STRATEGIES, METHODS AND TOOLS FOR SOLVING LONG-TERM TRANSMISSION EXPANSION PLANNING IN LARGE-SCALE POWER SYSTEMS

DESTA ZAHLAY FITIWI

MARCH 20, 2016

MADRID, SPAIN

STRATEGIES, METHODS AND TOOLS FOR SOLVING LONG-TERM TRANSMISSION EXPANSION PLANNING IN LARGE-SCALE POWER SYSTEMS

DESTA ZAHLAY FITIWI

MARCH 20, 2016

DOCTORAL THESIS SUPERVISORS (*EX AEQUO*):

PROF. DR. MICHEL RIVIER ABBAD   UNIVERSIDAD DE PONTIFICIA COMILLAS
PROF. DR. FERNANDO DE CUADRA   UNIVERSIDAD DE PONTIFICIA COMILLAS

MEMBERS OF THE EXAMINATION COMMITTEE:

Printed by

iv

Strategies, Methods and Tools for Solving Long-term Transmission Expansion Planning in Large-scale Power Systems

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen
op donderdag 15 oktober 2015 om 16:00 uur

door

DESTA ZAHLAY FITIWI

Master of Science in Electrical and Electronics Engineering

PETRONAS University of Technology

Tronoh, Malaysia

This dissertation has been approved by the promotors:

Composition of the doctoral committee:

Independent members:

Printed by

SETS JOINT DOCTORATE

The Erasums Mundus Joint Doctorate in Sustainable Energy Technologies and Strategies, SETS Joint Doctorate, is an international PhD program offered by six institutions in cooperation:

- Comillas Pontifical University, Madrid, Spain
- Delft University of Technology, Delft, the Netherlands
- Florence School of Regulation, Florence, Italy
- Johns Hopkins University, Baltimore, USA
- KTH Royal Institute of Technology, Stockholm, Sweden
- University Paris-Sud 11, Paris, France

The Doctoral Degree issued upon completion of the program are issued by Comillas Pontifical University, Delft University of Technology and KTH Royal Institute of Technology.

The Degree Certificates are giving reference to the joint program. The doctoral candidates are joined supervised, and must pass a joint examination procedure set up by the three institutions issuing the degrees.

This Thesis is a part of the examination for the doctoral degree.

The invested degrees are official in Spain, the Netherlands and Sweden, respectively.

SETS Joint Doctorate was awarded the Erasmus Mundus excellence label by the European Commission in year 2010, and the European Commission's Education, Audiovisual and Culture Executive Agency, EACEA, has partly supported the funding of this program.

The EACEA is not to be held responsible for contents of this Thesis.

This thesis is dedicated to the memory of my father.

ABSTRACT IN ENGLISH LANGUAGE

Driven by economic growth, changing life-styles and increasing penetration of electric vehicles, demand for electricity is continuously growing worldwide, and so are the global concerns of climate change, sustainability and energy security. The current energy production paradigm heavily depends on conventional energy sources i.e. fossil fuels. This is unsustainable because these sources of energy are quickly running out, leading to deep concerns of medium- and long-term energy security throughout the world. In addition, the heavy dependence on fossil fuels for power generation in the power industry has been substantially contributing to the increased level of greenhouse gas emissions. Because of all these and other techno-economic as well as structural issues, the electric energy industry is expected to undergo a paradigm shift with a considerably increased level of renewables (mainly, variable energy sources such as wind and solar), gradually replacing conventional power production sources. The scale and the speed of integrating such sources of energy are of paramount importance to effectively address the aforementioned concerns. As it is witnessed in recent years, wind and solar power have been attracting large-scale investments in many countries, especially in Europe. The favorable agreements of states in the recent climate conference in Paris (COP21), along with other driving factors, will further accelerate the renewable integration in power systems.

Renewable energy resources—RESs (wind and solar, in particular) are abundant almost everywhere on earth despite the fact that they are widely distributed and their energy intensities vastly differ from one place to another. Because of this, the global drive for high level integration of such energy sources can be realized by undergoing heavy investments in transmission infrastructures. In other words, transmission expansion planning (TEP) has to be carried out over geographically wide and large-scale networks. This helps to effectively accommodate the RESs and optimally exploit their benefits while minimizing the side effects. However, the stochastic nature of most of the renewable sources, along with the size of the network systems, results in a complex and combinatorial optimization problem, requiring a huge computational effort. The resulting problem can eventually become harder to solve, if not intractable. Thus, this demands that the models and the tools pursued to be computationally very efficient and reasonably accurate. At the same time, they should feature aspects that are believed to play a non-negligible role in TEP. To this end, this thesis presents solution strategies, tools and methods that collectively contribute to an effective and efficient resolution of such a complex problem within a finite simulation time.

From a modeling perspective, firstly, a new formulation is proposed for a long-term planning of transmission infrastructures under uncertainty with a multi-stage decision framework and considering a high level renewable integration. Secondly, recognizing the significant impacts network losses have on TEP solutions (which are often neglected in most TEP studies because of computational limitations), this thesis contributes new linear losses models, some of which strike the right balance between accuracy and

computational effort, particularly, in the context of medium to long-term TEP in large-scale power systems accommodating high level variable energy sources.

The integration of variable energy sources in the power systems introduces vast uncertainty and operational variability. This along with the uncertainty of electricity demand and other sources of uncertainty makes such a problem more complex. Hence, developing effective uncertainty and variability management tools is a very critical issue, especially in terms of computational requirements. A significant part of this uncertainty and variability is often handled by a set of operational states, here referred to as "snapshots", generation-demand patterns of power systems that lead to optimal power flow (OPF) patterns in the transmission network. A large set of snapshots, each one with an estimated probability, is then used to evaluate and optimize the network expansion. In long-term TEP of large networks, the number of operational states must be reduced. Hence, from a methodological perspective, this thesis shows how the snapshot reduction can be achieved by means of clustering, without relevant loss of accuracy, provided that a good selection of classification variables is used in the clustering process. The proposed method relies on two ideas. First, the snapshots are characterized by their OPF patterns (the effects) instead of the generation-demand patterns (the causes). This is simply because the network expansion is the target problem, and losses and congestions are the drivers to network investments. Second, the OPF patterns are classified using a "moments" technique, a well-known approach in Optical Pattern Recognition problems.

The entire TEP problem is kept as a stochastic mixed-integer linear programming (S-MILP) optimization, an exact solution method. This helps one to use effective off-the-shelf solvers and obtain expansion results within a finite simulation time, overall enhancing problem tractability. Furthermore, in order to significantly reduce the combinatorial solution search (CSS) space and hence facilitate the computation, a new heuristic solution strategy is devised. This approach works by primarily decomposing the problem into successive optimization phases. The foremost phases use relatively less complex optimization models than the following ones. And, each phase uses the results of the previous one. Hence, the main objective of this solution approach is to reduce the combinatorial solution search space, which in turn enhances tractability. Each optimization phase could be defined and solved as an independent problem, thus, allowing the use of specific decomposition techniques, or parallel computation when possible. A relevant feature of the solution strategy is that it combines both deterministic and stochastic modeling techniques on a multi-stage modeling framework with a rolling-window planning concept.

The planning horizon is divided into two sub-horizons: medium- and long-term, each having multiple decision stages. The first one is characterized by a set of investments which are good enough for all scenarios in the first sub-horizon while scenario-dependent decisions are sought in the second one.

The developed models, methods and solution strategies are tested on small-, medium- and large-scale network systems. This thesis also present numerical results of an aggregated 1060-node European network system obtained considering multiple RES development scenarios. Generally, test results show the effectiveness of the proposed TEP model, and the proposed methods and solution strategy are very effective in facilitating the solution process, and contribute to a significant reduction in computational effort while fairly maintaining optimality of the solutions.

RESUMEN (ABSTRACT IN SPANISH LANGUAGE)

SAMMANFATTNING (ABSTRACT IN SWEDISH LANGUAGE)

Samenvatting (Abstract in Dutch Language)

ACKNOWLEDGMENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

List of Abbreviations

2PEM	Two point estimate method
ACOPF	AC optimal power flow
ACTEP	Alternating current (AC) based transmission expansion planning
CbC	Clustering based on causes
CbE	Clustering based on effects
CL	Candidate line
COP21	Paris climate conference 2015
CSTEP	"Copper sheet" based transmission expansion planning
DC	Direct current
DCOPF	DC optimal power flow
DCTEP	"Direct current" (DC) based transmission expansion planning
ED	Economic dispatch
EL	Existing line
ENS	Energy not served
ENTSO-E	European network of transmission system operators for electricity
FOM	First order moment
FOR	Forced outage rate
FST	Fuzzy systems theory
GEP	Generation expansion planning
GHG	Greenhouse gas
HTEP	Hybrid based transmission expansion planning
IGDT	Info-gap decision theory
KCL	Kirchhoff's current law
KVL	Kirchhoff's voltage law
LinACTEP	Linearized AC based transmission expansion planning
LMP	Locational marginal price
MCS	Monte Carlo simulation
M-DCTEP	Modified DC based transmission expansion planning
MENA	Middle East and Northern Africa
MILP	Mixed integer linear programming
MINLP	Mixed integer linear programming
NPV	Net present value
NCUED	Network capacity unconstrained economic dispatch
NUTS-3	Nomenclature of territorial units for statistics region 3
ol	Overloaded lines
OPF	Optimal power flow
PC1	First Principal Component
PCA	Principal Component Analysis
PDF	Probability distribution function
PEM	Point estimate method

PTEP	Pipeline based transmission expansion planning
R-DCTEP	Relaxed DC based transmission expansion planning
RES	Renewable energy sources
RO	Robust optimization
RTS	Reliability test system
R&D	Research and development
SMILP	Stochastic mixed integer programming
SOM	Second order moment
SOS2	Special ordered sets of type 2
TEP	Transmission expansion planning

List of Sets and Indices

CL, Ω^{NL}	Set of candidate transmission lines
EL, Ω^{EL}	Set of existing transmission lines
KS_k	Set of snapshots clustered to cluster k
NS, Ω^{NS}	Set of system nodes
Ω^{GS}, GS	Set of generators in the system
Ω^{MS}, MS	Set of moment variables
Ω^{OS}, OS	Set of overloading snapshots
NOS, Ω^{NOS}	Set of non-overloading snapshots
Ω^{OL_s}, OL_S	Set of overloaded lines (masses) in a particular snapshot, s
Ω^{SAC}, SAC	Set of snapshots after clustering
Ω^{SBC}, SBC	Set of snapshots before clustering
0	Index for origin in Cartesian coordinate
c	Index for centroid
i, j	Bus indices
k	Index for clusters
l	Index for linear losses constraints
m	Index for moments
RES	Index for renewable generators ($wnd \cup sol$)
s	Snapshot index after clustering
sol	Index for solar generators
s'	Index for non-overloading snapshots
t	Snapshot index before clustering
wnd	Index for wind generators
x, y, z	Cartesian axes
$k/\Omega^{NL}, k/\Omega^{EL}$	Index/Set of new or existing lines
$\tau/\Omega^{T1}; \zeta/\Omega^{T2}$	Index/Set of stages in the first/the second period
w/Ω^w	Index/Set of snapshots
s/Ω^s	Index/Set of scenarios
a/Ω^a	Index/Set for transmission alternatives
t/Ω^t	Index/Set of time stages

List of Parameters

g_k, b_k	Susceptance of line k (per unit) (equivalent to inverse of reactance)
$d_{i,t}$	Demand level at node i for snapshot t (MW)
$d_{x,i}$	Distance from net load at node i to an axis of rotation (per unit)
$d_{x,k}$	Distance from an overloaded line i–j to an axis of rotation (per unit)
S_k^{max}	Power transmission capacity limit of line i–j (MW)
$g_{i,t}$	Generation level at node i for snapshot t (MW)
g_n^{max}, g_n^{min}	Generator's upper and lower power production capacity limits (MW)
ℓ_k	Length of line k (per unit)
u_{ns}, u_{nt}	Availability of generator n in snapshot s or t
v_{ci}	Cut-in wind speed (m/s)
v_{co}	Cut-out wind speed (m/s)
v_r	Rated wind speed (m/s)
v_t	Actual wind speed at time t (m/s)
x_{avg}	Average intra-cluster distance (per unit)
$x_{c,t}$	First order moment (centroid) of an overload pattern corresponding to snapshot t
x_c, y_c	Coordinate of center of a given mass (set of overloads or net loads)
$x_{k,c}, y_{k,c}$	Coordinate of the center of line i-j
Λ	Cost of energy not served (€/MWh)
IC_k	Investment cost of a candidate line i-j (million €)
D_{is}, D_{iw}	Electricity demand at node i in snapshot s or t (MW)
H	Planning horizon (years)
$I_{m,w}$	Computed moment m for snapshot w
$I_{x=x',w}$	Moment of inertia of overloads about a vertical axis $x = x'$, for snapshot w
K	Total number of clusters
L	Number of partitions in the linear losses modeling
M_k	Big-M parameter for line i–j
N_G	Total number of generators in the system
N_N	Total number of nodes in the system
N_{RES}	Total number of renewable generators in the system
N_{CL}	Number of candidate lines

N_{EL}	Total number of existing lines
N_k	Number of elements (snapshots) in cluster k
N_{ol}	Number of overloaded lines
$P_{wnd,w}$	Wind power output at time t (MW)
P_r	Rated power output of a wind or solar farm (MW)
$P_{RES,s}, P_{RES,w}$	Maximum available renewable generation output at snapshot s or t (MW)
$P_{sol,w}$	Solar power output at time w (MW)
R_c	A certain radiation point (often taken to be 150 W/m^2)
R_{std}	Solar radiation in standard condition (usually set to 1000 W/m^2)
R_w	Hourly solar radiation (W/m^2)
S	Total number of snapshots after clustering
S_B	Base power (MW)
T	Total number of snapshots before clustering
X	A coordinate of a snapshot in the moments' space
X_{avg}	Average inter-cluster distance (per unit)
$\hat{X}_{C_k}$	A center of mass (centroid) of cluster k in the moments' space
α_j	Weights to ensure that there is a correct balance between cost terms in the objective function
θ^{max}	Maximum allowable voltage angle (radians)
κ	Number of uncertain parameters
ρ_s	Probability of occurrence of snapshot s
σ	Interest rate (%)
υ_n	Variable (marginal) cost of generation of generator n (€/MWh)
ϕ_k	Angle that an overloaded line makes with the vertical axis (degrees)
ψ_k	Mass density of an overloaded line
Δp_k^{max}	Maximum step-size in losses linearization (MW)
ER_g^{NG}, ER_g^{EG}	Emission rates of new and existing generators, respectively
ER_g	Emission rate of a generator
g_k, b_k, S_k^{max}	Conductance, susceptance and flow limit of branch k
IC_k	Investment cost of line
$MC_{a,k}^{N1}, MC_{a,k}^{N2}$	Maintenance cost of candidate lines in the first and the second pools
$IC_{p,k}^{N1}, IC_{p,k}^{N2}$	Investment cost of lines in the first and the second pools
L	Total number of linear segments
$LT_k, LT_{a,k}$	Lifetime of line
MC_g^{NG}, MC_g^{EG}	Maintenance costs of new and existing generator per year
MC_k^{NL}, MC_k^{EL}	Maintenance costs of new and existing branch k per year

MP_k, MQ_k	Big-M parameters associated to active and reactive power flows through branch k, respectively
N_i	Number of buses
$OC_{g,i,s,w,t}^{NG}, OC_{g,i,s,w,t}^{EG}$	Operation cost of unit energy production by new and existing generators
$P_{sol,w}$	Hourly solar PV output (MW)
P_r	Rated power of a DG unit (MW)
$P_{wnd,w}$	Hourly wind power output (MW)
R_c	A certain radiation point (often taken to be 150 W/m^2)
R_w	Hourly solar radiation (W/m^2)
r_k, x_k	Resistance and reactance of branch k, respectively
R_{std}	Solar radiation in standard condition (usually set to 1000 W/m^2)
$TEEL_t$	Total expected emission level at stage t
v_{ci}	Cut-in wind speed (m/s)
v_{co}	Cut-out wind speed (m/s)
v_w	Observed/sampled hourly wind speed (m/s)
V_{nom}	Nominal voltage (kV)
V_{max}, V_{min}	Voltage limits
v_r	Rated wind speed (m/s)
$v_{s,w,t}$	Penalty for unserved power
Z_{ij}	Impedance of branch i-j (Ω)
$\alpha_l, \beta_l, \alpha_{k,l}, \beta_{k,l}$	Slopes of linear segments
$c_{k,l}$	Linear parameter (y-intercept) of a linear segment
$\lambda_g, \lambda_{g,s,t}$	Marginal cost of power production of generator
$\lambda_{s,w,t}^{CO_2e}$	Price of emissions (€/tons of CO_2 equivalent—€/tCO_2e)
ρ_s, π_w	Probability of scenario s and weight (in hours) of snapshot group w
$\Delta V^{min}, \Delta V^{max}$	Minimum and maximum permissible voltage deviation
Δp_k^{max}	Maximum step-size in losses linearization
S_B	Base power
LB_k, UB_k	Lower and upper bounds in losses linearization
$\Lambda, \Lambda_{s,w,t}$	Cost of unserved power
σ	Interest rate
Δ_b	Length of load block
$z_{k,max}$	Maximum investment limit in a corridor
r_k, x_k	Resistance and reactance of branch k
$S_{k,max}, S_k^{max}$	Capacity limit of branch k
$QG_{g,max}, QG_{g,min}$	Reactive power limits of a generator
$PG_{g,max}, PG_{g,min}$	Active power limits of a generator
PD_d, QD_d	Active and reactive power demand

List of Variables

PG_{ns}, PG_{nt}	A variable for the power output of the nth generator in snapshot s or t (in MW)
$p_{RES,s}, p_{RES,t}$	A generation output variable of renewable source in snapshot s or t (MW).
G_{is}, G_{it}	Total generated power at node i by all generators in snapshot s or t (MW)
z_k	Investment decision variable of candidate line k
$D_{s,w,t}^i, Q_{s,w,t}^i$	Active and reactive power demand at node i
$PG_{g,i,s,w,t}^N, PG_{g,i,s,w,t}^E$	Active power produced by new and existing generators
P_k, Q_k, θ_k	Active and reactive power flows, and voltage angle difference of link k, respectively
$\Delta p_{k,s,w,t,l}, \Delta q_{k,s,w,t,l}$	Step variables used in linearization of quadratic flows
$\Delta q_{k,l}, \Delta p_{k,l}$	Step variables in losses linearization
PL_k, QL_k	Active and reactive power losses, respectively
$QG_{g,i,s,w,t}^N, QG_{g,i,s,w,t}^E$	Reactive power consumed/produced by new and existing generators
$u_{g,i,t}, u_{k,t}$	Generator and line utilization variables (1 if is used, 0 otherwise)
V_i, V_j	Voltage magnitudes at nodes i and j
$z_{p,k,t}$	Investment variables for lines
$p_{i,s,w,t}$	Unserved power at node i
θ_i, θ_j	Voltage angles at nodes i and j
$u2_{k,s,\zeta}^{EL}$	Existing line's binary switching variable in the second period
$u1_{k,\tau}^{EL}$	Existing line's binary switching variable in the first period
$y_{a,k,s,\zeta}^{N2}$	Binary investment variable in the second period
$z'^{N1}_{a,k,s,\zeta}, z'_{s,\zeta}$	Binary recourse variable in the second stage
$z_{a,k,\tau}^{N1}, z_\tau$	Binary investment variable in the first period
$y_{s,\zeta}$	Binary investment variable in the second period.
$\phi_{2,k}, \phi_{1,k}$	Auxiliary variables used to linearize bilinear terms
Q_k^+, Q_k^-	Nonnegative auxiliary variables for reactive power flow
P_k^+, P_k^-	Nonnegative auxiliary variables for active power flow
ΔV_i	Voltage deviation at node i
P_k, Q_k	Active and reactive power flows of branch k
u_k	Utilization/switching variable of branch k
QG_g, PG_g	Actual reactive and active power generation of a generator
q_i, p_i	Unserved reactive and active power at node i

List of Functions

$EgyC_t^{NG}, EgyC_t^{EG}$	Expected cost of energy production by new and existing generators
$ENSC_t$	Expected cost of unserved power

$EmiC_t^{NG}, EmiC_t^{EG}$ Expected emission cost of power production using new and existing generators, respectively

$InvC_t^{LN}, MntC_t^{LN}$ NPV investment/ maintenance cost of a distribution line

z Objective function

1

I. INTRODUCTION

This chapter gives a brief introduction to the research topic, describes the scope and outlines the main as well as the specific objectives of this thesis. The thesis organization and structure is also presented at the end of this chapter.

1.1. BACKGROUND

Most of the energy that we consume today, in one form or another, comes from unsustainable energy sources. In particular, the electric industry is highly dependent on fossil fuels for power production. This has led to a series of questions from energy dependence and sustainability concerns to climate change issues, which are some of the major drivers of renewable energy source (RES) integrations in many power systems across the world. It is now widely recognized that integrating RESs in power systems brings about a lot of economic, environmental, societal and technical benefits to all stakeholders. Among the wide-range benefits of RESs is their significant contribution in combating climate change and abating its dire consequences. Most RES technologies (wind and solar PV, for instance) have very low carbon footprints, making them very suitable for solving such emission-induced health and environmental problems. Hence, integrating RESs in power systems partly replaces polluting (conventional) power generation sources, resulting in a "cleaner" energy mix i.e. one with lower emission levels.

