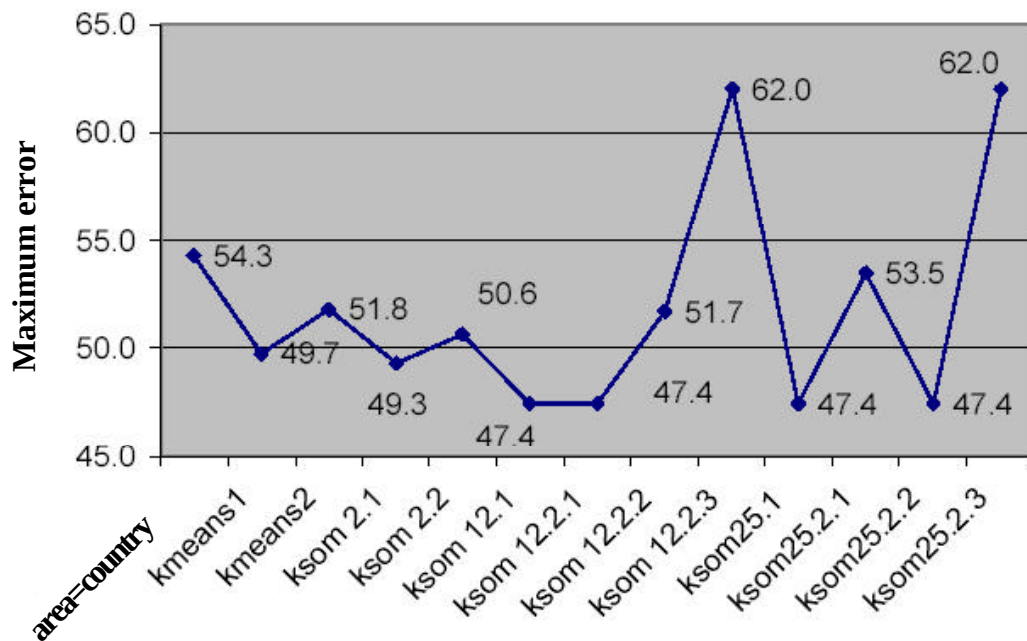


Figure 4.4-3: Maximum errors committed for the different classifications obtained corresponding to the division of the Iberian peninsula into two areas

Figure 4.4-4 shows the average errors committed for the different classifications obtained when dividing the Iberian peninsula into two areas. Numbers are expressed in MW. They have been computed for elementary transactions of 100MW. Thus, they mean the percentage of the power transacted that flows on the congested lines on average. For example, considering classification 'ksom12.2.2' one can learn from Figure 4.4-4 that, on average, 2.1% of the flow of any intra-area transaction uses the congested lines.

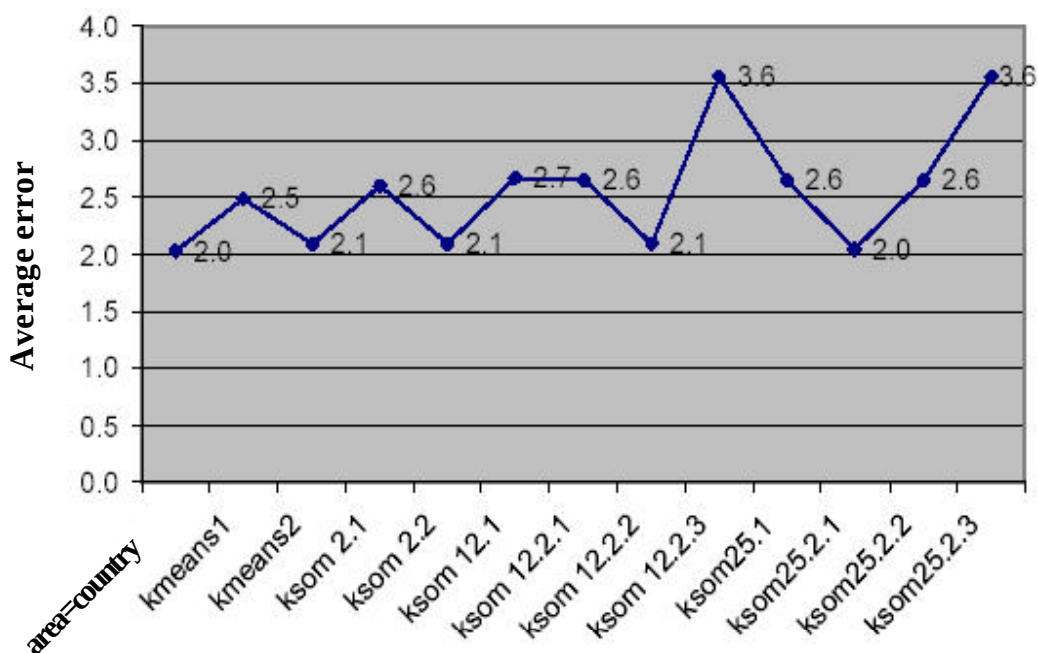


Figure 4.4-4: Average errors committed for the different classifications obtained corresponding to the division of the Iberian peninsula into two areas

When assessing a certain classification, one may focus on the transactions that are internal to each SPA when considered separately. Thus, the error committed within each SPA may also be a meaningful measure of the overall performance of a classification. Figure 4.4-5 and Figure 4.4-6 show the maximum and average transaction errors for each one of the SPAs in each of the classifications considered.

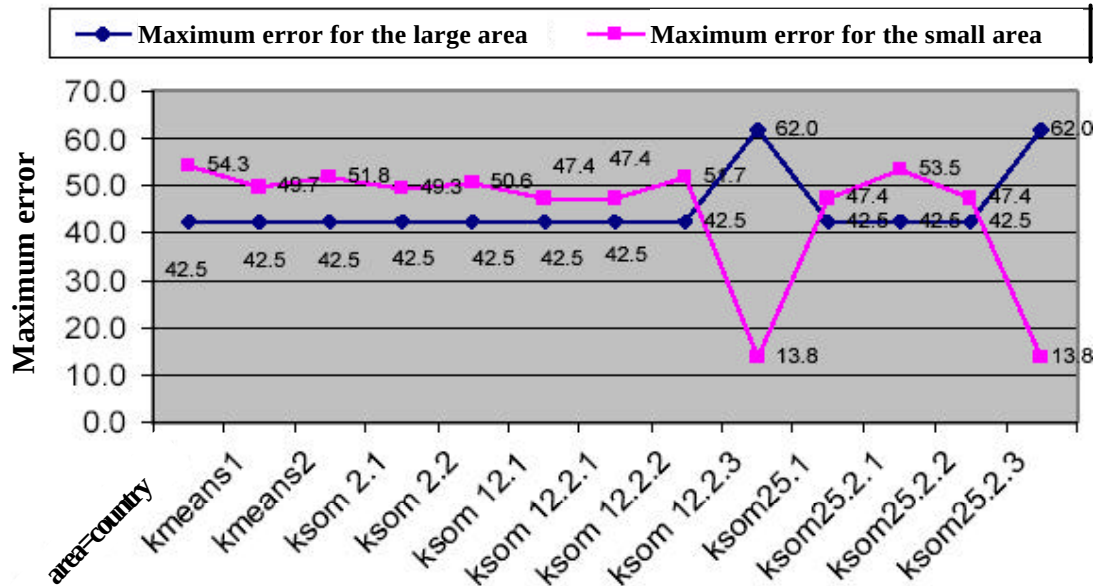


Figure 4.4-5: Maximum error committed within each SPA for all the classifications computed corresponding to the division of the Iberian peninsula into two SPAs

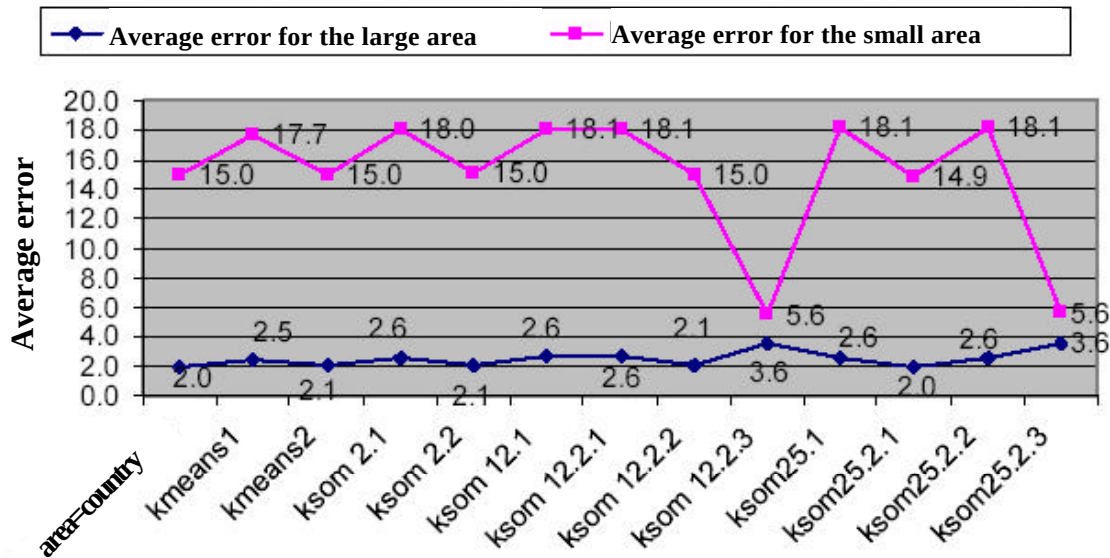


Figure 4.4-6: Average error committed over all the intra-area transactions within each SPA for all the classifications computed corresponding to the division of the Iberian peninsula into two SPAs

Lastly, Figure 4.4-7 shows the maximum and average errors corresponding to all the considered classifications.



Figure 4.4-7: Comparison of the maximum and average errors for all the classifications considered corresponding to the division of the Iberian peninsula into two SPAs

Based on the comparison that has been carried out, one may conclude that the best performing classification is that where SPAs are the same as countries. Transactions taking place between nodes where no generation is located will certainly not play any role in the management of congestion since all the demand will be served in normal conditions. Therefore, we can discard them. Then, the maximum transaction error committed in the chosen classification decreases from 54.3% to 54.1%.

Results corresponding to 3 areas

This section presents the results obtained when dividing the Iberian peninsula into three areas. All of the classifications but one include at least one area that is comprised of less than 5 nodes. According to the evaluation criteria employed, this is not acceptable. Hence, the remaining classification was chosen as the best possible division of the Iberian peninsula into three areas. Now we list the nodes included within each SPA:

- Area 1 includes all the Portuguese nodes.
- Area 2 includes 27 nodes in the western part of Salamanca, 21 nodes in Zamora, 4 in Salamanca and 3 in Valladolid.
- Area 3 includes the remaining Spanish nodes.

The maximum intra-area transaction error committed in area 1 is 18%, 54,3% in area 2 and 35,9% in area 3. The average error committed for area 1 is 1,3%, 15% for area 2 and 6,8% for area 3. The overall average error is 1,4%. This means that, on average, 1,4% of the power corresponding to an intra-area transaction flows through the congested lines. When we discard the transactions taking place between nodes where no

generation is located, the maximum intra-area transaction error decreases from 54,3% to 54,1%.

Results corresponding to 4 areas

This section provides the results obtained when dividing the Iberian peninsula into four SPAs. All the classifications but two include at least one area that is comprised of less than 5 nodes. Therefore, we restrict our analysis to the remaining two classifications. Table 4.4-1 compares the maximum and average errors committed in each of these two classifications.

	Maximum error		Average error	
	Class. 1	Class. 2	Class. 1	Class. 2
Area 1	35.9	8.7	6.8	1
Area 2	20.6	35.9	6.6	6.8
Area 3	47.4	54.3	24.8	15
Area 4	18.2	18.2	1.4	0.8
Overall	47.4	54.3	1.4	0.9

Table 4.4-1: Comparison of the maximum and average errors committed in each of the two classifications selected for dividing the Iberian peninsula into 4 areas. Values are expressed in %.

One may probably conclude from results in Table 4.4-1 that classification 2 is preferable to classification 1, since the maximum transaction error committed in the former is similar to that committed in the latter (though a bit larger), while the average error committed when implementing classification 2 is clearly below that committed when implementing classification 1. A list of the nodes belonging to each area in the classification finally chosen follows:

- Area 1 includes 39 nodes located in La Coruña, 4 located in Leon, 30 nodes in Lugo, 47 nodes in Orense, 13 in Pontevedra and 1 in the Northwest of Portugal. The total number of nodes in Area 1 is 134.
- Area 2 includes 16 nodes located in the centre east of Portugal, 30 in the centre west, 3 in the north west, 8 in the north and 3 in the south west. The total number of nodes in Area 2 is 60.
- Area 3 includes 4 nodes located in Salamanca, 27 in the west of Salamanca, 2 in Valladolid and 21 in Zamora. The total number of nodes in the area is 54.
- Lastly, area 4 includes the remaining 947 nodes of the Spanish system.

If we discard those transactions taking place between nodes where no generation is located the maximum transaction error decreases from 54,3% to 54,1%.

Final choice of a division of the Iberian peninsula into SPAs

We must determine which is the optimum number of SPAs to define. One must compare the best classifications obtained when dividing the system into 2, 3 and 4 areas. These 3 classifications have been assessed in the light of the criteria outlined in the introduction of this section. Figure 4.4-8 shows the maximum transaction errors committed in the three cases. Figure 4.4-9 compares the corresponding average transaction errors.

The maximum error seems to be unaffected by the number of areas considered while the average error is smaller the larger the number of areas defined. However, the average transaction error is quite small even in the worst case. On the other hand, as far as the practicability of each division into areas is concerned, defining two areas that are coincident with countries certainly is far less problematic than establishing three or four areas. Therefore, in the context of the Iberian electricity market, we recommend dividing the peninsula into two areas, each one coincident with a country. The maximum transaction error seems to be unacceptably high. This dictates that those transactions for which the error committed is highest should be singled out and dealt with separately from the rest (as node-to-node transactions instead of area-to-area transactions).

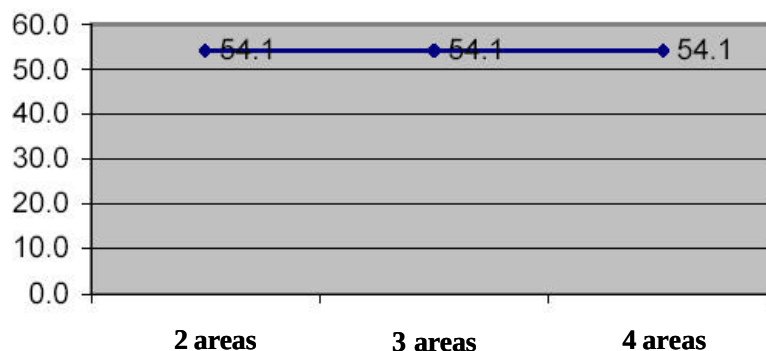


Figure 4.4-8: Comparison of the maximum transaction errors committed for the division of the Iberian system into 2, 3 and 4 areas

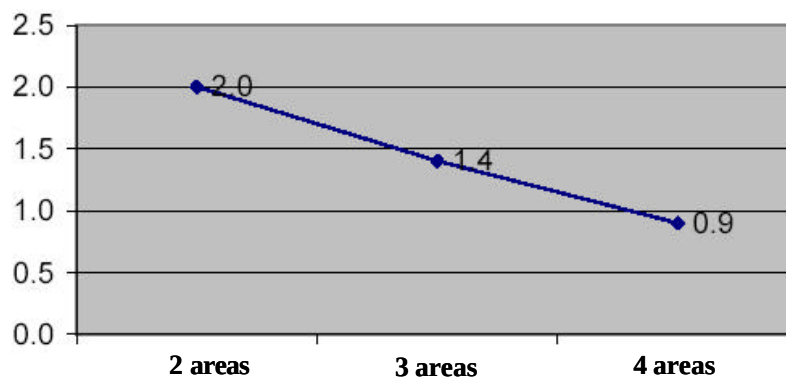


Figure 4.4-9: Comparison of the average transaction errors committed considering 2, 3 and 4 SPAs in the Iberian peninsula

Results for the identification of areas in the subsystem comprised of France, Spain and Portugal

Now we investigate the division into areas of the system comprised of France, Spain and Portugal. This time, not only the congested lines between Portugal and Spain must be taken into account but also those between the Spanish and the French systems.

The representation of the Iberian system considered here is simpler than the one studied in the previous subsection. In fact, the congested corridor between Spain and Portugal is modeled using two lines instead of three. The lines deemed to be congested are: the 220KV line between Aldeadávila (Spain) and Pocinho (Portugal), the 220KV one between Saucelle (Spain) and Pocinho (Portugal), the 220 KV line between Biescas (Spain) and Pragnere (France), the 220KV line between Vic (Spain) and Baixas (France) and the 220KV line between Hernani (Spain) and Cantegrit (France). We could have also considered the 400KV lines between France and Spain. However, results obtained when considering all cross-border lines should not differ from the ones presented here, since 400KV and 220KV cross-borderlines between France and Spain have been built in parallel.

The whole system is comprised of 1727 nodes. Each node has been assigned two different labels, one corresponding to the country where it is located and one corresponding to the area within the country. Zones in the Spanish and Portuguese systems are the same as those in the analysis carried out for the Iberian peninsula. The French system has been divided into 15 areas. These are represented in Figure 4.4-10. Note that these areas are not Single Price Areas. The system has been divided into geographical areas just to be able to easily locate each node within the corresponding country.

The number of SPAs in the system is an input to the method used to identify them. We have explored the possibility of defining 3, 4 and 5 areas. The remaining of this subsection is devoted to presenting and discussing the results obtained in each case. Finally, we choose the best possible division of the system into SPAs.

Results for the division into 3 areas

The process followed to identify Single Price Areas is the same used in the study for the Iberian peninsula. Two clustering methods have been used: *KSOM* and *kmeans*. Several classifications of the system nodes into areas have been obtained. Figure 4.4-11 illustrates the largest error committed for an intra-area transaction in each of the classifications computed. Labels in the horizontal axis relate to the clustering method used, the number of cells in the *Kohonen* map (only for the *KSOM* method) and the number of the experiment. Again, the first classification in the horizontal axis corresponds to the case where SPAs are the same as countries. Numbers are expressed in %.

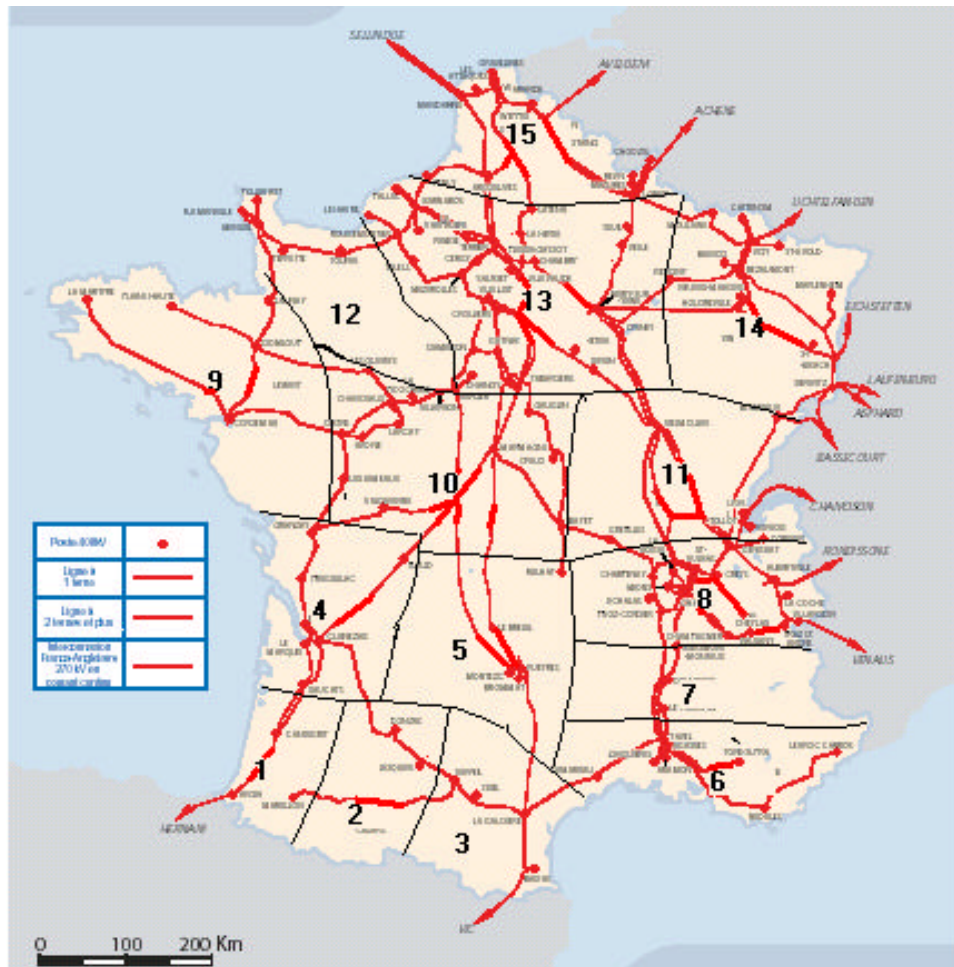


Figure 4.4-10: Map of the French system where the geographical areas considered have been represented

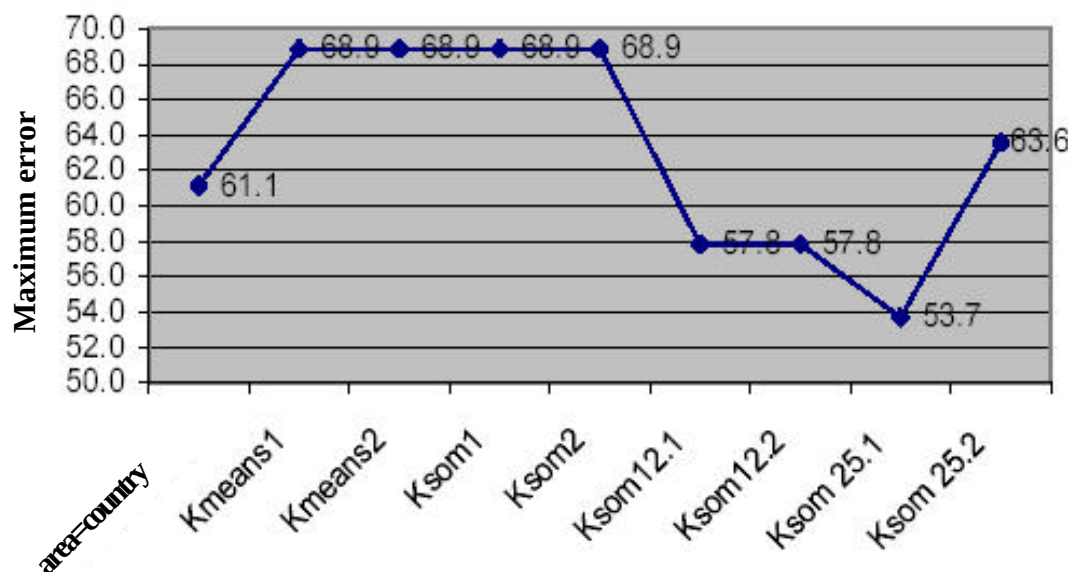


Figure 4.4-11: Maximum transaction errors committed for the different classifications corresponding to the division of the system comprised of France, Spain and Portugal into three areas

Figure 4.4-12 shows the average transaction error for each classification. This is the percentage of the power transferred in an intra-area transaction that flows through the congested lines, on average.

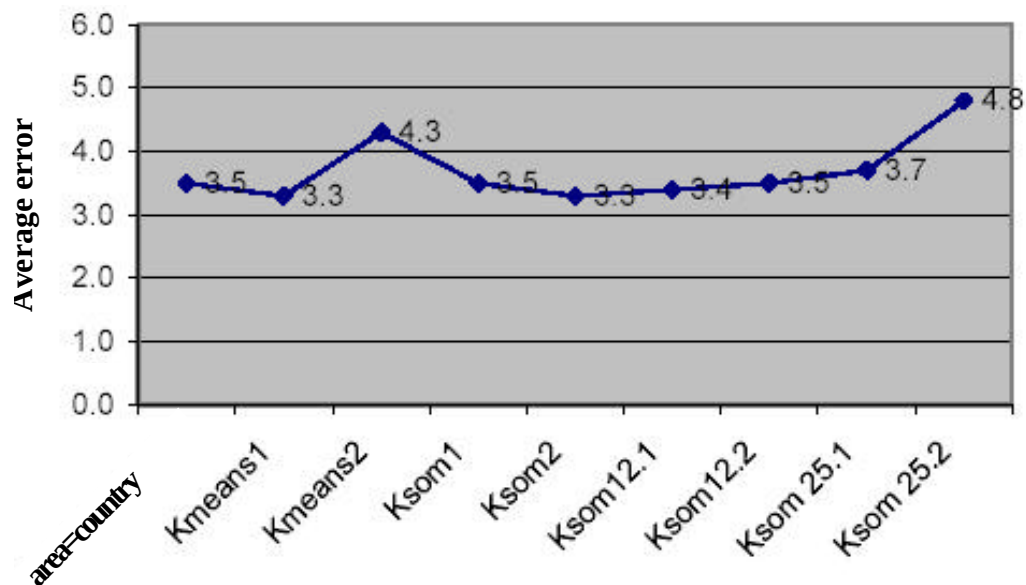


Figure 4.4-12: Average transaction errors committed for the different classifications corresponding to the division of the system comprised of France, Spain and Portugal into three areas

Now, maximum and average transaction errors are shown for each SPA within each classification, see Figure 4.4-13 and Figure 4.4-14.

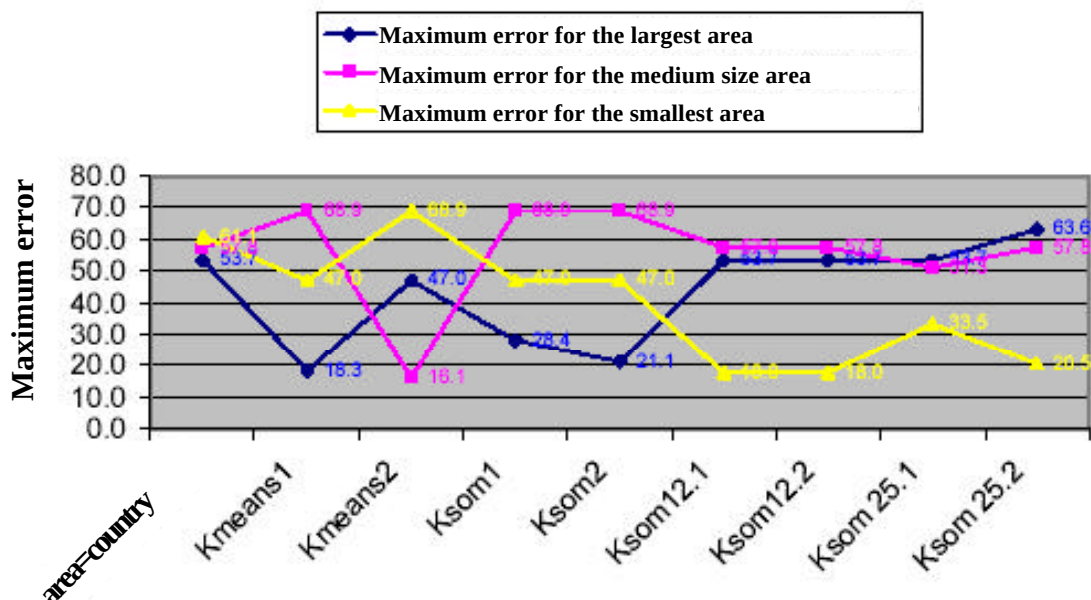


Figure 4.4-13: Maximum transaction error committed for each area and classification corresponding to the division into three areas of the system comprised of France, Spain and Portugal

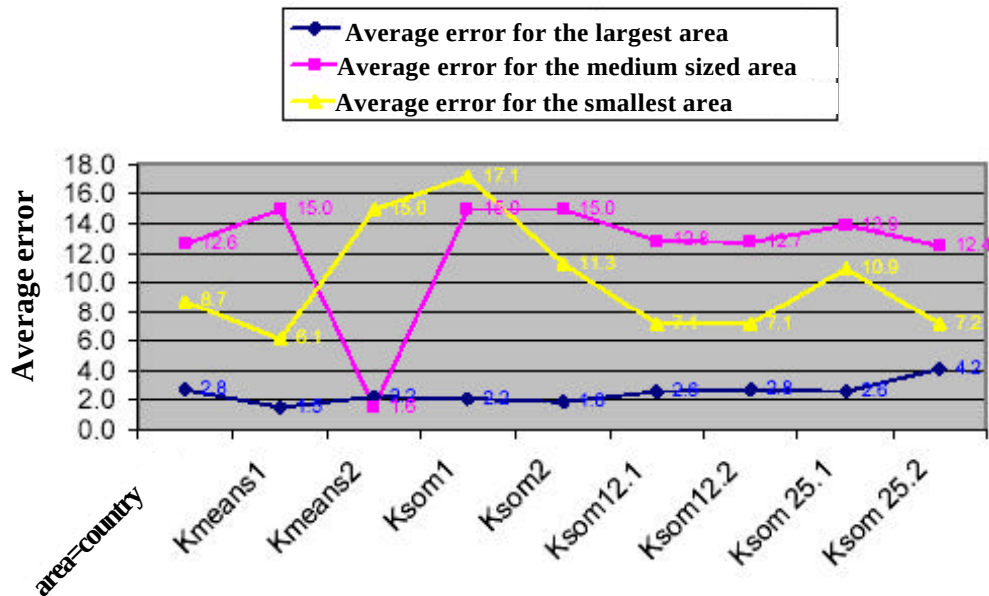


Figure 4.4-14: Average transaction error committed for each area and classification corresponding to the division into three areas of the system comprised of France, Spain and Portugal

Lastly, in Figure 4.4-15 the maximum and average transaction errors for all the classifications computed are jointly represented. According to the results obtained, the most efficient division into areas corresponds to first experiment carried out using the KSOM algorithm to create a *Kohonen* map of 25 cells ('ksom25.1'). These cells have been grouped into the following three areas: area 1 consists of 18 nodes. One of them is located in Palencia, 1 in Salamanca, 3 in the west of Salamanca, 7 in Valladolid, 5 in Zamora and 1 in the centre east of Portugal. Area 2 consists of 417 nodes including the rest of the Spanish and Portuguese nodes and two French nodes close to the border with Spain. Area 3 includes the remaining French nodes.