The potential of RESs is colossal because, in principle, they can meet several times the world demand. RESs such as wind, solar, hydro, biomass and geothermal can provide sustainable energy services based on available resources in all parts of the world. The transition to renewable energy based power systems tends to increase, while their costs continuously decline as gas and oil prices continue to oscillate. In the last half century, the demand for wind and solar energy systems has been continuously increasing, experiencing a reduction in capital costs and generated electricity costs. There have been continuous performance improvement and R&D undergoing in the sector in the past decades. As a result, the prices of renewable energy and fossil fuels, as well as social and environmental costs are to diverge in opposite directions. Economic and political mechanisms are expected to massively support the wide spread of sustainable markets for the rapid development of RESs. At this point, it is clear that the present and future growth will occur mainly in renewable energy and in some natural gas-based generation systems, and not common sources like coal or oil-fired power plants. The progress of RESs can increase diversity in the electricity markets, contributing to obtain long-term sustainable energy, helping to reduce local and global greenhouse gas (GHG) emissions, promote attractive trade options to meet specific energy needs, and create new economic growth opportunities.

With climate change, sustainability, energy security, continuously increasing demand for electricity and socio-economic factors as the main drivers, the level of global RES integration has been steadily growing during the past decades, as indicated in a 2015 report by the International Energy Agency (IEA) [1]. The report further shows that, in 2013 alone, an approximately 19.1% of global electric energy consumption came from

RESs, most of which was from hydropower [1], [2]. After several decades of efforts in research and continuous development in RES, the yearly growth in the capacity of these plants is becoming greater than the total investment capacity added in power plants based on coal, natural gas and oil all combined together [3]. Nowadays, RESs have reached a significant level of share in energy supply options, becoming one of the prominent global alternative power supply sources. The latest global trends in renewable energy investment status reports indicate that renewables represented a 58.5% of net additions to global power capacity in 2014, with significant growth in all regions, which represents an estimated 27.7% of the world's power generating capacity, enough to supply an estimated 22.8% of global electricity. Investments in wind and solar power sources continue to outpace other technologies. Figure 1.1 shows the trends cumulative wind and solar power additions in Europe as well as globally. These trends nothing but reflect the growing interest in developing renewables. The overall cost-cutting achieved to date has helped to ensure such a strong momentum in 2014, reaching an investment boom up to 29% in solar, and 11% in wind technologies globally [4]. These figures are further strengthened in 2015 [4] with more than 33% and 16% new investments made globally in solar and wind technologies.
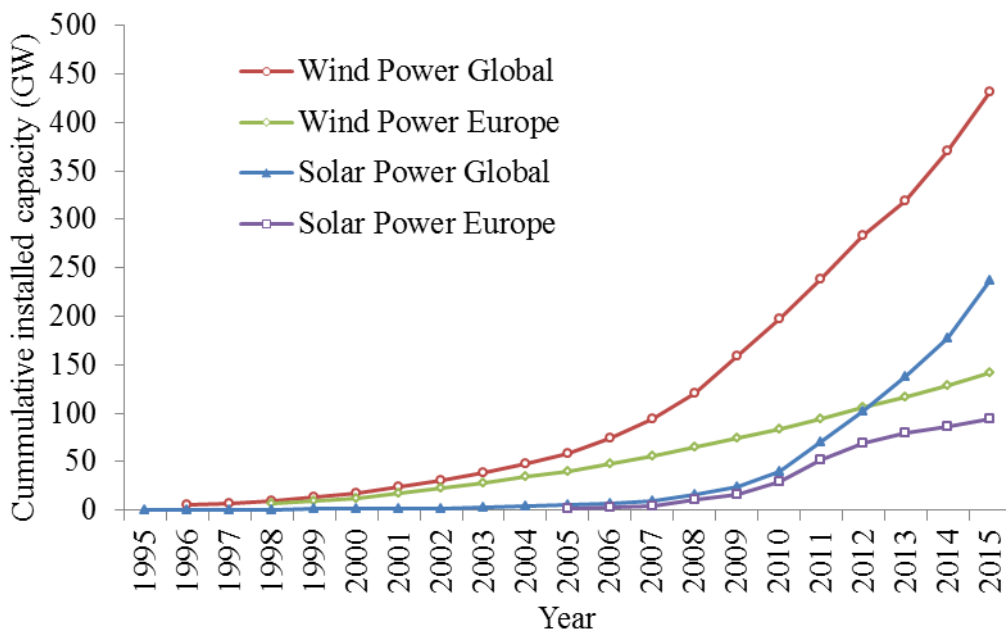
Fig. 1. 1 Cumulative installed capacity of wind and solar [4]–[7] .

These remarkable growths have been against a number of odds such as the recent global financial crisis, the dramatically falling fuel prices and the slowdown of increasing global electricity consumption that have been thought to decelerate or stall this trend [4]. The recent developments in the 2015 Paris climate conference (COP-21), overall trends in international policy on RESs, energy dependence concerns, the falling capital costs of several matured RES technologies, and other techno-economic factors are all favorably expected to further accelerate the level RES integrated into power systems. In general, there is a general consensus globally that RESs will cover a significant amount of electricity consumption in the years to come [2].

3

It can be inferred from Figure 1.1 that Europe, as the leading advocate of renewables, accounts for nearly half of the total installed capacities of these resources worldwide. European countries have set forth ambitious targets for emissions reductions and RES integrations. As in Figure 1.2, the renewable share in the final energy consumption in Europe is expected to reach 20%, 27% and 80% by 2020, 2030 and 2050, respectively. As a result, the integration of wind and solar is especially expected to increase significantly in the years to come.
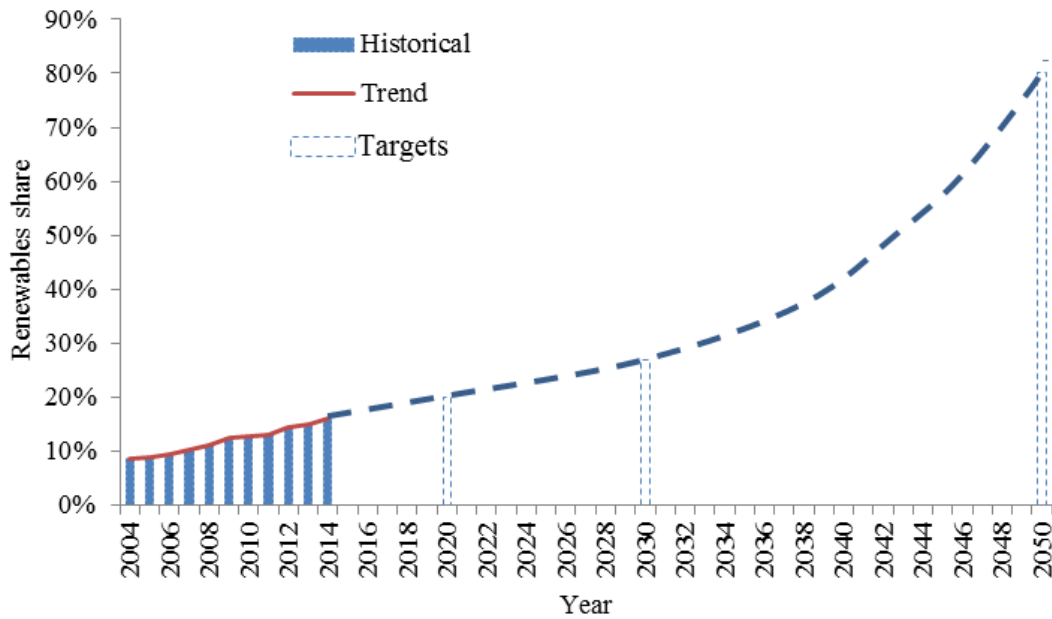
Fig. 1. 2 Historical and targeted trends of renewable energy share in gross final energy consumption in Europe [8].

Despite these interesting figures, several challenges remain in place pertaining to the tapping of large-scale RESs, their integrations and their efficient utilization. A growing effort in pursuing innovative approaches to increase RES participation is required to guarantee a clean energy future. Most of the challenges are related to the nature of such resources (especially wind and solar), which are abundant almost everywhere on earth despite the fact that they are widely distributed and their energy intensities vastly differ from one place to another. Their intermittent nature also poses significant challenge in operation and planning of power systems because of the vast uncertainty and variability such resources introduce to the system. In addition, power systems are subject to many more sources of uncertainties with different levels such as uncertainty in generation expansion/retirement, fuel prices, demand growth, component outages, carbon emissions, demand response, etc. The compound effect of all these creates considerable challenges, which increase with level of RES integration. The global drive for high level integration of such energy sources can be realized by undergoing heavy investments in transmission infrastructures among others, which help to even out the negative consequences of RES integration. In other words, because of their distributed nature, unprecedented transmission expansion planning (TEP) has to be carried over a

geographically wide area and large-scale networks to meet the short- to long-term objectives of integrating renewables (particularly, variable energy sources). This helps to effectively accommodate the RESs and optimally exploit their benefits while minimizing the side effects. In the context of Europe, for instance, there is an ambitious plan to develop massive RESs in the coming decades in a bid to curb GHG emissions, promote clean energy technologies and meet the increasing demand for electricity among other reasons. However, these energy sources are usually located in places of low demands. For instance, the West Coast and the North Sea are among the best locations in Europe for large-scale wind power developments. There are also initiatives to import huge amount of solar power from Middle East and North Africa (MENA) [9]. With these and other scenarios in mind, the European network will have to be adequately reinforced and expanded to support the integration of such developments [10], [11]. However, the high level uncertainty and variability inherent to such such resources [12], along with the size of such network systems and temporal scope, results in a complex and combinatorial optimization problem, requiring a huge computational effort.

Generally speaking, computational complexity of TEP problem dramatically increases with the size of the network dealt with. The resulting problem can eventually become harder to solve, if not intractable. Extensive literature survey reveals that existing planning models are not adequate to handle such a problem; they cannot seamlessly be extended to long-term TEP of large-scale networks with high level integration of variable energy sources mainly because they are not properly equipped with the necessary strategies and methods to systematically handle the vast uncertainty and operational variability inherent to such a problem. In addition, given the size of the problem, most traditional solution methods, which have been designed at most for national networks, have computational limitations, leading to tractability issues. All this explains the need for new strategies, tools and methods that effectively cope with a problem of this magnitude, which is the main theme of this thesis.

1.2. RESEARCH MOTIVATION AND PROBLEM DEFINITION

As introduced in the above background, the renewables' share in the total energy consumption will keep on increasing strongly. However, this will require tapping variable energy sources such as wind and solar in geographically wide and remote areas, far away from major demand centers and existing transmission infrastructures, leading to network expansion planning problems of exta-large network systems. This is highly needed to meet striving RES integration targets and global environmental-related obligations as well as balance out the negative effects of RES integration in power systems. However, the unprecedented uncertainty, temporal and spatial scopes of such a problem pose a significant computational challenge. This is the main motivation of the present work. Framed in this context, this thesis endeavors to address three main research questions emanating from a modeling and methodological perspectives of such a problem.

5

Network fidelity: *From the context of long-term planning of large-scale networks under high penetration of variable energy sources, what are the levels of details that can be included in a network expansion model that strikes the right balance between accuracy and computational requirement?* This question mainly relates to the modeling aspects of network systems. This is discussed in detail in Chapter 3, and modeling aspects of network losses is covered in our published work [13].

Uncertainty and variability: *From the same context, how should the different sources of uncertainty and variability be captured in such a way that ensures the right balance between problem tractability and solution accuracy?* To address this question, a new uncertainty and variability management tool is presented and thoroughly discussed in Chapter 4 and in [14].

TEP Model: The high temporal and geographical scope of the problem as well as the need for combining short- to long-term planning decisions demands a new TEP model. From this perspective, *how should the TEP model be formulated so that it meets the demands? What investments should be made now (in the short and medium term) and where should these be? What/where are the strategic investment decisions to be made considering different possible evolutions of the system?* These and other related issues are addressed in Chapter 5.

Managing combinatorial problem: *Given the size of the problem, how should the combinatorial solution search space be handled?* The present work proposes a heuristic solution method which includes a systematic way of decomposing the TEP problem into successive optimization phases. This is partly presented and discussed in [15].

1.3. THESIS OBJECTIVES

Main Objective—The main objective of this research is to develop mathematical optimization models, uncertainty and variability management methods, and solution strategies that support the complex decision-making process of long-term expansion planning of large-scale transmission grids under high level renewable integrations.

Specific Objectives —The specific objectives of this thesis are:

- To formulate a tractable long-term TEP model for very large-scale network under high level uncertainty and massive integration of variable energy sources;
- To propose methods for managing uncertainty and variability introduced by intermittent energy sources such as wind and solar power generators, electricity demand and price as well as component availabilities;
- To devise a new solution strategy for enhancing the tractability of the TEP problem in view of reasonably reducing computational time without significantly compromising the optimality of the solution;
- To test the proposed solution techniques on a realistic network under high penetration of renewables.

1.4. RESEARCH METHODOLOGY

In order to achieve the main research objective, this thesis develops simulation models, methods and solution strategies to analyze the long-term expansion of electricity grids under uncertainty and dramatically changing power generation scheme over time. In other words, the TEP problem is formulated from the perspective of long-term expansion planning and under high penetration level of renewables. Under these circumstances, the proposed model should sufficiently emulate the anticipated complex-decision making process planners have to face in relation to network expansion needs of especially large-scale power systems. This indicates that the tractability of such a model is of a paramount importance. On one hand, it is desired that the developed model embrace the inherent characteristics of the electrical systems in a reasonably accurate manner. On the other hand, the complex nature of the problem means that certain accuracy related issues should be compromised to ensure tractability. A tradeoff mechanism reconciles these two conflicting requirements. To this end, first, existing models are critically reviewed and compared in terms of their mathematical complexities, accuracy and possible applications to such a complex problem. Based on the results of the extensive analysis and comparisons, a tractable mathematical optimization model, based on an improved "DC" network flow model, is then proposed.

The main objective of the resulting TEP optimization is to meet the growing demand for electricity at the lowest cost possible (seen from the system perspective) while respecting all technical, economic and environmental constraints. This leads to a constrained optimization framework with an overall cost minimization as an objective. Hence, the resulting TEP model includes multiple cost terms such as investment, operation and maintenance, emission and reliability costs, which are combined to form a single objective function (the total cost in the system) in a stochastic programming and multi-stage planning framework.

In order to efficiently handle the uncertainty and the variability inherent to such problems, the thesis also introduces a new problem-specific methodology based on the theory of moments, which clusters operational situations based on their expected impact on expansion needs. Moreover, the research work proposes a new heuristic solution strategy that is proven to significantly facilitate the solution process. This strategy works by decomposing the original problem into successive optimization phases, which is structured in a manner that the output of given phase is the input for the subsequent phase.

The proposed optimization model as well as the solution strategy is implemented in GAMS$^{©}$ and mainly solved using CPLEX™ algorithm mostly with default parameters. Whereas, the clustering methodology is programmed in MATLAB$^{©}$ programming environment and Visual Basic™ with Excel$^{©}$ used as an interface for this purpose. The whole work here aims to provide a reliable expansion solution containing short-, medium as well as long-term decisions that can effectively cope with the rapidly changing environment in the power industry.

7

1.5. THESIS OUTLINE AND ORGANIZATION

This thesis is organized as follows. The first chapter presents a brief overview of the problem and motivation of the research work, and outlines the research objectives and methodology. The subsequent chapter presents an extensive review of the literature by organizing previous related works to highlight the research questions and objectives.

Chapter 3 begins by reviewing existing and modified TEP models, from the network representation perspective, and describing the modifications made in order to ultimately develop a reasonably accurate network representation which is to be used in the formulation a TEP model from the context of large-scale network applications. From computational requirement and accuracy standpoint, a comprehensive comparison of total of six TEP models, and thirteen variants of these models with different network fidelity levels, is discussed both theoretically and numerically to further motivate the need to develop a reasonably accurate TEP model for the stated problem. This chapter also presents a detailed modeling of network losses, which encompasses an extensive comparison of existing and novel losses models ones both from computational burden and accuracy viewpoints.

Chapter 4 introduces the novel methodology developed in this thesis for managing uncertainty and variability inherent to the problem at hand. The methodology is described in detail and its efficacy is demonstrated with a numerical example.

Chapter 5 presents detailed descriptions of the mathematical formulations of the TEP optimization problem in a multi-stage planning horizon and stochastic programming framework. This chapter also introduces the proposed solution strategy. In the subsequent chapter, the proposed strategy, tools and methods are verified by carrying out numerical studies on a realistic 1060-node European network system.

The last chapter gathers the main findings of this thesis in the form of conclusions, summarizes the main contributions of the thesis by revisiting the thesis objectives, and finally draws some directions for extending this work.

2

II. LITERATURE REVIEW

This chapter presents a comprehensive review of existing literature focusing on the relevant previous works in relation to transmission grid expansion planning.

2.1. CHAPTER OVERVIEW

In power systems, grid expansion planning is always one of the most critical issues that has to constantly be addressed for meeting the demand while maintaining system integrity and reliability. In other words, transmission expansion planning (TEP) is mandatory in every electric power industry which continually undergoes rapid changes in structure, management and operation [16] regardless of the electricity markets: traditional or competitive. The literature on the subject area of TEP includes several decades of research works, dating back to 1970. Recently, there have been a dramatic increase in the number of publications on this, especially in the past decade, indicating the growing concerns and challenges. This could be partly explained by the deregulation of power systems which increased the level of uncertainty in such systems, increasing the complexity of the problem, and increasing penetration level of non-conventional generation sources. A detailed review of existing literature on TEP as of 2003 is presented in [17], which has been recently complemented in [18], [19]. From the context of TEP, the following relevant issues define the structure of this literature review:

- fidelity of network representation (alternating current—AC, "direct current"—DC models, etc.),
- solution methods employed (mathematical, heuristic and meta-heuristic),
- nature of the electricity market (regulated vs. deregulated),
- objective function considered (investment cost, investment cost +operation costs, etc.),
- flexibility of expansion plans computed (static vs. dynamic), and
- methods adopted to handle uncertainty and variability inherent to the TEP problem (probabilistic, stochastic, etc.).

2.2. NETWORK REPRESENTATION FIDELITY

Power systems are characterized by their complex nature whose components are generally described by highly nonlinear and nonconvex models. The complexity of such systems are often systematically handled in complex power systems analysis, operation and planning problems mainly by using "proxy" models. Fidelity then refers to the level of details (i.e. actual physics describing the characteristics of the system) captured by such proxy models i.e. in relation to accuracy and complexity levels. This is especially the case in TEP problems, where the network is represented using various models such as the customary non-linear AC [20]–[22], the classic "direct current" (DC) [23]–[27], [28], "pipeline" [29], [30] and [31], hybrid (which combines the DC and the pipeline models) [30], or linear variants of these models and disjunctive models [10], [32], [33].

The AC network model is the most realistic model and it implicitly models network

losses but its mathematical complexity, nonlinear and nonconvex nature means its application in TEP problems is very limited. The DC model, which is derived from the AC model by making use of a number of simplifying assumptions, respects the physical laws that govern power flows in power systems. It is currently the most commonly used network model in TEP studies because of its appealing computational performance while delivering reasonably "accurate" solutions. However, the simplifying assumptions made in its formulation (which include flat voltage, zero resistance and very small voltage angles) leaves this model lossless. Moreover, the existence of reactive power flows is not acknowledged in the DC model. In a pipeline model, flows are only required to respect the capacity and nodal balance constraints. This model effectively treats the lines as flow networks. In other words, flows in a line can assume any value independent of its parameters and system variables (voltage and angles). Because of this, expansion solutions obtained from TEP models employing this model can be suboptimal or may sometimes be incompatible with the original network system. Hybrid network models combine both DC and pipeline models, and are generally better than the pipeline models in terms of accuracy. Some other network models, formulated by relaxing or linearizing the AC network model, have been proposed recently by researchers [34]–[36], [37], [38]. Authors claim that their models can bridge the AC and the DC network models, yet, their applicability in large-scale networks have not been demonstrated.

From computational viewpoint, the network models reviewed here have different computational requirements. Generally, the higher the fidelity level is, the more accurate the solution is but the higher the computational burden is. In Chapter 3, different TEP models formulated using these models and their variants are further reviewed and compared in terms of solution accuracy and their computational requirements, from which some conclusions are drawn.

2.3.   SOLUTION METHODS IN TEP

The solution methods employed in TEP can be generally classified as exact and non-exact methods.

2.2.1. Exact Solution Methods in TEP Optimization

The TEP problem is formulated into a constrained mathematical optimization with a certain objective function which is then solved by making use of pure mathematical procedures and algorithms. The solution obtained should therefore satisfy several technical, economic, and reliability criteria constraints imposed in the optimization process. As early as 1970, authors in [39] and [40] proposed mathematical optimization techniques using linear programming and dynamic programming, respectively, to solve the transmission expansion problem. The vast literaute on the TEP problem is dominated by mixed integer linear programming (MILP) optimization which embeds a DC network model, as in [41]–[43]. In general, the solution approaches adopted in such problems can be categorized as convex and nonconvex optimization techniques. The first category includes linear programming—LP [39], [44], [45], [46], mixed integer

11

linear programming [41]–[43], and quadratic programming—QP [47], [48]. Non-convex optimization techniques include nonlinear programming—NLP [49], [50] and mixed integer non-linear programming [51], [52]. There are also solution techniques that can exhibit characteristics of both categories such as dynamic programming [53], [40], [54], decomposition techniques [55], [55]–[57], and branch and bound methods [58], [59]. Other mathematical optimization techniques like Benders [60], [61], [62] and hierarchical decompositions [63], [64] have been also extensively used. Unlike others, LP- and NLP-based TEP optimization models do not take account of the lumpiness of investments because the investment variables are relaxed to assume continuous values.

A MILP model has been formulated to solve a long-term TEP problem in a competitive pool-based electricity market by maximizing social welfare and considering uncertainties in electricity demand [65]. The work utilizes a set of decision-making metrics such as changes in aggregate social welfare, generator surplus, demand surplus and merchandizing surplus to obtain an optimal TEP solution, as a guide to make investment decisions. Similarly, a static MILP for long-term TEP model based on disjunctive formulation, incorporating losses and N-1 security criteria has been developed in [66]. Authors in [66] characterize uncertainties due to contingencies and inflows to hydropower plants by using multiple scenarios.

Authors in [36] have introduced the concept of transmission expansion with 'redesign'. It is based on the notion of "*a transmission network may be efficient after cutting off some of its circuits*". Thus, a MILP TEP model based on disjunctive formulation has been developed where all transmission lines including existing ones are taken to be as candidates, while the cost of cutting off a line is considered to be zero. Authors have also presented a fair comparison of the disjunctive model with other variants of TEP mathematical formulations in terms of their performances. The analysis has also included N-1 contingency and a discussionon on how to handle uncertainties in demand and generation. TEP and network switching problems have been developed into a combined MILP problem in [67] and authors indicate that "*there can be some savings upon switching off some lines in a system*".

Acknowledging the complexity of the problem (and/or being motivated by the structure of the TEP problem), some researchers have resorted to the use of mathematical decomposition techniques to enhance its tractability and "speed up" the solution process. Benders decomposition technique is especially the most commonly used approach in TEP studies [55], [56], [61], [68]. Reference [68] presents a methodology to increase the robustness of TEP solution by incorporating a detailed contingency analyses (adequacy and N-1 security criteria), and considering uncertainties in load and wind generation via Monte Carlo simulations. Authors in [64], [69] develop a bi-level mathematical programming, where the TEP problem is split into two levels: upper and lower levels. The upper level minimizes the investment cost; whereas, the lower-level maximizes aggregate social welfare for a given investment decision (obtained from the upper level). The duality theory is employed to link the two levels.

The work in [70] proposes a mathematical method based on a network topological synthesis to investigate the impact of power-flow patterns on transmission planning in a competitive market environment. TEP based on some econ.omometric approaches has also been reported in the literature. Decision analysis scheme based on min-max regret criteria in future plan has been methodically employed to make the TEP solution robust and flexible enough in the face of uncertainty. The work in [71] adopts real options analysis. The main idea behind this approach stems from evaluating the worth and the risk of transmission expansions by constructing *binomial trees* (and *Monte-Carlo Simulation* as a second approach) to represent all possible paths for investments.

In general, the solution obtained by exact solution methods is usually accurate, which can be regarded as one of the advantages of using such solution methods. If the problem is fully convexified, global optimality is guaranteed within a finite simulation time. However, the use of such solution methods in complex power systems may be sometimes complicated.

2.2.2. Non-exact TEP Solution Methods

The complexity and combinatorial nature of the TEP problem prompted researchers to seek for various heuristic [59], [72] and meta-heuristic optimization methods [73]–[76] that can provide an expansion solution within a reasonable simulation time. Heuristic methods, mainly based on sensitivity analyses or invented engineeric methods, are often used when the structure or size of the problem makes it impossible or prohibitively expensive to use exact solution methods. Metaheuristics improve the performance of low level heuristic algorithms by employing higher level algorithms that increase the chance of avoiding or escaping locally optimal solutions.

In [39], a heuristic method is proposed to form fictitious overload paths in the network. Then, the approach makes use of guiding numbers to penalize those without initial transmission lines. Heuristic procedures based on sensitivity analysis are also proposed in [77]. This methodology has been later extended to multistage TEP with constructive heuristic algorithm applied to the problem [78], [79]. Least-effort algorithm has been also proposed in [80] where a heuristic index tries to identify the circuits that provide better power-flow distributions in the system. In [81], flow sensitivity-based TEP has been proposed where the expansion decision has been made based on a value given by the ratio of cost of a line and flow distribution factor (sensitivity across a corridor). In [82], a model for a static long-term TEP is developed, and possible investments are heuristically ranked in accordance with their effectiveness in increasing the system's load supplying capability or reducing unserved power. Similar approaches were applied to short-term TEP models in [20], [59].

In [83], a heuristic static TEP model with an objective of minimizing aggregate investment and operation costs has been developed. In this work, integer expansion variables have been represented by continuous sigmoid functions and the expansion

13

decision have been made using a heuristic model based on a sensitivity given by the values of the sigmoid function. In [72], an expansion decision has been made by analyzing the heuristic ratio of load shedding reduction as a result of investment and the investment cost of the line under consideration.

Meta-heuristic optimization methods have been also widely applied in TEP in a bid to further tackle the computational burden of TEP problems and improve the solution accuracy by avoiding local optima which is thought to be a common problem with heuristic solution techniques. These methods are often inspired by nature. The most common ones are simulated annealing (SA), genetic algorithm (GA), tabu search (TS), game theory (GT), expert systems (ES), fuzzy set theory (FS), ant-colony optimization (ACO), particle swarm optimization (PSO) and greedy randomized adaptive search procedure (GRASP) [84]. Meta-heuristic methods integrate the features of optimization and heuristic methods. Compared to heuristic methods, meta-heuristics usually yield high quality solutions within a relatively lower computational time.

The literature in this area includes neural network hybridized with genetic algorithm [85], genetic algorithm [86], [87], [88], [89] and [90], differential evolution algorithm [91]-[92], tabu search [93], [94], greedy randomized search algorithm [60], [76], simulated annealing [95], ant colony optimization[96], [74], particle swarm optimization [97]-[98], chaos quantum honey bee algorithm [99], expert systems [100] and scatter search [101].

The concept of object-oriented programming paradigm has been applied to model dynamic TEP in a deregulated environment [102]. Reference [103] reports a method based on evolutionary programming for solving a MINLP TEP problem which minimizes aggregate cost: investment, generation and penalty for unserved power. The proposed solution method has been compared with other methods such as GA, TS and SA.

Genetic algorithm has been applied to solve a least-cost and reliability base TEP problem in [104]. The work in [105] also uses GA to solve the same problem and proposes a methodology based on Taguchi's orthogonal arrays to handle uncertainty in renewable generation and demand. Authors in [106] propose a Niche GA (NGA) based algorithm for solving a stochastic MINLP TEP model. In [107], a combination of Benders decomposition and differential evolution algorithm (DEA) has been used to solve a multi-stage MILP TEP model based on a disjunctive formulation. Limited discrepancy local search (LDLS), a tree-search meta-heuristic optimization technique, has been proposed to solve TEP model in [108]. Here, the complex power system is encapsulated in a black-box which is then queried for information about the quality of a proposed expansion. Authors in [108] claim that the LDLS method can be applied flexibly to a power system of any size even if this has not been substantiated in their study. A GA-based 'overload minimization' instead of the classical 'unserved power minimization approach' has been proposed in [109] to solve TEP model. In this case, the fitness function includes investment cost, overload and underload penalties.

14

2.4. TEP in Regulated and Deregulated Power System Structures

Despite the fact that the main aim of expanding the power transmission network in both environments is to better serve growing demand for electricity while satisfying a number of economic and technical constraints, different ways are followed in order to achieve such an objective [110]. For instance, TEP in regulated environment is usually carried out in coordination with generation expansion hence the level of uncertainties is relatively low. Deregulation of the power systems has generally increased the level of uncertainty in the system, introduced additional objectives some of which can be conflictngr, and increased the requirements for transmission expansion problem. Under competitive market structure, the naturally regulated transmission utility needs to provide non-discriminatory access to all the market players and facilitate fair competition [111]. Because of these reasons, TEP in deregulated environment is more challenging than in regulated (traditional) environment. References in [12] and [112] present a detailed review of existing TEP models as well as methods adopted for incorporating uncertainties in a deregulated market environment. In addition, a comparison of centralized vs. deregulated expansion plan, and the need for new methodologies in the restructured power industry has been pointed out in [110].

The literature on TEP in regulated environment, whose objective is to meet the demand while satisfying certain reliability and quality standards, includes [39], [113], [46], [17], [114], [115], [60]-[61], [55], [77], [80], [30], [83], [116], [88] [100], [117] and [118]. Previous works on TEP in a deregulated environment include [62], [65], [70], [119] and [120] is not only very complex to solve but also usually accompanied by high uncertainties in load, generation and market associated with price volatilities. Methods employed for handling such uncertainties introduced as a result of deregulation fall either into deterministic or non-deterministic approaches. Deterministic approaches are usually based on the trivial worst-case-scenario analysis while non-deterministic approaches such as probabilistic load flow, probabilistic based reliability criteria, scenario techniques, decision analysis (a method for dynamic programming), fuzzy decision-making, etc have been employed [111]. A detailed review of issues related to uncertainty management in TEP is presented in Section 2.7.

Authors in [121] propose a mid-term transmission expansion model in a liberalized electricity market with an objective of maximizing the aggregate benefits of the whole system and considering power exchange deviations, N-1 security criteria and unsupplied power. Investment decisions are taken based on computing and analyzing investment sensitivities which are determined from dual variables and reduced costs as a result of the investment.

Authors in [122] propose a meta-heuristic based static TEP model in restructured power industry which included N-1 security criteria, and uncertainties in demand as well as in generation and consumer bids. The model minimizes investment and congestion costs and uncertainties are handled via Monte Carlo simulations [122]. A congestion driven TEP model is proposed in [123], [124] in the context of restructured markets. In [112]

15

and [125], a multi-stage TEP model is proposed for deregulated environments. Simulated annealing algorithm is then used to solve the resulting problem, and fuzzy models incorporate uncertainties due to load evolution along a planning horizon, and system component availabilities.

2.5. OBJECTIVE FUNCTION OF TEP

Traditionally, the objective of TEP has been to minimize the investment cost of lines subject to a number of operational and technical constraints [55], [126]–[128]. In other words, a centralized approach of TEP is mainly to meet current and future demand with adequate reliability and at a reasonable cost. However, the continuously changing environment of power systems is forcing reconsideration of this approach. In a deregulated environment, TEP has to satisfy multiple objectives set by the regulatory body and/or other stakeholders, some of which can be conflicting. In addition to minimizing the investment cost of lines, TEP in a deregulated environment should aim to provide non-discriminatory access, create a conducive environment for fair competition, increase network reliability, meet the demand at a minimum operation cost possible, mitigate transmission congestion, minimize risk, increase operation flexibility, and minimizing environmental impacts among others. From this context, the objective function used in TEP problems in the literature include minimizing operation and investment costs [22], [42], [56], [118], [129]–[131], costs of operation, investment and load shedding [16], [94], [125], [132], [133], congestion and load shedding costs [134], [135], and maximizing welfare [56] among others.

2.6. TEMPORAL SCOPE OF TEP

From the temporal scope of planning, TEP can be categorized as static and dynamic. Static planning framework does not recognize the dynamic nature of the problem; a single target year is instead considered, for which the optimal expansion solution is determined. All investments are assumed to be made in the same year. The fact that decisions can be postponed is not acknowledged in such framework, answers only the TEP questions of where and what investments are to be made on the system. Majority of the literature falls into this category, some examples are [16], [22], [25], [42], [55], [56], [94], [126], [128], [129], [134]–[137]. In contrast, dynamic planning involves a multi-year decision framework, emulating the dynamic nature of the problem. Dynamic planning obtains not only the type and the location of investments to be made but also the timing of each investment. Recent works on dynamic TEP include [23], [27], [118], [125], [127], [130], [132], [133], [138]–[140].

The dynamic planning is a more orthodox planning framework than the static counterpart, and it generally leads to a better expansion solution at a lower cost when compared to the static planning approach. However, dynamic planning is very complex, requiring higher computational effort. To overcome this, some researchers employ meta-heuristic approaches such as GA [118], SA [125], ordinal optimization [140] and others.

2.7.   TREATMENT OF UNCERTAINTY AND VARIABILITY IN TEP

Traditionally, TEP has been carried out deterministically, often for the worst-case scenario (peak-load) in many power systems in a centralized approach. Even if network investments have been often oversized to meet the worst case scenario, the deterministic approach has been operational in most cases where the system's conditions are relatively predictable. However, recent developments in the power industry (deregulation, increasing level of variable energy sources, etc.) have increased the level of uncertainty and variability in the system, and made it impossible to exactly distinguish what the worst-case scenario is. All this adds extra complexity to the decision-making process of grid expansions.

The vast literature on TEP is base on deterministic planning, but the review in this section is limited to the techniques applied to address some of the limitations of deterministic planning models by considering the effects of uncertainty on TEP solutions. So far, various methods have been employed for managing such uncertainty in network expansion planning problems. A comprehensive review of some of the techniques adopted for modeling uncertainty in such problems can be found in [18]. Authors in [141] also excellently present the uncertainty management techniques so far suggested or applied by researchers in the generic subject of energy systems. In the context of TEP, the techniques can be generally classified into probabilistic, stochastic and parametric methods, depending on how uncertainty is described in the input parameters.

The first category includes probabilistic-power-flow and probabilistic-reliability-based methods. Both methods are based on sensitivity analyses, which are often carried out by varying one uncertain input parameter at a time. But it may also include the combined variations of several uncertain inputs. Either way, to perform sensitivity analysis, all uncertain input parameters considered should have known PDFs so that some instances of the corresponding parameters can be sampled. The PDFs themselves are approximated from the respective historical data of uncertain parameters.

The principal goal of probabilistic approaches is to estimate the statistical parameters (e.g. mean values, variances, PDFs, etc) of relevant output variables such as the network expansion solution, the combined investment and operation cost, the loss of load probability, and the expected level of unserved energy. This can be achieved either numerically, analytically or a combination of both. Monte-Carlo simulation (MCS) is the most widely used numerical approach in estimating the PDFs of output variables. It involves an iterative process including generation of samples and running simulations. First, a sample containing realizations of the uncertain input parameters involved is generated using their respective PDFs. Second, considering this sample as an input, a deterministic optimization is run and corresponding values of random variables of interest are computed and recorded. This process is repeated until a sufficiently large number of samples are computed for the random output variables of interest so as to

estimate their PDFs. Note that during the iterative process, sampling can be carried out either sequentially or non-sequentially depending on the type of MCS used.

In particular, MCS has found wide applications in a TEP optimization framework to deal with various sources of operational uncertainty. For example, the authors in [12] developed a market-based TEP model which embeds MCS for generating different samples using the PDFs of various random inputs (load, component availability, generator bid prices and wheeling transactions), and then ultimately compute PDFs of locational marginal prices (LMPs). Similarly, uncertainties related to load, renewable power generation, fuel and emission allowance prices are considered in [142]. And the solution approach used in [142] combines optimization with MCS. In [129], uncertainty in CO_2 emission price is accounted for and simulated using MCS while other sources of uncertainties are largely ignored except demand uncertainty, which is represented by only two load levels. In [143], uncertainties associated to load and wind power generation are also simulated using MCS. In addition, the forced outage rates (FORs) of individual lines are used to randomly simulate line contingencies. Similarly, authors in [144] also consider uncertainties in load and wind power generation, as well as line outages using the so-called deterministic N-1 analysis. Even if correlations among most uncertain parameters naturally exist, they are not explicitly modeled in [143], while in [144], only a correlation factor of 0.75 is assumed between two wind speed regimes considered in the analysis. In [145], load uncertainty is considered and MCS is used to generate a large number of samples which are later reduced by employing a scenario reduction technique. In [146], MCS is used to simulate load and market price uncertainties and estimate the probability distribution of the adaptation cost of various candidate plans. The authors in [147] also use MCS to include wind power production uncertainty under a large-scale wind integration framework; and this work is further extended in [148] to include uncertainty in solar power production. Generator and transmission line availabilities are simulated using MCS in [149], and reference [150] applies a similar approach to handle uncertainties associated with load, generator and line outages. Since MCS is a generic approach in uncertainty handling, it has been widely applied in other fields than TEP. For instance, in a generation expansion planning framework, MCS has been used in [151] to capture uncertainties related to fuel prices, generator availability, and availability and price of electricity imports and exports. Also, in a unit commitment problem, [152] considers wind power generation uncertainty, and an MCS based on Latin hypercube sampling has been used to generate sufficiently large wind power output samples. Then, a conventional sample reduction algorithm is applied to reduce the size of the samples.

Generally, the MCS approach can be feasible in small-scale problems or when a small set of uncertain input is considered. However, it is worth noting here that the applicability of MCS-based analyses is limited in long-term TEP problems with high RES generation due to the following reasons. First, it naturally requires too many optimization runs which may cause long execution times before estimates of PDFs corresponding to the variables of interest are obtained. Second, the high level

uncertainty in such problems, and the correlations among uncertain parameters, further makes the MCS intricate and computationally expensive. Thus, the MCS approach may not be a practical and viable option for such huge problems. In an effort to overcome computational issues (convergence problem, in particular), variance reduction techniques such as importance and stratified sampling (e.g. see [146]) are sometimes applied to drastically reduce the samples required to estimate the PDFs of the output variables. However, since the variance reduction process is applied prior to the TEP optimization, it is very difficult to draw conclusions about whether the considered samples are reasonably good representatives of all the samples that are discarded, particularly from the network expansion strategy viewpoint. In some cases, in an effort to reduce the computational burden, all uncertain input parameters are simply replaced by their expected values, and subsequently, a deterministic mathematical problem (which can be stated as the expected value problem) is solved. However, this again could result in poor solutions, as it is very conservative and may not be well suited to extreme situations.

Analytical methods are also applied as alternatives or complements to MCS approaches. They are used to systematically approximate the statistical properties of random output variables of interest, which are themselves functions of one or more random input variables. These include methods such as cumulant, Gram-Charlier expansion, Taylor series expansion, first-order second-moment method, point estimation methods (PEMs), etc [153]. In comparison to MCS approaches, the analytical methods (the PEMs, in particular) are generally claimed to yield comparable results to MCS-based techniques with lower computational effort. But their merits highly depend on the dimension of the input uncertainty set considered. Intuitively, the higher this dimension is, the higher the computational cost will be, and it is harder to estimate the statistical behavior of random output variables. Furthermore, the assumptions and mathematical simplifications (e.g. linear approximations) required by most of these methods to simplify the problem may render non-negligible inaccuracy.

Among the aforementioned analytical methods, PEMs have been applied in TEP to the estimation of PDFs of certain output variables. For example, PEM is used in [23] to account for uncertainty associated with load and wind power output. A two point estimate method (2PEM), a variant of PEMs, is adopted in [154] to quantify uncertainty in transfer capacity by considering uncertainties in network parameters and bus injections. The same approach is further extended in [134] to handle wind power output uncertainty in a TEP problem incorporating large-scale wind power. The work in [135] presents an experimental analysis to show the versatility of the PEM-based approach theoretically developed in [134]. The authors in [155] use another variant of PEM approach, called 2-micro PEM, to handle uncertainties in load and wind power generation and estimate PDFs of desired output variables. The idea of PEM is to represent each uncertain input by its first statistical moments (e.g. mean, variance, skewness and kurtosis) and concentrations in either side of the mean value. In most cases, the mean and other two values (one below and another above this mean) are used,

which means this would require running 3^w deterministic optimizations (where w is the total number of uncertain parameters considered). The 2PEM even considers only two concentrations selected from either side of the mean (which may not necessarily be symmetric)For example, 2PEM is used to represent uncertainties related to transfer capacity [154] and wind power output [134]. This reduces the computational effort significantly, but with a relevant loss of accuracy. In the same way as in MCS approaches, the optimization problem is run a number of times (but with largely lower number of iterations than that required by MCS). This way, the expected value and higher order moments of output variables are determined, and using analytical methods, their PDFs are then estimated. Note that the number of iterations required in PEM-based approaches depends on how the uncertain input parameters are represented. For example, the 2PEM requires 2^w iterations where w is the total number of uncertain parameters.

While analytical methods based on PEM may be successfully applied in small (even medium-scale) problems with a few uncertain parameters, and may deliver useful estimates of PDFs, their application becomes of limited use (or computationally expensive) when the scale of the problem and the level of uncertainty under consideration increase (such as in long-term TEP problems with strong penetration of RESs on large-scale systems). Moreover, the existence of both spatial and temporal correlations among random input variables complicates the practical application of PEM in TEP.

Sometimes, a combination of MCS and analytical methods is used in power system analyses. In particular, the authors in [156] combine MCS with analytical methods to account for uncertainties in load and wind power output. They first use MCS to estimate the PDF of wind power output. Then, discrete samples of the wind power output are simulated by combining analytical and probabilistic methods in a chance-constrained TEP framework. The authors claim their approach is computationally more efficient than MCS. Similarly, the combination of MCS and PEM is also adopted in a two-paper work [134] and [135] when considering uncertainties in load and wind power generation. The authors in [157] develop an analytical methodology to consider uncertainty in wind power generation and generator availability. Their methodology is compared with MCS, and it is reported that the results obtained using both approaches largely coincide.

Stochastic methods, on the other hand, assume a given number of operational states is available, each one with a certain probability. All these operational states are then jointly considered in the analysis, the outcome being the expected values of relevant output variables. The quality of the solution based on this approach depends on how thoroughly the operational situations are explored and how representative the selected operational states are. In general, a good TEP solution is computed when a large number of operational states is considered. Nevertheless, this increases the computation burden. Because of this, the number of operational states must be significantly reduced before the stochastic programming model is run by using certain algorithms such as forward

and backward selection [158] (for example, see the previous works in the context of substation and TEP [145], joint generation and TEP [149], generic TEP [150] and power management [159] problems). Very often this number is predetermined. For instance, the authors in [160] represent the uncertainty related to pool price by considering three realizations of pool price per day corresponding to three periods, each with an 8-hours interval. But the number of operational states to be considered can also be iteratively estimated by monitoring some accuracy indices (for example, see [161]). The authors in [162] develop a methodology based on the roulette wheel technique to randomly generate a large number of samples with certain probabilities of occurrence, and employ backward scenario reduction algorithm before a stochastic optimization problem is solved. This technique uses the corresponding PDFs of load demand and wind power generation. In [163], uncertainties in demand and fuel price are modeled using a binomial Markov chain as a stochastic process. The work in [164] only considers uncertainty in CO_2 allowance price, and the "carbon" price uncertainty is modeled via samples generated from a set of PDFs obtained using Geometric Brownian Motion and MCS.