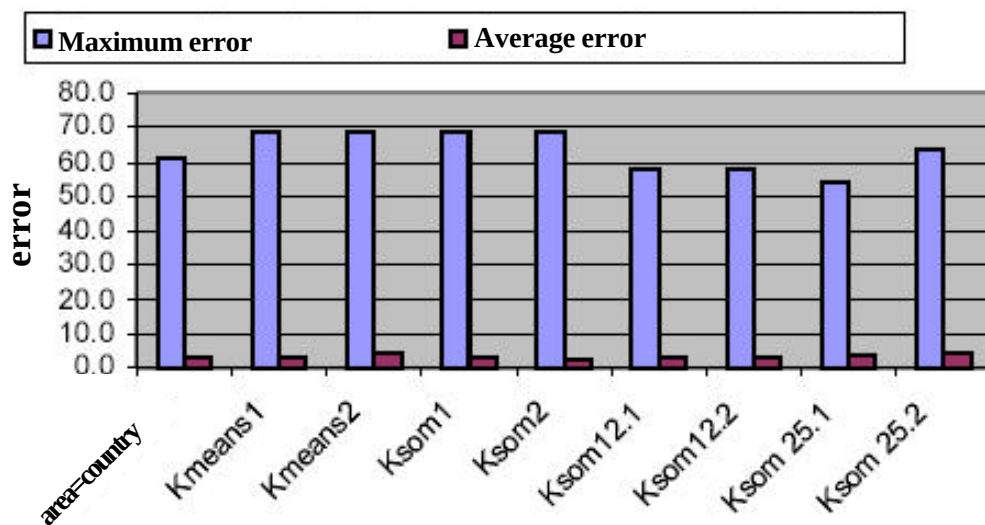


Figure 4.4-15: Comparison of the maximum and average transaction errors in each of the classifications obtained for the division into three areas of the system comprised of France, Spain and Portugal

If we discard those transactions that take place between nodes where no generation is located, the maximum transaction error for the chosen classification decreases from 53,7% to 47%. Figure 4.4-16 graphically shows the division into three areas adopted for the system.

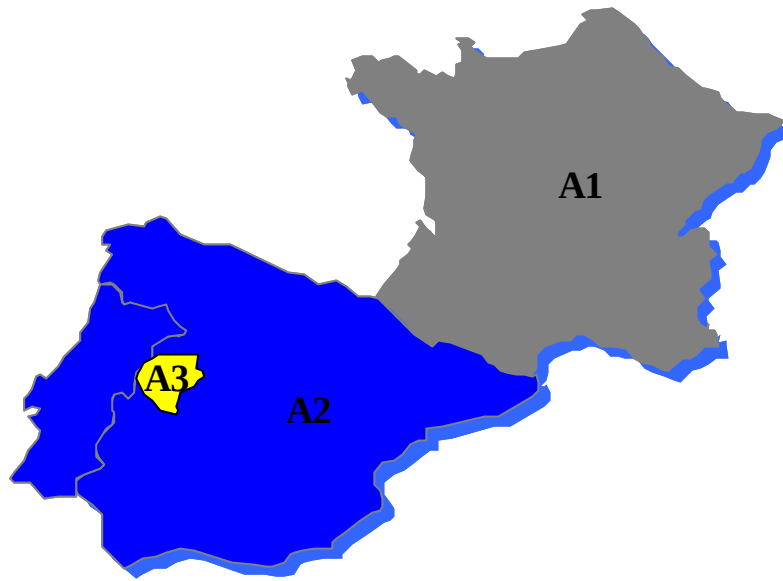


Figure 4.4-16: Division into three SPAs adopted for the system comprised of Spain, France and Portugal

Results for the division into 4 areas

We have computed several classifications of the nodes of the system into 4 areas. Both the *KSOM* and the *kmeans* algorithms have been applied. Figure 4.4-17 shows the maximum transaction error committed for each of the classifications that have been computed. This represents, for each classification, the maximum percentage of the power transacted between two points within the same area that flows through the congested lines. Labels used in the horizontal axis provide information on the algorithm applied, the number of cells in the map (only for *KSOM*) and the number of the experiment, in this order.

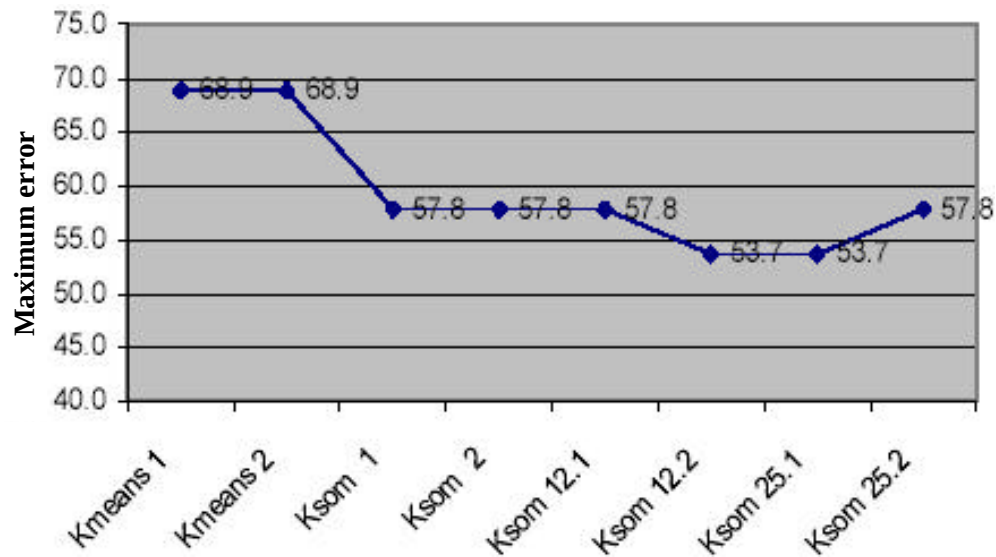


Figure 4.4-17: Maximum transaction error committed for the different classifications obtained corresponding to the division of the system comprised of France, Spain and Portugal into four areas

We now present in Figure 4.4-18 the average transaction error for each classification. The average transaction error is the mean percentage of the power transacted between two points within the same area that flows through the congested lines.

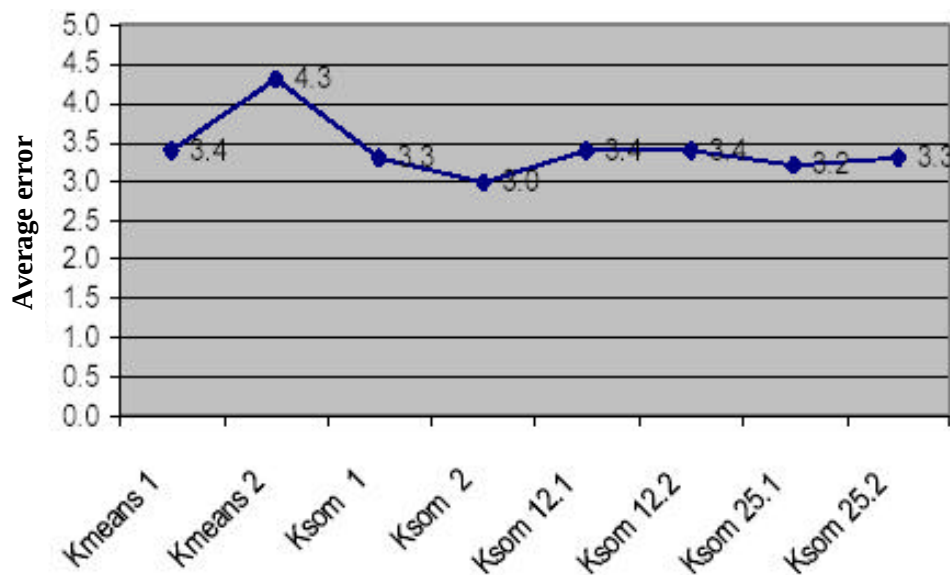


Figure 4.4-18: Average transaction errors committed for the different classifications computed corresponding to the division of the system comprised of France, Spain and Portugal into four areas

Figure 4.4-19 and Figure 4.4-20 represent the maximum and average transaction errors for each SPA within each classification.

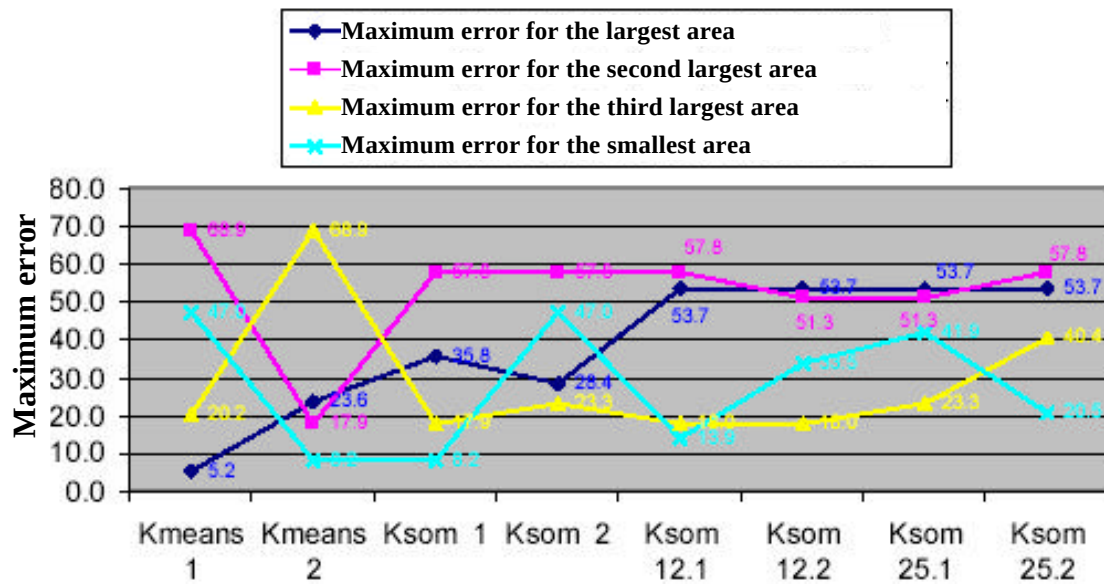


Figure 4.4-19: Maximum transaction error committed within each area and classification for the division into four areas of the system comprised of France, Spain and Portugal

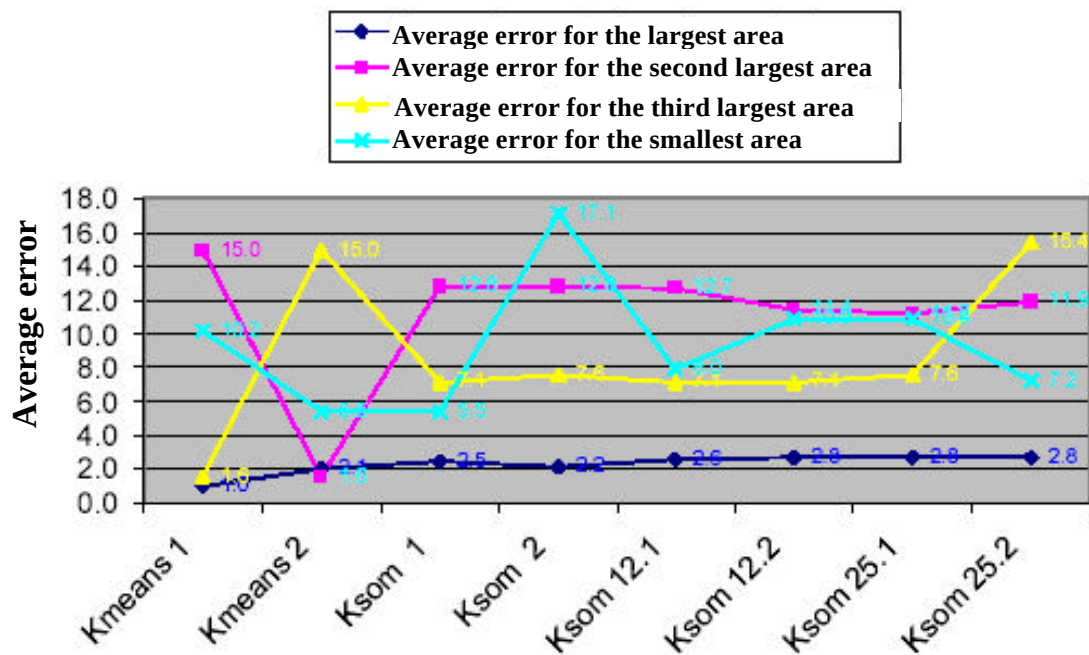


Figure 4.4-20: Average transaction error committed within each area and classification for the division into four areas of the system comprised of France, Spain and Portugal

As a summary, Figure 4.4-21 represents the maximum and average transaction errors for each classification.

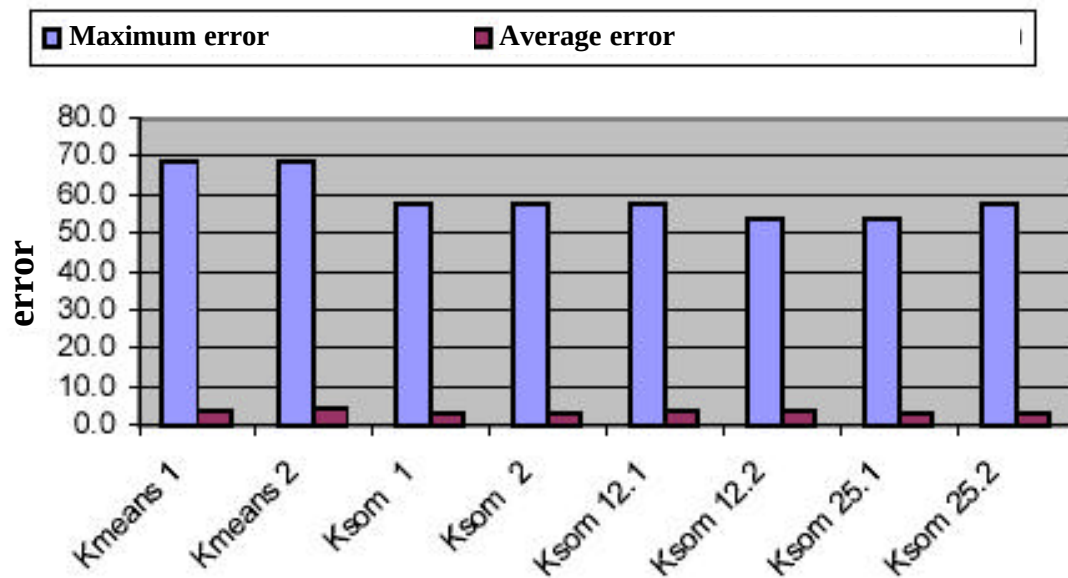


Figure 4.4-21: Maximum and average transaction errors for each of the classifications corresponding to the division into four areas of the system comprised of France, Spain and Portugal

The following classifications must be discarded because they include at least one SPA which is too small: 'Kmeans 1', 'Ksom 1', 'Ksom 2', 'Ksom 12.1' y 'Ksom 12.2'. Considering the remaining classifications, the best one is 'Kso 25.1' since it results in the best compromise between the maximum and the average transaction errors. The average transaction error committed in this classification is 3,2%. Once we discard those transactions taking place between nodes where no generation exists, the maximum transaction error committed is 47%. Figure 4.4-22 shows the distribution of SPAs on the map according to the classification that we have chosen.

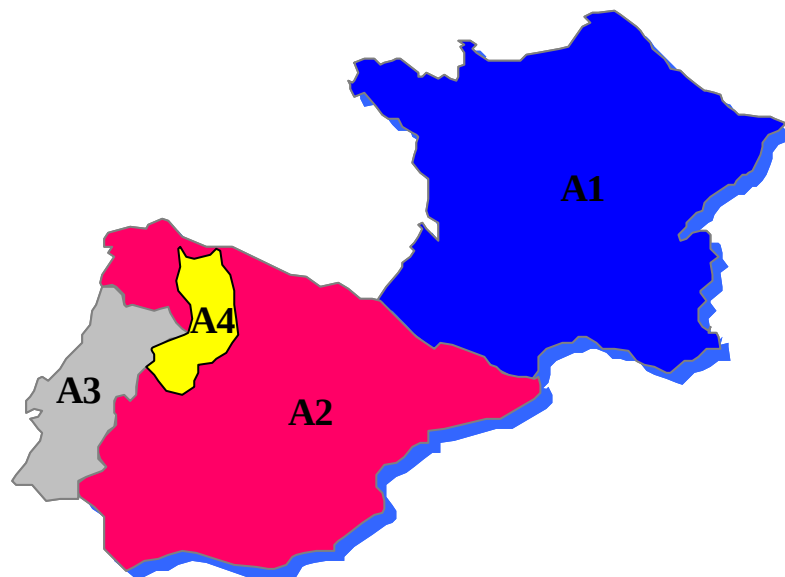


Figure 4.4-22: Division into four SPAs adopted for the system comprising Spain, France and Portugal

The number of nodes within each SPA as well as the geographical location of these nodes are provided next:

- Area 1 includes 306 nodes. All of them are in Spain.
- Area 2 consists of 57 nodes including 4 located in Asturias, 1 in Burgos, 4 in Cantabria, 1 in La Coruña, 10 in Leon, 2 in Lugo, 8 in Orense, 4 in Palencia, 1 in Salamanca, 2 in the west of Salamanca, 1 in Segovia, 7 in Valladolid and 8 in Zamora.
- Area 3 includes 1293 nodes. These are all the nodes located in France.
- Area 4 includes all the nodes located in Portugal.

Results for the division into 5 areas

Now we explore the possibility of defining 5 areas in the system. Both the *KSOM* and the *kmeans* algorithms have been used to classify the system nodes. The *KSOM* algorithm has been applied to produce maps of 5, 12 and 25 cells. Afterwards, we have grouped cells into 5 areas. Maps produced by the *kmeans* algorithm have 5 cells (as many as areas).

All the classifications studied but two include at least one SPA that is comprised of less than 5 nodes. Therefore, we have discarded them. Taking the remaining two, we have further studied the one where the maximum and average transaction errors committed are smaller. According to this classification, the following 5 areas must be defined:

- Area 1 consists of 18 nodes located in Palencia (1 node), Salamanca (1 node), west Salamanca (3 nodes), Valladolid (7 nodes), Zamora (5 nodes) and the Centre East of Portugal (1 node).
- Area 2 consists of 38 nodes located in Asturias (6 nodes), Burgos (1 node), Cantabria (4 nodes), La Coruña (1 node), Leon (10 nodes), Lugo (2 nodes), Orense (8 nodes), Palencia (3 nodes) and Zamora (3 nodes).
- Area 3 includes all the nodes in France.
- Area 4 consists of 70 nodes which include the remaining nodes in Portugal and 1 node in Caceres.
- Area 5 consists of 307 nodes. All of them are located in Spain.

Table 4.4-2 shows the maximum and average errors for the transactions taking place within each of those 5 SPAs.

	Area 1	Area 2	Area 3	Area 4	Area 5	Total
Maximum error	33.5	10.2	53.7	23.3	51.3	53.7
Average error	10.9	3.2	2.8	7.6	11.2	3.2

Table 4.4-2: Maximum and average error for each of the 5 areas identified in the system comprised of France, Spain and Portugal

Considering only those intra-area transactions where a generator may get involved, the maximum transaction error decreases from 53,7% to 47%. Figure 4.4-23 shows the geographical distribution of these 5 areas in the system.

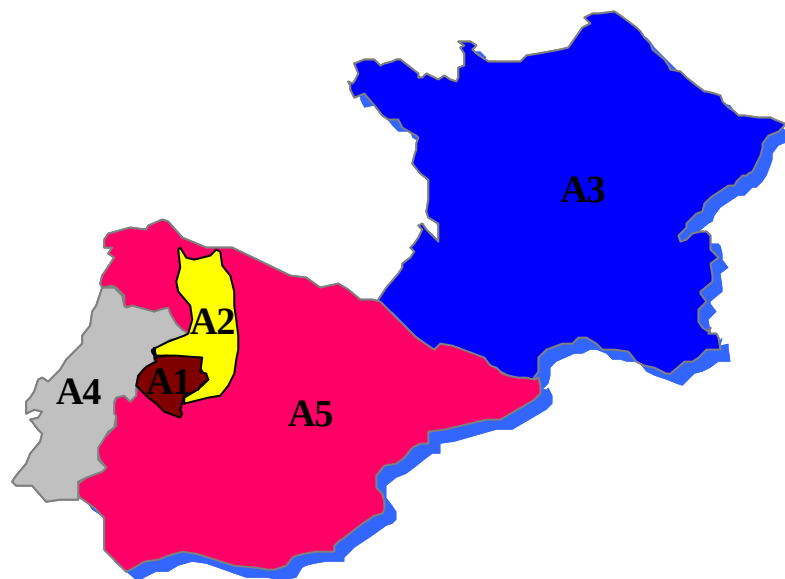


Figure 4.4-23: Geographical distribution of the 5 SPAs that have been defined in the system that is comprised of the French, Spanish and Portuguese ones

Final choice of a classification

At this stage of the process we are in the position to decide whether to divide the system into 3, 4 or 5 areas. We must compare the best classifications obtained in these three cases. Classifications must be assessed with respect to the criteria outlined in the introduction of this section. Figure 4.4-24 shows the maximum transaction errors committed when considering 3, 4 and 5 areas. The maximum intra-area transaction error committed decreases with the number of areas. However, this reduction is very modest. Figure 4.4-25 shows the average error corresponding to an intra-area transaction in each one of these classifications.

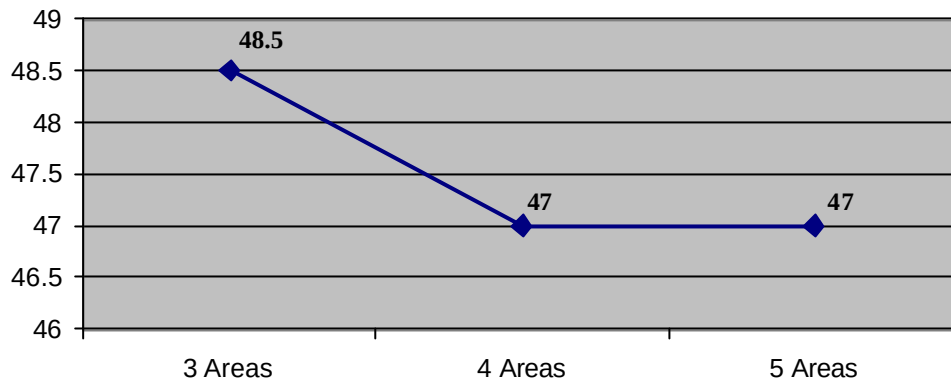


Figure 4.4-24: Maximum transaction errors committed when dividing the system comprised of France, Spain and Portugal into 3, 4 and 5 areas

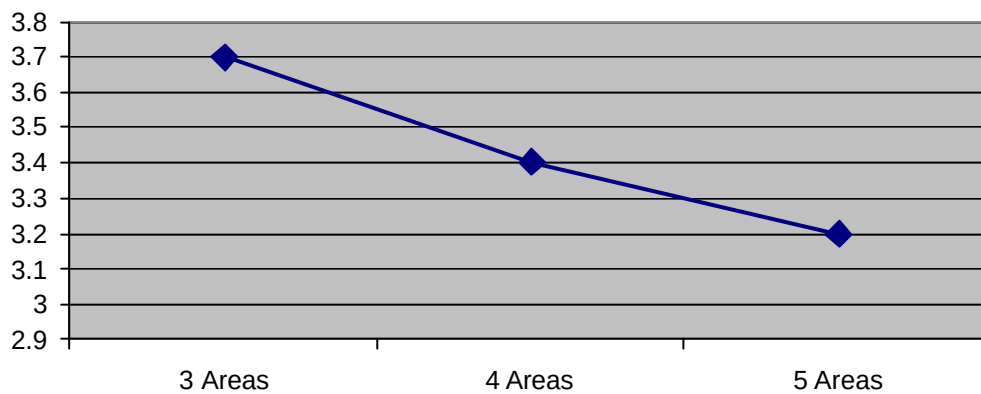


Figure 4.4-25: Average transaction errors committed when dividing the system comprised of France, Spain and Portugal into 3, 4 and 5 areas

According to the aforementioned evaluation criteria, the division into 4 areas probably is the one that ranks highest. Specifically: 1) the maximum transaction error for this classification is the lowest one, 2) its average transaction error is close to the lowest one and 3) it is easier to implement than the division into 5 areas. Finally we briefly describe this division into SPAs:

- Area 1 includes the majority of the Spanish system. The maximum error for an intra-area transaction within this area is 51,3%. The average error over all the transactions within the area is 11,3%.
- Area 2 consists of a group of nodes in the north west of Spain. The maximum error for an intra-area transaction within this area is 41,9%. The average error over all the transactions within the area is 10,8%.
- Area 3 consists of all the French nodes. The maximum error for an intra-area transaction within this area is 53,7%. The average error over all the transactions within the area is 2,8%.

- Area 4 consists of all the Portuguese nodes. The maximum error for an intra-area transaction within this area is 23,3%. The average error over all the transactions within the area is 7,6%.

Further work ahead. Identification and treatment of special transactions

Some useful conclusions may be drawn from the analysis carried out to identify SPAs both in the Iberian peninsula and in the system comprised of France, Spain and Portugal. The maximum transaction error committed is very large in all the considered classifications. Reducing the maximum committed error requires making some modifications to the process of identification of SPAs. Average committed transaction errors are much lower, i.e. only a small fraction of the power transacted between most pairs of nodes belonging to the same SPA flows through the congested lines.

As explained before in this section, transactions whose unit contribution to the flows on the congested lines is significantly different from the contributions of the rest of possible transactions in the system must be singled out. We have called them special transactions. Special transactions must be dealt with separately from the rest of transactions. We must compute specific PTDFs of the power flow on the congested lines with respect to the special transactions. Afterwards, these transactions must be taken into account specifically when managing congestion in the grid.

Any classification into SPAs must therefore be assessed in the light of new criteria. On the one hand, we must determine the number of transactions that deserve specific treatment. This number should be as small as possible. On the other hand, we must compute the error committed when considering a zonal system for the remaining transactions instead of a nodal one. Again, this error should be as small as possible.

We have defined the error committed corresponding to a certain transaction taking place between two SPAs (or within a single SPA) as the difference between the flows on the congested lines produced by this transaction and those produced by a typical transaction between the same two SPAs. According to the methodology laid out in section 4.4.3.1, we can identify special transactions as those for which the error committed is above a certain threshold. The threshold employed must strike a balance between the number of nodal (special) transactions we must consider and the total error committed corresponding to the remaining (typical) transactions. Different divisions into SPAs may result in different thresholds. However, the thresholds obtained for two different sets of SPAs should not differ much.

Once we set a threshold for a given classification, SPAS must be redefined since the fact that some transactions are no longer considered when defining the division into areas may affect the boundaries of these areas. The following figures show how the number of transactions requiring specific treatment and the average error for a classification change by changing the threshold. Results are shown both for the division into areas of the Iberian peninsula and for that of the system comprised of France, Spain and Portugal.

We now consider the division into SPAs that was finally chosen for the Iberian peninsula. Figure 4.4-26 shows the evolution of the number of special transactions in the Iberian peninsula with the level of the threshold. Figure 4.4-27 shows how the average transaction error that was committed changes when we change the level of the threshold. Threshold levels are expressed in % (the same units used to measure transaction errors). Average transaction errors in Figure 4.4-27 are expressed in %.

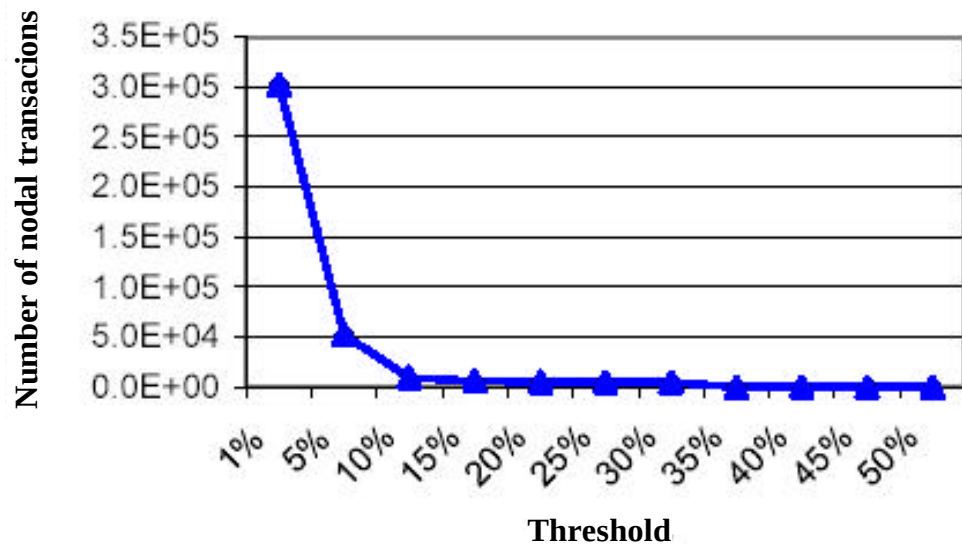


Figure 4.4-26: Number of nodal transactions in the Iberian system for different threshold levels

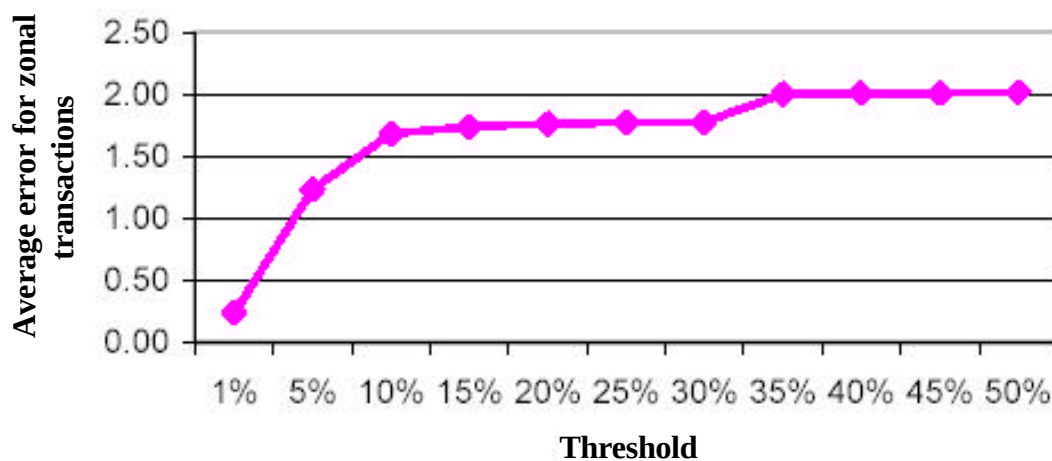


Figure 4.4-27: Average intra-area transaction error in the Iberian system for different threshold levels

According to this information, the threshold level used should be between 10% and 15%. As mentioned above, once we have chosen the threshold we should identify special transactions and recompute SPAs this time only taking into account typical (no special) transactions.