Under the auspices of stochastic methods, although not common, the variability of uncertain parameters may be individually aggregated to a predefined number of values with approximate probabilities or weights, such as the load aggregation technique in TEP [129] and joint substation and TEP problems [145]. In stochastic methods, data-mining techniques are also applied to drastically reduce an initially large number of operational states to be considered, prior to running the optimization. These include different supervised and unsupervised clustering techniques used in contingency and reliability [165] and electricity supply analyses [166]. In [161], authors use such techniques in order to take into account the uncertainties related to wind power output and load in a stochastic TEP model, a reduced number of clusters are formed from a two-dimensional random input dataset (i.e. containing load and wind power output series). Here, the input datasets themselves are generated using Gaussian copula, a multivariate probability distribution capable of describing the dependence of random variables. Instead of working with a fixed number of clusters (like in traditional clustering), the authors in [161] adopt a mechanism to iteratively determine the minimum number of clusters needed, by increasing the number of clusters until the marginal improvement in the objective value is sufficiently small. Note that, in addition to working on data series generated using joint PDFs, it is also possible to perform the clustering process on historical data samples (if available), forecast data series or samples generated from individual PDFs.

In general, the clustering algorithm uses uncertain parameters (the causes) as clustering variables (hereinafter, clustering based on causes or CbC). The entire clustering process involves grouping "similar" snapshots together, selecting representative snapshots per cluster and assigning probabilities to each one of them. Similarity is measured by the distance among snapshots in the uncertain input space. However, such clustering is not efficient because it is carried out without acknowledging the effects of the snapshots on

21

the target problem. This significantly conditions the outcome of the optimization, especially in the context of TEP. As it shall be explained in detail in the following Chapters, clustering based on effects (CbE), which is advocated by this work, is superior to CbC.

Both the probabilistic and the stochastic methods depend on the availability of historical data or PDFs of random variables. But sometimes the relevant random parameters may not have sufficient information (historical data) to formulate their PDFs. This is especially the case in deregulated power systems where information asymmetry is rampant. Inspired by such knowledge gap, parametric (non-probabilistic) methods [167] such as info-gap decision theory (IGDT), robust optimization (RO) and fuzzy systems theory (FST) are used to systematically account for random as well as non-random uncertainties. They all model uncertainty by characterizing the uncertain input parameters' space using parametric ranges, i.e. by forming parametric input datasets such as polyhedral (formed by upper and lower bounds of uncertain parameters), ellipsoidal (an approximate uncertainty space), etc.

As stated above, the IGDT tool is inspired by the severe lack of information about uncertain parameters. It requires only the definition of ranges of uncertain parameters over which the parameters may have certain values, which can be seen as an advantage over probabilistic and stochastic methods. In general, IGBT-based TEP models such as [168] seek robust solutions in the face of severe uncertainty, where the robustness of the solution is measured by "immunity" to a range of operational situations defined by the uncertainty set. But the theory itself has been the subject of strong criticism [169], citing its weaknesses such as conservativeness, localized and poor solution approximation, etc. Instead, robust optimization has been praised as a good alternative to decision making under severe uncertainty [170], [171].. The concept of RO is similar to IGDT. Like in IGDT, uncertainty comes from a known uncertainty set. In [172], uncertainties in renewable power generation and demand are considered and represented by their corresponding uncertainty sets in an RO-based TEP model. The solution obtained by RO should, in principle, be robust under the worst-case situation in the uncertainty set, which also makes RO highly conservative. Reference [173] presents a slight modification to ordinary RO by adding features to minimize conservatism i.e. by characterizing uncertainties using ellipsoidal constrained uncertainty sets and incorporate correlation factors of considered uncertain parameters by means of the variance-covariance matrix. Recently, there are also some ongoing research works (e.g. adjustable RO in [170]) to address the conservativeness of RO. These normally work by adjusting the uncertainty sets depending on how much uncertainty one desires to capture. But RO still remains to be a hot research area in mathematical optimization which requires further refining to solve robust problems.

On the other hand, FST is inspired by linguistic expressions such as "high", "medium", "low", etc. Each uncertain parameter is considered as a fuzzy variable and is represented by a certain membership function (often a trapezoidal membership function). For example, generator and consumer bid prices are modeled using this

22

method in [174]. Some of the major disadvantages with FST-based methods is the absence of clear guidelines to select appropriate membership functions, and the fundamental difficulty to prove the accuracy of the solutions.

A long-term TEP problem, with investment horizons spanning over 30 or more years and increasing RES penetration, demands extensive management of uncertainty and operational variability. This is the subject area of this thesis. The subsequent chapters describe in detail how the two types of uncertainties (random and nonrandom, according [12]) are handled. In general, high level uncertainties (also labeled as random in [12]) are modeled by considering a sufficiently large number of operational situations, also known as "snapshots" in this thesis. Then, a new clustering methodology based on moments technique [14], a tailor-made approach for TEP problems, is then used to substantially reduce the original set of snapshots by grouping them into a predefined number of clusters. The low level uncertainties on the other hand are characterized by a number of storylines (probable future scenarios), each with an estimated probability of realization.

2.8. SIGNIFICANCE OF THIS RESEARCH

There is no question about the importance of TEP in every electric power industry. It has always been mandatory, as a part of the changes needed to face the ever increasing demand for electricity within a reliable operational frame [16]. The power industry is expected to further experience rapid changes and transformations to meet environmentally-friendly, sustainable, secure and affordable energy to growing demand for electricity. Nowadays, there is a general consensus that this objective can be achieved by aggressively promoting the deployment of renewables, particularly variable energy sources, in particular. In the coming decades, because of the aforementioned techno-economic and environmental reasons, large amount of such energy resources are expected to integrated in power systems. However, such resources are often abundantly available in remote locations broadly dispersed across a geographically wide area, and far away from major demand centers. This will require huge transmission investment needs. This poses a huge challenge for network planners because the complexity of the resulting TEP problem. Both the size of the system and the level of uncertainty are huge. The intermittent nature of variable energy resources (such as wind and solar) also introduces significant uncertainty and variability to the system, further complicating the TEP problem. Exiting transmission networks should be largely reinforced and expanded to balance the extra operational uncertainty introduced by such energy sources. In general, solving a TEP problem for such a big system under high levels of uncertainty demands an exceptionally huge computational effort when using reasonably precise network expansion models.

In the European context, for instance, there is a huge potential of large-scale wind power in the North Sea, West Coast and Baltic areas. In addition, a huge amount of solar power is expected to be imported from the Middle East and North Africa (MENA) [9]. Under these circumstances, a pan-European electricity network expansion [9] has to

be adequately reinforced and expanded to support full integration of these large-scale RESs and fully ship the power generated from such sources to meet the ever increasing demand for electricity [11]. This is an extremely challenging task because of the unprecedented geographical, temporal and uncertainty scope, albeit there is a need for TEP tools to help in such a complex decision-making process.

Despite the extensive literature on TEP, the problem still remains very challenging especially with network instances of this magnitude. In other words, there has been little progress on a TEP problem which considers large-scale integration of RESs on a network of continental or intercontinental size. The literature is vastly composed of solving TEP problems in small- to medium-scale networks. In such networks, introducing any level of complexity (in modeling, solution strategy or both) may be affordable, but this is not the case with networks of a continental size. The size of the TEP problem have been getting more complex because of the ever-increasing size of the networks being dealt with, uncertainties growing from time to time, etc., increasingly becoming computationally demanding. As a result, traditional solution strategies and more detailed TEP models (such as the AC power flow based one) are no longer computationally affordable. Technically speaking, the geographically wide TEP optimization model should be as simple as possible to make sure that the problem is computationally and practically tractable but at the same time it should deliver reliable and robust solutions. In general, currently adopted TEP models and solution approaches cannot be seamlessly applied to such a huge problem, principally due to their computational limitations. Computational complexity of a TEP problem dramatically increases with the size of the network dealt with. Unless handled systematically, pursuing optimal expansion solutions in continental scale TEP problems such as the EU network deems to be impossible. This justifies the fact that the problem needs to be approached in a way different from the conventional ones. In general, new "*algorithmic and computational methods are needed to address (1) the high dimensionality of an optimization problem having a long decision horizon, large geographic scale and high uncertainty; (2) a need to provide solutions in terms of tradeoffs among multiple objectives; and (3) the discrete nature of the investment decisions*" [175].

In view of the complex nature of the problem, this thesis proposes a global strategy, methods and tools to solve this kind of problems, as outlined and discussed in [13], [14], [15]. This strategy comprises:

- Successive optimization problems, that reduce the space of combinatorial solution search while gradually using more detailed and accurate models.
- Multi-stage planning, to find short-term decisions that consider long-term scenarios.
- Two-period planning framework to combine short-term decisions and long-term strategies.
- Stochastic models combined with alternative deterministic storylines.
- Mathematical programming, empowered by heuristic and expert knowledge.

24

- Effective methods for handling the vast uncertainty and variability inherent to such a problem.
- Network models that adequately capture the physical characteristics of the network system.

3

III. MODELING ASPECTS

Both the tractability of a TEP problem and the accuracy of an expansion solution largely depend on the level of system details captured by the expansion model. This is associated with the characterization of physical network variables, in particular, flows and losses. This chapter thoroughly reviews a number of TEP models, commonly used in grid planning studies, in terms of accuracy and mathematical complexity levels. In addition to the systematic comparisons of various existing models both theoretically and numerically, this chapter contributes some improvements to the mathematical modeling of existing TEP models.

3.1. CHAPTER OVERVIEW

The transmission grid is the backbone of any power flow analysis, planning and operation. This is particularly indispensable in TEP studies because the expansion solutions are conditioned by the topology, the strength and the level of modeling details of the network which constitutes existing and candidate lines. Therefore, modeling the network (grid) should be an integral part of any TEP study. As extensively reviewed in Chapter 2, TEP models based on a number of network models, each with a different fidelity level, have been adopted. Here, the context of fidelity should be understood as the extent to which the physical characteristics of the system are captured. This Chapter reviews some of the most commonly used network models in the context of TEP studies, and discusses in detail the pros and the cons of each one from modeling complexity and computational performance.

3.2. TEP MODEL FIDELITY—THEORETICAL VIEW

The TEP problem can be considered as an optimal power flow problem consisting of a number of discrete constraints. A number of existing TEP models as well as improved and new ones are reviewed and discussed here. Theoretical and experimental comparisons of the different models are also presented. This is motivated by the conflicting accounts of existing network models in the literature [176]–[179] as well as by the need to build the right network model that balances the tradeoff between accuracy and computational requirement. Note that, for the sake of simplicity, a number of notations are suppressed.

3.3.1. An AC based TEP Model (ACTEP)

Current transmission networks are predominantly AC systems. The ACOPF, which is based on the customary AC power flow equations (1) and (2), employs the most accurate network model but the resulting optimization problem is highly nonlinear and nonconvex.

$$P_k = V_i^2 g_k - V_i V_j (g_k cos\theta_k + b_k sin\theta_k) \tag{1}$$

$$Q_k = -V_i^2 b_k + V_i V_j (b_k cos\theta_k - g_k sin\theta_k) \tag{2}$$

An ACOPF-based TEP (ACTEP) model minimizes a user-defined objective function as in (3) subject to a number of technical constraints given by Equations (4) through (15). Note that extending the ACOPF problem to a TEP problem only requires adding the

discrete variables to the power flow equations in (4)—(7) and corresponding capacity constraints in (8) and (9). Note that the flow equations related to existing lines are generalized to indicate the switching statuses/utilizations of the lines. Equations (4) and (5) represent the active and reactive power flows in existing lines, respectively; whereas, Equations (6) and (7) are the corresponding flows in candidate lines. The flow limits for existing and candidate lines are given by (8) and (9), respectively. Equations (10) and (11) represent the active and the reactive power balances at each node i.e. Kirchhoff's current law (KCL), respectively. Equations (12) and (13) provides the permissible bound for the active and the reactive power generation of a unit, respectively. Voltage and angle bounds as well the corresponding reference values are given by (14) and (15), respectively. As it can be seen, the resulting ACTEP model is a mixed integer programming (MINLP) problem, which is highly non-linear and non-convex. According to computational complexity theory, MINLP problems are regarded as NP-hard or even NP-complete problems [180], [181]. Generally, despite there are some advances in MINLP solvers in recent years, employing the AC flow equations in power system planning applications (especially for large-scale TEP problems) is yet increasingly difficult. For this reason, ACTEP is rarely employed in the literature. The few ACTEP models proposed in the literature are practically limited to small-scale systems, and often use heuristic and metaheuristic methods for solving the resulting problem. For instance, authors in [20] propose a constructive heuristic algorithm, guided by interior point method, for solving an ACTEP problem. Reference [182] proposes a genetic algorithm for solving a similar problem while Benders decomposition is applied to an ACTEP problem by decomposing it into a master involving only integer programming problem and a sub-problem with a nonlinear programming nature.

As mentioned in the previous Chapter, the heuristic and metaheuristic solution methods neither guarantee optimality nor give a measure to the optimal solution. Equations

$$Minimize \quad Z = Objective\ Function \tag{3}$$

Subject to:

$$P_k = u_k \left(V_i^2 g_k - V_i V_j (g_k cos\theta_k + b_k sin\theta_k) \right) \tag{4}$$

$$Q_k = u_k \left(-V_i^2 b_k + V_i V_j (b_k cos\theta_k - g_k sin\theta_k) \right) \tag{5}$$

$$P_k = z_k \left(V_i^2 g_k - V_i V_j (g_k cos\theta_k + b_k sin\theta_k) \right) \tag{6}$$

$$Q_k = z_k \left(-V_i^2 b_k + V_i V_j (b_k cos\theta_k - g_k sin\theta_k) \right) \tag{7}$$

$$P_k^2 + Q_k^2 \leq u_k S_{k,max}^2 \tag{8}$$

$$P_k^2 + Q_k^2 \leq z_k S_{k,max}^2 \tag{9}$$

$$\sum_{k \in i} P_k + \sum_{g \in i} PG_g + p_i - \sum_{d \in i} PD_d = 0 \tag{10}$$

$$\sum_{k \in i} Q_k + \sum_{g \in i} QG_g + q_i - \sum_{d \in i} QD_d = 0 \tag{11}$$

$$u_g PG_{g,min} \leq PG_g \leq u_g PG_{g,max} \tag{12}$$

$$u_g QG_{g,min} \leq QG_g \leq u_g QG_{g,max} \tag{13}$$

$$V_{min} \leq V_i \leq V_{max} \ ; \ V_{ref} = V_{nom} \tag{14}$$

$$\theta_{min} \leq \theta_i \leq \theta_{max} \ ; \ \theta_{ref} = 0 \tag{15}$$

In some cases, it is suggested that decoupling the products of binary (u_k and z_k) and continuous variables in (4)—(7) by means of disjunctive (also called big-M) formulation as in (16)—(19) may facilitate the computational process. However, choosing suitable values for the big-M parameters is not straightforward as there is no clear guideline so far used for selecting the right values to such parameters that ensure tight relaxations of the original equations. Inappropriate values may lead to numerical difficulties which can further impede the solution process.

$$\left| P_k - \left(V_i^2 g_k - V_i V_j (g_k cos\theta_k + b_k sin\theta_k) \right) \right| \leq MP_k(1 - u_k) \tag{16}$$

$$\left| Q_k - \left(-V_i^2 b_k + V_i V_j (b_k cos\theta_k - g_k sin\theta_k) \right) \right| \leq MQ_k(1 - u_k) \tag{17}$$

$$\left| P_k - \left(V_i^2 g_k - V_i V_j (g_k cos\theta_k + b_k sin\theta_k) \right) \right| \leq MP_k(1 - z_k) \tag{18}$$

$$\left| Q_k - \left(-V_i^2 b_k + V_i V_j (b_k cos\theta_k - g_k sin\theta_k) \right) \right| \leq MQ_k(1 - z_k) \tag{19}$$

Because of the computational issues associated with the AC-based TEP models, a number relaxed ACTEP models [34] have been proposed, and compared in terms of their computational requirement and solution quality. Even if the relaxed models are interesting and demand relatively lower computational effort when compared to ACTEP one, the authors concluded that they are not feasible for large-scale TEP applications. From this perspective, further reductions and mathematical simplifications are needed to solve such problems. Several computationally less-intensive linearized models, with different levels of fidelity and computational complexity, have been employed in TEP applications. They are derived from the AC power flow equations under simplifying assumptions. The most common used models are reviewed below.

3.3.2.　A Linearized AC based TEP Model (LinACTEP)

The formulation of this model, denoted as LinACTEP, includes the objective function (3) and constraints (8)—(15) as well as linearized forms of the AC power flow equations in (1) and (2). The linearization process is based on two practical assumptions, which is explained as follows. The first assumption is concerning the bus voltage magnitudes, which in power transmission systems are expected to be very close to the nominal value V_{nom}. Hence, without loss of generality, a flat voltage profile is assumed throughout the system. The second assumption is in relation to the angular difference θ_k across a line,

which is practically small because of stability reasons, leading to the trigonometric approximations $sin\theta_k \approx \theta_k$ and $cos\theta_k \approx 1$. Note that this assumption is valid in transmission systems, where the active power flow dominates the total apparent power in lines. The LinACTEP model, which is based on the two assumptions and a Taylor series expansion, is first introduced in [131] in the context of transmission expansion planning. In this model, the voltage magnitude at bus i can be expressed as the sum of the nominal voltage and a small deviation ΔV_i, as in (20).

$$V_i = V_{nom} + \Delta V_i, \ where \ \Delta V^{min} \leq \Delta V_i \leq \Delta V^{max} \tag{20}$$

Note that the voltage deviations at each node ΔV_i are expected to be very small. Substituting (20) in (1) and (2) and neglecting higher order terms, one gets:

$$P_k \approx (V_{nom}^2 + 2V_{nom}\Delta V_i)g_k - (V_{nom}^2 + V_{nom}\Delta V_i + V_{nom}\Delta V_j)(g_k + b_k\theta_k) \tag{21}$$

$$Q_k \approx -(V_{nom}^2 + 2V_{nom}\Delta V_i)b_k + (V_{nom}^2 + V_{nom}\Delta V_i + V_{nom}\Delta V_j)(b_k - g_k\theta_k) \tag{22}$$

Note that Equations (21) and (22) still contain nonlinearities because of the products of two continuous variables—voltage deviations and angle differences. However, since these variables (ΔV_i, ΔV_j and θ_k) are very small, their products can be neglected. Hence, the above flow equations become:

$$P_k \approx V_{nom}(\Delta V_i - \Delta V_j)g_k - V_{nom}^2 b_k\theta_k \tag{23}$$

$$Q_k \approx -V_{nom}(\Delta V_i - \Delta V_j)b_k - V_{nom}^2 g_k\theta_k \tag{24}$$

When the investment planning problem includes network switching, reinforcement, replacement and expansion of transmission lines, Equations (23) and (24) must be multiplied by the corresponding binary variables as in (25)—(28). This is to make sure that the flow through an existing or a new line is zero when the associated switching/investment variable is zero; otherwise, the flow in that line should obey the Kirchhoff's law. Note that the models here are generalized to include network redesign (switching) via the switching variable u_k i.e. existing network can be redesigned by cutting off some lines that improve the overall economic efficiency.

$$P_k \approx u_k\{V_{nom}(\Delta V_i - \Delta V_j)g_k - V_{nom}^2 b_k\theta_k\} \tag{25}$$

$$Q_k \approx u_k\{-V_{nom}(\Delta V_i - \Delta V_j)b_k - V_{nom}^2 g_k\theta_k\} \tag{26}$$

$$P_k \approx z_k\{V_{nom}(\Delta V_i - \Delta V_j)g_k - V_{nom}^2 b_k\theta_k\} \tag{27}$$

$$Q_k \approx z_k\{-V_{nom}(\Delta V_i - \Delta V_j)b_k - V_{nom}^2 g_k\theta_k\} \tag{28}$$

The bilinear constraints, involving products binary (u_k and z_k) with voltage deviation and angle difference variables, introduces undesirable nonlinearity to the problem. This nonlinearity can be avoided using the big-M formulation i.e. by reformulating the above

equations into their respective disjunctive equivalents as in (29)—(32). As a rule-of-thumb, the big-M parameter is often set to the maximum transfer capacity in the system.

$$\left|P_k - \{V_{nom}(\Delta V_i - \Delta V_j)g_k - V_{nom}^2 b_k \theta_k\,)\}\right| \le MP_k(1 - u_k) \qquad (29)$$

$$\left|Q_k - \{-V_{nom}(\Delta V_i - \Delta V_j)b_k - V_{nom}^2 g_k \theta_k\}\,\right| \le MQ_k(1 - u_k) \qquad (30)$$

$$\left|P_k - \{V_{nom}(\Delta V_i - \Delta V_j)g_k - V_{nom}^2 b_k \theta_k\}\right| \le MP_k(1 - z_k) \qquad (31)$$

$$\left|Q_k - \{-V_{nom}(\Delta V_i - \Delta V_j)b_k - V_{nom}^2 g_k \theta_k\}\right| \le MQ_k(1 - z_k) \qquad (32)$$

The apparent power flow S_k through a line is given by $\sqrt{P_k^2 + Q_k^2}$ and this has to be less than or equal to the rated value which is denoted as:

$$P_k^2 + Q_k^2 \le (S_k^{max})^2 \qquad (33)$$

Considering line switching/investment, Equation (33) can be rewritten as:

$$P_k^2 + Q_k^2 \le u_k \left(S_{k,max}\right)^2 \qquad (34)$$

$$P_k^2 + Q_k^2 \le z_k \left(S_{k,max}\right)^2 \qquad (35)$$

The quadratic expressions of active and reactive power flows in (34) through (35) can be easily linearized using piecewise linearization, considering a sufficiently large number of linear segments, L. There are a number of ways of linearizing such functions such as incremental, multiple choice, convex combination and other approaches in the literature [183], [13]. Here, the first approach (which is based on first-order approximation of the nonlinear curve) is used because of its relatively simple formulation. To this end, two non-negative auxiliary variables are introduced for each of the flow variables P_k and Q_k such that $P_k = P_k^+ - P_k^-$ and $Q_k = Q_k^+ - Q_k^-$, and by implication $|P_k| = P_k^+ + P_k^-$ and $|Q_k| = Q_k^+ + Q_k^-$. Note that these auxiliary variables (i.e. P_k^+, P_k^-, Q_k^+ and Q_k^-) represent the positive and the negative flows of P_k and Q_k, respectively. Expressing a variable as the difference of its positive and negative parts, which is called a bijection, is widely applied technique in linear programming problems. Bijection guarantees the equivalency the reformulated problem with the original problem, and a proof of this can be found in [184]. Bijection helps one to consider only the positive quadrant of the nonlinear curve, resulting in a significant reduction in the mathematical complexity and by implication the computational burden. In this case, the associated linear constraints are:

$$P_k^2 \approx \sum_{l=1}^{L} \alpha_{k,l} \Delta p_{k,l} \qquad (36)$$

$$Q_k^2 \approx \sum_{l=1}^{L} \beta_{k,l} \Delta q_{k,l} \qquad (37)$$

$$P_k^+ + P_k^- = \sum_{l=1}^{L} \Delta p_{k,l} \qquad (38)$$

$$Q_k^+ + Q_k^- = \sum_{l=1}^{L} \Delta q_{k,l} \qquad (39)$$

where $\Delta p_{k,l} \leq S_k^{max}/L$, $\Delta q_{k,l} \leq S_k^{max}/L$, $\Delta p_{k,l+1} \leq \Delta p_{k,l}$ and $\Delta q_{k,l+1} \leq \Delta q_{k,l}$.