Now we focus on the system comprised of the French, Spanish and Portuguese ones. We consider the division into SPAs that was finally chosen. Figure 4.4-28 shows the evolution of the number of special transactions with the level of the threshold. Figure 4.4-29 shows how the average transaction error that was committed changes when we change the level of the threshold.

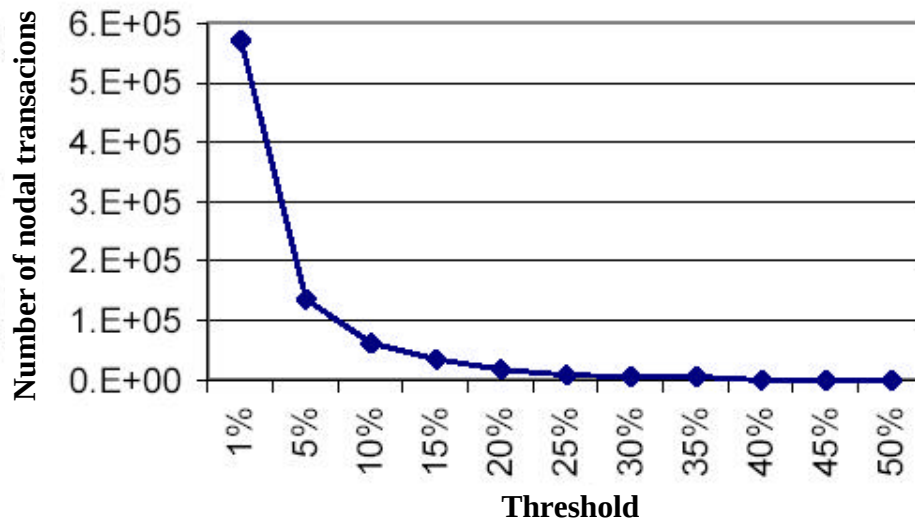


Figure 4.4-28: Number of nodal transactions in the system comprised of France, Spain and Portugal for different threshold levels

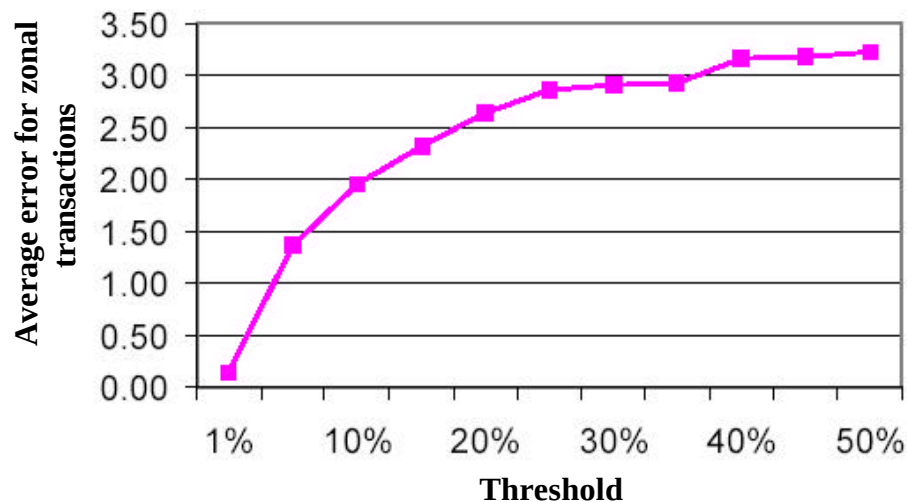


Figure 4.4-29: Average intra-area transaction error in the system comprised of France, Spain and Portugal for different threshold levels

According to this information, we may choose a threshold level of 15%. Again, now we should single out special transactions and redefine SPAs in the system.

In addition to dealing with special transactions individually, there are other ways of further improving the outcome of the classification process. We have only considered the grid topology corresponding to the base case scenario, where no contingency

occurs. However, active restrictions may correspond to an scenario where a line, a transformer or a generation plant has failed. When a contingency occurs, the topology of the system changes. Hence, flows over the lines also change. Lines that are not congested in the base case scenario may become congested in a contingency scenario.

The classification variable we have used is the set of Power Transfer Distribution Factors of the flows over the congested lines with respect to the elementary transaction between each node in the system and the reference node. PTDFs should be computed for the topology of the system corresponding to the scenario (or scenarios) where the corresponding line gets congested. Thus, we would obtain as many PTDFs per node as combinations of a congested line and a scenario where this line is congested. Nodes should be classified into areas according to the vector of PTDFs computed for each of them.

4.4.3.3 Conclusions of the work on the identification of SPAs

Defining Single Price Areas is central to the implementation of any zonal congestion management scheme. Coordinated congestion management methods must be employed to manage congestion on the borders between areas. SPAs have been successfully defined in some real life systems like the Nordel market. Many parties favor the implementation of a mechanism to implicitly manage the congestion caused by cross-border flows in Europe. Application of this mechanism will certainly be inefficient unless appropriate SPAs are defined (assuming they exist).

We have proposed a simple method to identify SPAs. In this method, nodes are classified into areas applying clustering techniques. The classification variable we have used is the set of Power Transfer Distribution Factors of the flows over the congested lines with respect to an elementary transaction between each node in the system and a common reference node. The only input to this method is the topology of the system grid.

Other methods may be superior to the one proposed here when we try to identify SPAs within a meshed network. However, these more refined methods require having access to economic information about the energy bids by market agents or about the variable costs of these agents. This information will probably not be made available to the institution in charge of defining SPAs.

Unfortunately, clear cut areas only exist in radial systems. In any case, there is still room for improvement in the application of the method proposed in this section. We may single out transactions that cannot be accurately characterized as an inter or an intra-area transaction. This transactions should be dealt with separately from the rest. Apart from this, contingency scenarios (in addition to the base case scenario) should be taken into account when computing the set of PTDFs used as the classification variable in the clustering algorithm.

4.4.4 Identification of a balancing point for electricity transmission contracts

Section 4.3.2.4 points out to the possibility of defining all transmission rights relative to a balancing point. This has been suggested as a measure to mitigate the exercise of market power by generators holding it, see (Gilbert, Neuhoff et al. 2004) and other technical works on the subject. In order for strategic generators not to find in transmission contracts an extra incentive to exercise their market power, all generators are restricted to buying transmission contracts that hedge sales of their output at the balancing point. Load buys the energy at the balancing point, and hedges the transmission risk from the balancing point to the location with a second transmission contract.

Of course, bilateral contracting among agents is perfectly compatible with the utilization of a balancing point, which only affects the definition of transmission rights to be acquired by generators and load. This approach allows hedging of all transmission price risk and provides the aforementioned simple rule to prevent agents from acquiring market power enhancing contracts. Under this rule, strategic generators are still allowed to acquire excessive amounts of transmission contracts to the balancing point. It would be difficult to monitor and enforce a rule that would require strategic generators to hold transmission contracts only corresponding to the energy they actually trade.

Some people in the academic community and industry have suggested the possibility of other entities than the TSOs organizing auctions for transmission rights. However, if transmission rights are defined as point-to-point contracts, rights bought and sold in an auction will not perfectly match each other in terms of the underlying transmission capacity these rights relate to. Thus, selling rights over transmission capacity that is not covered by previously issued contracts will certainly be necessary. Only the TSO of an area is in the position to issue rights over new transmission capacity if so required. Other entities would first have to buy this capacity from the TSO in order to put it up for sale. But then they would run the risk of not being able to sell this capacity. The perception of this risk would dissuade private entities from organizing capacity auctions. Hence, requiring TSOs to issue transmission contracts between any point in the region and the balancing point would be enough to make sure that all rights acquired by strategic generators do not encourage them to exercise their market power.

This section proposes a method to choose a balancing point so that market power of strategic generators is not enhanced even if they own more transmission contracts to the balancing point than required to hedge their trade positions. The work presented here was first described in (Olmos and Neuhoff 2006).

Detailed information about generation and demand structure and prices is rarely available. Therefore, the only input data required to identify the most suitable balancing point in the system is the price-responsiveness of local demand and

competitive generation as well as the network structure including the set of constrained lines. We furthermore assume nodal pricing.

The choice of the balancing point determines whether transmission contracts to the balancing point reduce or increase the incentive for generators to exercise market power. We now illustrate this effect and calculate equilibrium output choices in a simple *Cournot* model. Later in this section we derive a more general methodology to determine the optimal balancing point where we no longer calculate equilibrium output choices.

First, we consider the case where the generator selling its output does not own any transmission rights. In a simplified *Cournot* model representation, the generator maximizes the profits, which are a function of local price p_i , output q_i , and production costs $c(q_i)$:

$$\Pi = q_i p_i(q_i) - c_i(q_i) \quad (4.82)$$

The profit-maximizing output decision results from the first order optimality condition of the objective function in equation (4.82):

$$q_i = \frac{\frac{\partial c_i}{\partial q_i} - p_i}{\frac{\partial p_i}{\partial q_i}} \quad (4.83)$$

Now we assume that the same generator owns transmission contracts t between his location and a balancing node with price p_r . His profit function now is:

$$\Pi = q_i p_i(q_i) - c_i(q_i) + t(p_r(q_i) - p_i(q_i)) \quad (4.84)$$

The first order condition for equation (4.84) gives:

$$q_i = \frac{\frac{\partial c_i}{\partial q_i} - p_i}{\frac{\partial p_i}{\partial q_i}} + t \left(1 - \frac{\frac{\partial p_r}{\partial q_i}}{\frac{\partial p_i}{\partial q_i}} \right) \quad (4.85)$$

Comparing (4.83) and (4.85) shows that transmission contracts will increase the output q_i of the strategic generator and reduce its market power whenever $\frac{\partial p_r}{\partial q_i} > \frac{\partial p_i}{\partial q_i}$. Let us define the relative cross-price sensitivity of the price in the reference node r with respect to the output of generator i as $\frac{\partial p_r}{\partial q_i} / \frac{\partial p_i}{\partial q_i}$. The condition for transmission

contracts reducing market power is satisfied if this relative cross-price sensitivity is smaller than 1⁷.

Transmission contracts to the balancing node will not impact market power of a generator located at i if $\frac{\partial p_r}{\partial q_i} = \frac{\partial p_i}{\partial q_i}$. The third possibility is that $\frac{\partial p_r}{\partial q_i} < \frac{\partial p_i}{\partial q_i}$. Then, the price at the balancing point is more responsive to output changes of a generator located at i than the price at i . Such nodes are not suitable as balancing points.

From this analysis, an important conclusion must be drawn. The increase in the price of energy at the balancing point should be as small as possible when any agent in the system unilaterally decides to reduce its power output. We should bear in mind that our goal is to find a balancing point that reduces the exercise of market power by any generator in the system, not only by some of them. Ideally, the price at the balancing node would decrease when the output at any node in the system is reduced, but this is highly unlikely to happen. In reality, no matter which balancing node we choose, and as long as it is a physical node well-connected to the system, there will exist some other nodes in the vicinity so that the price of the balancing node increases with a decrease in the output at these neighboring nodes.

If we decide that the balancing node must be the physical node that has the smallest price increase with any output decrease in any of the remaining nodes in the system, we are likely to choose a node that is not extensively connected to the remaining system. However, the price differences between this node and any of the remaining nodes in the system would be very volatile. This implies that market participants would face significant price risk if they do not exactly match their trading positions with financial transmission contracts. In addition, the more the price at the balancing point differed from prices at other nodes, the higher the value of the corresponding transmission contracts would be and the more credit guarantees would be required from market participants.

Consequently, the balancing point chosen must comply with the following criteria:

1. It must not aggravate the problem of market power.
2. The energy price at the balancing point must be closely correlated with energy prices in the rest of the system.

As mentioned above, defining the balancing point as a physical node in the system is a possibility. However, it may be difficult to find a physical balancing node that has the aforementioned properties. Therefore, using a balancing node comprised of a set of physical nodes may be more appealing. The price at this reference node can be

⁷ Note that the own price sensitivity is always negative.

obtained by averaging the prices at all the nodes in this set. Thus, in PJM, trading hubs are defined to comprise many different physical delivery points distributed all over the grid. The price at this hub is the average of prices at all these delivery points, and the delivery points are chosen in such a way that the hub price is representative of larger areas and is stable. Assume a trading hub is defined as the average price of two nodes: X and Y . In order to hedge the transmission risk associated with sales at the hub, a generator buys financial transmission rights for half of his sales towards each node: X and Y . Likewise, the consumer buys financial contracts for half the volume of energy bought from each node: X and Y . This ensures full hedging of the transmission price risk for the amount of energy traded at the hub.

In any market design the volume of financial transmission contracts the TSO can issue is limited. According to revenue adequacy criterion, the network has to be capable of delivering the combined flow pattern implied by the transmission contracts. This ensures that the congestion revenue the TSO receives from nodal or zonal pricing suffices to reimburse the owners of financial transmission contracts.

If market participants want to change the flow pattern for which they hedge their transactions, then the TSO must reconfigure transmission contracts. Provided sufficient liquidity exists in the market, reconfiguration of contracts should not come at the expense of incurring some efficiency losses. Despite this, defining transmission contracts between any point or zone in the system and the transmission hub (reference or balancing node or zone) results in all generators and demands located in the same area buying and selling the same capacity product. This would allow agents in the same area to exchange transmission rights without having to resort to a centralized institution.

Now we present a method to compute cross-price sensitivities as defined above. Cross price sensitivities can then be used to define either a physical or a virtual hub since nodes comprising a hub should in principle have good properties. Here we shall investigate the possibility of using a physical node as the trading hub. Defining a trading hub comprised of those physical nodes that have the best properties according to criteria 1 and 2 above is always possible.

The methodology presented in section 4.4.4.1 is then implemented in section 4.4.4.2 to identify suitable physical balancing points in the European system. Finally, section 4.4.4.3 provides some final remarks.

4.4.4.1 Methodology for the computation of price sensitivities

Once we have determined the characteristics that the balancing point should have, we are in a position to decide which ones of the nodes of a large system such as the European one can qualify as balancing points. In contrast to the illustrative example in the introduction, we no longer calculate the equilibrium output choices. We did not obtain the demand, generation and contracting data that would be necessary to calculate the equilibrium. Even if these data were available to a regulator, commercial

confidentiality of data might limit publication and transparent replication of the calculation. Therefore, we develop an approach that relies only on network data and demand and generation volumes at each node. We computed relative cross-price sensitivities for this system and a particular set of operating conditions corresponding to available data of the real operation of the system on January 17, 2001 at 10.30 am. The term cross-price sensitivity refers to the change in price at a certain node when the output at another node, hereafter referred to as the generation node, decreases by one unit.

We start from the equations for market equilibrium under nodal pricing. We differentiate the corresponding equations with respect to the change of net demand (demand minus generation) at a network node. This gives us an expression for cross-price sensitivities. The benefit of this approach is that various input parameters that would be required to calculate the equilibrium output choice are no longer required. Without these input parameters, we cannot calculate the market equilibrium. However, assuming that the market equilibrium has been achieved, the equations derived from it are valid.

We made the following assumptions in our model. First, the set of active constraints does not change. There are several reasons to anticipate that this assumption does not create large distortions: first and most important, the remaining market agents will react to a decision by an agent A to unilaterally change its output. This reaction will probably counteract the increase or decrease of output by agent A . Second, marginal changes, as typically assessed for output optimization, are unlikely to affect whether constraints are binding. Third, in the large UCTE network, various constraints are simultaneously binding. Therefore, the change of one out of many binding constraints has limited impact. Lastly, strategic generators will not have access to complete information on network conditions, generation and output, and are therefore unlikely to understand exactly what constraint set is binding.

We furthermore assume that the net demand x_j at any node $j=1\dots N$ in the system reacts to the price of energy according to a linear model:

$$x_j = D_j - a_j p_j \quad (4.86)$$

where D_j is the intercept, a_j is the net demand slope and p_j is the price of energy at this node. We assume that net demand slope a_j is proportional to the power generation at node j . This is in accordance with the fact that most of the elasticity comes from the generation and not the demand side. We will show at the end of this proof that the proportionality factor assumed does not affect relative cross-price sensitivities. This independence facilitates the analysis, since the appropriate proportionality factor would be difficult to determine and crucially depends on the time scale we are assessing. We are only interested in marginal changes. Therefore, using the linear demand function in (4.86) does not restrict the results from being

equally valid for other functional forms. In fact, the actual output curves of generators are far from being linear. Most of them, though not all, either operate at their maximum capacity or do not produce any power. As we are assessing the impact of market power, we should also anticipate that generation companies submit bid schedules that provide less price responsiveness than the marginal cost curve of their generation assets would suggest. For example, a *Cournot* bid does not provide any price responsiveness of output. Such behavior could be represented in our model by adjusting demand slope with, for example, a regional measure of market power of generation companies like the HHI Index.

Further assumptions were necessary because the data set obtained from the TSOs did not provide any indication on which lines were actually congested. We therefore estimated the state of each line based on the ratio of the flow over the line to its capacity. We assume lines with a flow to capacity ratio above 0.7 are congested. Given the very limited amount of information about the system that is available, computing a different flow limit for each line based on the contingencies taken into account in the real dispatch and stability considerations is not possible. The coefficient 0.7 can be thought of as a fixed security margin used by TSOs when scheduling transactions (especially those between different countries or control areas) in the absence of sufficient data or the required level of coordination between areas. The absolute capacity of a line does not affect the results.

Finally, if several constraints are highly collinear among themselves, we must consider one of them and discard the rest. We assume that each one of these constraints represents the same limitation. Only after discarding highly collinear constraints were we able to invert the matrices described by the sets of linear equations (4.87), (4.88) and (4.89) in order to obtain cross-price sensitivities.

We now calculate cross-price sensitivities as resulting from traditional nodal pricing equations (Schweppe, Caramanis et al. 1988) corresponding to an optimal power flow (Wood and Wollenberg 1984)⁸. The global balance between load and generation in the system implies that the sum over all net-inflows x_j equals the transmission losses, which we assume to be zero in our analysis:

$$\sum_j x_j = 0 \quad (4.87)$$

The power flow distribution factor of the flow over line l with respect to a transaction between node j and the reference node r is here represented as g_{jl} . It gives the fraction of a transaction between these two nodes that passes along constraint l . The

⁸ An alternative market design of bilateral trading is described in [Chao, H.-P. and S. Peck (1996). "A Market Mechanism for Electric Power Transmission." Journal Of Regulatory Economics **10**: 25-29.

sum of the contributions of all the commercial transactions in the system to a binding constraint l is:

$$\sum_j x_j g_{jl} = K_l \quad \forall l \in C \quad (4.88)$$

where K_l represents the capacity of constraint l . The set of constraints C comprises all pre-contingency constraints on the line flows plus any kind of contingency constraints we may need to consider. Whenever a strategic generator reduces its output, competitive generators in the area will increase it and consumers will reduce their demand for energy due to an increase in the local price of energy. Hence demand and competitive generation will partly compensate for the decrease in output from the strategic generator. As we explained above, this is an argument in favor of assuming that marginal changes in the production of a generator will not change the set of binding constraints. This allows us to differentiate (4.88) with respect to the intercept of the net demand (and therefore also with respect to the changes of output initiated by strategic generators) to obtain:

$$\sum_j -\frac{\partial x_j}{\partial D_i} g_{jl} = 0, \quad \forall l \in C \quad (4.89)$$

Again, according to nodal pricing equations, the market-clearing price at node j equals the price at the reference node r plus the value of the marginal impact on all constraints of a transaction between nodes j and r . The last term is given by the sum, over all constraints, of the Lagrange multiplier associated with each constraint l (equation (4.88)) scaled with the marginal impact of a transaction between nodes j and r on constraint l ⁹:

$$p_j = p_r + \sum_{l \in C} g_{jl} r_l, \quad \forall j \quad (4.90)$$

Substituting (4.86) in equations (4.87), (4.89) and (4.90) and differentiating (4.87) and (4.90) with respect to the intercept of the net demand D_i at generation node i gives the following linear equation system for each generation node i that is considered:

$$\sum_j a_j \frac{\partial p_r}{\partial D_i} + \sum_{j, l \in C} a_j g_{jl} \frac{\partial r_l}{\partial D_i} = 1 \quad (4.91)$$

$$\sum_j a_j g_{jl} \frac{\partial p_r}{\partial D_i} + \sum_{j, m \in C} a_j g_{jl} g_{jm} \frac{\partial r_m}{\partial D_i} = g_{il} \quad \forall l \in C \quad (4.92)$$

⁹ r_l corresponds to flow gate prices or shadow price of transmission constraints under nodal pricing.

The equation system represented by (4.91) and (4.92) consists of $C+1$ linear equations with $C+1$ variables and thus can easily be solved by means of conventional algebraic matrix operations.

First, we compute $\frac{\partial r_l}{\partial D_i}$ and $\frac{\partial p_r}{\partial D_i} \forall l \in C$. This allows us to compute $\frac{\partial p_j}{\partial D_i} \forall j \in N$ using (4.90). The ratio $\frac{\partial p_j}{\partial D_i}$ is the cross-price sensitivities of the price at node j when the output at generation node i is marginally modified. Finally, relative cross-price sensitivities are obtained by dividing price sensitivities by that corresponding to generation node i , $\frac{\partial p_i}{\partial D_i}$. Assessing relative cross-price sensitivities has the advantage that they are independent of the assumptions about the proportionality factor that relates the net-demand responsiveness at each node with the power output at the same node.

The proof of the statement in the previous paragraph follows. Assume demand responsiveness is scaled by X . Then a_i is scaled by X and D_i changes but does not appear in (4.91) and (4.92). Equations (4.91) and (4.92) will continue to be satisfied if $\frac{\partial r_l}{\partial D_i}$ and $\frac{\partial p_r}{\partial D_i}$ are scaled by $1/X$. Hence any sum of both, namely $\frac{\partial p_j}{\partial D_i}$, will also be scaled by $1/X$, and the ratio of two cross-price sensitivities scaled by the same factor $1/X$ stays constant.

If we wish to implement transmission contracts defined towards a virtual hub, the sensitivity of the price at the hub with respect to the output at any node of the system could be computed by averaging the relative cross-price sensitivities of the nodes that the hub is comprised of.

Once we have developed a tool to compute relative cross-price sensitivities, we can ascertain whether the ownership of a transmission contract by an agent enjoying a certain amount of market power means an incentive for this agent to exercise it. All we must do is to determine whether the relative cross-price sensitivity of the balancing point relative to the node where the agent is located is greater or smaller than 1. In the first case, the ownership of this contract would encourage its owner to reduce output. In the latter, it would mitigate the exercise of market power by the agent. If the contract is defined between any two points of the system, we must consider the relative cross-price sensitivity of the destination node of the contract relative to the origin node.

Instead of defining transmission contracts with respect to the reference node, one might suggest allowing TSOs to issue transmission contracts between any two points in the region. Then, only those contracts likely to provide their owners with an incentive to exercise their market power would have to be monitored. Defining a balancing point that performs well under any circumstances may prove to be very difficult if congestion in the system changes much over time. However, in this case

almost every possible contract would be prone to exacerbate the exercise of market power by its owner in some cases. What is more, agents that are able to affect prices would also change from one scenario to another. Consequently, if using a balancing node to define contracts were not possible, every contract and agent would have to be subject to scrutiny.

4.4.4.2 Identification of a balancing point in the European system

This section identifies a set of suitable balancing points in the European system. In order to ascertain whether contracts defined with respect to a node encourage agents to exercise their market power, we apply the methodology presented in the previous section to compute cross-price sensitivities.

As discussed in the introduction of this section, the energy price at the balancing point must be highly correlated with energy prices in the rest of the system. Part B of this section illustrates how nodes that are well-connected with the remaining network can be identified. Additionally, rights defined towards the balancing node must not encourage agents to exercise market power. Part C of this section identifies points that exhibit relatively low cross-price sensitivities with respect to the energy price in the remaining nodes of the network. Both criteria are combined in a three-stage process. First, nodes are excluded from the set of possible balancing points if they are not well-connected. Second, nodes are excluded if they would significantly enhance market power for at least one generation node in the network. In a third step, the remaining nodes are sorted by how much they reduce market power at most of the generation nodes (95%). The locations of the best-suited nodes are presented. Various other methods may be imagined to combine the two criteria. Hence, investigating the robustness of the result for different combinations of the two selection criteria may be worthwhile.

A) The characteristics of the network

We study the synchronized European electricity network corresponding to the real system operation in 17 electrically connected UCTE countries on January 17, 2001 at 10.30 am. The UCTE model considered has 3,655 lines and 3,383 nodes, of which 708 are generation nodes. The total production is 244.00 GW, the total load is 240.86 GW and losses are 3.14 GW (Pérez-Arriaga, Olmos et al. 2002). Figure 4.4-30 shows a map of the 17 countries comprising the system considered in the study.

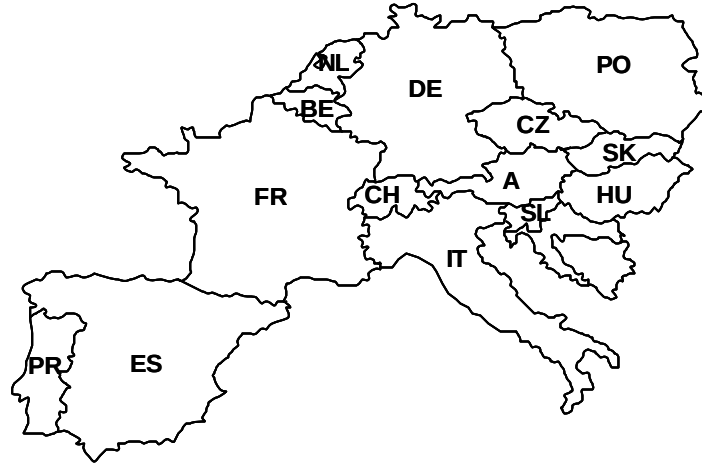


Figure 4.4-30: Geographical location of 17 UCTE countries considered in the study

Obtaining conclusive results regarding the physical location of the nodes that qualify as suitable balancing points requires taking into account a wide range of scenarios covering all possible system operating conditions. Thus, we would have to consider all different possible congestion patterns in the grid and compute cross-price sensitivities in all these cases. The maximum change in the energy price at the balancing node caused by a unilateral change in the power output from any agent should be small for any of the considered scenarios. Additionally, the selected balancing node should in any case be well-connected to the rest of the system. If the pattern of congestion in the grid were very variable, most nodes that would qualify as suitable balancing points in a particular scenario would not have good properties in other scenarios. Given that only one scenario was available for the study, we cannot make any indisputable statement about the identity of suitable balancing points. We cannot even be sure that a node exists that has good properties under all possible operating conditions.

Equations (4.91) and (4.92) provide the reaction of energy prices throughout the system to a change in the output at generation node i . Due to the fact that we carry out a marginal analysis, computing relative cross-price sensitivities does not require making any assumptions about the equilibrium price levels.

Network constraints affect the ability of generators to increase or decrease the energy price at the different nodes. This effect is illustrated for the case of a generation node located in central Germany (*KKPhili*). Figure 4.4-31 shows the relative cross-price sensitivity at various nodes of the European network as a result of an output change at *KKPhili*. Apparently, changing the output of a generator in *KKPhili* would significantly change the price in some nodes whereas the price in others would be barely affected. Due to the meshed nature of the grid, prices would change all over Europe, not only in the vicinity of the generation node. Thus, if we apply an efficient congestion management scheme (one that approximates the outcome of nodal pricing) we may find out that prices and hence operational decisions by agents in some areas are not independent of the operational decisions of those agents located in others. This may indicate that defining accurate SPAs in Europe, at least in the continental plate, may

not be possible. However, we should bear in mind that the pattern of flows that we have considered in our computations may probably not be optimal since, presently, the operation of the system is not coordinated at European level. A different pattern of flows may result in a different pattern of congestions from the one taken into account here. Hence, it is difficult to draw some definite conclusions about the existence and identity of SPAs in Europe from this piece of work. Carrying out an independent study focused on the identifications of SPAs in Europe remains to be necessary. Also related to this, the assessment of the sensitivity of relative cross-price sensitivities to the set of binding transmission constraints is a matter for future research as well.

In an unconstrained world, prices would only differ due to transmission losses. Hence, they would be highly correlated and relative cross-price sensitivities would be around 1. Changes in the set of binding constraints will likely have a significant effect on the results obtained for cross-price sensitivities and consequently on the selection of the balancing node. Using the reference node that has the best properties averaged over the different congestion patterns may be an option.

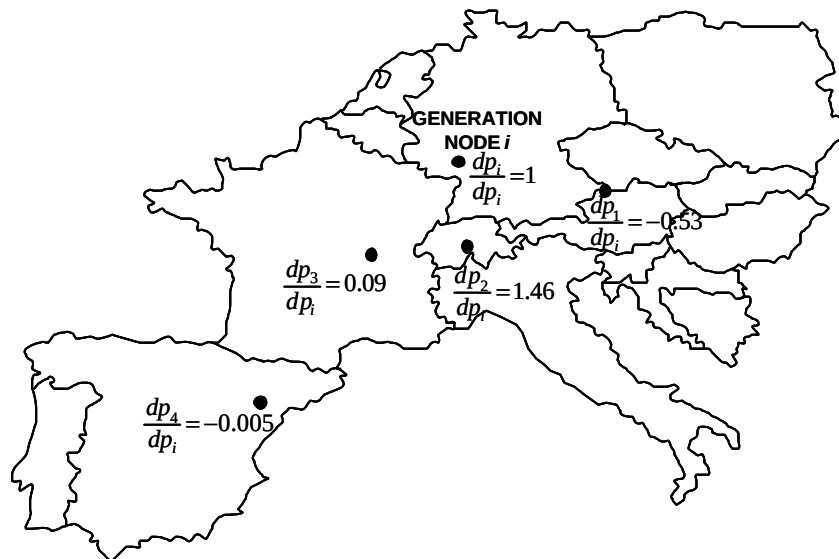


Figure 4.4-31: Impact on energy prices of a marginal decrease in the power output of an agent located in Germany (node i)

Going back to the example in Figure 4.4-31, the price of energy at the generation node increases as a result of a decrease in the power output at this node as it may be expected. The absolute increase depends on the assumptions made about the elasticity of demand. In any case, it is irrelevant in this analysis. However, other results deserve further clarification. Some of the remaining nodes in the network turned out to have negative relative cross-price sensitivities. Thus, when the power output at the German node *KKPhili* was reduced, the price at an Austrian node decreased instead of increasing. This effect might seem counterintuitive, but it has been anticipated in the literature using analysis of simple networks.

Assume that in the three-node network of Figure 4.4-32 the transmission line between nodes *A* and *B* is congested towards node *A*. In this case, decreasing output at node *A* will increase price at this node. Hence, the scarcity value of the transmission line will increase as well. Exporters from nodes *B* and *C* compete to use the transmission line. If the scarcity value of the transmission line increases, exporters from *C* face a competitive advantage because their exports to *A* require a smaller fraction of the transmission line per unit of energy exported than exports from *B*. Therefore, exports from *B* are reduced and the price drops at *B*. All in all, node *B* has a negative cross-price sensitivity to output changes at node *A*.

We have studied in detail the case of one of the many nodes with negative cross-price sensitivities. Some constrained lines in our real scale system play the same role as that of link *k* in the example of Figure 4.4-32. Claiming that this example explains why we obtained negative cross-price sensitivities for some of the nodes in our system would be absurd. However, this simple example demonstrates that negative relative cross-price sensitivities are perfectly possible.

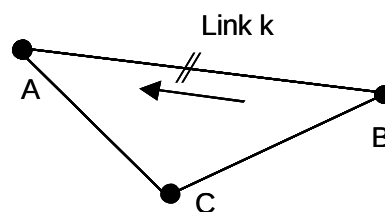


Figure 4.4-32: Example three-node network creating negative cross-price sensitivities

B) Selection of a suitable set of balancing nodes

Pre-selection of balancing points according to their level of connection to the rest of the system

This subsection discusses the process followed to select a set of suitable balancing points for the European system. This process is comprised of two main steps. First, we identify the subset of nodes that are well connected to the rest of the system. The rest of the nodes are discarded. Then, candidates to become the balancing point are selected as those, within the set of well connected nodes, that mitigate market power at all the generation nodes to the largest extent.

As discussed in the introduction of this section, having low relative cross-price sensitivities with respect to the generation nodes in the system is only one of the desirable features in a balancing point. In addition, the balancing point should be well-connected to the system as a whole. This ensures that the energy price at the balancing point is representative of the general evolution of energy prices in the system. Otherwise, price changes of transmission contracts would represent not only locational differences between energy prices but also an overall evolution of the price level. This would imply that transmission contracts could take both large positive and large negative values. Therefore, the counter-party risk that the holder of the contract would

default if the contract took a large negative value would increase significantly. Then, larger and more robust credit guarantees would be required, which increases costs and is likely to exclude some agents from acquiring transmission contracts.

We therefore need a criterion to judge how well-connected a node is to the network. Different measures of the level of connection of a node to the rest of the system may be used. The own price sensitivity of a node represents the change in price at the node per unit decrease in the output at the same node. The larger a node and the better-connected it is to the network, the lower the absolute value of its own price sensitivity will be (remember that the own price sensitivity of a node will always be negative). If no constraints were binding, every node would have the same own price sensitivity, since all the system would respond to a change in the output at any node. This common price sensitivity in the absence of constraints is a suitable reference value to decide whether a node is well-connected to the system. Only those nodes whose own price elasticity has an absolute value that is less than 10 times higher than the unconstrained system's own price elasticity have been considered well-connected nodes. Therefore, only these nodes are considered as potential balancing points in the analysis.

Selection of the final set of candidates to become balancing nodes based on their ability to mitigate market power at all generation nodes

For illustrative purposes, Figure 4.4-33 shows the relative cross-price sensitivities for each node of the UCTE system when power output at the node KKPhili in central Germany is changed. Nodes have been sorted by their relative cross-price sensitivity.

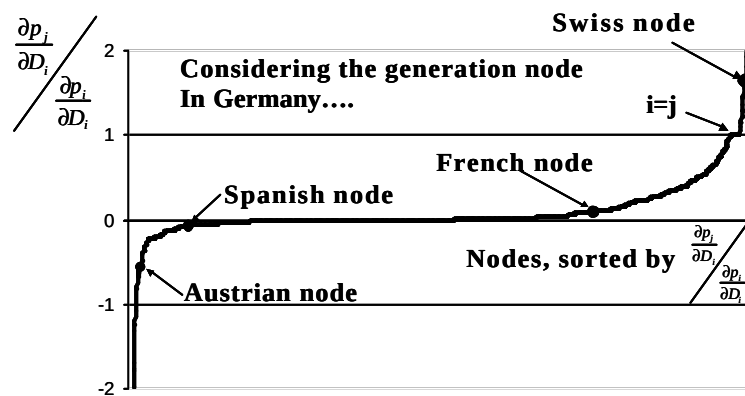


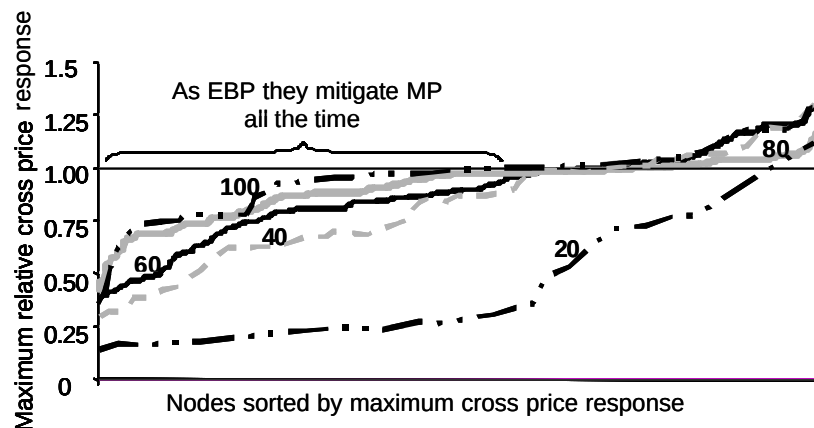
Figure 4.4-33: Relative cross-price sensitivities sorted by magnitude, with the generation node located in KKPhili (Germany)

As discussed in the introduction of this section, agents with a significant share in generation should be restricted to owning contracts that hedge their transmission risk. No generator should be allowed to buy contracts corresponding to an import to its node. We now establish the additional requirement that the relative cross-price sensitivity of the balancing point with respect to the power injection at the generation node be below 1. This ensures that a generator, by reducing output, will reduce the

value of his transmission contract. This, in turn, ensures that ownership of transmission contracts will reduce the exercise of market power at the generation node.

The balancing point must mitigate the exercise of market power at every node with generation assets. It does not suffice, therefore, to choose a balancing node that has a low cross-price sensitivity relative to one German node. Rather, the balancing point must exhibit low cross-price sensitivities relative to all generation nodes in the network. Then, we must test, for each potential balancing node, the relative cross-price sensitivities it experiences with respect to output changes at the remaining nodes of the network.

Figure 4.4-34 shows several curves representing the maximum relative cross-price sensitivity potential balancing nodes in the European system exhibit when we consider output changes at 1, 20, 40, 60, 80 and 100 generation nodes, respectively. Potential balancing nodes are listed on the horizontal axis sorted by their maximum price sensitivity when 100 nodes are considered.



(Lines are labelled according to the number of generation nodes considered)

Figure 4.4-34: Effect of the number of generation nodes considered on the shape of the curve representing the maximum relative cross-price sensitivity for each node in the system

All the nodes with a maximum cross-price sensitivity below 1 mitigate the exercise of market power at any of the generation nodes considered. As the number of generation nodes under consideration increases from 1 to 100, the number of potential balancing nodes whose maximum relative cross-price sensitivity is at least 1 increases. The more generation nodes we consider, the more probable it is that at least one of these generation nodes be very close to the balancing node and, therefore, has the same price. If the price at the balancing node is the same as that at a generation node A, increasing or decreasing the power output at this generation node will produce identical changes in the price at both nodes. Hence, the relative cross-price sensitivity computed will be 1.

We aim to leave out of the analysis those nodes that belong to the same price area as our balancing node (those that are in the immediate vicinity of the node) but still

consider relative cross-price sensitivities with respect to the remaining nodes in the system (most of the existing nodes). Generation nodes within the same price area as suitable balancing nodes exhibit the highest cross-price sensitivities in the distributions obtained for these balancing nodes. Therefore, we must get rid of the highest cross-price sensitivities in each distribution. Thus, instead of using the maximum relative cross-price sensitivity, we may consider that the 95 percentile of the distribution that is obtained for a certain potential balancing node is representative of this distribution¹⁰. Each one of the curves in Figure 4.4-35 depicts, for a different set of generation nodes whose size ranges between 1 and 100 nodes, the 95 percentile of the distribution of relative cross-price sensitivities of each potential balancing point in the system with respect to the considered set of generation nodes. Figure 4.4-35 illustrates that increasing the number of generation nodes considered beyond 20 does not result neither in an increase nor in a decrease of the 95 percentile of the distributions of relative cross-price sensitivities obtained for potential balancing points. This convergence suggests that using a representative set of nodes for the analysis is enough. Therefore, we would not need to compute relative cross-price sensitivities with respect to the output at every generation node. However, in order to achieve enough accuracy, results shown below have been computed using all 708 generation nodes in the snapshot.

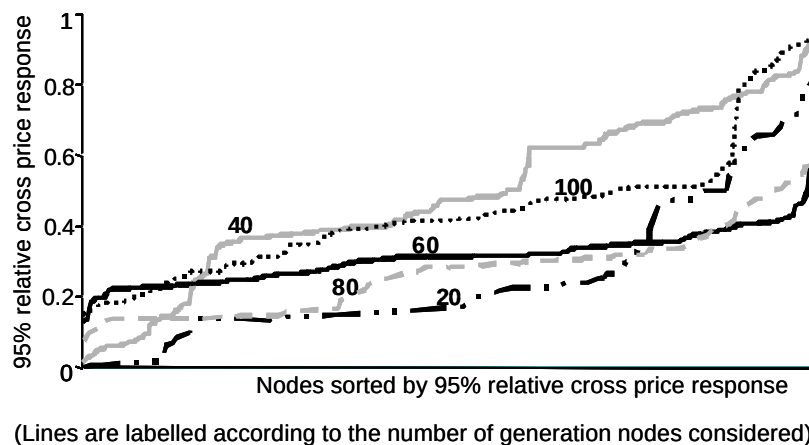


Figure 4.4-35: Curves representing the 95 percentile of the probability distribution of relative cross-price sensitivities for each node in the system. Each curve has been obtained for a different number of generation nodes

However, by purely relying on the 95 percentile of the distribution of cross-price sensitivities obtained for each candidate, we run the risk of selecting a balancing point that enhances market power at certain locations. Therefore, only nodes whose maximum relative cross-price sensitivity is below 1.1 have been considered as potential balancing points. Nodes within this group have been ranked according to the 95

¹⁰ The 95-percentile in a distribution is the value such that only 1/20 of the values in the distribution are higher.

percentile of their distribution. The best 20 nodes were chosen as our final candidates to become BPs. Figure 4.4-36 illustrates the combined selection process.

The dark line in Figure 4.4-36 represents the 95 percentile of the probability distribution of price-sensitivities for each candidate. This line has been divided into two separate segments. As explained in the graphic, the segment on the left corresponds to those nodes whose maximum relative cross-price sensitivity is below 1.1 (and therefore those which can be thought to mitigate market power in any case). The curve on the right side corresponds to the rest of the nodes. The light grey line represents the maximum value of the distribution. The most suitable balancing points in the figure would be those that are closest to the vertical axis since: they are well connected to the rest of the system (like any node represented in Figure 4.4-36), they exhibit the lowest maximum cross-price sensitivity with respect to the output at most nodes in the system (those that are supposed to belong to different price areas) and they have an absolute maximum relative cross-price sensitivity which is below 1.1 (and therefore they do not encourage the exercise of market power at any generation node in the grid). These are the nodes labeled 'Relevant' in Figure 4.4-36.

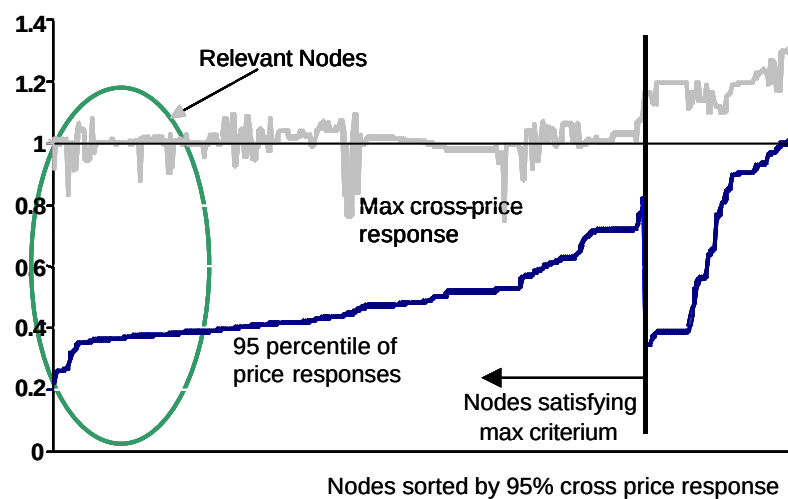


Figure 4.4-36: Election of the balancing point for the European system using both the maximum value of the distribution and the 95 percentile

Lastly, Figure 4.4-37 illustrates the geographical distribution of the twenty nodes that are well connected, exhibit maximum relative cross-price sensitivities below 1.1 and show in this group the lowest 95 percentile of relative cross-price sensitivities. Based on this analysis, the best candidate is the node labeled 'Langerak 132' in the Netherlands.

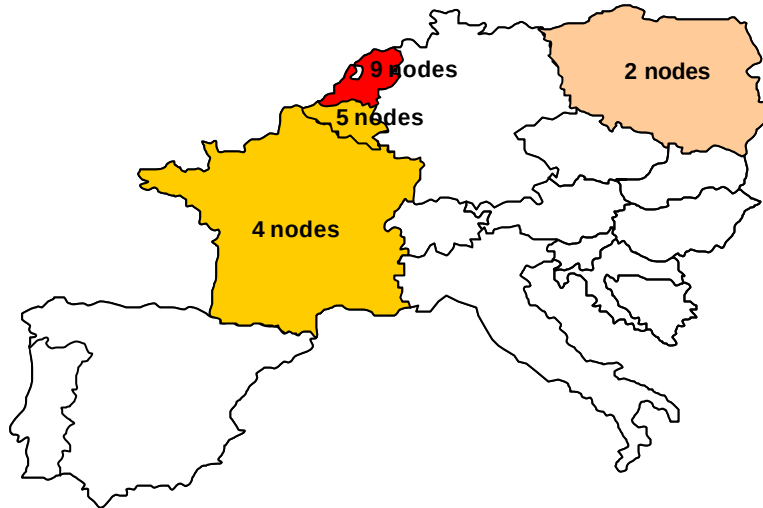


Figure 4.4-37: Geographical distribution of the 20 best candidates to become the European balancing point

C) Discussion of results

The fact that most of the suitable balancing points identified are located within Belgium or the Netherlands may result a bit surprising given that transmission constraints to these two countries are well documented. Apparently, these nodes do not comply with the criterion that balancing points should be well-connected to the remainder of the system. This seems to imply that criteria used to determine whether a node is well connected to the rest of the system are not rigorous enough. In fact, if the threshold used to decide whether a node is well-connected to the rest of the system were to be lowered to 5 times the reference value (instead of 10 times, as before) the 20 best performing nodes would be located in France.

It looks like the criteria used to locate congestion in the grid are unable to capture some of the constraints that TSOs take into account when operating their respective transmission systems. Both (n-1) type security criteria and the lack of efficient coordination in the management of congestion cause the importing capacity into the Netherlands to be reduced from more than 11,000 MW to less than 4,000 MW^{11,12}, see (IAEW 2002). Some of the binding constraints within Germany and some on the border between Germany and France have been considered. However, others on the borders between France and Belgium as well as between Germany and the Netherlands have not.

¹¹ This would provide an additional argument for integrating different markets to ensure the network is used more efficiently while retaining the same level of system security.

¹² The day-ahead prices for 10-11am on January 17, 2001, were 38.93 Euro/MWh at the LPX/EEEX for Germany and 50 Euro/MWh at the APX in the Netherlands. This price difference indicates that available transmission capacity into the Netherlands should have been used.

More than one scenario is probably necessary to identify the relevant network constraints. Changing regulations in most EU countries and the appearance of new trade opportunities between countries will probably significantly alter flow patterns in the European system. If competing authorities decided to define transmission contracts towards a balancing point, they should choose a balancing point considering not only the currently prevailing flow patterns but also the possible future ones. For example, the Dutch system operator, TenneT, states that expected market evolutions may transform the Netherlands into a heavily transited country sometime in the future (IAEW 2002). Therefore, properly selecting a balancing point would require repeating the analysis using better information. Our method is well-suited for this analysis since it requires minimal input data. Therefore, this method is appropriate to assess the impact of future flow patterns on the optimal choice of the balancing point.

In any case, results presented here show how closely prices are interlinked in the European network. Additionally, they give an empirically-based perspective on network interactions between regional electricity markets.

4.4.4.3 Some final remarks on the definition of a balancing point

Both market power considerations and the possibility of simplifying to some extent the secondary trading of transmission contracts make it advisable to define transmission contracts between any point in the system and a common balancing point. In order to make sure that transmission contracts mitigate market power, the energy price at the balancing point should never increase by more than the price at the node where a generator decreasing its output is located. Relative cross-price sensitivities have been computed using nodal pricing equations. Based on these sensitivities, we have chosen a preliminary set of candidates nodes to become the European Balancing Point. Although preliminary, the results obtained prove that the application of the methodology presented here to a real large-scale system is feasible. Results also show that, contrary to what might be expected, output decisions by generators affect prices all around Europe and not only in the area surrounding the agent. This fact can be attributed to the meshed nature of the European grid.

However, obtaining conclusive results requires further analysis. The effect that the number and identity of binding constraints have on the results remains unexplored. The assumption that the net demand is proportional to the amount of generation available at each node should be reconsidered. More data on real demand levels would improve accuracy. Exploring the case where agents own generation at different nodes would also be worthwhile. Agents may change their strategic decisions in the latter case. Finally, we have only explored the possibility of using one unique balancing point, which corresponds to a physical node, for the whole system. Using a virtual balancing node whose energy price is computed as the average of prices at many different nodes scattered all over the system is another possibility. Furthermore, different balancing points can be defined, depending on where each agent is trading its energy. A local balancing point for each national market may complement the European-wide balancing point for those agents buying or selling energy locally.

4.5 Conclusions

This chapter presents and discusses the problem of regulating access to the grid. Connection of agents to the transmission grid and the management of grid constraints are the two main aspects of the problem. Here we have focused on the second one, which has not been satisfactorily resolved in most systems, specially at the regional level.

Industry and academia have proposed diverse methods for the management of congestion in the grid. There seems to be wide agreement that the integrated dispatch of energy and capacity, where a different energy price is derived for each node, is the preferred approach from an efficiency point of view, at least when market power considerations are not taken into account. In any case, integrating the allocation of energy and capacity seems superior to dealing with both separately. However, this scheme requires a higher level of centralization than others, which may be a significant obstacle in the way towards its implementation, due to opposition from local authorities and institutions.

As far as it is compatible with the arrangements made in each region, we recommend implementing a regional implicit auction to allocate energy and capacity in the short term. If this is not possible because of the high level of centralization that the use of implicit auctions involves, we recommend implementing a coordinated explicit capacity auction at regional level followed by local implicit auctions, which must be security constrained if some grid restrictions within an area are binding. If grid constraints exist within an area but local authorities refuse to use implicit auctions, these auctions should be replaced by other market based mechanisms like explicit auctions where both offers to a pool and bilateral transactions should participate.

Apart from the design of the short term energy and capacity markets, other problems remain unresolved, such as the design of long and medium term capacity auctions, the format of the products that are exchanged in these auctions, the implication of the allocation of congestion rights on the exercise of market power or the allocation of congestion rents resulting from the dispatch. This thesis has attempted to contribute, at least partially, to the clarification of these issues.

Point-to-point rights confer agents an unambiguous right to execute the corresponding transaction or earn the corresponding congestion rents. Hence, they are clearly preferable for agents in order to shield against the risk of volatility in the transmission price. However, flow-gate rights defined for each corridor likely to be congested in a region allow decentralized trading.

We recommend allowing the joint trade of point-to-point and flow-gate rights in long and medium term capacity auctions.

Generally, financial rights are superior to physical rights, since they do not affect the short-term dispatch of energy in the absence of Market Power. Physical rights may affect the dispatch of energy if the capacity market is not liquid enough. In any case,

physical rights may be used for purposes other than risk hedging. Thus, they may provide a generator with the right to earn the capacity payment that is in place in the area where the withdrawal point of the right is located, or they may provide its owner with access to the green-certificates market. In this case, implementing them makes sense.

As a general rule, we recommend using financial rights in the long term. However, if energy and capacity are dealt with separately in the short-term, transmission rights owned by market agents just before the short-term energy dispatch must be physical. In addition, physical rights may exist in the long-term for agents to secure access to the physical transmission capacity if this is required.

There is room for trading both transmission capacity options and obligations. Obligations are preferable to options in markets that are not mature yet, since trading in obligations is easier to implement. Options may make sense in the long-term, when agents are uncertain about the commercial transactions that will engage in. Obligations are better suited for the short-term since they allow netting of flows when transmission rights are traded.

Both options and obligations should exist. Options should be issued and traded in the long term. They should be transformed into obligations as the operation of the system draws closer.

Transmission rights may represent an incentive for strategic generators to exercise their market power. Rules aimed at encouraging agents not to acquire market power enhancing rights may not be efficient in some cases. Thus, transmission capacity products available to agents should be carefully designed in order not to aggravate this problem. If rights are defined with respect to an appropriate physical or virtual balancing node, they may help hamper the exercise of Market Power by agents.

This subject is still in need of further research. Provisional results would suggest restricting the access of market agents to transmission rights. Agents might only be allowed to buy transmission rights that are defined towards a balancing point whose energy price is independent from the commercial decisions by these agents. However, finding a physical node with this characteristics may be difficult if not impossible in a large meshed system like the European one. Therefore, transmission rights might be defined with respect to virtual hubs with good properties.

As for the allocation of congestion rents, redistribution of flows caused by an increase in the capacity of a line should be taken into account when computing the marginal value of transmission capacity.

The true marginal value of transmission capacity must be used to compute the revenues of transmission owners. The transmission grid planner must use the true value of a line in the dispatch to decide about its construction. The value of a line should also guide the investment decisions by private entrepreneurs.

Zonal congestion methods may be superior to nodal ones in that the former may face less obstacles for their implementation due to the fact that they provide some degree of independence to the different areas in a region. However, a central problem to the implementation of any zonal congestion management method is the definition of areas of uniform prices or Single Price Areas (SPAs). Failing to define SPAs would mean that only nodal pricing techniques are compatible with the efficient operation of regional power systems. However, nodal pricing techniques are far from what seems achievable in most regional systems in the short and medium term.

We have conducted a preliminary study for the identification of SPAs in the Iberian peninsula and the system France-Spain-Portugal. A novel method is proposed to define these areas. This method is based on the application of clustering techniques combined with some expert knowledge about the functioning of the system. The definition of accurate enough SPAs in these regions requires further research. Some future lines of research have been outlined.

Zonal congestion management methods can only be implemented if accurate enough SPAs exist. If accurate enough SPAs cannot be identified, those inter and intra-area transactions whose impact on the flows of congested lines is significantly different from that of typical inter or intra-area transactions must be singled out and dealt with separately. When dividing a system into SPAs, the SO must choose the division so that the loss of efficiency in the dispatch with respect to the nodal optimal one is kept to a minimum.

4.5.1 Interaction between the regulation of the access to the grid by market agents and other aspects of transmission regulation

The management of congestion is not isolated from other aspects of transmission regulation. Mutual influences exist among the investment carried out in transmission capacity, the management of congestion and the tariff setting process that has been implemented. Here, we will comment on the relation that exists between congestion management and the two remaining topics.

Access rents due to transmission capacity scarcity are a clear indicator of the need to invest in new lines. In fact, assuming the relationship between the capacity of a line and its admittance is linear, the value of the capacity of the line equals the difference between the energy prices on both sides of the line times the flow through it, according to marginal theory. Thus, marginally, the higher the difference in energy prices at both ends of a line, the larger the benefits that the users of the network can obtain from reinforcing this line. In any case, carrying out an independent cost-benefit analysis to determine the optimal investment plan seems unavoidable.

Additionally, only under strongly stylized conditions, the revenues from the management of congestion would be enough to recover the cost of building the grid¹³. In practice, congestion rents prove to be clearly insufficient to completely pay for grid costs in a well developed network. In any case, to the extent transmission lines are built through regulated investment projects, congestion rents may contribute to the partial recovery of the sunk cost of the grid. Alternatively, these rents could be used to cover the losses incurred by the SO when solving those grid restrictions that remain after the final organized energy auction takes place (normally through redispatch).

Apart from that, changes in the grid topology that are produced by new investments result in a different distribution of flows in the grid. Consequently, the value of the transmission capacity of each line also changes. For this reason, owners of merchant transmission lines, who earn the congestion rents produced by the line as we explain in chapter 2, may oppose the construction of new lines that (partially) remove the congestion in the system. As mentioned above, only a benefit-cost analysis carried out at system level (for all the region in the case of international markets) can determine the extent to which congestion in the grid must be addressed through new investments in transmission facilities. In this sense, transmission and generation may be seen as substitutes of one another since, structurally speaking, systematic congestion can be solved by either installing new generation capacity or building new transmission lines between the exporting and the importing zones.

Single Price Areas that are defined for congestion management purposes mainly depend on the system grid topology. Hence, expanding the grid may affect the number and boundaries of the optimal set of SPAs. In general, defining single price areas in radial systems is obvious, whereas doing so in meshed networks may turn out to be impossible.

The recovery of the sunk grid cost should not affect the operation of the system. Transmission tariffs should not affect the allocation of scarce transmission capacity. Thus, it is essential that grid charges are not perceived by the agents as a variable cost. Otherwise, the agents will internalize these costs in the bids and offers they submit to contract energy and capacity and the dispatch will not result in the set of transactions that are most efficient for the system. Ill designed transmission tariffs that modify the variable costs of agents will certainly alter the value that these agents place on transmission capacity. In this case, transmission capacity would not be priced at its true marginal value.

¹³ In order for this to be true the transmission grid must be perfectly adapted to the system needs at all times, which is not possible due to the fact that the load and generation patterns as well as the grid topology vary over time. Additionally, investments in transmission facilities should be continuous, which is not possible (they are discrete). Finally, existing reliability constraints also invalidate this result.

Bibliography

- Aguado, J. A., V. H. Quintana, et al. (2004). "Coordinated Spot Market for Congestion Management of Inter-Regional Electricity Markets." IEEE Transactions on Power Systems **19**(1): 180-187.
- Alomoush, M. I. and S. M. Shahidehpour (2000). "Contingency-constrained congestion management with a minimum number of adjustments in preferred schedules." Electrical Power and Energy Systems **22**: 277-290.
- Baldick, R. (2003). "Variation of Distribution Factors with Loading." IEEE Transactions on Power Systems **18**(4): 1316-1323.
- Baran, M. E., V. Babunarayanan, et al. (2000). "Equitable allocation of congestion relief Cost to Transactions." IEEE Transactions on Power Systems **15**(2): 579-585.
- Bartholomew, E. S., A. S. Siddiqui, et al. (2003). "The New York Transmission Congestion Contract Market: Is it Truly Working Efficiently?" The Electricity Journal: 14-24.
- Bautista, G. and V. H. Quintana (2005). "Screening and Mitigation of Exacerbated Market Power Due to Financial Transmission Rights." IEEE Transactions on Power Systems **20**(1): 213-222.
- Bertrand, J. (1883). "Theorie Mathematique de la Richesse Sociale." Journal des Savants: 499-508.
- Biskas, P. N. and A. G. Bakirtzis (2002). "Decentralised Congestion Management of Interconnected Power Systems." IEE Proceedings: 432-438.
- Björndal, M. and K. Jörnsten (2001). "Zonal Pricing in a Deregulated Electricity Market." The Energy Journal **22**(1): 51-73.
- Bompard, E., P. Correia, et al. (2003). "Congestion-Management Schemes: A Comparative Analysis Under a Unified Framework." IEEE Transactions on Power Systems **18**(1): 346-452.
- Borenstein, S., J. Bushnell, et al. (2000). "The Competitive Effects of Transmission Capacity in a Deregulated Electricity Industry." Rand Journal of Economics **31**(2): 294-325.
- Boucher, J. and Y. Smeers (2001). Towards a common European Electricity Market- Paths in the right direction... still far from an effective design. Harvard Electricity Policy Group, Web page:

http://www.ksg.harvard.edu/hepg/Standard_Mkt_dsgn/Smeers_Interconnections1_4jni_3.do1.pdf.