Note that at most one of the two auxiliary variables introduced per active and reactive flows through a line should be zero at a time. This condition is implicitly enforced by the theory of optimality because, as it can be inferred from (38), network losses are a function of $(P_k^+ + P_k^-)$, and should be minimized. Setting both of them to be greater that zero does not only make sense but contradict with the notion of optimality. A small penalty can alternatively be included in the objective function to ensure at most one of them is zero at a time. As shall be described in the following section, this losses model can in some situations result in "fictitious" losses [13]. Several existing and proposed losses models are compared theoretically as well as numerically in Section 3.5 and [13].

The active and reactive power losses in line k can be approximated as follows:

$$PL_k = P_{k,ij} + P_{k,ji} \approx 2V_{nom}^2 g_k(1 - cos\theta_k) \approx V_{nom}^2 g_k \theta_k^2 \tag{40}$$

$$QL_k = Q_{k,ij} + Q_{k,ji} \approx -2V_{nom}^2 b_k(1 - cos\theta_k) \approx -b_k V_{nom}^2 \theta_k^2 \tag{41}$$

Clearly, Equations (40) and (41) are nonlinear and nonconvex functions, making the problem nonconvex and more complex to solve. This can be overcome by having the quadratic angle differences piecewise-linearized, as it is done in [131] by introducing additional binary variables and big-M formulation to avoid unnecessary constraints on the angle differences when binary variable associated to an existing or candidate line is zero. A major disadvantage of the linear models of (40) and (41) in [131] is that the additional binary variables required as well as the introduction of the big-M method increase the complexity to the TEP problem. Instead of doing this, this thesis proposes flow-based losses, which has substantial benefits from the computational point of view, which will be explained shortly. The angle-based losses models in (40) and (41) are expressed in terms of the active and the reactive power flows as in (42) and (43). Note that Equation (42) can be easily obtained by multiplying the squared expressions of both sides of the equations in (23) and (24) by the resistance of the branch, combining the resulting equations, neglecting higher order terms and reordering both sides of the resulting equation. Equation (43) can also be obtained in a similar fashion but by multiplying the squared expressions by the reactance the line. More details about the derivation of Equations (42) and (43) is provided in Appendix A.

$$PL_k = r_k\{P_k^2 + Q_k^2\}/V_{nom}^2 \tag{42}$$

$$QL_{kt} = x_k\{P_k^2 + Q_k^2\}/V_{nom}^2 \tag{43}$$

Note that expressing the losses as a function of flows has two advantages. First, doing so reduces the number of nonlinear terms that has to be linearized, which in turn results in a model with a reduced number of equations and variables. For example, if Equations (41) and (42) are used instead, in addition to the quadratic power flow terms P_k^2 and Q_k^2, the quadratic angle differences θ_k^2 need to also be linearized to make the problem linear and convex. On the contrary, when Equations (43) and (44) are used, one is only required to

linearize P_k^2 and Q_k^2. Second, it avoids unnecessary constraints on the angle differences when a line between two nodes is not connected or remains not selected for investment. This is often avoided by introducing binary variables and using a so-called big-M formulation [131]. However, this adds extra complexity to the problem.

Losses are often treated as "virtual" loads connected to the buses. In this respect, the losses in a given line are equally distributed to the nodes connecting the line. The load balance equations in (10) and (11) should be slightly modified to take account of these changes as in (44) and (45). The line capacity constraints in (34) and (35) may also be extended as in (46) and (47). The quadratic terms in these equations can be linearized in the same way as in the quadratic flow functions. However, even if this is an elegant approach, the additional linear constraints needed to do this leads to further computational complexity. Because of this reason, Equations (34) and (35) are adopted in the analysis throughout this work. Note that the absolute value flow terms in (46) and (47) are replaced by the linear expression $|P_k| = P_k^+ + P_k^-$.

$$\sum_{k \in i} P_k + \sum_{g \in i} PG_g + p_i - \sum_{d \in i} PD_d + 0.5 \sum_{k \in i} PL_k = 0 \tag{44}$$

$$\sum_{k \in i} Q_k + \sum_{g \in i} QG_g + q_i - \sum_{d \in i} QD_d + 0.5 \sum_{k \in i} QL_k = 0 \tag{45}$$

$$(|P_k| + 0.5PL_k)^2 + (|Q_k| + 0.5QL_k)^2 \leq u_k (S_{k,max})^2 \tag{46}$$

$$(|P_k| + 0.5PL_k)^2 + (|Q_k| + 0.5QL_k)^2 \leq z_k (S_{k,max})^2 \tag{47}$$

Computationally speaking, the LinACTEP problem (either lossy or lossless) is relatively less complex when compared with the full ACTEP model. The entire LinACTEP model is a MILP optimization problem, for which efficient and of-the-shelf solvers are available, and optimal solution is guaranteed in a reasonable simulation time.

3.3.3. A "DC" based TEP Model (DCTEP)

This model, which is denoted as DCTEP, is the most commonly used model in technical and economic analyses of complex power systems [32], [51], [79], mainly because of its relatively lower computational requirement compared to the models discussed previously. This TEP model often minimizes a certain objective function (48), and is based on the classic "direct current" (DC) branch flow model in (49) [178]. It is derived from the well-known AC network flows under the simplifying assumptions (i.e. the assumptions related to the unity voltages, and small angular differences across lines) described above in Subsection 3.3.2 and zero resistance. Further details of the DC network model including its full derivation and related details can be found in [178].

The DCTEP model respects constraints related to the Kirchhoff's voltage law (KVL) of existing (51) and candidate (52) lines and the corresponding network capacity limits given by (53) and (54), nodal active power balance (55), the generation limits (56) and the voltage angle related constraints (57). However, the assumptions means that information regarding reactive power and voltage magnitude variations among nodes are not provided.

$$Minimize \quad Z = Objective \; Function \tag{48}$$

Subject to:

$$P_k = -V_{nom}^2 b_k u_k \theta_k \; ; \; where - b_k = 1/x_k \tag{49}$$

$$P_k = -V_{nom}^2 b_k z_k \theta_k \; ; \; where - b_k = 1/x_k \tag{50}$$

$$|P_k + V_{nom}^2 b_k \theta_k| \le M_k(1 - u_k); \; where - b_k = 1/x_k \tag{51}$$

$$|P_k + V_{nom}^2 b_k \theta_k| \le M_k(1 - z_k) \; ; \; where - b_k = 1/x_k \tag{52}$$

$$-u_k S_{k,\max} \; \le P_k \le u_k S_{k,\max} \tag{53}$$

$$-z_k S_{k,\max} \; \le P_k \le z_k S_{k,\max} \tag{54}$$

$$\sum_{k \in i} P_k + \sum_{g \in i} PG_g + p_i - \sum_{d \in i} PD_d = 0 \tag{55}$$

$$u_g PG_{g,min} \le PG_g \le u_g PG_{g,max} \tag{56}$$

$$\theta_{min} \le \theta_i \le \theta_{max}; \; \theta_{ref} = 0 \tag{57}$$

As described in the preceding Section, the bilinear terms in Equations (49) and (50) are separated by the method of disjunctive formulation as in (51) and (52). Sufficiently large values should be selected for the big-M parameters involved in this formulation to make sure that reformulated problem is tight enough and that numerical problems are avoided. The approach presented in [184] can be used to approximate the minimum value for each corridor.

 Basically, the underlining assumptions make the DC model lossless. However, losses are often approximated by the quadratic expression in (40) [178], or some proxy of it, and combined with the DC power flow model. Some of the existing linear losses models (presented in the next subsection) are derived from (40).

Notice that Equation (40) is both nonlinear and nonconvex. In complex problems such as large-scale TEP, linear models are welcome. The expression in (40) could be linearized in order to include losses in DCTEP models. The most common approach in this case is to perform a piecewise linearization of the expression in (40) as proposed in [42], and further applied in formulating a long-term TEP problem of deregulated power systems [65]. As explained before (see Section 3.3.2), the main drawback of such linearization when used in TEP problems, is that angular differences between nodes are inappropriately constrained to be zero for those nodes connected by lines selected for contingency screening or candidate new lines that are not built, since in common piecewise-linearized models, angle differences are formulated in terms of the line flows (zero flow implies equal angles). To avoid this problem, the corresponding linear constraints are reformulated into their respective disjunctive equivalents (described as the big-*M* approach as in [42]) to guarantee that these constraints are not binding for lines that are not built or not operative. However, the big-*M* approach creates some numerical difficulties during the OPF solution process, such as ill-conditioning of matrices

representing the system topology. To avoid the use of the big-*M* approach, losses can be expressed as a function of flows (as in the case of LinACTEP) instead of angle differences. The relationship between a line flow and its losses can be readily derived using the DC flow model equations or directly from Equation (42) by simply neglecting the reactive power flow, which leads to:

$$PL_k = r_k P_k^2 / V_{nom}^2 \tag{58}$$

Unlike angle differences, line flows are bound to be zero in lines that are not built (candidate lines) or not operative (because of contingency screening or maintenance). Another advantage of Equation (58) is its possible application to model losses in HVDC lines or, generally, in lines where flows are independent of the voltage angles at the buses they are connected to. The linearization of the quadratic flow function in (58) is as described in the preceding Section, and includes the constraints given in (36) and (38).

Like in the case of lossy LinACTEP model formulation, losses in each line are treated as "virtual" loads connected to the two end nodes of the line. In other words, losses in a given line are equally distributed to the nodes connecting the line. When formulating a lossy DCTEP model, the line capacity constraints (53) and (54) as well as the load balance equation in (55) need to be slightly modified to take account of the losses as in (59)—(61), respectively.

$$|P_k| + 0.5 PL_k \le u_k S_{k,\max} \tag{59}$$

$$|P_k| + 0.5 PL_k \le z_k S_{k,\max} \tag{60}$$

$$\sum_{k \in i} P_k + \sum_{g \in i} PG_g + p_i - \sum_{d \in i} PD_d + 0.5 \sum_{k \in i} PL_k = 0 \tag{61}$$

In Equations (59) and (60), the absolute flow terms $|P_k|$ are easily linearized by introducing two non-negative continuous auxiliary variables P_k^+ and P_k^- such that $P_k = P_k^+ - P_k^-$. This implies $|P_k| = P_k^+ + P_k^-$. These two auxiliary variables correspond to the forward and the backward flows in a line. Note that at most one of them will be zero at a time. This condition is implicitly enforced by the theory of optimality because network losses are a function of $(P_k^+ + P_k^-)^2$ and should be minimized. Setting both of them to zero does not only make sense but contradict with law of optimality. A small penalty can alternatively be included in the objective function to ensure at most one of them is zero at a time. Computationally speaking, the DCTEP problem (either lossy or lossless) is relatively less complex when compared with the LinACTEP model. Since the entire formulation keeps the problem linear, like in the case of LinACTEP, commercially available solvers can solve problems of this type efficiently.

3.3.4. A Modified "DC" based TEP Model (M-DCTEP)

It has been stated that the formulation of DC network model is anchored on the basic assumption that the voltage magnitudes are close to the nominal one, which effectively leads to a flat voltage profile in the system. This assumption is valid in most cases especially in electrical networks spanning over small geographical areas because in such

networks, the transmission lines are often short and low impedances, leading to low voltage drops. However, in bigger networks, which is the subject of this thesis, the voltage drops may be very high as long lines are very common in such networks. Moreover, since RESs are often available in remote areas, very far from major demand centers, long lines are expected to be constructed to tap the available sources. Because of these reasons, it can be appealing to modify the DC model to include some of the interesting features of LinACTEP model. In this model, denoted as M-DCTEP, the customary flow equations in the DC model, which solely depend on angular differences, are replaced with the following equations:

$$\left|P_k - \{V_{nom}(\Delta V_i - \Delta V_j)g_k - V_{nom}^2 b_k \theta_k)\}\right| \leq MP_k(1 - u_k) \qquad (62)$$

$$\left|P_k - \{V_{nom}(\Delta V_i - \Delta V_j)g_k - V_{nom}^2 b_k \theta_k\}\right| \leq MP_k(1 - z_k) \qquad (63)$$

where Equations (62) and (63) stand for the disjunctive flow models in existing and candidate lines, respectively, and $\Delta V^{min} \leq \Delta V_i \leq \Delta V^{max}$. The remaining constraints in DCTEP are also retained here. The full list of constraints can be found in Appendix C.

3.3.5. Relaxed "DC" based TEP Model (R-DCTEP)

The relaxed DC TEP (R-DCTEP) model can be considered as an alternative formulation of the DCTEP model. As explained before, the DC model relies on disjunctive formulations for decoupling bilinear terms. Selecting appropriate big-M parameters can be problematic in most cases, and this directly influences the solution process. Unlike the DCTEP, this model does not require big-M formulation in the case of candidate lines, which can be regarded as a significant computational advantage. Instead of using the disjunctive model, the DC model is relaxed by replacing the bilinear terms with new continuous auxiliary variables. In other words, the proposed R-DCTEP model is linearized by transforming the bilinear terms in the DC power flow equations into separable functions [185]. Here, we show how this is done for the bilinear terms in Equation (46). First, two auxiliary continuous variables $\phi_{1,k}$ and $\phi_{2,k}$ are introduced such that $\phi_{1,k} \coloneqq (z_k + \theta_k)/2$ and $\phi_{2,k} \coloneqq (z_k - \theta_k)/2$. This means the product of discrete and continuous variables $u_k \theta_k$ appearing in the DC flow equation (46) can be transformed into separable functions given by $\phi_{1,k}^2 - \phi_{2,k}^2$ as in (66). The linearization of these quadratic terms is straightforward; the incremental approach (described in the preceding Section) is adopted here. Like in the DCTEP, this model minimizes a given objective function (64) subject to a number of technical constraints. Equation (65) corresponds to the big-M equivalent formulation of the DC power flow model in existing lines while the relaxed form of such a model for candidates is shown in (66). Constraints (67)—(70) form the set of additional constraints required to make the linearization approach complete. Further details of the linearization technique adopted here can be found in an optimization modeling book [185]. The rest of the constraints correspond to the power flow limits (71) and (72), load balance (73), generation capacity limits (74) and voltage angle bounds (75).

$$Minimize \quad Z = Objective \ Function \tag{64}$$

Subject to:

$$|P_k + V_{nom}^2 b_k \theta_k| \le M_k(1 - u_k); \ where \ -b_k = 1/x_k \tag{65}$$

$$P_k = -V_{nom}^2 b_k \left(\phi_{1,k}^2 - \phi_{2,k}^2 \right); \ where \ -b_k = \frac{1}{x_k} \tag{66}$$

$$\phi_{1,k} = \frac{z_k + \theta_k}{2} \quad ; \quad \phi_{2,k} = \frac{z_k - \theta_k}{2} \tag{67}$$

$$0 \le z_k \le z_{k,\max} \tag{68}$$

$$\frac{\theta_{min}}{2} \le \phi_{1,k} \le \frac{z_{k,\max} + \theta_{max}}{2} \tag{69}$$

$$\frac{0 - \theta_{max}}{2} \le \phi_{2,k} \le \frac{z_{k,\max} - \theta_{min}}{2} \tag{70}$$

$$-u_k S_{k,\max} \le P_k \le u_k S_{k,\max} \tag{71}$$

$$-z_k S_{k,\max} \le P_k \le z_k S_{k,\max} \tag{72}$$

$$\sum_{k \in i} P_k + \sum_{g \in i} PG_g + p_i - \sum_{d \in i} PD_d = 0 \tag{73}$$

$$u_g PG_{g,min} \le PG_g \le u_g PG_{g,max} \tag{74}$$

$$\theta_{min} \le \theta_i \le \theta_{max}; \ \theta_{ref} = 0 \tag{75}$$

When it is desired to include network losses in the TEP study, they can be modeled in a similar way as in the DCTEP model described before i.e. by including the constraints (58)—(61).

Note that the investment variables in this model need not be only be discrete variables; this model equally works for continuous as well as discrete variables. As one of its salient features, this model avoids big-M formulation; and hence, demands relatively less computational effort when compared with its DC counterpart. Unlike the DC model, the investment variables can be relaxed to hold continuous values instead of discrete ones while respecting the physical laws of flows, which is another feature of this model. This is relevant because, sometimes, a first-hand estimate of the network expansion needs may be required. In such cases, it is desirable that such information be made available as fast as possible to deliver the results for carrying out the required analysis. One way to do this is by relaxing the discrete investment variables to continuous ones. Thus, the R-DCTEP model with continuous investment variables can be used in this regard. We will demonstrate the usage of such models in the following Chapters.

3.3.6. A Hybrid TEP Model (HTEP)

Due to the computationally intensive nature of the problem, researchers have resorted to further simplify the DC model. A hybrid TEP model (HTEP), which has been used in

network expansion problems [10], [30], is formulated by exempting the candidate lines from obeying KVL. In other words, candidate lines only respect load balance and capacity limits. On the other hands, flows in existing lines are governed by both Kirchhoff's laws. Since the DC flow equations are not in this model, the discrete variables can be relaxed to continuous ones. Although the simplifications and assumptions made in this model lead to a more manageable TEP model (computationally speaking), it has a major drawback associated with "reverse" flows (i.e. flows in a direction opposite to that determined by the law of physics). In power systems, physical laws dictate that power always flows from high potential to low potential. In the case of DC models, this should be from nodes with high voltage angles to those with low voltage angles. However, when HTEP is used as a transmission investment model, the flows in the newly added lines (i.e. candidates) could unfortunately be in the opposite direction in certain circumstances, violating the physical laws that govern power flows in AC systems. To further clarify this problem, consider the system in Figure 3. 1. All corridors can be reinforced with the same line characteristics as the existing one. Assume the generator connected to node 1 is renewable type with very low cost of power production. As we can see, there are two electrical paths to the high-load node 5 namely 1-2-3-4-5 and 1-6-5. Suppose the former path is congested, with all lines along the path reaching their respective maximum capacity, and suppose the latter path has sufficient capacity for sending more power to the load node. However, the congestion in the parallel path (1-2-3-4-5) makes it impossible to send more power to this node. When the system is expanded by making use of the HTEP model, instead of investing in all lines in this path, the model may instead result in investment in corridor 1-2 to allow reverse flow in the new line. This temporarily relieves the congestion and enables to send more power to node 5. The sum of investment cost of line 1-2 and operation of cost of the power injected by G2 (in the form of counter flow in the newly added line) may in the end be lower than the overall investment cost the four lines along the path (1-2-3-4-5). This phenomenon is detected in the numerical analysis of all models, which will be discussed shortly.
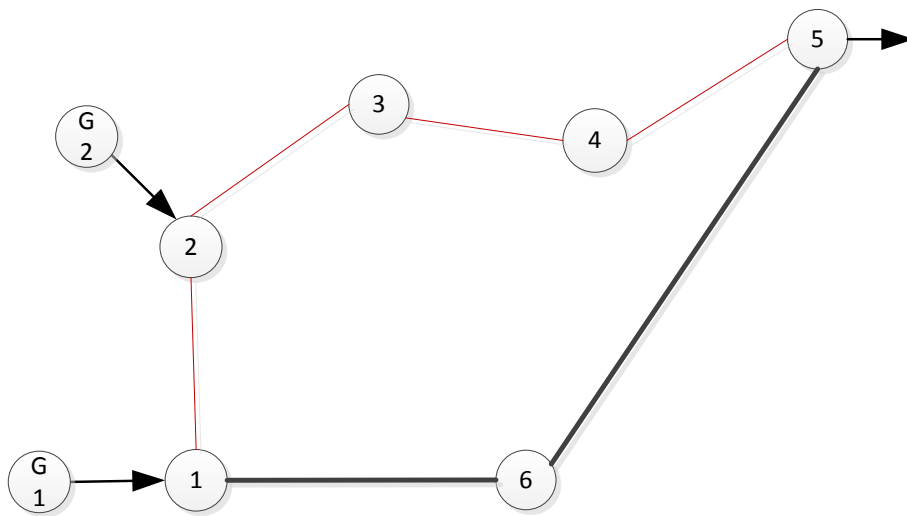
Fig. 3. 1 Illustrative example of counter flows

3.3.7.   A "Pipeline" TEP Model (PTEP)

This model, denoted as PTEP, is sometimes referred to as the "flow" or transportation model which has been used in TEP studies in [31]. The lines are regarded as pipelines, which respect only the capacity limits and nodal balance. PTEP does not obey the Kirchhoff's voltage law. This means that a particular line can carry any desired amount of power flow independent of the impedance of that line and the angular differences. A PTEP model can be formed with any of the models presented and discussed before by excluding the KVL constraints related to both existing and candidate lines. As an example, a lossless PTEP model can be formulated from the DCTEP model in Subsection 3.3.3 by considering only the constraints in (53)—(57). The PTEP model is mathematically less complex and computationally less-intensive when compared to any other model discussed so far. However, given the overly simplified network model to form the PTEP model, the expansion solutions obtained by employing this model can be largely suboptimal. Like HTEP model, this may also be prone to problems of reverse flows.