- Cardell, J., C. C. Hitt, et al. (1997). "Market Power and Strategic Interaction in Electricity Networks." Resource and Energy Economics **19**(1-2): 109-137.
- Chao, H.-P. and S. Peck (1996). "A Market Mechanism for Electric Power Transmission." Journal Of Regulatory Economics **10**: 25-29.
- Chao, H.-p. and S. Peck (2000). "Flow-based Transmission Rights and Congestion Management." Electricity Journal **13**(8): 38-59.
- Chen, L., H. Suzuki, et al. (2002). "Components of Nodal Prices for Electric Power Systems." IEEE Transactions on Power Systems **17**(1): 41-49.
- Christie, R. D. and I. Wangensteen (1998). "The Energy Market in Norway and Sweden: Congestion Management." IEEE Power Engineering Review: 61-63.
- Consentec/Frontier (2004). Analysis of Cross-Border Congestion Management Methods for the EU Internal Electricity Market, Consentec and Frontier Economics. Prepared for the European Commission. Directorate-General Energy and Transport. Web page: http://europa.eu.int/comm/energy/electricity/publications/index_en.htm.
- Cournot, A. (1838). "Reserches sur les Principes Mathematics de la Theorie des Richesses."
- Deladreu, S., F. Brouaye, et al. (2003). "Using Two Multivariate Methods for Line Congestion Study in Transmission Systems Under Uncertainty." IEEE Transactions on Power Systems **18**(1): 353-358.
- EOR (2005). Reglamento del Mercado Eléctrico Regional (RMER). Libro III del RMER. De la Transmisión., Ente Operador Regional (EOR).
- EP (2003). Regulation (EC) No 1228/2003 of the European Parliament and of the council of 26 June 2003 on conditions for access to the network for cross-border exchanges in electricity. Web page: http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_176/l_17620030715en00010010.pdf: 10.
- ERGEG (2005). The Creation of Regional Electricity Markets. Web page: http://www.ergeg.org/portal/page/portal/ERGEG_HOME/ERGEG_DOCS/ERGEG_DOCUMENTS_NEW/ELECTRICITY_FOCUS_GROUP/: 123.
- ETSO (1999). Evaluation of congestion management methods for cross-border transmission, European Transmission System Operators (ETSO). Web page: <http://www.etsa-net.org/>.

- ETSO (2001). Co-ordinated Auctioning: A market-based method for transmission capacity allocation in meshed networks, Association of European Transmission System Operators (ETSO). Web Address: <http://www.etsonet.org/upload/documents/Coordinated%20Auctioning.pdf>: 22.
- ETSO (2001). Key Concepts and Definitions for Transmission Access Products, European Transmission System Operators (ETSO). Web page: <http://www.etsonet.org/>.
- ETSO (2002). Reconciliation of market splitting with coordinated auction concepts, European Transmission System Operators (ETSO). <http://www.etsonet.org/>.
- ETSO (2004). An Overview of Current Cross-border Congestion Management Methods in Europe, European association of Transmission System Operators (ETSO). Web address: http://www.etsonet.org/upload/documents/Current_CM_methods_final_20040908.pdf: 31.
- ETSO / EuroPEX (2004). Flow-based Market Coupling: A Joint ETSO-EuroPEX Proposal for Cross-Border Congestion Management and Integration of Electricity Markets in Europe, Association of European Transmission System Operators (ETSO) and association of European Power Exchanges (EuroPEX). Web address: http://www.etsonet.org/upload/documents/ETSO-EuroPEX_Interimreport_Sept-2004-.pdf: 26.
- EuroPEX (2003). Using implicit auctions to manage cross-border congestion: Decentralized Market Coupling. Rome, European association of Power EXchanges (EuroPEX). 10th Meeting of the European Electricity Regulatory Forum. Web page: <http://www.europex.org/pages/943372/index.htm>.
- Fang, R. S. and A. K. David (1999). "Optimal Dispatch under Transmission Contracts." IEEE Transactions on Power Systems **14**(2): 732-737.
- FERC (2002). Notice of Proposed Rulemaking, Remedying Undue Discrimination through Open Access Transmission service and Standard Electricity Market Design. Washington, Federal Energy Regulatory Commission. Web page: <http://www.ferc.gov/industries/electric/indus-act/smd/nopr.asp>.
- FERC (2003). White Paper: Wholesale Power Market Platform, Federal Energy Regulatory Commission (FERC). Web page: <http://elibrary.ferc.gov/idmws/nvcommon/NVViewer.asp?Doc=9686957:0>: 32.
- Galiana, F. D. and M. Ilic (1998). "A Mathematical Framework for the Analysis and mangement of power transactions under Open Access." IEEE Transactions on Power Systems **13**(2): 681-687.

- Gan, D. and D. V. Bourcier (2002). "Locational Market Power Screening and Congestion Management: Experience and Suggestions." IEEE Transactions on Power Systems **17**(1): 180-185.
- Gilbert, R., K. Neuhoﬀ, et al. (2004). "Allocating Transmission to Mitigate Market Power in Electricity Networks." Rand Journal of Economics **35**(4): 691-709.
- Glachant, J.-M. and V. Pignon (2002). Nordic electricity congestion's arrangements as a model for Europe: Physical constraints or operators' opportunism? ISNIE International Conference, MIT, Boston.
- Grijalba, S., P. W. Sauer, et al. (2003). "Enhancement of Linear ATC Calculations by the incorporation of reactive Power Flows." IEEE Transactions on power Systems **18**(2): 619-624.
- Hamoud, G. (2000). "Assessment of Available Transfer Capability of Transmission Systems." IEEE Transactions on Power Systems **15**(1): 27-32.
- Hartigan, J. (1975). Clustering Algorithms. New York, John Wiley & Sons.
- Henney, A. (2002). "What the US Could Learn from Western Europe and Elsewhere." Electricity Journal **15**(10): 53-64.
- Hogan, W. (1996). "A Wholesale Pool Spot Market Must Be Administered by the Independent System Operator: Avoiding the Separation Fallacy." The Electricity Journal: 26-37.
- Hogan, W. W. (1992). "Contract Networks for Electric Power Transmission." Journal of Regulatory Economics **4**(3): 211-242.
- Hogan, W. W. (2002). "Electricity Market Restructuring: Reforms of Reforms." Journal of Regulatory Economics **21**(1): 103-132.
- Hogan, W. W. (2002). Financial Transmission Right Formulations. Cambridge. <http://www.ksg.harvard.edu/hepg>, Center for Business and Government, John F. Kennedy School of Government, Harvard University.
- IAEW (2001). Analysis of Electricity Network Capacities and Identification of Congestion. Aachen, Institute of Power Systems and Power Economics (IAEW) and CONSENTEC Consulting fur Energiewirtschaft und -technik. Report for the European Commission, Directorate-General Energy and Transport.
- IAEW (2002). Analysis of Electricity Network Capacities and Identification of Congestion, Institute of Power Systems and Power Economics (IAEW) and CONSENTEC Consulting fur Energiewirtschaft und technik. Prepared for the European Commission, Directorate-General of Energy and Transport.

- Ilic, M., E. Hsieh, et al. (2003). "Transmission Pricing of Distributed Multilateral Energy Transactions to Ensure System Security and Guide Economic Dispatch." IEEE Transactions on Power Systems **18**(2): 428-434.
- Joskow, P. and J. Tirole (2000). "Transmission rights and market power on electric power networks." Rand Journal of Economics **31**(3): 450-487.
- Kavicky, J. A. and S. M. Shahidehpour (1997). "Determination of Generator Siting and Contract Options Based on Interutility Tie Line Flow Impacts." IEEE Transactions on Power Systems **12**(4): 1649-1653.
- KEMA (2002). Estudios para el Desarrollo del Reglamento de Operación Técnica y Comercial del MER, Report prepared by KEMA Consulting and ISA Interconexión Eléctrica for the Unidad Ejecutora del Proyecto SIEPAC, in Spanish: 117.
- Kench, B. T. (2004). "Let's Get Physical! Or Financial? A Study of Electricity Transmission Rights." Journal of Regulatory Economics **25**(2): 187-214.
- Khairuddin, A. B., S. S. Ahmed, et al. (2004). "A Novel Method for ATC Computations in a Large-Scale Power System." IEEE Transactions on Power Systems **19**(2): 1150-1158.
- Klemperer, P. D. and M. A. Meyer (1989). "Supply Function Equilibria in Oligopoly under Uncertainty." Econometrica **57**(6): 1243-1277.
- Kohonen, T. (1995). Self-Organizing Maps, Springer.
- Kristiansen, T. (2004). "Congestion management, transmission pricing and area price hedging in the Nordic region." International Journal of Electrical Power & Energy Systems **Vol. 26**(No. 9): 685-695.
- Kumar, A., S. C. Srivastava, et al. (2004). "A Zonal Congestion Management Approach Using Real and Reactive Power Rescheduling." IEEE Transactions on Power Systems **19**(1): 554-562.
- Latorre-Bayona, G. and I. Pérez-Arriaga (1994). "CHOPIN, A heuristic Model for Long Term Transmission Expansion Planning." IEEE Transactions on Power Systems **9**(4): 1886-1894.
- Méndez, R. and H. Rudnick (2004). "Congestion Management and Transmission Rights in Centralized Electric Markets." IEEE Transactions on Power Systems **19**(2): 889-896.
- Neuhoff, K. (2002). Integrating Transmission and Energy Markets Mitigates Market Power, Cambridge, DAE. Web page: <http://www.econ.cam.ac.uk/electricity/publications/wp/index.htm>.

- Neuhoff, K., J. Barquin, et al. (2002). "Network-constrained Cournot models of liberalized electricity markets: the devil is in the details." Energy Economics **27**(3): 495-525.
- Newbery, D., E. v. Damme, et al. (2003). Benelux market integration: Market power concerns, Dutch energy regulator (DTe). Web page: http://www.dte.nl/nl/Images/12_13216.pdf.
- Olmos, L. and K. Neuhoff (2006). "Identifying a Balancing Point for Electricity Transmission Contracts". IEEE Transactions on Power Systems **21** (4): 91-98.
- O'Neill, R. P., U. Helman, et al. (2002). "A Joint Energy and Transmission Rights Auction: Proposal and Properties." IEEE Transactions on Power Systems **17**(4): 1058-1067.
- OATI (2003). Open Access Technology International (OATI). Web page: <http://www.oatiinc.com>. **2003**.
- Oren, S. S. and A. M. Ross (2002). "Economic Congestion Relief Across Multiple Regions Requires Tradable Physical Flow-gate Rights." IEEE Transactions on Power Systems **17**(1): 159-165.
- Ott, A. L. (2003). "Experience with PJM Operation, Design and Implementation." IEEE Transactions on Power Systems **18**(2): 528-534.
- Ou, Y. and C. Singh (2002). "Assessment of Available Transfer Capability and Margins." IEEE Transactions on Power Systems **17**(2): 463-468.
- Pérez-Arriaga, I. J. and L. Olmos (2005). "A plausible congestion management scheme for the internal electricity market of the European Union." Utilities Policy **Vol. 13**: 117-134.
- Pérez-Arriaga, J. I., L. Olmos, et al. (2002). Cost components of cross border exchanges of electricity, Directorate-General for Energy and Transport / European Commission. Web page: http://europa.eu.int/comm/energy/electricity/publications/index_en.htm.
- Rassenti, S. J., V. L. Smith, et al. (2003). "Discriminatory Price Auctions in Electricity Markets: Low Volatility at the RExpense of High Price Levels." Journal of Regulatory Economics **23**(2): 109-123.
- Rau, N. S. (2000). "Transmission Loss and Congestion Cost Allocation: An Approach Based on Responsibility." IEEE Transactions on Power Systems **15**(4): 1401-1409.
- Rau, N. S. (2003). "Issues in the Path Toward an RTO and Standard Markets." IEEE Transactions on Power Systems **18**(2): 435-443.

- Rivier, M. and J. I. Pérez-Arriaga (1993). Computation and decomposition of spot prices for transmission pricing. 11th PSCC Conference, Avignon, France.
- Rothkopf, M. H. (2001). Control of Market Power in Electricity Auctions. Piscataway, RUTCOR, Rutgers University.
- Rothkopf, M. H. (2002). "Control of Market Power in Electricity Auctions." The Electricity Journal: 15-24.
- Rubio, F. J. and J. I. Pérez-Arriaga (2000). "Marginal Pricing of Transmission Services: A Comparative Analysis of Network Cost Allocation Methods." IEEE Transactions on Power Systems **VOL. 15**(NO. 1): 448-454.
- Ruff, L. E. (1999). Competitive Electricity Markets: Why They Are Working and How to Improve Them. <http://www.ksg.harvard.edu/hepg, n / e / r / a>.
- Ruff, L. E. (2000). Flowgates vs. FTRs and options vs. Obligations. <http://www.ksg.harvard.edu/hepg>.
- Ruff, L. E. (2001). "Flowgates, Contingency-Constrained Dispatch and Transmission Rights." The Electricity Journal **14**(1): 34-55.
- Schweppe, F., M. Caramanis, et al. (1988). Spot pricing of electricity, Kluwer Academic Publishers.
- SEEEC (2005). South East Europe Energy Community Treaty. 2005, 25 October. Web page: http://europa.eu.int/comm/energy/electricity/south_east/treaty_en.htm.
- Singh, H., S. Hao, et al. (1998). "Transmission Congestion Management in Competitive Electricity Markets." IEEE Transactions on Power Systems **13**(2): 672-680.
- Stoft, S. (1998). "Congestion Pricing with Fewer Prices than Zones." The Electricity Journal **11**(4): 23-31.
- Stoft, S. (1999). "Financial Transmission Rights Meet Cournot: How TCCs Curb Market Power." The Energy Journal **20**(1): 1-23.
- Tabors, R. and M. Caramanis (2000). Real Flow, A Preliminary Proposal for a Flow-based Congestion Management System. Cambridge, MA, Web page: <http://www.ksg.harvard.edu/hepg/flowgate/Real%20Flow.pdf>.
- Tao, S. and G. Gross (2002). "A Congestion Management Allocation Mechanism for Multiple Transaction Networks." IEEE Transactions On Power Systems **17**(3): 826-833.

- Tobin, V. J. (1990). "Transmission Access: A Voluntary System for Resolution of Transmission Access and Pricing Problems." IEEE Power Engineering Review **10**(4): 18-21.
- Wang, P., Y. Xiao, et al. (2004). "Nodal Market Power Assessment in Electricity Markets." IEEE Transactions on Power Systems **19**(3): 1373-1379.
- Wei, P., Y. Ni, et al. (2002). "Decentralised approach for congestion management and congestion price discovering." IEE Proceedings on Generation, Transmission and Distribution **149**(6): 645-652.
- Wood, A. J. and B. F. Wollenberg (1984). Power Generation, Operation and Control. New York, John Wiley & Sons.

Annex 1: Equivalence, under ideal conditions, between regional explicit auctions followed by local implicit ones and regional implicit auctions

We begin by presenting the mathematical formulation of the optimization problems when solved under both approaches. Approach I corresponds to the combination of regional explicit auctions for transmission capacity and local implicit auctions. Approach II corresponds to the case of regional implicit auctions. The formulation of both problems has been simplified for the sake of clarity.

Under approach I, the problem corresponding to the allocation of capacity at regional level can be stated as:

$$\begin{aligned} \max \quad & \sum_j B_{j,a} p_{j,a} \\ \text{st.} \quad & \\ & \sum_{j,a} p_{j,a} PTDF_{j,a,k} \leq \overline{F}_k^+ \quad , g_{e,k} \quad \forall k=1, \dots, n \text{ number of corridors} \end{aligned} \quad (4.93)$$

$$\sum_{j,a} p_{j,a} PTDF_{j,a,k} \geq \overline{F}_k^- \quad , m_{e,k} \quad \forall k=1, \dots, n \text{ number of corridors} \quad (4.94)$$

while the problem corresponding to the implicit auction taking place within each area is:

$$\begin{aligned} \max \quad & \sum_i B_{i,a} p_{i,a} \\ \text{st.} \quad & \\ & \sum_{i,a} p_{i,a} = 0 \quad , p \quad \text{global balance equation} \end{aligned} \quad (4.95)$$

$$p_{i,a} \leq \overline{p}_{i,a} \quad , q_{i,a} \quad \forall a=1, \dots, A \text{ number of agents} \quad (4.96)$$

The notation used here is analogous to the one used in section 4.3 when presenting the different regional congestion management mechanisms. Here, j is the set of bids for capacity in the explicit auction whereas i is the set of energy bids in the subsequent implicit auctions. The power scheduled corresponding to bid i by agent a is $p_{i,a}$ whereas the capacity scheduled for bid j by agent a is $p_{j,a}$. Therefore, $\overline{p}_{i,a}$ is the maximum amount of power that agent a is willing or able to transact as part of transaction i .

Under approach II, the formulation for the implicit regional auction is:

$$\begin{aligned} \max \quad & \sum_e B_{e,a} p_{e,a} \\ \text{s.t.} \quad & \sum_{e,a} p_{e,a} = 0 \quad ,p \quad \text{global balance equation} \end{aligned} \quad (4.97)$$

$$\sum_{e,a} p_{e,a} PTDF_{e,a,k} \leq \overline{F}_k^+ \quad ,g_{e,k} \quad \forall k=1,...,n \quad \text{number of corridors} \quad (4.98)$$

$$\sum_{e,a} p_{e,a} PTDF_{e,a,k} \geq \overline{F}_k^- \quad ,m_{e,k} \quad \forall k=1,...,n \quad \text{number of corridors} \quad (4.99)$$

$$p_{e,a} \leq \overline{p_{e,a}} \quad ,q_{e,a} \quad \forall a=1,...,A \quad \text{number of agents} \quad (4.100)$$

Notation used in this case is the same as under Approach I. Only indexes i and j have been substituted with index e to refer to the regional implicit auction. In this formulation of the regional implicit auction bids for capacity corresponding to bilateral contracts have not been included in order to make the formulation more compact. However, dealing with these bids would not be a problem. In fact, this would not add any complexity to the proof from a conceptual point of view.

As mentioned in the main text of this chapter, for Approaches I and II to be equivalent we must assume that agents are rational, they have access to complete information and no changes in the system state and configuration takes place between the explicit and the implicit auctions in Approach I. Provided this stringent conditions apply, we can assure that:

$$B_{i,a} = B_{e,a} - B_{j,a} \quad \forall \text{ agent and for corresponding transactions} \quad (4.101)$$

$$\overline{p_{e,a}} = \overline{p_{i,a}} \quad \forall \text{ agent and for corresponding transactions} \quad (4.102)$$

$$PTDF_{j,a,k} = PTDF_{e,a,k} \quad \forall \text{ agent and for corresponding transactions} \quad (4.103)$$

$$p_{j,a} = p_{i,a} \quad \forall \text{ agent and for corresponding transactions} \quad (4.104)$$

Equation (4.101) assumes that agents are rational and all of them have access to complete information. Then, the offer agent a makes in the local implicit auction held in the area where it plans to trade its energy equals its marginal cost or utility (the amount it would bid in the regional implicit auction under Approach II) less the marginal value of the capacity required to access this area (the amount this agent would bid and pay for this capacity in the regional explicit auction). Equation (4.102) states that the maximum amount of energy an agent is willing to transact would be the same in both cases. Equation (4.103) assumes that the configuration of the grid is the same under both approaches. Finally, equation (4.104) assumes that all the capacity bought by an agent in the explicit auction will be used later to buy or sell power in the corresponding area. Provided capacity products are obligations, this agent would be penalized if it does not comply with this requirement.

If we derive first order optimality conditions for the implicit auction in Approach I we have:

$$B_{i,a} + p_i - q_{i,a} = 0 \quad \forall i, a = 1, \dots, I, A \quad \text{number of agents and transactions} \quad (4.105)$$

$$\sum_i p_{i,a} = 0 \quad \text{global balance equation} \quad (4.106)$$

$$q_{i,a} (p_{i,a} - \overline{p_{i,a}}) = 0 \quad \forall i, a = 1, \dots, I, A \quad \text{number of agents and transactions} \quad (4.107)$$

$$q_{i,a} \geq 0 \quad \forall i, a = 1, \dots, I, A \quad \text{number of agents and transactions} \quad (4.108)$$

First order conditions for the explicit auction under Approach I are:

$$B_{j,a} - \sum_k PTDF_{j,a,k} (g_{j,k} - m_{j,k}) = 0 \quad \forall j, a \text{ number of agents and transactions} \quad (4.109)$$

$$g_{j,k} (\sum_{j,a} p_{j,a} PTDF_{j,a,k} - \overline{F_k^+}) = 0 \quad \forall k = 1 \dots K \text{ number of corridors} \quad (4.110)$$

$$m_{j,k} (\sum_{j,a} p_{j,a} PTDF_{j,a,k} - \overline{F_k^-}) = 0 \quad \forall k = 1 \dots K \text{ number of corridors} \quad (4.111)$$

$$g_{j,k} \geq 0 \quad \forall k = 1 \dots K \text{ number of corridors} \quad (4.112)$$

$$m_{j,k} \leq 0 \quad \forall k = 1 \dots K \text{ number of corridors} \quad (4.113)$$

Apart from that, first order conditions for the regional implicit auction in Approach II are:

$$B_{e,a} + p_e - \sum_k PTDF_{e,a,k} (g_{e,k} - m_{e,k}) - q_{e,a} = 0 \quad \forall e, a = 1, \dots, E, A \text{ number of agents and transactions} \quad (4.114)$$

$$\sum_e p_{e,a} = 0 \quad \text{global balance equation} \quad (4.115)$$

$$q_{e,a} (p_{e,a} - \overline{p_{e,a}}) = 0 \quad \forall e, a = 1, \dots, E, A \text{ number of agents and transactions} \quad (4.116)$$

$$g_{e,k} (\sum_{e,a} p_{e,a} PTDF_{e,a,k} - \overline{F_k^+}) = 0 \quad \forall k = 1 \dots K \text{ number of corridors} \quad (4.117)$$

$$m_{e,k} (\sum_{e,a} p_{e,a} PTDF_{e,a,k} - \overline{F_k^-}) = 0 \quad \forall k = 1 \dots K \text{ number of corridors} \quad (4.118)$$

$$g_{e,k} \geq 0 \quad \forall k = 1 \dots K \text{ number of corridors} \quad (4.119)$$

$$m_{e,k} \leq 0 \quad \forall k = 1 \dots K \text{ number of corridors} \quad (4.120)$$

$$q_{e,a} \geq 0 \quad \forall e, a = 1, \dots, E, A \text{ number of agents and transactions} \quad (4.121)$$

It is easy now to check that optimality conditions under both approaches coincide. Thus, adding (4.105) and (4.109) we obtain:

$$B_{i,a} + B_{j,a} + p_i - \sum_k PTDF_{j,a,k} (g_{j,k} - m_{j,k}) - q_{i,a} = 0 \quad \forall e, a \quad (4.122)$$

If we substitute (4.101) in equation (4.122) we end up obtaining equation (4.114). At the same time, if equation (4.114) is true, then it is always possible to decompose $B_{e,a}$ into two variables $B_{i,a}$ and $B_{j,a}$ such that (4.105) and (4.109) are also satisfied.

The remaining optimality conditions for Approach I are the same as those for Approach II substituting $\overline{p_{i,a}}$ with $\overline{p_{e,a}}$ (equation (4.102)), $PTDF_{j,a,k}$ with $PTDF_{e,a,k}$ (equation (4.103)) and $p_{j,a}$ with $p_{i,a}$ (equation (4.104)).

Annex 2: Equivalence between the net revenue from the application of nodal prices and that associated with the aggregate marginal cost of restrictions limiting the flow on lines

This annex proves the equivalence between the sum of the marginal values of all the lines in the system and the aggregate cost in the dispatch of all the restrictions on the flow over the transmission lines. For the sake of simplicity, we shall assume a dispatch where there are only restrictions of this type. See the formulation below, which is identical to that used for a regional implicit auction in annex 1.

$$\begin{aligned} \max \quad & \sum_e B_{e,a} p_{e,a} \\ \text{s.t.} \quad & \sum_{e,a} p_{e,a} = 0, p \quad \text{global balance equation} \end{aligned} \quad (4.123)$$

$$\sum_{e,a} p_{e,a} PTDF_{e,a,k} \leq \overline{F}_k^+ \quad , g_{e,k} \quad \forall k=1, \dots, n \quad \text{number of corridors} \quad (4.124)$$

$$\sum_{e,a} p_{e,a} PTDF_{e,a,k} \geq \overline{F}_k^- \quad , m_{e,k} \quad \forall k=1, \dots, n \quad \text{number of corridors} \quad (4.125)$$

$$p_{e,a} \leq \overline{p_{e,a}} \quad , q_{e,a} \quad \forall a=1, \dots, A \quad \text{number of agents} \quad (4.126)$$

From nodal pricing equations we know that the marginal energy price at any node i can be expressed as:

$$r_i = p + \sum_k PTDF_{i,r,k} (g_{e,k} + m_{e,k}) \quad (4.127)$$

where p is the system energy price, $g_{e,k}$ and $m_{e,k}$ are the dual variables of restrictions on the flow over line k in both directions, r_i is the marginal price in node i and $PTDF_{i,r,k}$ is the Power Transfer Distribution Factor of the flow over line k with respect to a transaction between the reference node r , where the system energy price is computed, and node i .

Then the sum over all nodes i of the corresponding nodal price times the net demand at the node is:

$$\sum_i r_i p_i = \sum_i \left(p + \sum_k PTDF_{i,r,k} (g_{e,k} + m_{e,k}) \right) p_i \quad (4.128)$$

Taking into account that $\sum_i p_i = \sum_{e,a} p_{e,a} = 0$ and operating equation (4.128), we have:

$$\sum_i r_i p_i = \sum_k \left((g_{e,k} + m_{e,k}) \sum_i PTDF_{i,r,k} p_i \right) \quad (4.129)$$

At the same time, we have that the sum of the shadow prices of the restrictions on the flow over line k may be different from zero only if this line is congested. Otherwise, the corresponding restrictions in equations (4.124) and (4.125) are not binding and their marginal cost is zero. Thus, we have:

$$\sum_i r_i p_i = \sum_{k'} (g_{e,k'} + m_{e,k'}) \overline{F}_{k'} \quad (4.130)$$

where k' is the set of congested lines in the system. We have demonstrated that the net revenue from the application of nodal prices is equal to the sum over all congested lines of the marginal cost of each line in the dispatch times the capacity of the line.

Now, we aim to prove that the net result of applying nodal prices is equal to the sum of the marginal values of all the lines in the system. Section 4.4.1 concludes that, under certain conditions, the marginal value of a line is equal to the difference between the nodal prices at both ends of the line times the flow over the line. Then we have that:

$$\sum_k MV_k = \sum_k (r_{k,1} - r_{k,2}) f_k \quad (4.131)$$

where k is an index representing the set of lines in the system, $r_{k,1}$ is the energy price at the node where f_k arrives and $r_{k,2}$ is the energy price at the node where f_k departs from. MV_k is the marginal value of line k in the dispatch. Operating the right hand term in equation (4.131), we obtain:

$$\sum_k (r_{k,1} - r_{k,2}) f_k = \sum_i r_i \sum_k f_k x_{i,k} \quad (4.132)$$

where i is the set of nodes in the system, r_i is the nodal energy price at node i and $x_{i,k}$ is a parameter whose value is 1 if f_k arrives at node i , -1 if f_k departs from node i and 0 otherwise. Since for every node i in the system inflows must be equal to outflows, we have:

$$\sum_k f_k x_{i,k} = p_i \quad (4.133)$$

Hence, substituting (4.133) in (4.132) and (4.132) in (4.131), we finally obtain:

$$\sum_k MV_k = \sum_i r_i p_i \quad (4.134)$$

Finally, combining equations (4.134) and (4.130) yields:

$$\sum_k MV_k = \sum_{k'} (g_{e,k'} + m_{e,k'}) \overline{F_{k'}} \quad (4.135)$$

As we tried to demonstrate.

5 Conclusions

5.1 Summary and conclusions

The regulation of the transmission activity must be approached as a whole, instead of as a collection of separate topics. The three main aspects of transmission regulation, namely investments in new infrastructure, computation of transmission tariffs and access to the transmission grid, are highly interrelated. The following interrelationships can be observed between the three aforementioned topics:

- **Investments and tariffs:** as explained in chapter 2 of this thesis, the revenues of the transmission activity, to be recovered through the application of transmission tariffs, critically depend on the regulatory paradigm adopted for the expansion of the grid. At the same time, locationally differentiated transmission charges may influence the decision by market agents on whether to install a new generation plant or load and where to install it. Load and generation patterns, which result from the individual investment decisions made by market agents, must, in turn, be taken into account by the entity in charge of achieving an optimal expansion of the grid, when identifying the most efficient grid reinforcements.
- **Investments and congestion:** implementing an efficient mechanism of congestion management results in energy prices at different locations being different. Differences in energy prices are one of the major indicators of the need to carry out investments in new transmission facilities. The difference in energy prices between two points of the grid is related to the decrease in the cost of congestion that would result from building a line between these two points, as proved in chapter 4. Therefore, congestion costs are one of the major drivers behind transmission investments. At the same time, network reinforcements may significantly change the amount of congestion in the system, therefore benefiting or hurting the economic interests of the network users. The optimal development of the grid results in an adequate balance between the cost of congestion and that of network investments.
- **Tariffs and congestion:** as it has been proved in previous research work, see (Rubio 1999), if certain ideal conditions were met, congestion rents would suffice to recover the total cost of transmission. In practice, revenues related to congestion resulting from the application of nodal or zonal energy prices fall far short of the total regulated revenues of the transmission activity. In any case, congestion rents should contribute to the partial recovery of the cost of the grid. Hence, the higher these rents are, the lower the transmission tariffs would need to be. Investing in new transmission facilities normally reduces congestion rents and, therefore, increases the fraction of the cost of the existing grid to be recovered from transmission tariffs.

On the other hand, transmission tariffs should not affect the operation of the system, in order not to distort the optimal short-term signals sent by nodal energy prices. Transmission tariffs should not be a variable charge since these would alter the operational decisions made by agents. Therefore, the congestion in the grid resulting from the system operation should not be affected by the way the cost of the grid is allocated among agents.

As a consequence of all this, the models and rules that are implemented to deal with a certain aspect of transmission regulation must be consistent with those that are used to deal with the other aspects. What is more, we should bear in mind the consequences that applying a certain rule, aimed at dealing with a problem, may have on the remaining aspects of the functioning of the transmission activity.

5.1.1 Reference regulation of the transmission activity

Now we focus on the case where a system is perfectly integrated (a national market, for example), and outline rules and mechanisms to regulate the functioning of the transmission activity in this system. We may draw the following general conclusions from the previous analysis of the three main aspects of transmission regulation.

5.1.1.1 Investments in new transmission facilities

After having discussed the main paradigms that exist for regulating the expansion of the grid in the new liberalized environment, we have concluded that both the public and the private initiative must have a role to play in this process. Central planning must be given priority over merchant investments and investments at risk since only the former model will yield a transmission grid that is sufficiently developed. Merchant investors, and associations of network users, can only be relied upon to promote or actually perform the construction of very specific lines. Private investments could include those that are directly associated with significant new generation investments, those that are perceived by the central planner as very risky and those whose construction would be delayed if they had to be included in the indicative grid expansion plan.

Both the public and the private initiative must be given some role in the expansion of the grid. However, centralized expansion planning must take precedence over merchant lines and investments at risk.

Accepting or rejecting a particular investment or a set of them (investment plan) requires applying a set of rules in order to determine whether the proposed reinforcement is beneficial for the system and how it compares to alternative reinforcements. This set of rules is commonly called Regulatory Test. Adequately designing the Regulatory Test is central to the successful implementation of any mechanism for the expansion of the grid. The regulatory test must be able to evaluate the benefits that each one of the investment proposals will yield in the considered time

horizon. The test to be applied to regulated reinforcements must be different from that used for merchant projects or investments at risk since private promoters and not consumers pay the cost of the latter ones. In any case, negative benefits or losses incurred by grid users as a result of the construction of a line or lines should be taken into account when deciding about their construction.

Under the new regulatory paradigm, investments in grid facilities should be aimed at maximizing the net aggregate benefit of market agents in the region. Both positive and negative benefits (losses) incurred by market agents as a consequence of the construction of each line must be taken into account. Contrary to the case of regulated investments, the regulatory test to be applied in order to decide on the construction of lines financed by private investors must not consider the cost of the investment. Starting from a set of possible and previously identified reinforcements, the SO must propose those that are expected to result in the highest increase in the social welfare. The regulator¹ must check that the investment proposals by the SO are beneficial for the system. This entity may also ask the SO to evaluate regulated reinforcements proposed by third parties, which should be compared to the ones already proposed by the SO itself. Investments at risk should be approved as long as they are not detrimental to the system and no alternative regulated project has been already proposed.

The ensuing loss of efficiency due to the unbundling of investment decisions in generation, transmission and distribution may be partly compensated by coordinating locational economic signals at the different interfaces. Efficient investment in generation can be encouraged by the use of nodal energy prices and transmission tariffs with locational content. These signals should be designed on the basis of responsibility on the incurred network costs. The usefulness of network-related locational signals for transmission investment critically depends on the adopted regulatory approach for transmission and, in particular, on which is the entity in charge of the expansion of the grid: the regulatory authority, the transmission system operator, merchant investors or the network users themselves. Different regulatory paradigms for the expansion of the grid should result in different remuneration schemes. Hence, the economic incentives for the construction of new lines will probably differ from one organizational structure to another. Other useful locational signals of non economic nature, whose purpose is to decrease the level of uncertainty that both potential generation investors and transmission network planners must face, involve providing reliable information on the future investment decisions (generators) and the expected network conditions (System Operator).

Economic locational signals that are based on cost causality are always useful to guide the siting in the network of new generation and load facilities. The issue here is

¹ In some countries these functions of supervision and authorization by the regulator are not explicit and they are only exercised indirectly when determining the remuneration of the transmission activity. This thesis advocates a more clear role of the regulator in the authorization of new transmission investments.

whether these locational signals may or may not be strong enough to have any actual impact on the location of new network users or even on the decision to retire a power plant or a load from a given site. The thesis has provided criteria and numerical examples showing the nature of these signals and what can and cannot be expected from their application. A novel scheme to design transmission charges with a special emphasis on cost causality has been proposed. Locational signals may be also useful at the interface between transmission and distribution networks in order to minimize the distortions that result from having separate network operators under widely different regulatory schemes.

The transmission firms must be paid according to the costs they incur when building, operating and maintaining regulated transmission facilities including some incentives related to availability performance. Regarding system operation, some moderate performance-based incentives related to the reduction of losses and congestion and the improvement of reliability may be included in its remuneration scheme. Private owners of grid facilities must earn the marginal value of their lines in the dispatch or revenues resulting from any transmission capacity contracts they may sign.

In order to encourage efficient investment in new generation, energy prices must be efficient (they must include the marginal effect of losses and congestion). Additionally, transmission nodal charges to be paid by new network users must reflect the extra grid costs that can be associated to the installation of a new generator or load in any given node or area. Relevant information must be systematically exchanged between market agents and the central planner.

Coordinating signals should be sent so that the joint development of transmission and distribution grids is efficient. These signals are specially important at the interface between both types of grids.

5.1.1.2 Computation of transmission tariffs

Nodal prices are optimal short term energy signals. They also send long term signals to agents and help recover part of the cost of the transmission grid.

Revenues that might be obtained by the system from the sale of transmission rights in the long term and the short-term dispatch with nodal energy prices must be used to partly pay the cost of the existing transmission grid or to finance new grid investments.

Typically, revenues from the application of nodal prices (or any other dispatch model) only collect a small fraction of the cost of the grid. Therefore, additional transmission charges are necessary. These charges must promote efficiency in investments in the long run and allocative efficiency in the short run. Conceptually, grid tariffs should be based on the benefits that agents obtain from the network (or the cost they cause the system to incur). In practice, this cost allocation can only be performed without ambiguity for connection lines (dedicated to the physical connection of an agent to the grid) and for specific network investments whose beneficiaries are not very much dispersed. Connection lines must be defined as those transmission lines that connect an agent to the rest of the system and which are less than 30 Km long. Longer lines that

connect an agent to the system not only perform a connection function but also a transmission one. In other words, they do not only benefit the agent that gets connected to the rest of the system but also the consumers that are served by this agent (if the former agent is a generator) or the generators that serve this agent (if it is a load). However, when building a new line, we may restrict ourselves to consider the benefits that network users will obtain from it in the few years immediately following the time when the line goes into operation. This facilitates the identification of the beneficiaries of this line.

The cost of connection lines must be paid by the agents that get connected to the rest of the system through these lines. Lines of this type that are longer than a certain length (30 Km) cannot be considered connection lines, since they benefit both the agent that gets connected and the ones that are served by it (or the ones that supply this agent with power). In general, new lines whose beneficiaries can be clearly identified should be paid by them.

If identifying the beneficiaries of a line or set of lines is not possible, and sending long-term locational signals is necessary (due to the fact that the decision by agents on where to install new generation or load may cause the system to build new transmission lines), tariffs should be based on the usage that each agent is making of the grid. An in depth discussion of the technical, economic and legal aspects of the problem of allocation of sunk grid costs has led us to conclude that the Average Participations method is better suited than other usage-based methods to compute transmission tariffs. Properties of AP include the following:

1. AP yields very reasonable results under any circumstances. Results seem to be quite stable. They do not vary excessively from one scenario to another within the same year, nor from one year to another.
2. The AP method is very easy to implement. Applying AP only requires having available the power injections and withdrawals in the system as well as the flows over the lines for a number of representative snapshots of the system operation during a given year.
3. The AP algorithm agrees with the basic economic and engineering principles that must rule the allocation of sunk grid costs. According to the physics laws governing the functioning of power systems, it computes the true average use that generators and loads make of the transmission grid. Additionally, the proportionality branching rule applied in the method can be justified from the point of view of the economics theory. In fact, the solution provided by the AP method coincides with that of a cooperative game where the inflows into each node of the network are evenly distributed among the outflows.
4. This method provides an answer to other aspects of transmission regulation including the allocation of the cost of new network investments or the computation of locational signals.

The AP method computes the average use that each agent makes of the grid. Thus, locational signals obtained from the AP method may not reflect the incremental cost incurred by the system as a result of the installation or retirement of a generator or load. However, charges provided by AP may be modified in order to strengthen the relation between the cost caused by an agent and the transmission charges it faces. Another possibility that this thesis recommends to explore makes use of the Cooperative Game Theory to provide efficient transmission charges. However, these methods, which provide solutions that are more difficult to dispute, are still in the development phase.

If identifying without much ambiguity the beneficiaries of regulated lines is not possible, and sending locational signals is a priority, usage-based methods must be applied to compute transmission tariffs. The AP method is best suited to this task, among the usage-based methods considered in this thesis. Usage-based methods based on cooperative game theory are a promising alternative.

The structure of network charges is an issue of utmost importance. If transmission charges are computed according to the use that each agent makes of the transmission grid, the charge to be paid by each agent will be affected by the short-term decisions by this agent. This is not a desirable feature of transmission tariffs since network users will not only base their operational decisions on variable production costs but also on network costs, which is not efficient. The relative importance of grid charges when compared to energy price differences determines which is the best tariff structure to use. If transmission charges, when expressed in €/KWh, are of the same order of magnitude as energy prices, structuring grid tariffs as capacity charges is preferable. Otherwise, the dispatch of energy in the system would be significantly affected by the allocation of sunk network costs. In this case, we may first compute the total transmission charge to be paid by all the generators or loads of the same type (profile) that are located within the same area. Then, we would allocate this charge to the agents within this group proportionally to the installed capacity of each agent. On the other hand, if transmission tariffs expressed in €/KWh are small compared to energy prices, making these tariffs part of the variable cost faced by agents is acceptable from the point of view of efficiency distortion and it is easier to implement.

In order not to distort operational decisions by network users, the structure of transmission charges must be made dependent on the relative importance of these charges when compared to locational energy price differences. If transmission tariffs, when expressed in €/KWh, are significantly smaller than energy prices, structuring these tariffs as energy charges is preferable. If transmission tariffs, when expressed in €/KWh, are of the same order of magnitude as energy prices, the former should be structured as capacity charges. In this case, a set of price areas may be defined. Price areas defined for the purpose of computing transmission tariffs should comprise groups of agents that make a similar use of the network per unit of (generation / load) capacity (installed/ contracted). Agents with the same profile that are located within the same price area should pay the same transmission charge per unit of generation capacity installed, in the case of generators, or per unit of contracted capacity, in the case of loads.

The internal split of the total amount of charges between generation and demand in a country is also very relevant to our discussion. It is widely accepted that generators should pay a smaller part of the cost of the grid than loads. However, harmonizing the way in which transmission tariffs (to generators mainly) are computed may be central to achieving a leveled playing field where competition among agents in different countries can develop.

The fraction of the cost of the transmission grid within each state or country that is paid by generators should be approximately the same in all the countries included a regional market.

5.1.1.3 Access to the grid

Connection of agents to the transmission grid and the management of grid constraints are the two main aspects of the regulation of the access to the transmission grid. Here we have focused on the second one, which has not been satisfactorily solved in most systems, specially at the regional level.

Industry and academia have proposed diverse methods for the management of congestion in the grid. There seems to be wide agreement that the integrated dispatch of energy and transmission capacity, where a different energy price is derived for each node, is the preferred approach from an efficiency point of view, at least when market power considerations are not taken into account. In any case, integrating the allocation of energy and capacity seems superior to dealing with both separately. However, this scheme requires a higher level of centralization than others, which may be a significant obstacle in the way towards its implementation, due to opposition from local authorities and institutions.

We recommend implementing implicit auctions for the allocation of transmission capacity in the short-term markets, since they minimize the loss of efficiency resulting from the lack of coordination between the transmission capacity and energy allocation procedures. Specifically, they ensure that the scheduled transactions maximize the value of the available transmission capacity. This regulatory recommendation is valid for the case of the reference market model (one of a perfectly integrated system). When considering the implementation of a congestion management scheme in the existing regional markets, the type of auction adopted should be well suited to the political situation and institutional arrangements that are in place in each region.

Apart from the design of the short term energy and transmission capacity markets, other problems remain unsolved, such as the design of long and medium term capacity auctions, the format of the products that are exchanged in these auctions, the implication of the allocation of congestion rights on the exercise of market power or the allocation of congestion rents resulting from the dispatch. This thesis has attempted to contribute, at least partially, to the clarification of these issues.

Point-to-point rights confer agents an unambiguous right to execute the corresponding transaction or earn the corresponding congestion rents. Hence, they are clearly preferable for agents in order to shield against the risk of volatility in the transmission

price. However, flow-gate rights that are defined for each corridor that is likely to be congested in a region allow decentralized trading.

In general, we recommend allowing the joint trade of point-to-point and flow-gate rights in long and medium term capacity auctions. However, there are some regional markets where using only point-to-point rights is enough. This is the case of the MER in Central America.

Generally, financial rights are superior to physical rights since they do not affect the short-term dispatch of energy in the absence of market power. Physical rights may affect the dispatch of energy if the capacity market is not liquid enough. In any case, physical rights may be used for purposes other than risk hedging. Thus, they may provide a generator with the right to earn the capacity payment that is in place in the area where the withdrawal point of the right is located, or they may provide its owner with access to the green-certificates market. In the latter cases, implementing physical rights makes sense.

As a general rule, we recommend using financial rights in the long term. However, if energy and transmission capacity are dealt with separately in the short-term, transmission rights owned by market agents just before the short-term energy dispatch must be physical. In addition, physical rights may exist in the long-term for agents to secure access to the physical transmission capacity, if this is required.

There is room for trading both transmission capacity options and obligations. Obligations are preferable to options in markets that are not mature yet, since trading in obligations is easier to implement. Options may make sense in the long-term, when agents are uncertain about the commercial transactions that they will engage in. Obligations are better suited for the short-term since they allow netting of flows when transmission rights are traded.

Both options and obligations should exist. Options should be issued and traded in the long term. They should be transformed into obligations as the operation of the system draws closer.

Transmission rights may represent an incentive for strategic generators to exercise their market power. Rules aimed at discouraging agents from acquiring transmission rights that provide these agents with an extra incentive to exercise their market power may not be efficient in some cases. Thus, the transmission capacity products that are available to agents should be carefully designed in order not to aggravate this problem. If rights are defined with respect to an appropriate physical or virtual balancing node, they may help hamper the exercise of market power by agents.

If the exercise of market power by agents in a regional market is a serious concern, we recommend restricting the access of these agents to transmission rights. In this case, agents must only be allowed to buy transmission rights that are defined towards a balancing point whose energy price is independent from the commercial decisions by these agents. Finding a physical node with this characteristics may be difficult if not impossible in a large meshed

system like the European one. Therefore, we recommend defining transmission rights towards virtual hubs with good properties.

As for the allocation of congestion rents, the redistribution of flows that is caused by an increase in the capacity of a line should be taken into account when computing the marginal value of the transmission capacity of the line.

The true marginal value of transmission capacity must be used to compute the revenues of transmission owners. The transmission grid planner must use the true value of a line in the dispatch to decide about its construction. The value of a line should also guide the investment decisions by private entrepreneurs. Assuming that the relation between the transmission capacity and the admittance of a line is linear, the marginal value of this transmission line can be computed as the difference between the nodal energy prices at both ends of the line times the flow over the line.

5.1.2 Regulation of transmission in regional markets

The set of rules to be applied in a perfectly integrated system must be adapted to the political and institutional framework existing in each regional market. The competent authorities must develop the most efficient regulatory framework that is compatible with the political, regulatory and institutional arrangements that exist in the considered region.

The features of the model chosen for the regulation of the expansion of the grid must adapt to the particular situation in the regional market where we plan to implement it. Thus, we have analyzed how the hybrid approach (which leaves room both for the public and the private initiative) can be adapted to best fit the characteristics of three of the most relevant regional markets that exist, namely the IEM of the European Union, the regional markets being created in the US and the Regional Electricity Market in Central America. In order to attract investment in new transmission facilities, the set of rules implemented in any regional market to regulate the expansion of the grid must reduce, to the extent possible, the regulatory risk perceived by promoters.

In those regions where the transmission grid is clearly underdeveloped and central planning is unlikely to succeed in getting all the required lines built, private promoters should be encouraged to undertake some of these investments by implementing efficient pricing schemes (nodal or zonal marginal pricing). These private promoters will be interested in building lines so that the improvement in the nodal energy prices that they experience compensates for the investment costs and any changes in transmission charges. Providing owners of (physical) transmission rights with extra benefits, in addition to the collection of congestion rents, should also be considered.

Regarding the allocation of grid costs in the context of regional markets, in many of them transmission charges are computed by the local authorities in each country or state (the IEM, the market in Central America, some regional markets in the US, etc.).

Local charges are modified according to the net compensation that is faced by each country, resulting from the application of some allocation method at regional level. This allocation method must be efficient. However, if, despite using an efficient method at regional level, local charges within each country are not computed according to a sound criterion (for example, uniform unit transmission charges are applied), most of the locational content of transmission tariffs will be lost.

Transmission lines whose construction has been decided by the countries that are expected to benefit from them should be paid by these countries. In principle, the whole cost of these lines should be allocated in proportion to the electrical usage that each of these countries makes of them. For regional transmission lines whose construction is unilaterally decided by the country where the line is located, other countries should only contribute to the recovery of the cost of these lines to the extent that the latter countries benefit from it. Then, assuming that electrical usage is a suitable proxy to economic benefits, each country, other than that where the line is located, should only pay the cost of the fraction of the line that it uses. However, if building the latter line is efficient from a regional point of view, despite the fact that the line is only partially used, the total cost of this line should be allocated among the countries that use it. Thus, lines beneficial for the region should be paid by the countries that use them, regardless of the identity of the party that proposed their construction.

Therefore, as long as sending locational signals is a priority, harmonizing some basic features of transmission tariffs computed within each country or state is essential in order for transmission tariffs to encourage agents to make efficient investment decisions. Jointly with regulated transmission tariffs computed at regional level, national (local) transmission tariffs must pay the whole cost of transmission lines, both existing and new, that are relevant to cross-border trade. Additionally, regional tariffs must pay the cost of only the used fraction of those lines whose construction has been unilaterally decided by countries or states and are not of special interest to market agents trading their energy regionally. In principle, regulated transmission investments should not be discriminated based on the party that has proposed them (regional institutions versus countries).

When choosing a congestion management mechanism for its implementation in a regional market, one must bear in mind the fact that coordinated implicit auctions are the most efficient solution from a conceptual point of view. However, they require a high level of centralization that may be incompatible with the institutional arrangements and distribution of responsibilities adopted in this region.

As far as it is compatible with the arrangements made in each region, we recommend implementing a regional implicit auction to allocate energy and transmission capacity in the short term market. If this is not possible because of the high level of centralization that the use of implicit auctions involves, we recommend implementing a coordinated explicit capacity auction at regional level followed by local implicit auctions, which must be security constrained if some grid restrictions within an area are binding. If grid constraints exist within an area but local authorities refuse to use implicit auctions, these auctions should be replaced by other market

based mechanisms like explicit auctions where both offers to a pool and bilateral transactions should participate.

Zonal congestion methods may be superior to nodal ones in that the former may face less obstacles for their implementation due to the fact that they provide some degree of independence to the different areas in a region. However, a central problem to the implementation of any zonal congestion management method is the definition of areas of uniform prices or Single Price Areas (SPAs). Failing to define SPAs would mean that only nodal pricing techniques are compatible with the efficient operation of regional power systems. However, nodal pricing techniques are far from what seems achievable in most systems in the short and medium term.

We have conducted a preliminary study for the identification of SPAs in the Iberian peninsula and the system France-Spain-Portugal. A novel method is proposed to define these areas based on the application of clustering techniques combined with some expert knowledge about the functioning of the system. The definition of accurate enough SPAs in these regions requires further research. Some future lines of research have been outlined.

Zonal congestion management methods can only be implemented if accurate enough SPAs exist. If accurate enough SPAs cannot be identified, those inter and intra-area transactions whose impact on the flows of congested lines are significantly different from those of typical inter or intra-area transactions must be singled out and dealt with separately. When dividing a regional transmission system into SPAs, the SOs must choose the division so that the loss of efficiency in the dispatch with respect to the nodal optimal one is minimum.

Now, we provide some recommendations for the regulation of the transmission service in the IEM of the European Union, the Central American Market and the regional markets in the US.

5.1.2.1 The Internal Electricity Market of the European Union

Congestion management

Implementing a centralized implicit auction at European level seems far from what is achievable in the medium to short term. However, there are some sub-regions, like the Iberian Peninsula, the north-western part of the continental plate or the Nordic countries, that have achieved a high level of integration. Organizing implicit auctions of energy and capacity within these sub-regions may be a realistic objective. A coordinated explicit auction taking place at European level could be used to allocate the interconnection capacity between sub-regions. Market and institutional arrangements within these areas should gradually adapt to a common design in order to achieve, in the long term, the complete integration of all these sub-systems.

Financial transmission rights should be available in the long term for agents to shield from the volatility in transmission prices. Point-to-point obligations are the easiest to

implement though options and flow-gate rights should also be offered if requested by market participants.

Investments

Contrary to what happens in other markets, no single institution is in charge of regulating the expansion of the regional grid. Therefore, a relatively complex system, based on the cooperation among countries, must be put in place so that those transmission lines that are relevant to the trade in the region are built. Regional investment projects must be proposed by one or more national TSOs. Afterwards, either the investment is unilaterally approved and paid by the TSO that has proposed it, or TSOs affected by its construction negotiate the approval of the project and compute the allocation of its cost (with the help of a standard cost allocation method). Lastly, regulatory authorities should give their approval to the project after having checked that it is not detrimental to the region.

The inter-TSO payment scheme that is in place in the region automatically allocates the cost of any new network reinforcement. Usage based methods, like AP, are best suited to compute these compensations. If the country where a new reinforcement will be located considers that the allocation of the cost of this line resulting from the inter-TSO compensation mechanism is unfair, a predefined allocation procedure that is based on the estimation of the economic benefits produced by the line may have to be applied. Apart from that, merchant investors may find it attractive to build some inter-connectors among countries. For instance, some submarine cables in the Nordic sub-region have been constructed as merchant projects.

The procedure used to collect the regulated revenue of regional investments must reduce, to the minimum extent possible, the risk for the promoters of these lines of not recovering their investment costs plus a reasonable rate of return. Thus, the construction, operation and maintenance of regulated regional lines should be auctioned and paid as bid to the winner party.

Regarding the development of national grids, the paradigms implemented to regulate the expansion of the grid in the different countries differ widely. Consequently, taking some regulatory measures may be necessary in order to prevent, to the extent possible, overinvestment by passive TSOs and underinvestment by the active ones.

Transmission tariffs

Instead of implementing a system of pan-European transmission tariffs, parties in the region have agreed to establish a system of compensations among countries for the use that each country makes of the grid of the others. Authorities in each country must allocate the net compensation corresponding to the country to the local generators and load, thus modifying national transmission tariffs. Payment of the corresponding national transmission tariff grants access to the whole European grid. Countries have authority over the design of national tariffs. However, they must harmonize the basic

structure of the tariffs as well as the split of transmission charges between generation and load. Otherwise, competition in the region would be distorted.

An efficient method must be used to compute inter-TSO compensations (for example the AP method or the Aumann-Shapley method). In addition to contributing to the partial recovery of the cost of national grids, inter-TSO compensation methods implicitly allocate among the different countries the cost of any new network reinforcement. Therefore, choosing an efficient method is essential to achieving an adequate development of the regional grid. If countries perceive the method as an unfair one, they may not agree to pay their share of the cost of new reinforcements and some of these reinforcements will not get built. If sending locational signals is necessary, as it seems to be, transmission charges resulting from the AP method may be modified in order to strengthen the relation between the grid costs the system incurs, as a consequence of the decision of an agent to install or to retire a power plant or load, and the transmission charges this agent has to pay (both for the grid costs this agent causes in its country and for those caused in other countries).

5.1.2.2 The regional market in Central America (MER)

Congestion management

Achieving a high level of integration is possible in this region since local institutions and arrangements are not deeply rooted. However, local markets in some countries have not been developed yet. Load and generation patterns in the latter countries are determined through a traditional minimum cost dispatch.

In the short term, an optimal economic dispatch allocates the transmission capacity of the regional grid to those transactions that place the highest value on it. Thus, regionally, energy and transmission capacity are dealt with in an integrated manner. This regional implicit auction results in a different energy price for every node of the regional grid. The congestion management scheme employed in the region is, theoretically speaking, the most advanced one. Therefore, no modifications to it are necessary.

Given the low level of development of some national markets, implementing locally the same congestion management scheme that is used at regional level is not possible. However, local authorities must work to create conditions within their national markets that are consistent with the regional one. In any case, discrimination between local and regional transactions must be avoided. Thus, regardless of whether energy prices within each country present some form of locational differentiation, or a single price is computed for all the power injections and withdrawals resulting from the national market, all the energy consumed or produced at a node must be valued at the same price, without taking into account the market where this energy is traded.

A system of long-term transmission rights complements the short-term regional economic dispatch. Financial rights exist that are aimed at protecting agents from

volatile transmission prices. However, those generators signing a long-term firm supply contract must acquire physical rights over the transmission capacity they would use if they had to effectively deliver the contracted energy. Otherwise, firm supply contracts would not be valid. These contracts allow generators to earn the capacity payment of the country where their counterpart in the contract (the load) is located. Generators in a mainly exporting area may probably build new lines to get the corresponding physical rights and earn the capacity payments that are in place in other areas. Thus, long-term firm supply contracts efficiently promote the construction of new lines of interest for the region.

Investments

The EOR is in charge of the development of the regional grid (RTR). An adequate system is in place to determine whether a line should be considered to be part of the RTR or not. Periodically, the EOR must produce indicative expansion plans for this grid, which have to be approved by the regional regulatory body (the CRIE). Countries in the region cannot oppose the construction of a certain line that has been approved by the CRIE. Thus, centralized transmission planning plays a very important role in the construction of regional grid reinforcements. Merchant investors may also promote the construction of new lines to appropriate the corresponding congestion rents resulting from the regional dispatch. Besides, as we have just mentioned, generators willing to get the capacity payments implemented in other areas may construct and pay new lines in order to get hold of the physical rights that are required to access these areas. All these arrangements make sense, as long as central planning is given priority over investments by private entrepreneurs. The latter type of investments are necessary in order to build those lines that are beneficial and the central planner fails to build.

Within each country, national authorities must work to achieve an adequate development of the national grid. Again, central planning should be given priority since the private initiative is unlikely to promote the construction of a significant number of local reinforcements.

Transmission tariffs

As we have already mentioned, the remuneration of the transmission activity depends on the adopted paradigm. If regulated investments and investments at risk coexist, as it is the case in Central America, the revenues of each line owner must be dependent on the considered type of line.

Application of nodal prices resulting from the regional market results in some revenues for the owners of regional lines. These are the so-called congestion rents. Owners of merchant lines must only earn the corresponding congestion rents but regulation must not place a limit on their revenues. If merchant investors want to secure some amount of revenues they may sell long-term rights over the capacity of their lines. Rents corresponding to a regulated line may contribute to the recovery of the cost of the existing grid or they may finance the construction of new regulated

reinforcements. Regardless of whether a regulated line belonging to the RTR has been promoted by the EOR or by the planner in a country, its owner must recover the amount he has invested, and earn a reasonable return on it. However, his revenues should not exceed a reasonable limit. The cost of the unused fraction of regulated lines must be socialized to demand.

The usage-based method that is applied in the region to compute the complementary charge corresponding to regulated lines is a combination of the Average Participations method (AP) and the Dominant Flow Method (MFD), see the description of this method in section 3.3.3.4 of chapter 3 of the thesis and (EOR 2005). The MFD method decomposes the line flows into two parts: one associated with the corresponding national super-transaction and the other one associated with the regional one. A national super-transaction is comprised of all the power injections and withdrawals that take place within the corresponding national (local market). The regional super-transaction consists of the injections and withdrawals scheduled in the regional market. The AP method is used to trace the fraction of the flow over each line that is associated with the regional super-transaction. Therefore, AP determines how much each agent trading its energy regionally uses the RTR. Lastly, the participations of regional power injections and withdrawals in the use of the RTR are aggregated at country level, thus obtaining the use that regional injections and withdrawals within each country make of the grids of the different countries. Consequently, the results produced by this method are dependent on the commercial transactions taking place, which is not a desirable feature. Besides, it discriminates between power injections and withdrawals that are scheduled in national markets and those scheduled in the regional market. We recommend employing a method that agrees with the market philosophy in the region (one regional market superimposed on the national ones) but which does not discriminate between injections of one type or another. One possibility is resorting to the AP method, tracing separately the power injections and withdrawals scheduled in each market, but using always the set of line flows that actually exist to trace these injections and withdrawals. In any case, a single method should be applied to compute the complementary charge of both regional and local lines. This would contribute to achieving a level playing field in the region.

5.1.2.3 Development of regional markets in the US

Congestion management

Federal authorities proposed the adoption of a common congestion management system by regional markets across the country. This scheme is based on the computation of nodal short term marginal prices, which are optimal short term signals. In addition, market agents would be able to exchange long-term financial transmission rights to manage the risk related to the volatility of congestion costs. Conceptually speaking, this scheme is very efficient. However, its implementation may not be possible in those regions where market structures are less mature or they do not exist at all. What is more, this scheme requires centralizing the dispatch of energy in the

region, which may be against the market and institutional arrangements that are in place in some parts of the country.

In our view, market authorities in a region should choose between applying the SMD scheme, backed by the FERC, and making use of a less centralized, less intrusive market based solution, like zonal pricing or bilateral trading. Energy trade that is taking place between different regions might be managed using a coordinated explicit auction of a national scope, whereby the inter-regional transmission capacity would be auctioned. The utilization of long-term financial transmission rights (FTRs) is far less problematic than that of nodal prices, since the former do not condition the short-term physical energy dispatch and they are compatible with the existence of any short-term congestion management mechanism. Hence, FTRs should become a standard across the country to manage congestion in the long term.

Investments

Regional transmission lines should be identified as those that are significantly used by regional transactions. RTOs and SOs in each region must promote the development of regional grids while FERC should be responsible for deciding over the construction of lines that cross at least one border between states and local authorities should decide over the construction of the rest of regional lines. The decisions by FERC regarding regional cross-border investment projects should not be contestable by the states. In order for SOS and RTOs to look after the interest of the whole region, ownership of the transmission activity should be unbundled from that of other activities. The development of local, intra-area, grids shall be promoted by local SOs and approved by local authorities.