3.3.8.   A "Copper Sheet" TEP Model (CSTEP)

The copper sheet TEP (CSTEP) model regards existing lines as if they did not have flow limits i.e. by relaxing the flow limits. This model can be alternatively understood as a TEP model without flow limit constraints. CSTEP can be formulated with any of the TEP models presented and discussed so far by excluding the capacity constraints of existing lines or relaxing the binary switching variable associated to these lines to have continuous values with no bounds imposed. For instance, a lossless CSTEP form of the DCTEP model would include the constraints in (51), (52) and (55)—(57) as well as the constraints in (53), where $u_k \in \mathbb{R}$ and $u_k \geq 0$. Such a model can be a very handy tool in quickly analyzing corridors that are prone to congestion so that preventive measures can be undertaken. In addition, it can be very useful in identifying corridors that may need reinforcements/investments. This application is especially relevant when carrying out TEP on large-scale networks, where the huge geographical scope makes it difficult to short-list candidate lines for investments. In such network systems, planners cannot rely on expert knowledge (unlike in small- to medium-scale systems) for the candidate selection procedure. In the following chapters, we will further show its application in this regard.

3.3.   TEP MODEL FIDELITY—NUMERICAL COMPARISONS

3.3.1.   Input Data and General Description

The TEP models briefly described and discussed under Subsection 3.2 have been compared numerically by running case studies constituting of the Garver's 6-bus, IEEE 24- and 118-bus test systems in terms of computational requirement as well as solution accuracy. As mentioned earlier, the motivation of such a comparative analysis is to find the model that strikes the right balance between accuracy and computational demand in the context of large-scale TEP applications.

39

For the analysis here, a deterministic model with an objective function given by (76) is considered which is subject to the constraints corresponding to each model. Equation (76) is composed of the net present values (NPV) of investment cost, operation as well as load shedding costs.

The investment cost of a line is amortized in fixed annual installments throughout its lifetime LT_k, which is considered to be 30 years here. It should be noted that operation and load shedding costs are incurred every year during and after the planning horizon, leading to infinite payments of these costs annually. To further clarify this, consider the illustrative example in Figure 3.2. It is understood that investments are made in a specific year within the planning horizon (the second year in this case) and the investment costs are amortized throughout its lifetime. However, the operation and load shedding costs are incurred every year within and after the planning horizon. To balance these cost terms and to take account of the long-term impact of network investments, a perpetual planning horizon, i.e. an endless payment horizon of fixed installments is assumed here. In other words , the concept of perpetuity described in detail in [186] is adopted. Based on the finance theory in [186], the present value of perpetuity, which is the sum of the net worth of infinite annual fixed payments, is determined by dividing the fixed payment at a given period by the interest rate r. Based on this, the operation and load shedding costs include the associated annual costs within (part I) and outside the planning horizon (part II). The latter (part II) are determined by the perpetuity of the costs in the last planning stage updated by NPV factor in this case $(1 + \sigma)^{-3}$. Note that after the lifetime of the line elapses, it is assumed that investments will be made in the same lines with the same cost and technical characteristics in agreement with the concepts of a perpetual planning horizon.

$$
\begin{aligned}
\min_{z_{k,t}, PG_{g,b,t}, p_{i,b,t}} & \quad Z \\
= & \sum_t \sum_k (1+r)^{-t} \frac{\sigma(1+\sigma)^{LT_k}}{(1+\sigma)^{LT_k} - 1} z_{k,t} IC_k / \sigma \\
& + \underbrace{\sum_t \sum_g \sum_b (1+\sigma)^{-t} \Delta_b PG_{g,b,t} \lambda_g}_{I} \\
& + \underbrace{\sum_g \sum_b (1+\sigma)^{-1} \Delta_b PG_{g,b,t} \lambda_g / \sigma}_{II} \\
& + \underbrace{\sum_t \sum_g \sum_b (1+\sigma)^{-t} \Delta_b p_{i,b,t} \Lambda}_{I} \\
& + \underbrace{\sum_t \sum_g \sum_b (1+\sigma)^{-1} \Delta_b p_{i,b,t} \Lambda / \sigma}_{II}
\end{aligned}
\tag{76}
$$
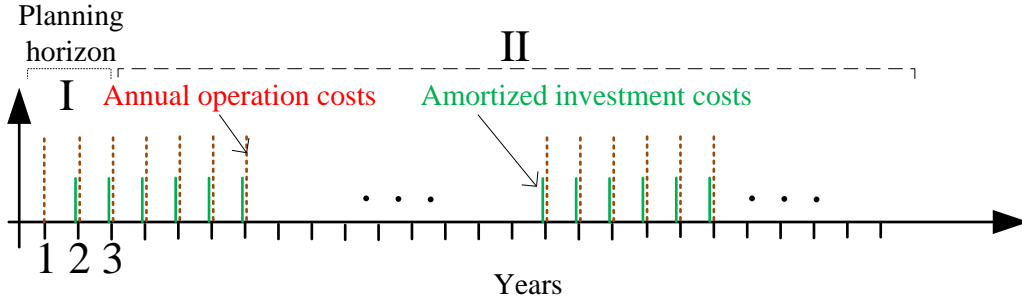
Fig. 3. 2 Illustration of cost components within and outside the planning

For the sake of simplicity, the duration of planning horizon is assumed to be one year. Hence, Equation (70) becomes:

$$
\min_{z_k, PG_{g,b}, p_{i,b}} Z \; = \; \sum_k (1+\sigma)^{-1} \frac{\sigma(1+\sigma)^{LT_k}}{(1+\sigma)^{LT_k}-1} z_k IC_k/\sigma
$$

$$
+ \underbrace{\sum_g \sum_b (1+\sigma)^{-1}\Delta_b PG_{g,b}\lambda_g}_{I} + \underbrace{\sum_g \sum_b (1+\sigma)^{-1}\Delta_b PG_{g,b}\lambda_g/\sigma}_{II} \qquad (77)
$$

$$
+ \underbrace{\sum_g \sum_b (1+\sigma)^{-1}\Delta_b p_{i,b}\Lambda}_{I} + \underbrace{\sum_g \sum_b (1+\sigma)^{-1}\Delta_b p_{i,b}\Lambda/\sigma}_{II}
$$

The constraints of lossy TEP models presented before can be extended to a multi-load level planning framework. For quick reference, a summary of each of the models is presented in Appendix B. All simulations are carried out in HP Z820 Workstation with E5-2687W processor, clocking at 3.1 GHz. GAMS 24.0™ is used to code and run the optimizations. Throughout the analysis in this section, CPLEX 12.0™ is called to solve the problems with default parameters. A 5% interest rate is considered, and the number of partitions for all sorts of linearization is set to 10 but five segments are sufficient according our extensive analysis on this issue [13]. The range of permissible node voltage deviations is between +10% and -10% of the nominal voltage; voltage angles are allowed to vary 1.5 and -1.5 radians.

An hourly demand series for one year is aggregated dividing the load duration curve into 30 load blocks, as shown in Figure 3.3. The duration (in hours) of each load block is indicated in Figure 3.4. Further input data used in the simulations can be found in Appendix C.
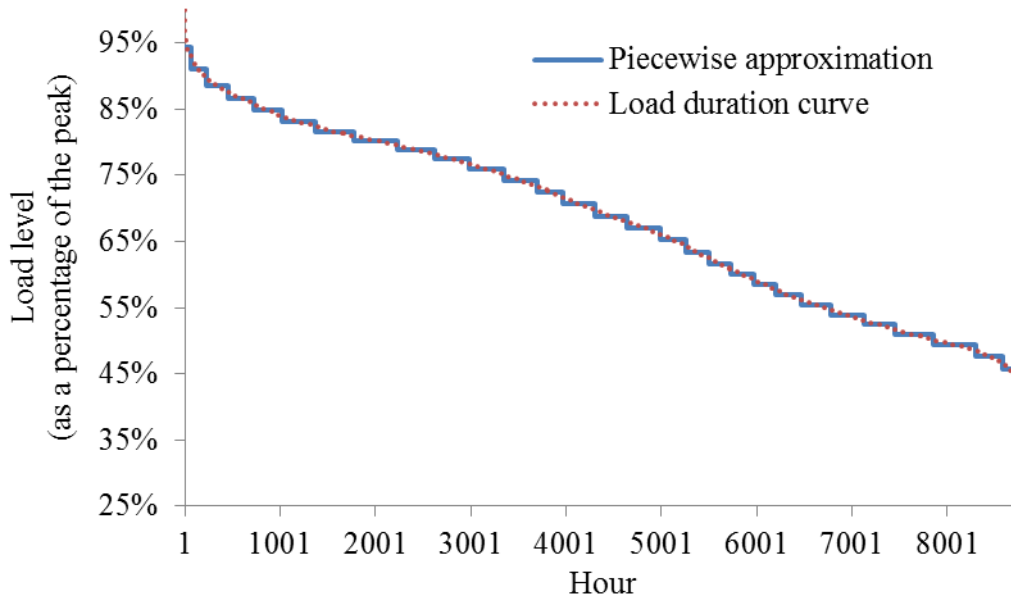
Fig. 3. 3 Aggregation of a load duration curve
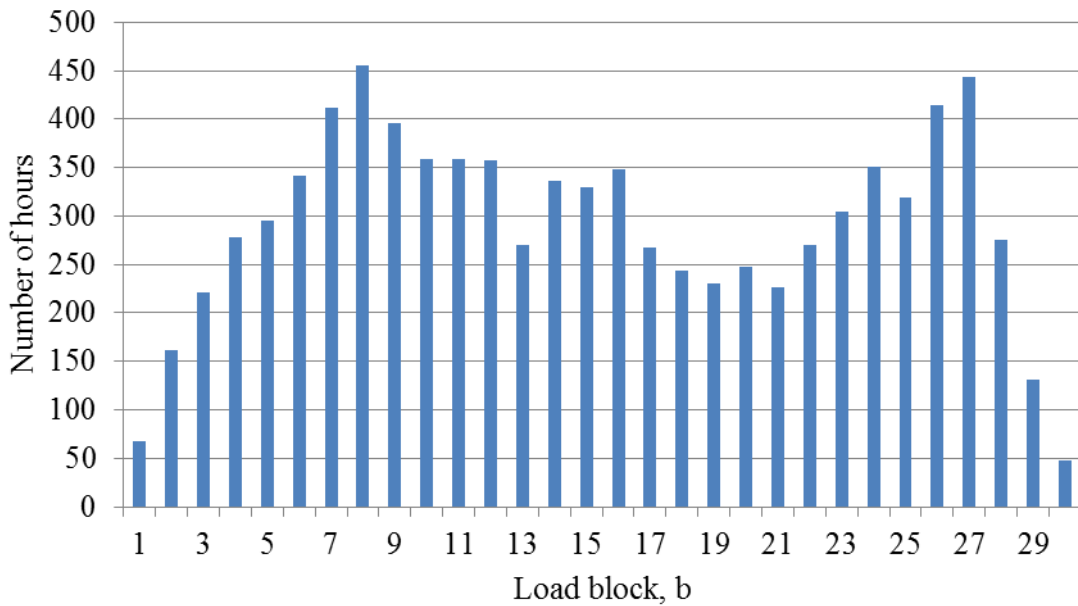
Fig. 3. 4 Number of hours for each load block Δ_b

3.3.2. Numerical Results and Comparisons

Numerical performance of each model is assessed by carrying out simulations on the aforementioned test systems. Simulations results are summarized in Tables 3.1 through 3.4 and Appendix D. Table 3.1 presents the investment decisions obtained by each model with and without losses. As can be observed, neglecting losses generally leads to underinvestment or even a different expansion solution. The overall costs for the lossless cases seem lower than those computed with losses. However, these are unrealistic because of the cost and impact of losses are unaccounted for. In the following section, we shall see a detailed analysis of losses and their influences in expansion results.

Comparing the expansion results with that of an AC solution reveals that all lossy models provide very similar results. This is especially the case with LinACTEP and DCTEP as well as the modified DCTEP models. The expansion results of lossy LinACTEP model only differ by one from the solutions of AC and lossy DC-based TEP models. Yet, the total investment costs for the three models are the same in all three models, as can be seen in Table 3.2. Figure 3.5 compares the losses in each load level computed by each TEP model. It can be inferred from this figure that the DCTEP results in the lowest losses, followed by LinACTEP and the remaining models.

Table 3. 1 Network expansion solutions for different TEP models – 6-bus case

			Investment solution															
		From	1	1	1	1	1	2	2	2	2	3	3	3	4	4	5
		To	2	3	4	5	6	3	4	5	6	4	5	6	5	6	6
Discrete	PTEP	Lossy					1	1			1			1		1	1
Discrete	PTEP	Lossless					1				1			1		1	1
Discrete	HTEP	Lossy					1	1			1			1		1	1
Discrete	HTEP	Lossless					1				1			1		1	1
Discrete	R-DCTEP	Lossy					1	1			1		1	1		1	1
Discrete	R-DCTEP	Lossless					1				1			1		1	1
Discrete	DCTEP	Lossy					1	1			1	1	1	1		1	1
Discrete	DCTEP	Lossless					1	1			1	1		1		1	1
Discrete	M-DCTEP	Lossy					1	1			1	1	1	1		1	1
Discrete	LinACTEP	Lossy				1	1	1			1	1		1		1	1
Discrete	LinACTEP	Lossless					1	1			1	1		1		1	1
Discrete	ACTEP*						1	1			1	1	1	1		1	1
Continuous	PTEP	Lossy				0.1	1	0.17			1		0.1	1		1	1
Continuous	PTEP	Lossless					1				1			1		1	1
Continuous	HTEP	Lossy				0.1	1	0.26			1		0.2	1		1	1
Continuous	HTEP	Lossless					1				1			1		1	1
Continuous	R-DCTEP	Lossy				0.1	1	0.26			1		0.2	1		1	1
Continuous	R-DCTEP	Lossless					1				1			1		1	1

* Best solution after a number of restarts

Computational burden generally increases with model fidelity i.e. PTEP, HTEP, R-DCTEP, DCTEP, M-DCTEP, LinACTEP and ACTEP. Despite its solution accuracy, the LinACTEP demands nearly 5 times more computational effort to solve the problem

than the DCTEP model. Interestingly, the modified lossy DCTEP models (R-DCTEP, M-DCTEP) perform well. The computational requirement of R-DCTEP is significantly lower than that DCTEP while the increase in simulation time when using the M-DCTEP is marginal compared with the simulation time of DCTEP. Figure 3.6 demonstrates this phenomena. In general, from the simulation results, one can see that the models which strike the right balance between accuracy and computational demand are lossy DCTEP and its derivative M-DCTEP.

Another observation in Tables 3.1 and 3.2 is that the models, whose investment variables are converted to continuous ones, yield interesting expansion outcomes. The values of those lines make up optimal solution set are significant, which is very relevant information which can exploited in reducing the combinatorial solution search space, which will be discussed in detail in Chapter 5.

Table 3. 2 Costs and simulation times for different TEP models – 6-bus case

			Investment cost (€)	Total cost (€)	CPU time (s)
Discrete	PTEP	Lossy	284987239.4	1316109950.0	0.764
		Lossless	260205740.3	1073807336.9	0.078
	HTEP	Lossy	284987239.4	1316174242.8	2.855
		Lossless	260205740.3	1073807337.0	0.125
	R-DCTEP	Lossy	311007813.4	1316174242.8	1.372
		Lossless	260205740.3	1073807337.0	0.234
	DCTEP	Lossy	382874160.8	1919502705.2	34.991
		Lossless	358092661.7	1661650011.6	3.697
	M-DCTEP	Lossy	382874160.8	1844384854.2	36.442
	LinACTEP	Lossy	382874160.8	1844384838.0	166.375
		Lossless	358092661.7	1658033570.6	116.532
	ACTEP	Lossy	358092661.7	1658033601.9	434087.045
Continuous	PTEP	Lossy	269495228.7	1297679459.3	0.187
		Lossless	260205740.3	1073807336.9	0.015
	HTEP	Lossy	269767517.4	1298477815.1	0.172
		Lossless	260205740.3	1073807337.0	0.031
	R-DCTEP	Lossy	269767517.4	1298477815.1	0.359
		Lossless	260205740.3	1073807337.0	0.094
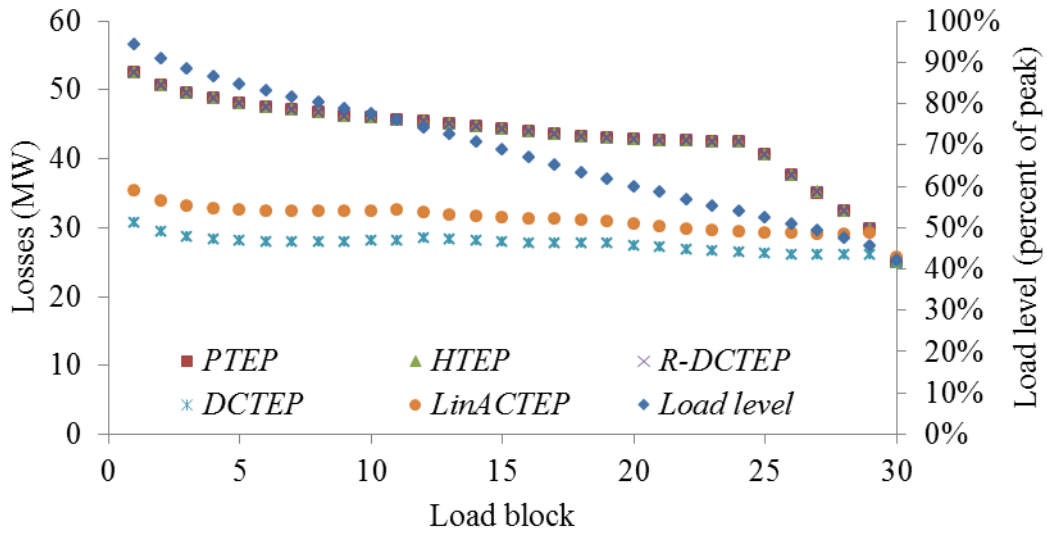
Fig. 3. 5 Losses computed by selected TEP models – 6-bus case
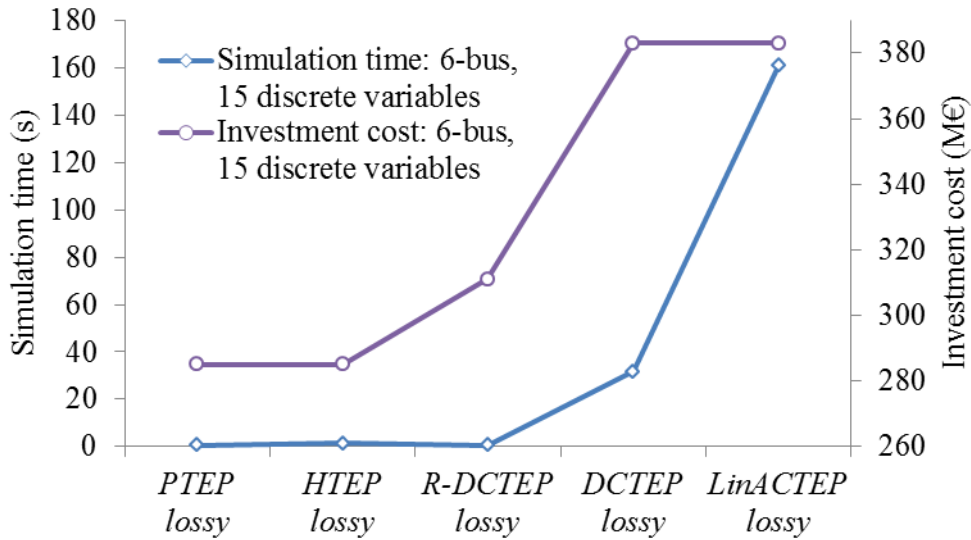
Fig. 3. 6 Performance comparison of selected TEP models – 6-bus case

Table 3. 3 Network expansion solutions for different TEP models – 24-bus case

			Investment solution (values shown in brackets)
Discrete	PTEP	Lossy	2-4 (1), 2-8 (1), 4-9 (1), 16-17 (1)
		Lossless	2-4 (1), 4-9 (1)
	HTEP	Lossy	2-4 (1), 2-8 (1), 4-9 (1), 16-17 (1)
		Lossless	2-4 (1), 2-8 (1), 4-9 (1), 16-17 (1)
	R-DCTEP	Lossy	2-4 (1), 2-8 (1), 4-9 (1), 16-17 (1)
		Lossless	2-4 (1), 2-8 (1), 4-9 (1), 16-17 (1)
	DCTEP	Lossy	1-2 (1), 2-4 (1), 2-8 (1), 4-9 (1), 10-11 (1), 16-17 (1)
		Lossless	1-2 (1), 1-8 (1), 2-4 (1), 4-9 (1), 16-17 (1)

LinACTEP	Lossy	1-2 (1), 2-4 (1), 2-8 (1), 4-9 (1), 10-11 (1), 16-17 (1)
	Lossless	1-2 (1), 1-8 (1), 2-4 (1), 4-9 (1), 16-17 (1)
ACTEP*		1-2 (1), 2-4 (1), 2-8 (1), 4-9 (1), 10-11 (1), 16-17 (1)
Continuous	PTEP	Lossy	1-2 (0.2), 1-8 (0.2), 2-4 (1.0), 2-8 (0.3), 4-9 (1.0), 14-16 (0.2), 16-17 (0.6), 16-19 (0.1), 17-18 (0.1)
		Lossless	2-4 (1.0), 4-9 (1.0)
	HTEP	Lossy	1-2 (0.2), 1-8 (0.2), 2-4 (1.0), 2-8 (0.3), 4-9 (1.0), 9-10 (0.03), 10-11 (0.02), 14-16 (0.2), 15-21 (0.07), 16-17 (0.8), 16-19 (0.13), 17-18 (0.11)
		Lossless	1-2 (1.0), 2-4 (1.0), 2-8 (0.47), 4-9 (1.0), 16-17 (0.77)
	R-DCTEP	Lossy	1-2 (0.2), 1-8 (0.2), 2-4 (1.0), 2-8 (0.3), 4-9 (1.0), 9-10 (0.03), 10-11 (0.02), 14-16 (0.2), 15-21 (0.07), 16-17 (0.8), 16-19 (0.13), 17-18 (0.11)
		Lossless	1-2 (1.0), 2-4 (1.0), 2-8 (0.47), 4-9 (1.0), 16-17 (0.77)

Best solution after multiple restarts

The simulation results pertaining to the 24-bus case, shown in Tables 3.3 and 3.4, and Figure 3.7, largely support the analysis and conclusions made in the 6-bus case. The expansion outcome of lossy LinACTEP and DCTEP models exactly match with AC expansion solution; however, the simulation times of these models significantly differ. Like in the 6-bus case, LinACTEP is a lot more computationally demanding than any other model. about five times and 41 times more expensive computationally than DCTEP in the 24-bus and 118-bus cases. Hence, DCTEP balances well accuracy with computational requirement. Figure 3.8 plots the losses computed by selected models. It can be observed that the difference in these losses curves is not significant mainly because of the similarity in the expansion outcomes.