The cost of the new regional reinforcements must be paid by market agents in the different areas of the region (or the different regions of the country) according to the results of a cost-benefit analysis. This analysis should be carried out by the regional system operator and the resulting transmission rates should be approved by FERC. As explained for the European system, if electrical usage is accepted as a proxy to the economic benefits produced by a line, and transmission tariffs in the different areas within a region (and the different interconnected regions) are computed using a common usage-based method, this method will tell us how to allocate the cost of the whole regional grid, including any new lines, among grid users. Therefore, this same method will allocate the cost of regional reinforcements among the different areas and regions. Authorities in a state may oppose the construction of a regional line that is internal to the state if they feel the share they would have to pay of the cost of this reinforcement is not in line with the benefit they would obtain from it. Forcing authorities in a state to approve the construction of a regional line located within the state will probably not be possible. Therefore, a system of side payments between states may help get this line built. Once a line is built, FERC has authority to approve or not regional transmission tariffs aimed at recovering the cost of this line.

The construction of an approved reinforcement could be assigned through a competitive bidding process whose winner would be paid the amount he requested when making the offer. This mechanism is thought to be superior to the computation of standard unit construction costs based on the utilization of benchmarks.

Participation of private entrepreneurs in the expansion of the grid is possible, though centralized planning should take a predominant role. Experience in the country suggests that investments by private entrepreneurs are unlikely to significantly contribute to the optimal expansion of the grid. Thus, RTOs, whenever they exist, as well as SOs should conduct analysis in order to identify the most interesting grid upgrades. If authorities want to leave the private initiative a more predominant role in the expansion of the grid, they could give private promoters the opportunity to undertake the construction of the previously spotted most interesting reinforcements. Associations of grid users may propose and construct some very specific lines (mainly local retailers representing their customers) but their contribution is likely to be minor.

Implementing efficient schemes to price the energy traded within each area and region may be necessary to signal the advisability of installing new lines. Computing marginal nodal or zonal energy prices appears to be the best option in this regard.

Transmission tariffs

As it is proposed by FERC in the SMD, transmission tariffs computed within a region should be unique, i.e. the same method must be used to compute the transmission tariffs to be paid by all grid users in the region. What is more, in order not to distort competition between regions and provide useful long-term locational signals, the tariff setting process should be harmonized across the country. If this is not possible, authorities in all the regions and areas should, at least, agree to harmonize some of the basic features of the tarification system, such as the global split of transmission charges between generation and load.

Contrary to what is proposed by FERC in its SMD, both generators and loads should pay transmission charges in proportion to their responsibility in these costs. Connection charges would be employed to recover the cost of those lines that are necessary to physically connect a new generator or load. Apart from them, both generators and loads should pay access charges. Access charges should be computed based on the electrical usage that each agent makes of the grid in the region. The method of Average Participations should be applied for this. If applying nodal access charges is not possible, zonal G and L unit charges might be computed by averaging the charges obtained for all the agents of the same type (generators or loads with the same profile) located within the same area. Obviously, several price areas should be defined a priori. If the regional grid were meshed enough and no major grid investments were necessary, postage stamp tariffs might be the best option and sending locational signals would not be a priority. However, this is not the situation of regional markets in the US. The necessity to apply locationally differentiated

transmission charges, both to generators and load, is a consequence of the vast amount of new transmission capacity required because of the change in the trade patterns within the country. The decision by market agents on the location of new generation plants or load may probably have a significant impact on the grid costs. Thus, charging agents according to the grid costs they cause will make them internalize these costs. Charging the load only is clearly a sub-optimal solution.

Apart from that, transmission tariffs should reflect the fact that agents in a regional market benefit from the grid in other regions because of the changes that are taking place in trading patterns. Thus, if transmission tariffs in different regions (or states) are computed separately, as it seems likely to occur, these regions (or states) should compensate each other for the use they make of others' grids. A system of compensations for the usage that agents in each regional market make of the grids of other markets is necessary. This inter-regional compensation mechanism would resemble that implemented in the IEM and it should take into account the pattern of cross-border flows between the existing regional markets. Additionally, compensations may be necessary in order to get built some lines that are essential for the regional commerce. Ideally, a system of nodal or zonal transmission tariffs, based on network usage as a proxy to economic benefits, should apply at country level.

5.2 Original contributions

This thesis has analyzed the regulatory framework of the activity of transmission of electricity in a comprehensive way. Based on both the existing regulatory mechanisms and designs, as well as on the ones that have been developed in the context of this work, we have produced a set of recommendations and rules aimed at achieving an optimal functioning of the transmission service in the context of regional markets.

The major task that this thesis has achieved is to examine the soundest regulatory approaches in transmission regulation and to extend them to the much more complex situation of a regional market. It has been found that it is possible to define some basic regulatory guidelines for transmission that appear to be valid under any kind of institutional and regulatory arrangements in the participating countries. These rules have been described first in the context of a reference or ideal case which, for the sake of simplicity has been defined as a perfectly integrated market.

Then, these ideal rules have been adapted to obtain the most efficient regulatory designs that are compatible with the political and institutional arrangements of three of the most representative regional markets that presently exist in the world: the IEM of the European Union, the regional markets in formation in the US and the MER in Central America. In fact, the author of this thesis and the thesis supervisor have actively participated in the decision making process leading to the development of the regulatory solutions that have been implemented in the European and Central American regional markets. In order to come up with a satisfactory regulatory scheme, we have deeply examined the interactions that exist among the different aspects of transmission regulation, so that the solutions proposed to deal with each one of these aspects do not interfere in any negative form with the ones that we have proposed to deal with the others. In addition to the design of a global regulatory framework for transmission, which is a central contribution of the thesis, we have produced a number of developments and designs that are related to specific aspects of transmission regulation. These specific contributions can be classified into three main categories, according to the concerned regulatory aspects. Some contributions are related to the process of setting transmission tariffs. Others are related to the allocation of transmission capacity among market agents. Lastly, a few of them are associated with the regulation of the expansion of the grid.

5.2.1 Contributions related to the design of transmission tariffs

The author of this thesis, together with the thesis supervisor, has advised the European Commission, the Council of European Energy Regulators and some Transmission System Operators (like the Italian or the Spanish ones) on the selection of the most efficient method to compute compensations among countries (ITCs) within the IEM of

the EU. As a result of this research and consultancy work, we have produced several reports, see (Pérez-Arriaga, Olmos et al. 2002; Pérez-Arriaga and Olmos 2003; 2005). Throughout this thesis work, several methods that are aimed at computing ITCs have been designed and implemented in actual regional markets. These methods include the Average Participations applied to Transits method (APT), the Simplified Average Participations method (SAP) and the Modified Average Participations method (MAP). We have compared qualitatively and quantitatively the proposed ITC methods to others that had been proposed previously. The AP method is one of the methods considered for the computation of inter-TSO compensations. We have deeply analyzed the theoretical justification of the Average Participations method presented by Bialek et al. in (Bialek and Kattuman 2002). The discussion provided in this thesis shows that the aforementioned justification is applicable only after having accepted some critical assumptions. Based on the thorough and systematic analysis that has been carried out, we have concluded that the AP method - among the ones that have been considered- is the preferred one for computing inter-TSO compensations and, whenever needed, nodal transmission tariffs.

5.2.2 Contributions related to the management of congestion

As for the allocation of the potentially scarce transmission capacity among the agents that bid for it, we have conceptually analyzed and compared the most relevant congestion management methods that exist, in order to determine which one should be applied in the context of the IEM of the European Union. When assessing the different methods, we have not only taken into account how efficient the provided solution is, but also their compatibility with the political situation and institutional arrangements in place in the EU. In the short to medium term, we propose the utilization of a coordinated explicit auction at European level, together with implicit auctions taking place within a number of pre-defined sub-regions. We have published an article related to this subject, see (Pérez-Arriaga and Olmos 2005). In addition to this, the thesis has contributed several developments that deal with specific topics of the management of congestion. A list of them follows:

- Design of a method to define Single Price Areas in a system, for congestion management purposes. The initially proposed method has been refined by using the experience of its application to two actual power systems (the Iberian system and the one comprised of Spain, France and Portugal). Further improvement upon the proposed method is a topic of future research.
- Using a novel method based on marginal pricing equations, we have investigated the identification of the most suitable balancing point for the definition of transmission contracts in a large scale system (the European one). The results obtained in the thesis are still provisional but this research line is a promising one. The main findings of this research work can be found in (Olmos and Neuhoff 2006).

- Conceptual analysis of some specific aspects of the design of auctions for transmission capacity. The addressed issues include a discussion of the marginal value of a line in the dispatch, the treatment of some specific reliability restrictions and the equivalence between implicit and explicit auctions under ideal conditions.

5.2.3 Contributions related to the regulation of the expansion of the grid

These contributions mostly refer to conceptual issues in the decision-making process of investing in reinforcements or extensions of transmission grids. The most relevant topics are:

- Theoretical design of a method that is aimed at strengthening the causal relationship that exists between the grid costs that each agent causes the system to incur and the transmission charge it has to pay. This method takes into account the fact that new lines are mainly the responsibility of new generators and load being installed in the system. The method proposed has been applied, in combination with the AP method, to compute transmission tariffs in the Iberian Peninsula.
- Analysis of the interactions that exist among economic signals related to investment and involving different agents at the interfaces between generation, transmission and distribution. The different investment signals that may be received by a typical CCGT power plant have been numerically compared in a realistic case-example. This analysis shows that grid investment signals may be comparable in size to locational differences in energy prices. Therefore, the market agents should also receive economic signals related to their location in the grid. Complete details are provided in another publication that has derived from this thesis, see (Pérez-Arriaga and Olmos 2005).
- A detailed theoretical discussion is provided regarding the regulatory test to be applied in order to decide over the construction of a new transmission line. A different set of rules should be applied in the case of regulated lines or when a merchant transmission investment is being evaluated.

5.3 The way forward

The number of future lines of research that may expand on the topics that have been presented in this thesis work is very large. Here we shall mention some of the most relevant ones but the reader will probably be able to identify many others. Developments are classified into the three main areas of knowledge that have consistently been used throughout this document, namely, investments in new facilities, management of congestion and computation of transmission tariffs.

5.3.1 Allocation of the cost of construction of the grid

In addition to the several methods that have been proposed and analyzed in this thesis, alternative methods could be proposed and compared to the existing ones. None of the methods that have already been considered is completely satisfactory. Therefore, room for improvement exists. In this regard, a research group at Power Systems Research (PSR), located in Brazil, has devised a new implementation strategy for a cost allocation method based on game theory. In this game, each market agent must contribute to the grid costs according to the Aumann-Shapley value obtained for this agent when allocating the cost of developing and maintaining the grid. This implementation strategy is believed to result in a significant decrease in the computational effort (and time) required to obtain the solution of the game in a large scale system.

5.3.2 Management of congestion in the grid

The method proposed for the identification of Single Price Areas should be refined. As outlined in chapter 4, 'special', problematic, transactions should be singled out and dealt with separately. Besides, the single objective of the classification process should be identifying the minimum set of SPAs, such that the net aggregated social benefit, resulting from the zonal system dispatch, is close enough to the result of the optimum nodal dispatch.

The effects that different transmission capacity allocation methods have on the exercise of market power by strategic generators should be further investigated. Some research groups Cambridge University, as well as in other places, have already been exploring this line of research. Market power is present in every power system that has been opened to competition. The potentially devastating consequences of the massive exercise of market power have been well documented. It is widely known that market power cannot completely be eliminated using behavioral measures. In addition to them, some structural measures are necessary. However, mechanisms and rules employed should not aggravate this problem.

Also related to the previous point, additional measures should be developed in order to eliminate the incentive for the exercise of market power that the ownership of a transmission rights portfolio may represent. In particular, the case where strategic generators own power plants in different locations scattered all around the system should be further investigated.

5.3.3 Investments in new transmission facilities

Exploring the possibility of devising a set of regulatory incentives for the System Operator to perform its function efficiently, including the planning of the expansion of the grid, would certainly be useful (these are the so-called performance based incentives). These incentives have proved to be very problematic in the past. An adequate system of incentives could prevent the SO from over-investing, thus avoiding some grid investments that are ill-justified from a strict reliability-economic point of view.

Lastly, one may think of applying to a real-scale system the set of locational signals that has been developed in this thesis. These signals are aimed at market agents considering the installation, or the retirement, of a new power plant or load and relate to the costs that the system incurs as a consequence of their decision to install a new power plant in one part of the grid or another. Locational signals should include: the cost of building transmission lines (the causal relationship between responsibility in grid costs and charges should be strengthen), transmission losses caused by each agent, locational differences among the value of energy in different parts of the system and the cost of the infrastructures required to transport the fuel that a generation facility uses.

Bibliography

- Bialek, J. W. and P. A. Kattuman (2002). Proportional Sharing Assumption in Tracing Methodology.
- EOR (2005). Reglamento del Mercado Eléctrico Regional (RMER). Libro III del RMER. De la Transmisión., Ente Operador Regional (EOR).
- FSR (2005). A study on the inter-TSO compensation mechanism, Florence School of Regulation. Web page: <http://www.iue.it/RSCAS/ProfessionalDevelopment/FSR/>: 96.
- Olmos, L. and K. Neuhoﬀ (2006). "Identifying a Balancing Point for Electricity Transmission Contracts." IEEE Transactions on Power Systems **Vol. 21**(No. 1): 91-98.
- Pérez-Arriaga, I. J. and L. Olmos (2005). "A plausible congestion management scheme for the internal electricity market of the European Union." Utilities Policy **Vol. 13**: 117-134.
- Pérez-Arriaga, J. I. and L. Olmos (2003). Extension of the project on cost components of cross border exchanges of electricity, Directorate-General for Energy and Transport / European Commission. Web page: http://europa.eu.int/comm/energy/electricity/publications/index_en.htm.
- Pérez-Arriaga, J. I. and L. Olmos (2005). Compatibility of Investment Signals in Distribution, Transmission and Generation. Electricity Infrastructure Investment Workshop. To be published as a chapter in the book 'Investments in Generation and Transmission in Electricity Liberalized Regimes', CRE, Paris, France.
- Pérez-Arriaga, J. I., L. Olmos, et al. (2002). Cost components of cross border exchanges of electricity, Directorate-General for Energy and Transport / European Commission. Web page: http://europa.eu.int/comm/energy/electricity/publications/index_en.htm.
- Rubio, F. J. (1999). Metodología de Asignación de Costes de la Red de Transporte en un Contexto de Regulación Abierta a la Competencia. Doctoral Thesis. Escuela Técnica Superior de Ingeniería. Madrid (in Spanish), Universidad Pontificia Comillas.

General Bibliography

- Abraham, S. (2002). "National Transmission Grid Study". Washington DC, US Department of Energy. Web page: <http://www.energy.gov/engine/content.do:108>.
- ACCC (1999). "Regulatory Test for New Interconnectors and Network Augmentations", Australian Competition and Consumer Commission, <http://www.accc.gov.au/content/index.phtml/itemId/54367/fromItemId/54368>.
- ACCC (2004). "Review of the Regulatory Test for Network Augmentations", Australian Competition and Consumer Commission (ACCC). Web page: <http://www.accc.gov.au/content/index.phtml/itemId/532002/fromItemId/54368>.
- ACCC Australian Competition and Consumer Commission (ACCC). Web page: <http://www.accc.gov.au/content/index.phtml/itemId/142>.
- Aguado, J. A., V. H. Quintana, et al. (2004). "Coordinated Spot Market for Congestion Management of Inter-Regional Electricity Markets." IEEE Transactions on Power Systems 19(1): 180-187.
- Agurto, R. (1992). Cargos por uso de la transmisión en Chile: Análisis del concepto de áreas de influencia. 6º Congreso de América Latina y el Caribe sobre Economía de la energía eléctrica, Mendoza, Argentina.
- Alguacil, N., A. L. Motto, et al. (2003). "Transmission Expansion Planning: A Mixed-Integer LP Approach." IEEE Transactions on Power Systems 18(3): 1070-1077.
- Alomoush, M. I. and S. M. Shahidehpour (2000). "Contingency-constrained congestion management with a minimum number of adjustments in preferred schedules." Electrical Power and Energy Systems 22: 277-290.
- Anderson, K. P. and A. McCarthy (1999). "Transmission pricing and expansion methodology: lessons from Argentina." Utilities Policy Vol. 8(No. 4): 199-211.
- Bakirtzis, A., P. Biskas, et al. (2001). "Comparison of two methods for long-run marginal cost-based transmission use-of-system pricing." IEE Proceedings on Generation, Transmission and Distribution 148(4): 477-481.
- Baldick, R. (2003). "Variation of Distribution Factors with Loading." IEEE Transactions on Power Systems 18(4): 1316-1323.
- Baldick, R. and E. Kahn (1993). "Network Costs and the Regulation of Wholesale Competition in Electric Power." Journal of Regulatory Economics 5: 367-384.
- Baran, M. E., V. Babunarayanan, et al. (2000). "Equitable allocation of congestion relief Cost to Transactions." IEEE Transactions on Power Systems 15(2): 579-585.
- Baran, M. E., V. Banunarayanan, et al. (1999). "A Transaction Assessment Method for Allocation of Transmission Services." IEEE Transactions on Power Systems Vol. 14(No. 3): 920-927.
- Bartholomew, E. S., A. S. Siddiqui, et al. (2003). "The New York Transmission Congestion Contract Market: Is it Truly Working Efficiently?" The Electricity Journal: 14-24.
- Baughman, M. L. (1997). "Pricing of open-access transmission services in Texas." Utilities Policy 6(3): 195-201.

- Bautista, G. and V. H. Quintana (2005). "Screening and Mitigation of Exacerbated Market Power Due to Financial Transmission Rights." *IEEE Transactions on Power Systems* 20(1): 213-222.
- Bertrand, J. (1883). "Theorie Mathematique de la Richesse Sociale." *Journal des Savants*: 499-508.
- Bialek, J. (1996). "Tracing the flow of electricity." *IEE Proceedings on Generation, Transmission and Distribution* 143(4): 313-320.
- Bialek, J. (1997). "Topological Generation and Load Distribution Factors for Supplement Charge Allocation in Transmission Open Access." *IEEE Transactions on Power Systems* Vol. 12(No. 3): 1185-1190.
- Bialek, J. W. and P. A. Kattuman (2002). "Proportional Sharing Assumption in Tracing Methodology".
- Bialek, J. W. and P. A. Kattuman (2004). "Proportional sharing assumption in tracing methodology." *IEE Proceedings on Generation, Transmission and Distribution* Vol. 151(No. 4): 526-532.
- Bialek, J. W. and S. Ziemianek (2003). "Tracing based transmission pricing of cross-border trades: fundamentals and circular flows". *IEEE Power Tech Conference*, Bologna, Italy.
- Binato, S. and G. C. Oliveira (1995). "A heuristic procedure to cope with multi-year transmission expansion planning". *IEEE/KTH Stockholm Power Tech Conference*, Stockholm, Sweden.
- Binato, S., M. V. F. Pereira, et al. (2001). "A New Benders Decomposition Approach to Solve Power Transmission Network Design Problems." *IEEE Transactions on Power Systems* 16(2): 235-240.
- Biskas, P. N. and A. G. Bakirtzis (2002). "Decentralised Congestion Management of Interconnected Power Systems." *IEE Proceedings*: 432-438.
- Bjorndal, E., G. C. Stamtzis, et al. (2005). "Finding core solutions for power system fixed cost allocation." *IEE Proceedings on Generation, Transmission and Distribution* Vol. 152(No. 2): 173-179.
- Björndal, M. and K. Jörnsten (2001). "Zonal Pricing in a Deregulated Electricity Market". *The Energy Journal* 22(1): 51-73.
- Bompard, E., P. Correia, et al. (2003). "Congestion-Management Schemes: A Comparative Analysis Under a Unified Framework". *IEEE Transactions on Power Systems* 18(1): 346-452.
- Borenstein, S., J. Bushnell, et al. (2000). "The Competitive Effects of Transmission Capacity in a Deregulated Electricity Industry". *Rand Journal of Economics* 31(2): 294-325.
- Boucher, J. and Y. Smeers (2001). "Towards a common European Electricity Market-Paths in the right direction... still far from an effective design". *Harvard Electricity Policy Group*, Web page: http://www.ksg.harvard.edu/hepg/Standard_Mkt_dsgn/Smeers_Interconnections1_4jni_3.do1.pdf.
- Brown, R. E. (2002). "Electric Power Distribution Reliability". New York, Marcel Dekker, Inc.
- Brunekreeft, G. (2003). "Market-Based Investment in Electricity Transmission Networks: Controllable Flow". *CMI electricity project working paper*. Web page: <http://www.econ.cam.ac.uk/electricity/publications/wp/index.htm>.
- Bushnell, J. and S. Oren (1997). "Transmission pricing in California's proposed electricity market." *Utilities Policy* 6(3): 237-244.
- Bushnell, J. and S. Stoft (1996). "Grid Investment: Can a Market do the Job?" *The Electricity Journal* 9(1): 74-79.

- Bushnell, J. B. and S. E. Stoft (1996). "Electric Grid Investment Under a Contract Network Regime." *Journal of Regulatory Economics* 10: 61-79.
- Cardell, J., C. C. Hitt, et al. (1997). "Market Power and Strategic Interaction in Electricity Networks." *Resource and Energy Economics* 19(1-2): 109-137.
- Carrillo-Caicedo, G. and I. Pérez-Arriaga (1995). Optimal reconfiguration of distribution networks for a diversity of regulatory frameworks. IEEE/KTH Stockholm Power Tech Conference, Stockholm, Sweden.
- CEER (2003). CEER Guidelines on Regulatory Control and Financial Reward for Infrastructure, Council of European Energy Regulators (CEER). Web page: <http://www.ceer-eu.org/>.
- CER (2004). Electricity Tariff Structure Review: International Comparisons, Commission for Energy Regulation (CER). Web page: <http://www.cer.ie/cerdocs/cer04101.pdf>.
- Chang, Y.-C. and C.-N. Lu (2001). "An electricity tracing method with application to power loss allocation." *International Journal of Electrical Power and Energy Systems* Vol. 23(No. 1): 13-17.
- Chao, H.-P. and S. Peck (1996). "A Market Mechanism for Electric Power Transmission." *Journal Of Regulatory Economics* 10: 25-29.
- Chao, H.-P. and S. Peck (1996). "A Market Mechanism for Electric Power Transmission." *Journal Of Regulatory Economics* 10: 25-29.
- Chao, H.-p. and S. Peck (2000). "Flow-based Transmission Rights and Congestion Management." *Electricity Journal* 13(8): 38-59.
- Chen, L., H. Suzuki, et al. (2002). "Components of Nodal Prices for Electric Power Systems." *IEEE Transactions on Power Systems* 17(1): 41-49.
- Chicco, G., G. Gross, et al. (2002). "Allocation of the Reactive Power Support Requirements in Multitransection Networks." *IEEE Transactions on Power Systems* Vol. 17(No. 4).
- Chisari, O. O., P. Dal-Bó, et al. (2001). "High-tension electricity network expansions in Argentina: decision mechanisms and willingness-to-pay revelation." *Energy Economics* 23: 697-715.
- Choi, J., A. R. A. El-Keib, et al. (2005). "A Fuzzy Branch and Bound-Based Transmission System Expansion Planning for the Highest Satisfaction Level of the Decision Maker." *IEEE Transactions on Power Systems* 20(1): 476-484.
- Christie, R. D. and I. Wangensteen (1998). "The Energy Market in Norway and Sweden: Congestion Management." *IEEE Power Engineering Review*: 61-63.
- Coase, R. H. (1990). "The firm, the market and the law", The University of Chicago Press.
- Consentec/Frontier (2004). "Analysis of Cross-Border Congestion Management Methods for the EU Internal Electricity Market", Consentec and Frontier Economics. Prepared for the European Commission. Directorate-General Energy and Transport. Web page: http://europa.eu.int/comm/energy/electricity/publications/index_en.htm.
- Consentec/Frontier (2005). "Study on the further issues relating to the inter-TSO compensation mechanism, European Commission". Directorate-General Energy and Transport. Web page: http://europa.eu.int/comm/energy/index_en.html. Report prepared by Consentec and Frontier Economics.: 72.
- Consultants, C. E. and H. B. Putnam (1995). "Transmission Services Costing Framework. Volume 2: Framework Description and Application", EPRI. Ref. EPRI TR-105121-V2.
- Contreras, J. and F. F. Wu (1999). "Coalition Formation in Transmission Expansion Planning." *IEEE Transactions on Power Systems* 14(3): 1144-1152.

- Contreras, J. and F. F. Wu (2000). "A Kernel-Oriented Algorithm for Transmission Expansion Planning." IEEE Transactions on Power Systems 15(4): 1434-1440.
- Cournot, A. (1838). "Reserches sur les Principes Mathematics de la Theorie des Richesses".
- CRIE (2002). "Reglamento Transitorio del Mercado Eléctrico Regional, Comisión Regional de la Interconexión Eléctrica", Web page: <http://www.omca.net/documentos/Rtmer01.pdf>.
- Daniel, J. S., R. S. Salgado, et al. (2005). "Transmission loss allocation through a modified Ybus." IEE Proceedings on Generation, Transmission and Distribution Vol. 152(No. 2): 208-214.
- Danitz, F., H. Rudnick, et al. (2002). "Use Based Allocation Methods for Payment of Electricity Transmission Systems". PowerCon 2002 IEEE-PES/CSEE International Conference on Power System Technology, Kunming, China.
- Deladreue, S., F. Brouaye, et al. (2003). "Using Two Multivariate Methods for Line Congestion Study in Transmission Systems Under Uncertainty." IEEE Transactions on Power Systems 18(1): 353-358.
- Dismukes, D. E., R. F. C. III, et al. (1998). "Capacity and economies of scale in electric power transmission." Utilities Policy 7(3): 155-162.
- EC (2004). "Medium Term Vision for the Internal Electricity Market". DG Energy and Transport Working Paper. Strategy paper. Brussels, DGTREN. European Commission.
- EC (2005). "Annual Report on the Implementation of the Gas and Electricity Internal Market". Brussels, European Commission (EC): 10.
- EOR (2005). "Reglamento del Mercado Eléctrico Regional (RMER). Libro III del RMER. De la Transmisión"., Ente Operador Regional (EOR).
- EP (2003). "Directive 2003/54/EC of the European Parliament and of the Council concerning common rules for the internal market in electricity and repealing Directive 96/92/EC".
- EP (2003). "Regulation (EC) No 1228/2003 of the European Parliament and of the council of 26 June 2003 on conditions for access to the network for cross-border exchanges in electricity". Web page: http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/L_176/L_17620030715en00010010.pdf: 10.
- ERGEG (2005). "The Creation of Regional Electricity Markets". Web page: http://www.ergreg.org/portal/page/portal/ERGEG_HOME/ERGEG_DOCS/ERGEG_DOCUMENTS_NEW/ELECTRICITY_FOCUS_GROUP/:123.
- "Escobar, A. H., R. A. Gallego, et al. (2004). "Multistage and Coordinated Planning of the Expansion of Transmission Systems." IEEE Transactions on Power Systems 19(2): 735-744.
- ETSO (1999). "Evaluation of congestion management methods for cross-border transmission", European Transmission System Operators (ETSO). Web page: <http://www.ets-net.org/>.
- ETSO (2001). "Coordinated Auctioning: A market-based method for transmission capacity allocation in meshed networks", Association of European Transmission System Operators (ETSO). Web Address: <http://www.ets-net.org/upload/documents/Coordinated%20Auctioning.pdf>: 22.
- ETSO (2001). "Key Concepts and Definitions for Transmission Access Products", European Transmission System Operators (ETSO). Web page: <http://www.ets-net.org/>.

- ETSO (2002). "Reconciliation of market splitting with coordinated auction concepts", European Transmission System Operators (ETSO). <http://www.etsonet.org/>.
- ETSO (2004). "An Overview of Current Cross-border Congestion Management Methods in Europe", European association of Transmission System Operators (ETSO). Web address: http://www.etsonet.org/upload/documents/Current_CM_methods_final_20040908.pdf: 31.
- ETSO/EuroPEX (2004). "Flow-based Market Coupling: A Joint ETSO-EuroPEX Proposal for Cross-Border Congestion Management and Integration of Electricity Markets in Europe", Association of European Transmission System Operators (ETSO) and association of European Power Exchanges (EuroPEX). Web address: http://www.etsonet.org/upload/documents/ETSO-EuroPEX_Interimreport_Sept-2004-.pdf: 26.
- EuroPEX (2003). "Using implicit auctions to manage cross-border congestion: Decentralized Market Coupling". Rome, European association of Power EXchanges (EuroPEX). 10th Meeting of the European Electricity Regulatory Forum. Web page: <http://www.europex.org/pages/943372/index.htm>.
- Evans, F., J. M. Zolezzi, et al. (2003). "Cost Assignment Model for Electrical Transmission System Expansion: An Approach Through the Kernel Theory." IEEE Transactions on Power Systems. Vol. 18(No. 2): 625-632.
- Fang, R. and D. J. Hill (2003). "A New Strategy for Transmission Expansion in Competitive Electricity Markets." IEEE Transactions on Power Systems 18(1): 374-380.
- Fang, R. S. and A. K. David (1999). "Optimal Dispatch under Transmission Contracts." IEEE Transactions on Power Systems 14(2): 732-737.
- Fang, W. L. and H. W. Ngan (2002). "Succinct method for allocation of network losses." IEE Proceedings on Generation, Transmission and Distribution Vol. 149(No. 2): 171-174.
- Faria, H., S. Binato, et al. (2005). "Power Transmission Network Design by Greedy Randomized Adaptive Path Relinking." IEEE Transactions on Power Systems 20(1): 43-49.
- Farmer, E. D., B. J. Cory, et al. (1995). "Optimal Pricing of Transmission and Distribution Services in Electricity Supply." IEE Proceedings on Generation, Transmission and Distribution 142(1): 1-8.
- Farmer, E. D., B. L. P. P. Perera, et al. (1995). "Optimal Pricing of Transmission Services: Application to Large Power Systems." IEE Proceedings on Generation, Transmission and Distribution 142(3): 263-268.
- FERC (2002). "Notice of Proposed Rulemaking, Remedying Undue Discrimination through Open Access Transmission service and Standard Electricity Market Design". Washington, Federal Energy Regulatory Commission. Web page: <http://www.ferc.gov/industries/electric/indus-act/smd/nopr.asp>.
- FERC (2003). "White Paper: Wholesale Power Market Platform", Federal Energy Regulatory Commission (FERC). Web page: <http://elibrary.ferc.gov/idmws/nvcommon/NVViewer.asp?Doc=9686957:0>: 32.
- Fernández, J. L. and M. Remacha (2005). "Comparison on transmission pricing in Europe: Synthesis 2004", Tariffs Task Force, European Association of Transmission System Operators (ETSO): 10.
- Fradi, A., S. Brignone, et al. (2001). "Calculation of Energy Transaction Allocation Factors." IEEE Transactions on power Systems Vol. 16(No. 2): 266-72.