Table 3. 4 Costs and simulation times for different TEP models – 24-bus case

			Investment cost (€)	Total cost (€)	Simulation time (s)
Discrete investement variable	PTEP	Lossy	1375373.199	3916428514.8	6.973
		Lossless	743444.9723	3913114138.0	0.219
	HTEP	Lossy	1375373.199	3916554450.2	16.069
		Lossless	1375373.199	3913746068.8	1.95
	R-DCTEP	Lossy	1375373.199	3916554450.2	15.741
		Lossless	1375373.199	3913746068.8	1.529
	DCTEP	Lossy	1660360.438	3917018556.0	62.556
		Lossless	1437326.947	3913885370.5	4.368
	LinACTEP	Lossy	1660360.438	3916449214.4	362.562
		Lossless	1437326.947	3913873912.4	148.591
	ACTEP	Lossy	1660360.438	3916450118.3	-
Continuous	PTEP	Lossy	1195667.085	3916117546	0.952

investment variable		Lossless	743444.9723	3913114138	0.063
	HTEP	Lossy	1314158.391	3916253139	3.775
		Lossless	1145970.474	3913516667	0.156
	R-DCTEP	Lossy	1314158.391	3916253139	2.746
		Lossless	1145970.474	3913516667	0.39

As explained in the preceding sections, a major concern with the hybrid and pipeline models is the occurrence of reverse flows. These are not observed in the first two case studies but they are detected in the IEEE 118-bus system in corridors 34—37, 84—85, 85—89, 88—89, which have been part of the expansion solution. This is corrected by excluding the candidate lines in these corridors. Alternatively, this can be avoided by including a small penalty in the objective function. The penalty factor should however be selected carefully not to influence the outcome.
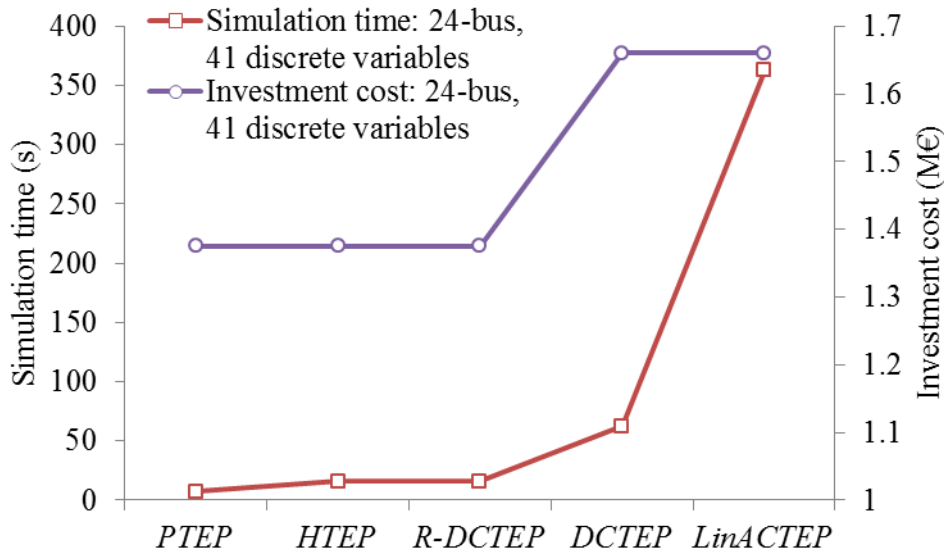
Fig. 3. 7 Performance comparison of selected TEP models – 24-bus case

It has been stated from the outset that the main motivation of doing this analysis is to identify and/or propose an improved TEP model that balances the tradeoff between accuracy and computational requirement from the context of large-term TEP problems under uncertainty in large-scale networks. From this perspective, the computational requirement can be roughly estimated from the simulation results in this section. As the plots in Figures 9 and 10 show the simulation times appear to follow polynomial trends. Holding other parameters the same, the expected simulation times for a system with 1000 nodes or 1000 candidates are computed. These are depicted in Table 3.5. Note that these values only give rough estimates. Yet, the figures show the stark differences in the computational complexity of the models. With the same computing machine, in a 1000-node system, LinACTEP would likely take astoundingly 85 and a half days (nearly three months) before it returns the solution; whereas, the DCTEP model would finish within

47

approximately 2 days. The observations with the number of candidates is the same. This strengthens the previous argument that DCTEP or its "equivalent" formulations, R-DCTEP and M-DCTEP, are the most feasible models that can be extended to TEP problems of a significant network size. Based on the comprehensive analysis made in this section, these models strike the right balance between computational requirement and solution accuracy.
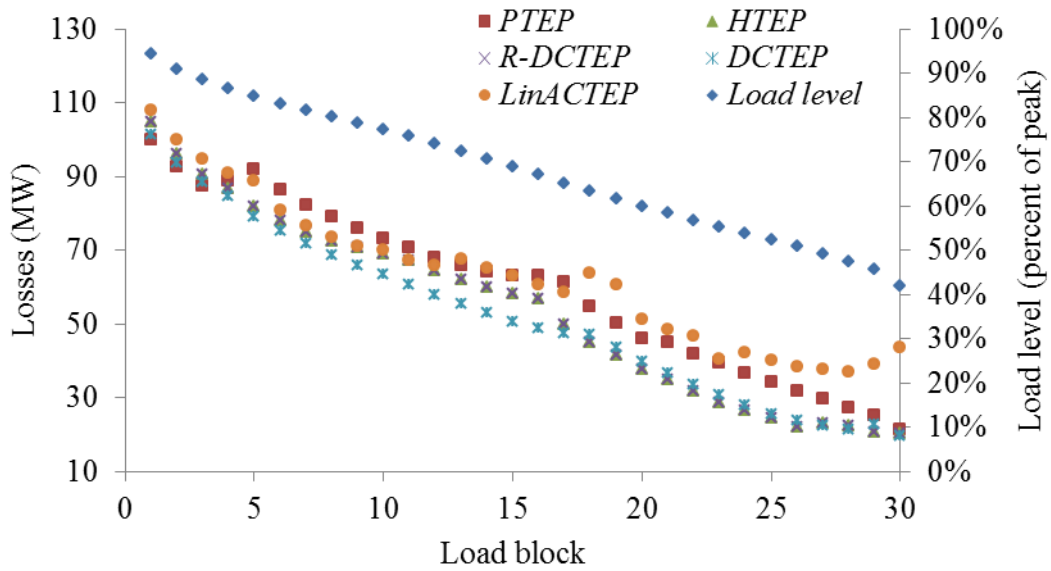
Fig. 3. 8 Losses computed by selected TEP models – 6-bus case
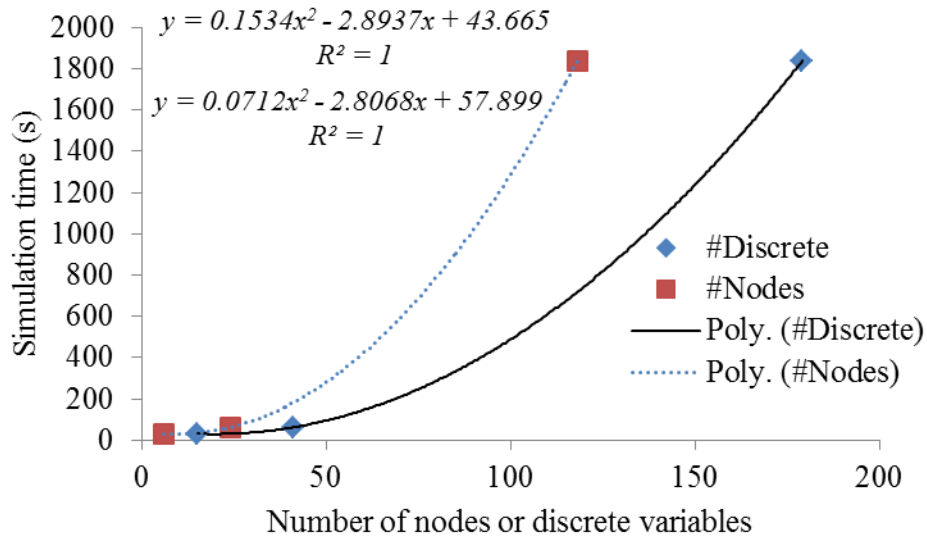
Fig. 3. 9 Simulation time trends as a function of system parameters—DCTEP
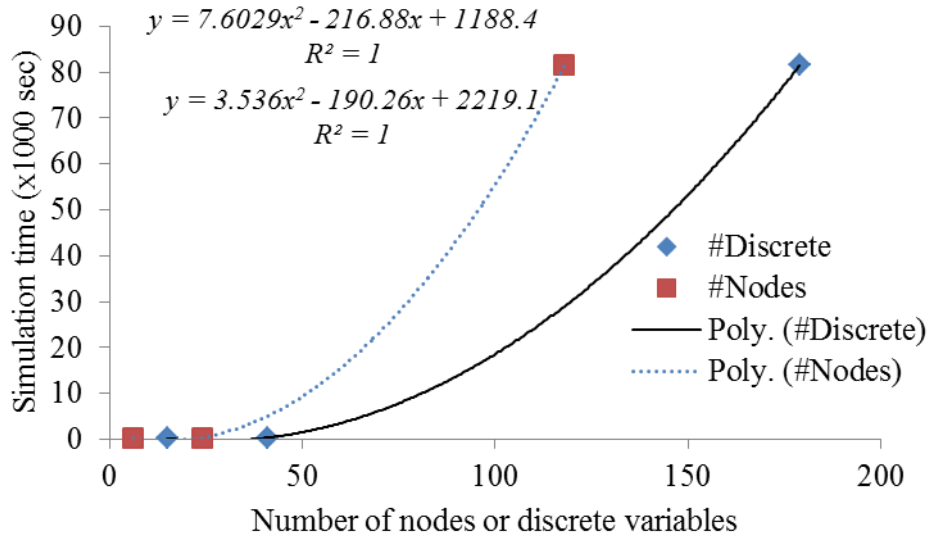
Fig. 3. 10 Simulation time trends as a function of system parameters—LinACTEP

Table 3. 5 Estimating the computational burden (measured in days) of selected lossy TEP models

	System size	
TEP Model	1000 Nodes	1000 Candidates
PTEP	0.025	0.013
HTEP	1.794	0.814
R-DCTEP	1.219	0.553
DCTEP	1.742	0.792
LinACTEP	85.500	38.750

3.4. REPRESENTATION OF TRANSMISSION LOSSES

Most of the existing losses models fall into one of the categories reviewed in the following subsection. It should be noted here that we have slightly modified the common formulations of those models. First, the flow-based losses expression in (58) is used instead of the angle-based one in (40) when formulating the linear models. Second, additional features and constraints are included in some of these models to improve their computational performance and accuracy in representing losses. We also subsequently present some alternative losses models.

3.4.1. Motivation and Overview

The global push for the integration of renewable energy sources (RESs) involves planning the expansion of the transmission grid over geographically wider areas. Moreover, the expected high penetration of RESs introduces significant uncertainties in the development and operation of the system, which need to be accounted for. In most cases, large-scale renewable generation projects will be located far away from major demand centers. Due to the intermittency of their production, ensuring an acceptable level of guarantee of supply in systems with very high RES penetration will require a well-developed

49

transmission network with sufficient capacity to transport the renewable power produced at remote areas to any other area where renewable production is very low. Depending on the availability of RESs, the power flow patterns of the system are expected to undergo dramatic changes over time.

As a result, to properly address a TEP study, a large number of operational states (snapshots) and network investment candidates must be considered, together with several timeline scenarios (or storylines) to represent the uncertainty about the evolution of the system in the future. This leads to a very complex combinatorial TEP optimization problem, requiring a large number of optimal power flow (OPF) computations, which can eventually become intractable. The common practice of considering only the OPF for the peak demand scenario is no longer valid in such power systems, particularly in the context of TEP, where operational states stressing different parts of the network may be largely different. Thus, the OPF formulation considered in TEP should be computationally very efficient to ensure tractability while delivering results with an acceptable level of accuracy. For instance, using a full alternating current optimal power flow (AC-OPF) model, similar to the model used in [187], is not computationally affordable for such a problem, while the classic direct current optimal power flow (DC-OPF) [178] may not be a good solution either because it neglects transmission losses. In general, the OPF formulation should feature all aspects that are believed to play a non-negligible role in TEP, especially in large-scale systems.

Network losses may change the economic generation dispatch and affect optimal solutions for the development of the network; see in [30] and more thorough analyses in [42]. In spite of this, losses are frequently neglected in TEP models or treated in an overly simplified way, mainly to reduce the computational burden when dealing with systems of a significant size. Finding an appropriate representation of losses is critical when the scope of the considered system becomes as wide as the full European transmission network [188]. Moreover, as mentioned previously, large power flows are expected in large-scale network of systems with high penetration of RESs, leading to higher losses which could in turn play a more relevant role in TEP.

When using the conventional AC-OPF model, network losses (both active and reactive) are implicitly modeled because such a model includes all network parameters. However, the resulting problem is highly nonlinear and non-convex which makes computing the optimal solution very demanding. Acknowledging the complexity of the AC-OPF problem, distributed and parallel computation schemes are proposed in [189]. But in some cases, the AC-OPF problem is directly solved via mathematical optimization techniques (for example, the interior-point method in [190]). Due to the nature of the problem, such techniques often rely on a series of approximations to reduce its complexity. Moreover, the nonlinear and non-convex nature of the problem means global optimality could be highly compromised because the solution algorithm could get stuck at local optima. This limitation, combined with the complexity of the AC-OPF problem, led researchers to resort to different heuristic and meta-heuristic solution methods which are based on different nature-inspired algorithms such as: harmony search [191], evolutionary

50

programming [192], imperialist competitive [187], chaotic invasive weed optimization [193], particle swarm optimization [194], shuffle frog leaping [195] and many others [196]. Such solution approaches are claimed to find "good" solutions within an acceptable computational time but provide no guarantee of achieving global optimality. Generally, even if the AC-OPF network model is the most detailed and accurate modeling approach, its practical application is only limited to flow analysis pertaining to single or very few system snapshots due to its mathematical complexity. In other words, it is computationally expensive, if not impossible, to carry out multi-faceted analysis using an AC-OPF based network model and given the sheer size of current power system networks with a high level uncertainty (for example, long-term TEP problems). Therefore, a full modeling of losses (i.e. using an AC power flow model) is not computationally affordable, especially in the TEP context. Therefore, a tradeoff between accuracy in losses representation and efficiency (in computational terms) of the OPF model becomes critical to address TEP studies with high renewable generation penetration scenarios and large-scale networks. This work addresses this objective and contributes losses formulations and a strategy to solve the resulting problem that best achieves this trade-off. The proposed losses models and other existing ones are compared in terms of accuracy in losses representation and computational efficiency.

A review of some of the existing linear modeling approaches of losses is provided in [197]. A losses model based on mixed integer linear programming is reported in [42], applying a piecewise linear approximation of the quadratic expression of losses. And, the same model is applied in TEP studies in [65]. An iterative way of adding linear constraints is adopted in [197] using a dynamic piecewise linear model. In this case, the fully accurate expression of losses is iteratively approximated by adding linear cuts of actual transmission losses. A further extension of this iterative approach is reported in [198], where losses are approximated by progressively adding linear cuts of equally distributed nodal losses, instead of line losses. The node-based approach in [198] is reported to take advantage of the fact that there are fewer nodes than lines in a typical power system. Iterative or dynamic methods to compute losses are feasible in small to medium-scale systems, but in very large-scale systems, performing several iterations may be computationally unaffordable.

In some cases, a single linear losses equality constraint determined by curve fitting is used [199], but this may either overestimate or underestimate transmission losses, depending on the parameters of the constraint (i.e. slope and intercept). In a similar manner, the authors in [142] simply represent losses in a given line as a certain percentage of its flow. In other cases, a quadratic function of losses is merely added to a DC branch flow model to account for losses in TEP [83]. But this adds nonlinearity to the problem, thus, negatively influencing the convergence speed of the computation process. Elsewhere, in problems other than TEP such as locational marginal price calculations, transmission losses are modeled by a fictitious load either concentrated at a single node (often the slack bus) or distributed among all nodes of the system. The distribution of losses is based on either predefined [200] or adaptive coefficients (alternatively termed as distribution

factors of losses) [201]. In reference [200], the entire system losses are distributed among all nodes based on fixed losses distribution factors obtained from an AC power flow analysis; whereas, the authors in [201] assume the losses in each line are distributed as additional loads between its terminals. In the latter case, the distribution factors are computed by means of a DC power flow analysis and losses are iteratively estimated. A further extension of the work proposed in [200], with adaptive coefficients instead of fixed ones, is presented in [202], and authors in [203] combine and extend the methods developed in these works, i.e. an iterative linear approximation of losses with adaptive coefficients is employed in [203]. These coefficients are modified iteratively based on information obtained from an AC power flow analysis, the operational system states, the operation point of generators and the network parameters.

In generation expansion planning (GEP) frameworks, transmission losses and hence their associated impacts on the system are mostly neglected because GEP is often carried out without considering transmission networks. A few works in the GEP subject area incorporate losses by using certain loss allocation methods. For example, losses in transmission and distribution networks are simply considered to be a certain percentage of the demand to be supplied at each node in [204]. The authors in [205] account for losses by multiplying the total power generation at each node with a predefined coefficient (which ranges from 1.08 to 1.10). Similarly, power injections at each node are assumed to comprise a certain ratio of losses [151]. Losses estimated using such approaches may be sufficient in the GEP context; however, such a rough estimation method cannot be extended to TEP, which must consider the entire network system.

Another losses modeling approach, mostly common in economic dispatch (ED) problems, is Kron's loss formula [206], which is based on the concept of marginal transmission losses allocation. Here, losses are represented as a function of levels of power injections (i.e. power generation levels of generating units). This can be understood as an approach which calculates the marginal increase in transmission losses due to an increase in the load or generation level. The so-called B-loss coefficients [206] capture such sensitivity factors i.e. the transmission loss coefficients. These coefficients are determined once using power flow analysis and often considered to remain unchanged over a large set of operational situations, which seems to be a very conservative assumption. In [207], Kron's loss formula is used to estimate losses in an ED problem which minimizes the total cost of power generation. Transmission losses are also modeled using the same formula in a stochastic [208] and a deterministic [209] multi-objective ED optimization framework considering wind power generation. The differences between these two works lie in the solution algorithms employed and the level of details in handling uncertainty. Reference [208] presents a stochastic programming framework to better handle uncertainties in load and wind power generation. And, particle swarm optimization is used to solve the resulting problem; whereas, reference [209] uses a variant of firefly algorithm for the same purpose. Most recently, the authors in [210] embed the Kron's loss formula in a reliability constrained unit commitment problem to estimate the total transmission losses. The application of Kron's loss formula in the subject area of ED is

not limited to the aforementioned works. In [211], this formula is embedded in an ED optimization model which has the cost of power generation as an objective function, and an imperialist competitive algorithm is employed as a solution method to the resulting problem. Authors in [212] use a different heuristic method (charged system search algorithm) to solve the same problem as in [211], but including emission costs. Other works related to ED, incorporating Kron's loss formula, employ point estimate method (analytical) [213] and a derivative of genetic algorithm (meta-heuristic) [214] to solve the ED problem. Both works consider wind power integration, and an objective function that jointly minimizes the costs associated to power generation and emissions.

In [215], with an apparently different strategy, losses are represented by incorporating penalty functions in the objective function. Similarly, the authors in [216] include a linear cost term in the objective function in order to account for the cost of losses and solve a constrained TEP optimization problem which is based on a modified DC-OPF. The authors in [187] also extend this concept by considering a nonlinear formulation for the cost of losses, which is to be minimized in a multi-objective TEP framework based on an AC-OPF. The penalty method may significantly reduce overall losses computed in the system if a large penalty factor is used. Finding an appropriate penalty factor is not easy. Hence, there is a tendency to over-condition the system through the application of large factors, which may lead to sub-optimal results. In many DC-OPF based TEP problems, transmission losses are altogether neglected (for instance, see in [136]), mainly for computational reasons.