- FSR (2005). "A study on the inter-TSO compensation mechanism, Florence School of Regulation". Web page: <http://www.iue.it/RSCAS/ProfessionalDevelopment/FSR/>: 96.
- Galiana, F. D. and M. Ilic (1998). "A Mathematical Framework for the Analysis and management of power transactions under Open Access." *IEEE Transactions on Power Systems* 13(2): 681-687.
- Galiana, F. D., A. J. Conejo, et al. (2003). "Transmission Network Cost Allocation Based on Equivalent Bilateral Exchanges." *IEEE Transactions on power Systems* Vol. 18(No. 4): 1425-1431.
- Gan, D. and D. V. Bourcier (2002). "Locational Market Power Screening and Congestion Management: Experience and Suggestions." *IEEE Transactions on Power Systems* 17(1): 180-185.
- Gans, J. S. (2001). "Regulating Private Infrastructure Investment: Optimal Pricing for Access to Essential Facilities." *Journal of Regulatory Economics* 20(2): 167-189.
- Gilbert, R., K. Neuhoff, et al. (2004). "Allocating Transmission to Mitigate Market Power in Electricity Networks." *Rand Journal of Economics* 35(4): 691-709.
- Glachant, J.-M. and V. Pignon (2002). "Nordic electricity congestion's arrangements as a model for Europe: Physical constraints or operators' opportunism?" *ISNIE International Conference*, MIT, Boston.
- Green, R. (1997). "Electricity transmission pricing: an international comparison." *Utilities Policy* 6(3): 177-184.
- Green, R. (1997). "Transmission Pricing in England and Wales." *Utilities Policy* 6(3): 185-193.
- Grijalba, S., P. W. Sauer, et al. (2003). "Enhancement of Linear ATC Calculations by the incorporation of reactive Power Flows." *IEEE Transactions on power Systems* 18(2): 619-624.
- Haffner, S., A. Monticelli, et al. (2001). "Specialised branch-and-bound algorithm for transmission network expansion planning." *IEE Proceedings on Generation, Transmission and Distribution* 148(5): 482-488.
- Hamoud, G. (2000). "Assessment of Available Transfer Capability of Transmission Systems." *IEEE Transactions on Power Systems* 15(1): 27-32.
- Happ, H. H. (1994). "Cost of wheeling methodologies." *IEEE Transactions on Power Systems* Vol. 9(No. 1): 147-156.
- Hartigan, J. (1975). "Clustering Algorithms". New York, John Wiley & Sons.
- Helm, D. (2003). "Auctions and Energy Networks." *Utilities Policy* 11(1): 21-25.
- Henney, A. (2002). "What the US Could Learn from Western Europe and Elsewhere." *Electricity Journal* 15(10): 53-64.
- Hogan, W. (1996). "A Wholesale Pool Spot Market Must Be Administered by the Independent System Operator: Avoiding the Separation Fallacy." *The Electricity Journal*: 26-37.
- Hogan, W. W. (1992). "Contract Networks for Electric Power Transmission." *Journal of Regulatory Economics* 4(3): 211-242.
- Hogan, W. W. (2002). "Electricity Market Restructuring: Reforms of Reforms." *Journal of Regulatory Economics* 21(1): 103-132.
- Hogan, W. W. (2002). "Financial Transmission Right Formulations". Cambridge. <http://www.ksg.harvard.edu/hepg>, Center for Business and Government, John F. Kennedy School of Government, Harvard University.
- Hur, D., C.-I. Yoo, et al. (2004). "Transmission embedded cost allocation methodology with consideration of system reliability." *IEE Proceedings on Generation, Transmission and Distribution* Vol. 151(No. 4): 427-432.

- IAEW (2001). "Analysis of Electricity Network Capacities and Identification of Congestion". Aachen, Institute of Power Systems and Power Economics (IAEW) and CONSENTEC Consulting fur Energiewirtschaft und -technik. Report for the European Commission, Directorate-General Energy and Transport.
- Ilic, M., E. Hsieh, et al. (2003). "Transmission Pricing of Distributed Multilateral Energy Transactions to Ensure System Security and Guide Economic Dispatch." IEEE Transactions on Power Systems 18(2): 428-434.
- Jong, J. d. (2004). "The regional approach in establishing the internal EU electricity market", Clingendael International Energy Programme. Netherlands Institute of International Relations. Web page: http://www.clingendael.nl/publications/2004/20041200_ciep_paper_dejong.pdf.
- Joskow, P. and J. Tirole (2000). "Transmission rights and market power on electric power networks." Rand Journal of Economics 31(3): 450-487.
- Joskow, P. and J. Tirole (2002). "Transmission Investment: Alternative Institutional Frameworks. Wholesale Markets for Electricity. November 22-23., Toulouse, France.
- Joskow, P. L. (2005). "Transmission Policy in the United States." Utilities Policy Vol. 13: 95-115.
- Joskow, P. L. (2005). "Making Transmission Owners Accountable". Roundtable on "Investments in Transmission", Conference on "The Economics of Electricity Markets", Toulouse.
- Joskow, P. L. (2005). Patterns of Transmission Investment. Electricity Infrastructure Investment Workshop. To be published as a chapter in the book 'Investments in Generation and Transmission in Electricity Liberalized Regimes', CRE, Paris, France.
- Joskow, P. L. and J. Tirole (2005). "Merchant Transmission Investment." Journal of Industrial Economics Vol. 53(2): 233-264.
- Kandil, M. S., S. M. El-Debeiky, et al. (2001). "Hybrid mathematical and rule-based system for transmission network planning in open access schemes." IEE Proceedings on Generation, Transmission and Distribution 148(5): 455-462.
- Kattuman, P. A., R. J. Green, et al. (2001). A Tracing Method for Pricing Inter-Area Electricity Trades, Department of Applied Economics, University of Cambridge, Cambridge Working Papers in Economics with number 0107. Web page: <http://ideas.repec.org/p/cam/camdae/0107.html>. June.
- Kavicky, J. A. and S. M. Shahidehpour (1997). "Determination of Generator Siting and Contract Options Based on Interutility Tie Line Flow Impacts." IEEE Transactions on Power Systems 12(4): 1649-1653.
- Keller, K. and J. Wild (2004). "Long-term investment in electricity: a trade-off between co-ordination and competition?" Utilities Policy 12(4): 243-251.
- KEMA (2002). "Desarrollo Institucional y Organización de la Comisión Regional de Interconexión Eléctrica" - CRIE y el Ente Operador Regional - EOR, Report prepared by Kema Consulting and ISA Interconexión Eléctrica for the Unidad Ejecutora del proyecto SIEPAC, in Spanish: 75.
- KEMA (2002). "Estudios para el Desarrollo del Reglamento de Operación Técnica y Comercial del MER", Report prepared by KEMA Consulting and ISA Interconexión Eléctrica for the Unidad Ejecutora del Proyecto SIEPAC, in Spanish: 117.
- Kench, B. T. (2004). "Let's Get Physical! Or Financial? A Study of Electricity Transmission Rights." Journal of Regulatory Economics 25(2): 187-214.

- Khairuddin, A. B., S. S. Ahmed, et al. (2004). "A Novel Method for ATC Computations in a Large-Scale Power System." *IEEE Transactions on Power Systems* 19(2): 1150-1158.
- Khator, S. K. and L. C. Leung (1997). "Power Distribution Planning: A Review of Models and Issues." *IEEE Transactions on Power Systems* 12(3): 1151-1158.
- Kim, B. H. and M. L. Baughman (1997). "The Economic Efficiency Impacts of Alternatives for Revenue Reconciliation." *IEEE Transactions on Power Systems* Vol. 12(No. 3): 1129-1135.
- Kirschen, D. and G. Strbac (1999). "Tracing Active and Reactive Power between Generators and Loads Using Real and Imaginary Currents." *IEEE Transactions on Power Systems* Vol. 14(No. 4): 1312-1318.
- Kirschen, D., R. Allan, et al. (1997). "Contributions of individual generators to loads and flows." *IEEE Transactions on Power Systems* 12(1): 52-60.
- Klemperer, P. D. and M. A. Meyer (1989). "Supply Function Equilibria in Oligopoly under Uncertainty." *Econometrica* 57(6): 1243-1277.
- Kohonen, T. (1995). "Self-Organizing Maps", Springer.
- Kristiansen, T. (2004). "Congestion management, transmission pricing and area price hedging in the Nordic region." *International Journal of Electrical Power & Energy Systems* Vol. 26(No. 9): 685-695.
- Kumar, A., S. C. Srivastava, et al. (2004). "A Zonal Congestion Management Approach Using Real and Reactive Power Rescheduling." *IEEE Transactions on Power Systems* 19(1): 554-562.
- Latorre, G., R. D. Cruz, et al. (2003). "Classification of Publications and Models on Transmission Expansion Planning." *IEEE Transactions on Power Systems* 18(2): 938-946.
- Latorre-Bayona, G. and I. Pérez-Arriaga (1994). "CHOPIN, A heuristic Model for Long Term Transmission Expansion Planning." *IEEE Transactions on Power Systems* 9(4): 1886-1894.
- Léautier, T.-O. (2000). "Regulation of an Electric Power Transmission Company." *The Energy Journal* 21(4): 61-92.
- Léautier, T.-O. (2001). "Transmission Constraints and Imperfect Markets for Power." *Journal of Regulatory Economics* 19(1): 27-54.
- Lima, J. W. M., M. V. F. Pereira, et al. (1995). "An Integrated Framework for Cost Allocation in a Multi-owned Transmission System." *IEEE Transactions on Power Systems* Vol. 10(No. 2): 971-977.
- Littlechild, S. C. and C. J. Skerk (2004). "Regulation of transmission expansion in Argentina Part I: State ownership, reform and the Fourth Line", Working paper ep61. CMI electricity project. Department of Applied Economics, University of Cambridge. Web page: <http://www.econ.cam.ac.uk/electricity/publications/wp/75>.
- Littlechild, S. C. and C. J. Skerk (2004). "Regulation of transmission expansion in Argentina Part II: Developments since the Fourth Line", Working paper ep62. CMI electricity project. Department of Applied Economics, University of Cambridge. Web page: <http://www.econ.cam.ac.uk/electricity/publications/wp/75>.
- Llarena, D. G. (2004). *Un nuevo modelo mercantil para la determinación de cargos de transmisión en mercados eléctricos competitivos con acceso abierto*. Buenos Aires, Argentina, Mercados Energéticos.
- Lo, K. L. and S. P. Zhu (1993). "Wheeling and marginal wheeling rates: theory and case study results." *Electric Power Systems Research* Vol. 27(No. 1): 11-26.

- Marangon, J. W. and E. J. d. Oliveira (1998). "The Long Term Impact of Transmission Pricing." IEEE Transactions on Power Systems Vol. 13(No. 4): 1514-1520.
- McDaniel, T. (2003). "Auctioning access to networks: evidence and expectations." Utilities Policy 11(1): 33-38.
- ME (2005). "Plan Estratégico para Modernización del Marco Regulatorio", Final report prepared by Mercados Energéticos for Osinerg (Peruvian Energy Regulatory Agency). Web page: <http://www.osinerg.gob.pe/>.
- Méndez, R. and H. Rudnick (2004). "Congestion Management and Transmission Rights in Centralized Electric Markets." IEEE Transactions on Power Systems 19(2): 889-896.
- Muñoz, A. (1996). "Aplicación de técnicas de redes neuronales artificiales al diagnóstico de procesos industriales." Departamento de electrónica y automática. School of engineering. Madrid, Universidad Pontificia Comillas.
- Mutale, J., G. Strbac, et al. (2000). "Allocation of losses in distribution systems with embedded generation." IEE Proceedings on Generation, Transmission and Distribution Vol. 147(No. 1): 7-14.
- Neuhoff, K. (2002). "Integrating Transmission and Energy Markets Mitigates Market Power", Cambridge, DAE. Web page: <http://www.econ.cam.ac.uk/electricity/publications/wp/index.htm>.
- Neuhoff, K., J. Barquin, et al. (2002). "Network-constrained Cournot models of liberalized electricity markets: the devil is in the details." Energy Economics 27(3): 495-525.
- Newbery, D. M. (2003). "Network capacity auctions: promise and problems." Utilities Policy 11(1): 27-32.
- Newbery, D., E. v. Damme, et al. (2003). "Benelux market integration: Market power concerns", Dutch energy regulator (DTe). Web page: http://www.dte.nl/nl/Images/12_13216.pdf.
- O'Neill, R. P., U. Helman, et al. (2002). "A Joint Energy and Transmission Rights Auction: Proposal and Properties." IEEE Transactions on Power Systems 17(4): 1058-1067.
- OATI (2003). Open Access Technology International (OATI). Web page: <http://www.oatiinc.com>. 2003.
- Olmos, L. and K. Neuhoff (2006). "Identifying a Balancing Point for Electricity Transmission Contracts." IEEE Transactions on Power Systems Vol. 21(No. 1): 91-98.
- Oren, S. S. and A. M. Ross (2002). "Economic Congestion Relief Across Multiple Regions Requires Tradable Physical Flow-gate Rights." IEEE Transactions on Power Systems 17(1): 159-165.
- Ott, A. L. (2003). "Experience with PJM Operation, Design and Implementation." IEEE Transactions on Power Systems 18(2): 528-534.
- Ou, Y. and C. Singh (2002). "Assessment of Available Transfer Capability and Margins." IEEE Transactions on Power Systems 17(2): 463-468.
- Outhred, H. R. and R. J. Kaye (1996). "Incorporating Network Effects in a Competitive Electricity Industry: an Australian Perspective. Issues in Transmission Pricing and Technology". M. Einhorn and R. Siddiqui, Kluwer Academic Publishers: 207-228.
- Pan, J., Y. Teklu, et al. (2000). "Review of Usage-Based Transmission cost Allocation Methods under open Access." IEEE Transactions on Power Systems Vol. 15(No. 4): 1218-1224.

- Pantos, M. and F. Gubina (2004). "Ex-Ante Transmission-Service Pricing Based on Load-Flow Patterns." *IEEE Transactions on Power Systems* 19(2): 796-801.
- Park, Y. M., J.-B. Park, et al. (1998). "An Analytical Approach for Transaction Costs Allocation in Transmission System." *IEEE Transactions on Power Systems* Vol. 13(No. 4): 1407-1412.
- Peco, J. P. (2001). "Modelo de Cobertura Geográfica de una Red de Distribución Eléctrica." Doctoral thesis. Instituto de Investigación Tecnológica. Madrid (in Spanish), Universidad Pontificia Comillas.
- Perán, F. L., I. J. Pérez-Arriaga, et al. (2001). "Benchmark of electricity transmission tariffs in the countries of the internal electricity market of the European Union." *Utilities Policy* Vol. 10: 47-56.
- Pereira, M. V. F. and B. G. Gorenstin (1996). "Transmission pricing methods". Seminar: "Remuneration and Pricing of Power Services", Bogotá, Colombia.
- Pereira, M. V. F., M. F. McCoy, et al. (2000). "Managing risk in the new power business." *IEEE Comput. Applicat. Power*: 18-24.
- Pérez-Arriaga, I. J. and L. Olmos (2005). "A plausible congestion management scheme for the internal electricity market of the European Union." *Utilities Policy* Vol. 13: 117-134.
- Pérez-Arriaga, J. I. (2002). "Methodology for cross-border tarification in the Internal Electricity Market of the European Union." *Power Systems Computing Conference (PSCC)*, Seville.
- Pérez-Arriaga, J. I. (2005). "Libro Blanco sobre la reforma del marco regulatorio de la generación eléctrica en España". Madrid, Ministry of Industry, Tourism and Commerce: 579.
- Pérez-Arriaga, J. I. and F. J. Rubio (1995). "Marginal pricing of transmission services: An analysis of cost recovery". *IEEE Transactions on Power Systems* Vol. 10(No. 1): 546-553.
- Pérez-Arriaga, J. I. and L. Olmos (2003). "Extension of the project on cost components of cross border exchanges of electricity", Directorate-General for Energy and Transport / European Commission. Web page: http://europa.eu.int/comm/energy/electricity/publications/index_en.htm.
- Pérez-Arriaga, J. I. and L. Olmos (2005). "Compatibility of Investment Signals in Distribution, Transmission and Generation". Electricity Infrastructure Investment Workshop. To be published as a chapter in the book 'Investments in Generation and Transmission in Electricity Liberalized Regimes', CRE, Paris, France.
- Pérez-Arriaga, J. I. and L. Omos (2005). "Compatibility of Investment Signals in Distribution, Transmission and Generation". Electricity Infrastructure Investment Workshop. To be published as a chapter in the book 'Investments in Generation and Transmission in Electricity Liberalized Regimes', CRE, Paris, France.
- Pérez-Arriaga, J. I. and Y. Smeers (2003). "Guidelines on tariff setting. Transport Pricing of Electricity Networks". F. Levesque, Kluwer Academic Publishers.
- Pérez-Arriaga, J. I., F. L. P. Montero, et al. (2002). "Benchmark of Electricity Transmission Tariffs", Directorate-General for Energy and Transport / European Commission. Web page: http://europa.eu.int/comm/energy/electricity/publications/doc/bench_trans_tarif_en.pdf.
- Pérez-Arriaga, J. I., L. Olmos, et al. (2002). "Cost components of cross border exchanges of electricity", Directorate-General for Energy and Transport / European Commission. Web page: http://europa.eu.int/comm/energy/electricity/publications/index_en.htm.

- Pérez-Arriaga, J. I., T. G. S. Román, et al. (1987). "Situación del estado del arte en la planificación de redes de transporte de energía eléctrica", Report prepared for Red Eléctrica de España (in Spanish).
- Petrucelli, M. (2003). "Estimación del uso del sistema de transporte por los agentes del mercado de electricidad". Madrid, Comillas University.
- PHB-HAGLER/SYNEX (2000). "Diseño General del Mercado Eléctrico Regional". Buenos Aires, Informe elaborado por Mercados Energéticos S.A., PHB-HAGLER BAILLY y SYNEX Ingenieros Consultores para Unidad Ejecutora del Proyecto SIEPAC.
- Poletti, S. (2005). "Optimal Time-of-use Access Charges." *Journal of Regulatory Economics* Vol. 27(No. 2): 203-224.
- PSR (2002). "Alocacao de Custos de Transmissao com Base no Uso da Rede", *Comite de Revitalizacao/MME*: 23.
- PSR (2005). "Allocation of Transmission Service Costs. Methods and Tools for Transmission Costing", *Power Systems Research PSR*. Web page: <http://www.psr-inc.com.br/>: 25.
- Ramsey, F. (1927). "A contribution to the theory of taxation." *Economic Journal* 32: 47-61.
- Rassenti, S. J., V. L. Smith, et al. (2003). "Discriminatory Price Auctions in Electricity Markets: Low Volatility at the Expense of High Price Levels." *Journal of Regulatory Economics* 23(2): 109-123.
- Rau, N. S. (1989). "Certain Considerations in the Pricing of Transmission Service." *IEEE Transactions on Power Systems* Vol. 4(No. 3): 1133-9.
- Rau, N. S. (2000). "Transmission Loss and Congestion Cost Allocation: An Approach Based on Responsibility." *IEEE Transactions on Power Systems* 15(4): 1401-1409.
- Rau, N. S. (2003). "Issues in the Path Toward an RTO and Standard Markets." *IEEE Transactions on Power Systems* 18(2): 435-443.
- Rau, N. S. and F. Zeng (2004). "Adequacy and Responsibility of Locational Generation and Transmission- Optimization Procedures." *IEEE Transactions on Power Systems* 19(4): 2093-2101.
- Reta, R. and A. Vargas (2001). "Electricity tracing and loss allocation methods based on electric concepts." *IEE Proceedings on Generation, Transmission and Distribution* Vol. 148(No. 6): 518-522.
- Riechmann, C. (2000). "Strategic pricing of grid access under partial price-caps-electricity distribution in England and Wales." *Energy Economics* Vol. 22(No. 2): 187-207.
- Rivier, M. and J. I. Pérez-Arriaga (1993). "Computation and decomposition of spot prices for transmission pricing". 11th PSCC Conference, Avignon, France.
- Rivier, M. and J. I. Pérez-Arriaga (1993). "Computation and decomposition of spot prices for transmission pricing". 11th PSCC Conference, Avignon, France.
- Romero, R., C. Rocha, et al. (2003). "Analysis of heuristic algorithms for the transportation model in static and multistage planning in network expansion systems." *IEE Proceedings of Generation, Transmission and Distribution* 150(5): 521-526.
- Romero, R., C. Rocha, et al. (2005). "Constructive heuristic algorithm for the DC model in network transmission expansion planning." *IEE Proceedings on Generation, Transmission and Distribution* 152(2): 277-282.
- Rothkopf, M. H. (2002). "Control of Market Power in Electricity Auctions." *The Electricity Journal*: 15-24.

- Rubio, F. J. (1999). "Metodología de Asignación de Costes de la Red de Transporte en un Contexto de Regulación Abierta a la Competencia". Doctoral Thesis. Escuela Técnica Superior de Ingeniería. Madrid (in Spanish), Universidad Pontificia Comillas.
- Rubio, F. J. and J. I. Pérez-Arriaga (2000). "Marginal Pricing of Transmission Services: A Comparative Analysis of Network Cost Allocation Methods." IEEE Transactions on Power Systems VOL. 15(NO. 1): 448-454.
- Rudnick, H., R. Palma, et al. (1995). "Marginal Pricing and Supplement Cost Allocation in Transmission Open Access." IEEE Transactions on Power Systems Vol. 10(No. 2): 1125-1142.
- Rudnick, H., R. Varela, et al. (1996). "Evaluation of alternatives for power system coordination and pooling in a competitive environment". IEEE / PES Winter Meeting, Baltimore.
- Ruff, L. E. (1999). "Competitive Electricity Markets: Why They Are Working and How to Improve Them". <http://www.ksg.harvard.edu/hepg, n/e/r/a>.
- Ruff, L. E. (2000). "Flowgates vs. FTRs and options vs. Obligations". <http://www.ksg.harvard.edu/hepg>.
- Ruff, L. E. (2001). "Flowgates, Contingency-Constrained Dispatch and Transmission Rights." The Electricity Journal 14(1): 34-55.
- Saphores, J.-D., E. Gravel, et al. (2004). "Regulation and Investment under Uncertainty: An Application to Power Grid Interconnection." Journal of Regulatory Economics 25(2): 169-186.
- Schweppe, F., M. Caramanis, et al. (1988). "Spot pricing of electricity", Kluwer Academic Publishers.
- SEEEC (2005). South East Europe Energy Community Treaty. 2005, 25 October. Web page: http://europa.eu.int/comm/energy/electricity/south_east/treaty_en.htm.
- Shirmohammadi, D., C. Rajagopalan, et al. (1991). "Cost of transmission transactions: an introduction." IEEE Transactions on Power Systems Vol. 6(No. 4): 1546-1560.
- Shirmohammadi, D., P. R. Gribick, et al. (1989). "Evaluation of transmission network capacity use for wheeling transactions." IEEE Transactions on Power Systems 4(4): 1405-1413.
- Shirmohammadi, D., X. Vieira, et al. (1996). "Some Fundamental Technical Concepts about Cost Based Transmission Pricing." IEEE Transactions on Power Systems Vol. 11(No. 2): 1002-1008.
- Shrestha, G. B. and P. A. J. Fonseka (2004). "Congestion-Driven Transmission Expansion in Competitive Power Markets." IEEE Transactions on Power Systems 19(3): 1658-1665.
- Silva, E. L. d., H. A. Gil, et al. (2000). "Transmission Network Expansion Planning Under an Improved Genetic Algorithm." IEEE Transactions on Power Systems 15(3): 1168-1175.
- Silva, E. L., S. E. C. Mesa, et al. (1998). "Transmission Access Pricing to Wheeling Transactions: a Reliability Based Approach." IEEE Transactions on Power Systems Vol. 13(No. 4): 1481-1486.
- Singh, H., S. Hao, et al. (1998). "Transmission Congestion Management in Competitive Electricity Markets." IEEE Transactions on Power Systems 13(2): 672-680.
- Skokljek, I. and V. Maksimovic (2004). "Transaction allocation symbolic analysis method." European Transactions on Electrical Power(No. 14): 261-275.
- Smeers, Y. (2005). "Long term locational prices and investment incentives in the transmission of electricity". Electricity Infrastructure Investment Workshop.

- To be published as a chapter in the book 'Investments in Generation and Transmission in Electricity Liberalized Regimes', CRE, Paris, France.
- SMI (2005). "Orden ITC/102/2005 por la que se establece la retribución de las actividades reguladas del sector gasista", Spanish Ministry of Industry, Tourism and Trade. Web page: <http://www6.mityc.es/energia/archivos/Orden%20Retribuciones%202005.pdf>.
- SNC/Statkraft (2004). "Reglamento de Transmisión y Calidad de Servicio del MER", Report prepared by SNC Lavalin and Statkraft Groner for the Unidad Ejecutora del Proyecto SIEPAC, in Spanish: 120.
- Stamtsis, G. C. and I. Erlich (2004). "Use of cooperative game theory in power system fixed-cost allocation". IEE Proceedings Generation, Transmission and Distribution Vol. 151(No. 3): 401-406."
- Stern, J. and R. Turvey (2003). "Auctions of capacity in network industries." Utilities Policy 11(1): 1-8.
- Stoft, S. (1998). "Congestion Pricing with Fewer Prices than Zones." The Electricity Journal 11(4): 23-31.
- Stoft, S. (1999). "Financial Transmission Rights Meet Cournot: How TCCs Curb Market Power." The Energy Journal 20(1): 1-23.
- Stoft, S. (2005). "Problems of Transmission Investment in a Deregulated Power Market". Electricity Infrastructure Investment Workshop, CRE, Paris, France.
- Strbac, G., D. Kirchen, et al. (1998). "Allocating transmission system usage on the basis of traceable contributions of generators and loads to flows." IEEE Transactions on Power Systems 13(2): 527-534.
- Styczynski, Z. A. (1999). "Power network planning using game theory". 13th Power Systems Computation Conference, Trondheim.
- Sun, H. and D. C. Yu (2000). "A multiple-objective optimization model of transmission enhancement planning for independent transmission company (ITC)". Proceedings IEEE Power Engineering Society Summer Meeting.
- Tabors, R. and M. Caramanis (2000). "Real Flow, A Preliminary Proposal for a Flow-based Congestion Management System". Cambridge, MA, Web page: <http://www.ksg.harvard.edu/hepg/flowgate/Real%20Flow.pdf>.
- Tan, X. and T. T. Lie (2002). "Application of the Shapley Value on transmission cost allocation in the competitive power market environment." IEE Proceedings on Generation, Transmission and Distribution Vol. 149(No. 1): 15-20.
- Tao, S. and G. Gross (2002). "A Congestion Management Allocation Mechanism for Multiple Transaction Networks." IEEE Transactions On Power Systems 17(3): 826-833.
- Tobin, V. J. (1990). "Transmission Access: A Voluntary System for Resolution of Transmission Access and Pricing Problems." IEEE Power Engineering Review 10(4): 18-21.
- TP-NZ (1990). "Transmission Pricing. Calculation of Network Charges", Trans Power NZ Limited.
- Turvey, R. (2000). "Infrastructure access pricing and lumpy investments". Utilities Policy Vol. 9(No. 4): 207-218.
- Usaola, J. (2005). "A transaction-based method for allocation of transmission grid cost and losses". Electric Power Systems Research 76: 395-403.
- Vazquez, C., L. Olmos, et al. (2002). "On the choice of slack bus in mechanisms for transmission network cost allocation that are based on network utilization". Power Systems Computation Conference, PSCC, Seville.

- Vázquez, C., M. Rivier, et al. (2005). "Adecuación del marco regulatorio para la integración regional". Madrid, Report prepared for the Peruvian energy regulator (OSINERG-GART) by the Instituto de Investigación Tecnológica from the Universidad Pontificia Comillas. In Spanish: 123.
- Vlachogiannis, J. G. (2000). "Accurate model for contribution of generation to transmission system effect on charges and congestion management." IEE Proceedings Generation, Transmission and Distribution Vol. 147(No. 6): 342-8.
- Vogelsang, I. (2001). "Price Regulation for Independent Transmission Companies". Journal of Regulatory Economics 20(2): 141-165.
- Wang, P., Y. Xiao, et al. (2004). "Nodal Market Power Assessment in Electricity Markets". IEEE Transactions on Power Systems 19(3): 1373-1379.
- WB (2000). "Interconexiones Regionales de Mercados Eléctricos en Sudamerica". Propuesta para Facilitar el Incremento de los Intercambios Energéticos en la Región. Buenos Aires, The World Bank, ESMAP, Department of Energy US, CIER.: 151.
- Wei, P., Y. Ni, et al. (2002). "Decentralised approach for congestion management and congestion price discovering". IEE Proceedings on Generation, Transmission and Distribution 149(6): 645-652.
- Wei, P., Y. Ni, et al. (2002). "Load flow tracing in power systems with circulating power". International Journal of Electrical Power and Energy Systems Vol. 24(No. 10): 799-885.
- Wood, A. J. and B. F. Wollenberg (1984). "Power Generation, Operation and Control". New York, John Wiley & Sons.
- Wu, F. F., Y. Ni, et al. (2000). "Power Transfer Allocation for Open Access Using Graph Theory—Fundamentals and Applications in Systems Without Loopflow." IEEE Transactions on Power Systems Vol. 15(No. 3): 923-9.
- Yarrow, G. (2003). "Capacity auctions in the UK energy sector". Utilities Policy 11(1): 9-20.
- Yu, C. W. (2002). "Transmission embedded cost allocation using sensitivity analysis". IEE Proceedings Generation, Transmission and Distribution Vol. 149(No. 5): 627-632.
- Yu, C. W. and A. K. David (1997). "Pricing Transmission Services in the Context of Industry Deregulation". IEEE Transactions on Power Systems Vol. 12(No. 1): 503-510.
- Zolezzi, J. M. and H. Rudnick (2002). "Transmission Cost Allocation by Cooperative Games and Coalition Formation". IEEE Transactions on Power Systems VOL. 17(NO. 4): 1008-15.