The main motivation of our study is as follows. As we shall explain in more detail in the subsequent sections, most of the linear losses models currently used in TEP applications have certain accuracy and/or computation related drawbacks. Of a particular interest here is the estimation accuracy of losses. Most of the linear losses models in TEP do not have the capability to effectively limit "artificial losses" (i.e. extra losses which do not exist in reality, but computed by some models to increase the economic efficiency of the optimal solution under specific circumstances). This means that the computation of such losses leads to an artificial increase in cheap power generation, yet reduce the overall operation cost in the system. Models that do not appropriately limit "artificial" losses normally rely on linear inequality constraints that mainly form an unbounded feasible losses space.

Artificial losses normally involve spilling cheap energy produced in an exporting area to ease network congestion between this area and an importing one, thus allowing some extra demand in the importing area to be supplied with the remaining cheap energy. Congestion occurs because there are several parallel paths between the exporting and the importing area, one of which has very low transfer capacity. Then, spilling some cheap energy along the constrained path in the form of artificial losses eases the network congestion and allows more power to be transported along the other paths while complying with Kirchhoff's laws in AC systems. Generally, "artificial" losses are computed when overly simplified losses models with an unbounded feasible solution space are used in the OPF analyses of a system with a lot of cheap generation from RESs.

53

Based on what has been explained, one can conclude that artificial losses are not exclusive of systems where RES generation exists. However, given that these losses normally make economic sense only when large amounts of cheap power production are available, some of which are to be spilled (in the form of artificial losses), the computation of artificial losses is especially worrisome when there is abundant, intermittent, and non-controllable RES generation in the system.

Another important factor is the computational complexity of the resulting linear losses model. As mentioned earlier, this work is written in the context of large-scale and long-term network expansion planning under high penetration of renewable generation, where there is no room for detailed or complex models of losses. Long term network expansion planning problems are of a huge size when formulated for large systems like the European one or the eastern or western interconnections in the USA. Therefore, our main goal in the present study is to seek a losses representation that is accurate enough to appropriately address problems like the avoidance of artificial losses, while not imposing a significant computational burden. All in all, the main purpose of our study is to find a good linear model for losses in such problems, considering computational efficiency, accuracy of losses estimation, and especially effective limits to "artificial losses".

In our work, two novel linear losses models that represent an alternative to currently existing ones, and two variants of existing models, are compared to one another. The losses models considered here are compared in terms of their accuracy and the increase in the computational time as a result of including them in the OPF formulation; always from the perspective of their application to large TEP problems. Case studies including small, medium and large-scale networks are used to illustrate the performance of the models.

3.4.2. Transmission Network Losses in TEP

3.4.2.1. Impact of Losses on TEP Results

As mentioned earlier, neglecting transmission losses in TEP studies significantly reduces the computation burden of the problem. However, this can jeopardize the accuracy of TEP solutions, especially in large-scale power systems (where power may flow over long distances).

According to the analyses in [30] and more thoroughly in [42], accounting for losses in a TEP problem influences expansion decisions, often resulting in a higher number of line investments. The following three points summarize the impact of losses on optimal transmission expansion.

- *"Free" power transfer*: Neglecting network losses in TEP problems involves ignoring the operational cost of transporting power. Therefore, a lossless TEP results in a network configuration with a lower expansion cost but higher network losses. Considering network losses allows balanced expansion plans that minimize overall costs.

- *Hiding congested lines*: Actual losses imply not only additional generation, but also additional power flows all over the network. If losses are neglected, lines that would in fact be congested may seem to be uncongested, resulting in such lines being excluded from the set of possible expansion decisions and; hence, leading to a different expansion solution.
- *Changes to generator dispatch profile*: Network losses can considerably affect the dispatch order of generators. In large systems or those with large flows, this may in turn affect the network expansion solution.

The aforementioned effects have been verified in two case studies as we shall present in the results section. Generally, a TEP model with losses leads to a network configuration with lower overall system costs, where a trade-off is achieved between all considered cost components. When losses are considered, some extra investments may be undertaken to reduce congestion and losses.

3.4.2.2.   Modeling Aspects: Artificial Losses and Their Consequences

In losses models, another important aspect that should be appropriately handled is the presence of so-called "artificial losses". The term "artificial losses" is used here to refer to the amount of losses exceeding the real ones which may be computed if the losses model (used to solve economic dispatch -ED- or TEP problems) does not provide an appropriate upper bound for the estimation of losses. Therefore, the word "artificial" is used here to indicate that such losses do not occur in reality, and the related power flows are not realistic. Such inaccuracy in modeling losses is due to the use of an overly simplified formulation of losses, with the purpose of making the problem computationally tractable (i.e. by keeping all formulations linear).

In a convex cost-minimizing optimization problem such as the ED or the TEP problems, computed losses in each line should normally be very close to their real values even if the "feasible" region defined for losses is unbounded. This is because minimizing losses normally makes economic sense. However, under special circumstances, an artificial increase of losses may result in a reduced operation cost. A simple example of this case is provided in the system shown in Figure 1. Note that, in linear ED and TEP problems, losses in each line are effectively treated as demand by equally distributing them to both extreme ends of the line (i.e. nodes).

Artificial losses may appear in areas where there is power production available at a very low cost, such as solar or wind power, and network congestion prevents this cheap energy from being exported to other areas. There may be several parallel paths to transport power from the exporting to the importing areas, while the capacity of one of these parallel paths is significantly lower than that of the remaining paths. Note that congestion in power systems is caused by the physical limitations of the grid, i.e. power transmission capacity limits. A power transmission line, for example, has a maximum level of power carrying capability that should not be exceeded for its healthy operation. Otherwise, it could get overheated due to the resistive losses in the line. This may eventually lead to not only malfunctioning and irreparable damages to the line but also operational and technical

problems in the system. In interconnected AC systems, power flows predominantly depend on angle differences between nodes, when the system tries to increase the power flow in a particular line, other power flows increase as side-effects. We clearly demonstrate this by using a three-node AC system, shown below.

To understand how losses models with unbounded feasible losses space can result in artificial losses, let us assume an AC network, so power-flows must comply with both Kirchhoff's laws. Then, some artificial losses along the congested path would result in some extra amount of power being shifted along the remaining ones without violating the power flow capacity constraint for the former path. Recall that losses in every line are treated as demand by equally distributing the losses to both extreme ends of the line (i.e. effectively considered as "virtual loads"). Since the power produced in the exporting area is very cheap (i.e. energy produced at zero or very low cost from intermittent, renewable, energy sources), consuming extra power at some nodes to supply an artificial demand would make economic sense i.e. in reducing the overall operation cost in the system. This occurs when the incremental supply costs computed at these nodes turn out to be negative. Then, creating an artificial demand in the form of artificial losses in lines connected to these nodes would be efficient from an economic point of view (unrealistically lowering operation costs, and/or avoiding network investments).

Note that artificial losses are higher in congested paths than in uncongested ones because this is a means to artificially "reduce" or "control" the amount of power flowing in congested paths so that the capacity limit and flow constraints are not violated. In other words, spilling energy in the congested paths in the form of artificial losses would keep (though this is not realistic) the flow in these paths within the limits set by the capacity of the congested lines, while still complying with the 2nd Kirchhoff's law, which rules the distribution of power flows in the system. Thus, by reducing the amount of power flowing in congested paths, more power can be transferred over uncongested paths while complying with the laws of physics and keeping flows within line capacity limits.

Therefore, inaccurate losses models may result in artificial losses. Additional losses in lines adjacent to congested ones may relax some active constraints, and thus lower the overall system cost [217]. Negative incremental supply costs rarely happen in properly developed networks. However, in TEP problems, the currently existing network will be exposed to demand and generation scenarios only occurring in the (long term) future, sometimes including much higher demand and generation levels than now. As a result, the original network system may be very heavily loaded and stressed, and therefore, not well-adapted to the operation situations being represented in the TEP problem. In some scenarios, losses might be artificially increased to reduce operation costs while avoiding certain network investments. Therefore, the losses model used in a TEP problem should prevent artificial losses by setting appropriate upper bounds or relevant constraints.

It has been already stated that some existing losses models do not properly limit artificial losses. Here, we use a fictitious three node AC system [217] to demonstrate that such losses may exist in expansion planning studies if not properly handled. We shall present

below how some of the existing losses models deal with such losses. The system considered includes a low-cost (renewable wind power) generator at node 1 whose installed capacity is1000 MW and an expensive (conventional) generator at node 2 with a capacity of 400 MW, as shown in Figure 3.5 (a). Their associated marginal costs of power production are depicted in Figure 3.5.The demand at each node is also shown in Figure 3.5. The power transfer capacity limits of the lines 1-2, 1-3 and 3-2 are 1000, 500 and 200 MW, respectively The data for this system, including impedances of the three lines, can be found in [217].
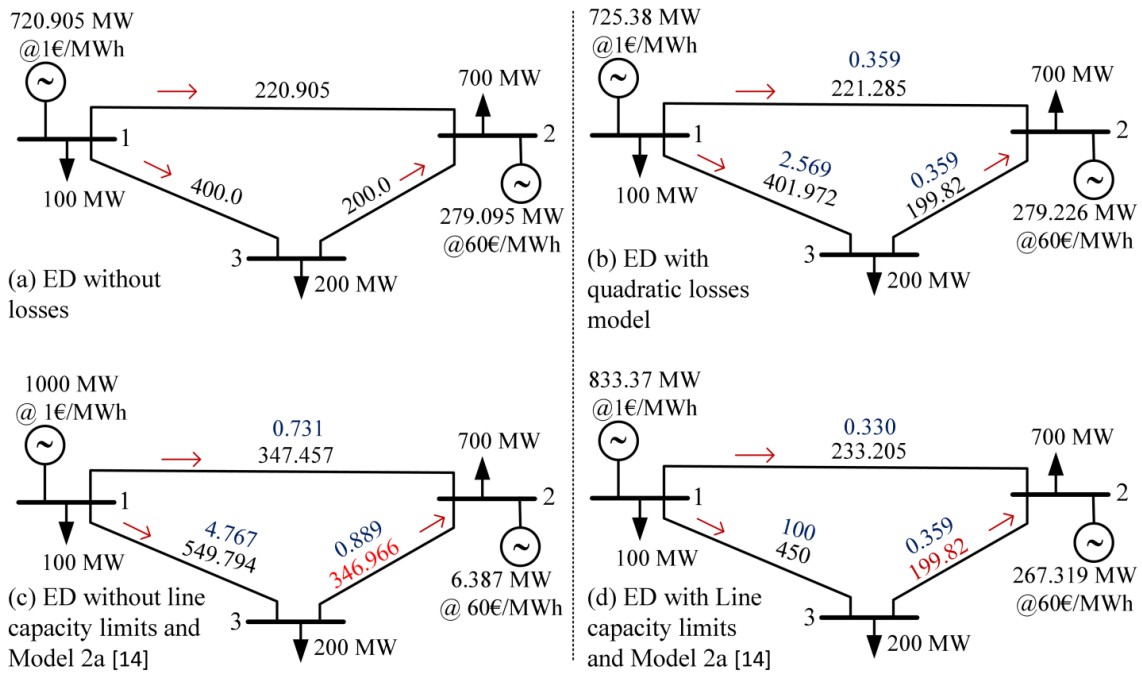
Fig. 3. 11 An illustrative three-node system.

Suppose the economic dispatch (ED) of this AC system, which is based on a DC-OPF model, minimizes the total cost of generation while meeting the following technical constraints: both Kirchhoff's laws and generators' minimum and maximum power production limits. First, the resulting DC-OPF based ED problem is solved by neglecting losses. And, Figure 3.5(a) shows the economic dispatch results corresponding to this case. Here, one can easily observe that line 3-2 is loaded to its full capacity. Second, a quadratic losses model is embedded in the DC-OPF based ED problem. This is needed for comparing the losses computed by an existing losses model [197]. The ED problem (which encompasses the aforementioned constraints and the nonlinear losses model) is then solved by including the transmission capacity constraints. The corresponding OPF results (i.e. the actual power generations, line flows and corresponding losses) are presented in Figure 3.5(b). Note that the two values associated with each line, shown in Figure 3.5, correspond to losses (upper) and power flow (lower) in MW through the line, respectively. It should also be noted that the mismatches in load balance at each node in Figure 3.5 correspond to the losses in the lines connected to the node, which are represented by a quadratic losses model.

Third, the quadratic losses model is replaced by the linear losses model in [197]. In this case, the resulting ED problem is solved excluding and including transmission capacity constraints. The dispatch solutions corresponding to these cases are shown in Figure 3.5 (c) and (d), respectively. Clearly, the solution in Figure 3.5(c) dictates that if there were no transmission capacity constraints, the cheaper generator at node 1 would produce its maximum allowable power (1000 MW) and literally cover all the demand. The expensive generator at node 2 would only contribute a small amount of power to cover the remaining balance (in this case example, the losses in the system). In addition, artificial losses would not be computed in the transmission system. This is because it would be possible to increase flow through link 3-2 beyond its rated capacity, which would remove the congestion (or, alternatively speaking, the bottleneck), and also allow more flows to go through the parallel path 1-2. In other words, it would not make sense to spill power in the form of artificial losses when it is possible to send as much power through the lines as needed to the other side of the network (i.e. node 2). However, the ED results (considering line capacity constraints) show that this link is congested in reality. Hence, an ED model which excludes the transmission capacity constraints does not lead to a realistic OPF solution. In other words, given the physical limitations of the lines, i.e. the power transfer limits shown in Figure 3.5(a), the dispatch solution in Figure 3.5(c) is not practically feasible. This is because, as shown in Figure 3.5(c), lines 1-3 and 2-3 are loaded above their physical limits. And, this is not acceptable because of the previously stated reasons. In order to correct this, generator 1 should step down its power production while generator 2 should step up power production so that a feasible dispatch solution as in Figure 3.5(b) is obtained. As it can be seen in Figure 3.5(b), such rearrangement of nodal injections increases the operation cost of the system but this is required if feasible solution is pursued in such instances. In fact, it does not make sense at all to run ED neglecting line capacity limits. We brought this argument here only to demonstrate the relationship between congestion and artificial losses. The consideration of line capacity limits is always crucial to obtain a realistic solution. But the underlying point here is that artificial losses will be computed if improper losses models are used in systems where congestion and massive low cost generation are present.

Generally, when low cost generators are unable to deliver power to a particular consumer because of congestion, other expensive generators located elsewhere on the grid are dispatched. For instance, as shown in Figure 3.5(b), the power production by the expensive generator is increased by nearly 70% of its rated capacity while that of the cheaper generator is reduced by about 275 MW. As a result, this temporarily relieves the congestion in the lines, and effectively avoids its consequences. The re-dispatching process, explained here, is one of the tools commonly used for congestion management in real power systems. Others remedies of congestion include line reinforcements or switching. In general, most of the congestion management tools rely on ED or TEP optimization models, and the transmission losses model embedded in such models plays a relevant role in the final solution. We will now explain why some of the existing linear losses models do not behave well in certain circumstances.

As stated earlier, the dispatch solution in Figure 3.5(d) corresponds to the solution of an ED problem, embedding the linear losses model proposed in [197] and considering the line capacity limits. Here, it can be observed that the losses computed in line 1-3 are artificially high, nearly 40 times higher than the actual value in Figure 3.5 (b). This can be explained as follows. As in Figure 3.5 (b), line 2-3 is fully loaded to its rated capacity (200 MW) while the other lines are only partially loaded. One can see that the demand at nodes 1 and 3 is easily met making use of the output of the cheaper generator at node 1. The problem arises when we try to supply the demand at node 2 with cheap energy produced at node 1. Since line 2-3 is loaded to its full capacity, it is not possible to transport as much cheaper power from generator 1 as we would like to fully serve the load at node 2. This is due to the Kirchhoff's laws that govern the distribution of energy flows among lines. According to Kirchhoff's law, part of each MW of power injected at node 1 to supply the load at node 2 would flow through the other parallel path connecting nodes1, 3 and 2 to reach node 2. As a result, the flow through line 2-3 would increase beyond its rated capacity. Hence, it is not technically possible to send more flow through line 1-2 instead of using lines 1-3 and 2-3, trying to avoid the congestion at the latter. Thus, the only feasible solution here is to dispatch the more expensive generator located at node 2 to supply part of the load at the same node, as shown in Figure 3.5(b). However, some losses models with an unbounded "feasible" solution space of losses (see the models reviewed in Section 3, especially those based on linear inequalities), may result in artificial losses.

For instance, in the considered example, using the losses model in [197] leads to losses as high as 100 MW in line 1-3, as illustrated in Figure 3.5 (d). From an economic point of view, this reduces the overall dispatch cost (even if it is not technically possible). This is because extra losses come at a nearly zero cost (i.e. the cost of producing power from primary wind energy is zero), allowing the congestion in line 2-3 to be (artificially) relieved; and thus paving the way to transfer about 12 MW of cheaper power through line 1-2 to meet the demand at node 2. Losses computed with the aforementioned model for line 1-3 amount to 20% of the capacity of this line, and are nearly 40 times higher than the losses that would actually exist in the line (which should be about 2.5 MW). As a result, the overall system cost is lowered by nearly 14% with respect to the situation where losses computed in the ED for line 1-3 are limited to their actual level. This is a feasible solution from a mathematical point of view, but it makes no sense from a physical point of view (this is why these losses are called "artificial"). As already pointed out, such unrealistic results arise from the imperfect modeling of losses, and this is not unique to the losses model in [197]. It happens with many of the commonly used existing losses models, where the "feasible space" considered for losses in the OPF problem is not effectively bounded (i.e. because of the "bigger than" linear inequality constraints commonly used in the losses models). The results of the entire economic dispatch for this small illustrative example can be found in Table 3.6, in the following section.

Furthermore, if artificial losses are allowed in the solution, conducting the TEP on this system would not lead to reinforcing line 2-3. However, reinforcing line 2-3 might in fact

reduce the overall system cost when real losses are considered because this could relieve the existing network congestion and allow the expensive power produced by the generator at node 2 to be replaced with the low-cost power produced at node 1.

The following section reviews some of the formulations of existing losses models, focusing on their modeling accuracy (artificial losses, in particular) and computational requirements. Since existing models do not achieve an adequate compromise between accuracy and computational efficiency in the context of TEP, this thesis contributes two alternative losses models, which are able to adequately deal with this problem. The performance of different models is analyzed and compared in the subsequent sections.

3.4.3. Review of Existing Linear Transmission Losses Models

3.4.1.1. *Model 1—Single Linear Equality Constraint*

A rare, but possible, option is assuming losses to be proportional to flows, i.e. representing them using a single equality constraint. The parameters (i.e. slope and intercept) of such a linear constraint can be determined by minimizing the mean squared error (MSE) for values of losses that range from zero to the maximum flow capacity of the line. This results in the expression in equation (78), which is similar to that of the model proposed in [199] apart from the fact that a non-zero intercept is assumed here. Note that the coefficients included in (78), 1 and -0.165, correspond to the optimized slope and intercept parameters of the linear losses equality constraint that best "fits" the scaled quadratic function of losses $(P_k/S_k^{max})^2$ achieving an MSE value as low as 0.006.

$$PL_k = r_k * (S_k^{max})^2 \left\{ 1 * \frac{|P_k|}{S_k^{max}} - 0.165 \right\} \tag{78}$$

In equation (5), S_k^{max} denotes the capacity of the line k connecting nodes i and j. Considering the absolute value of flow in line k, $|P_k|$, in (78) may seem to add non-linearity to the problem, but this can be easily linearized by introducing two non-negative auxiliary variables, representing the flow in the positive and the negative direction for the line, as explained in the preceding sections.

This model avoids artificial losses, which is a relevant feature. However, representing losses with a single equality constraint is not accurate enough for TEP problems, since it results in a significant underestimation or overestimation of losses depending on its parameters (i.e. the slope and the intercept). At least four to five linear constraints are needed for the error in losses estimation to be acceptable. Numerical examples will be given at the end of this section to justify this argument.

3.4.1.2. *Model 2a—Tangent or Traversing Linear Inequality Constraints*

This is a linearization method in which a series of lines are defined as linear constraints which set a lower limit to losses, as shown in Figure 2a [197]. This model can be formulated either using L tangent lines or L lines traversing the quadratic losses curve whose equations are given by the right hand side of the linear constraints in (73) and (74),

respectively. The parameters of each of these lines (the slope and the intercept) are determined using the values of flow, "real" losses and derivate at the intersection points of the quadratic losses curve and the corresponding line representing a linear constraint. For example, for the first tangent line, its slope is given by evaluating the first derivative of the quadratic losses function (4) at $P_k = \Delta p_k^{max}$ which becomes $2r_k \Delta p_k^{max}$; while its intercept can be determined by substituting the values of P_k and PL_k at the Cartesian coordinate $(\Delta p_k^{max}, r_k(\Delta p_k^{max})^2)$ in the linear equation of the line. In general, the linear expressions of the lth constraint (where $l \in (1, 2, ..., L)$), which corresponds to the lth tangent or traversing line, are given by (73) and (74), respectively.

$$PL_k \geq r_k\{2l\Delta p_k^{max}|P_k| - (l\Delta p_k^{max})^2\} \tag{79}$$

$$PL_k \geq r_k\{(2l - 1)\Delta p_k^{max}|P_k| - (l^2 - l)(\Delta p_k^{max})^2\} \tag{80}$$

where $|P_k| = P_k^+ + P_k^-$; $\Delta p_k^{max} = S_k^{max}/L$ is the maximum step-size used in the representation of losses, and L is the number of linear constraints. Note that, for the sake of simplicity, Figures 2a and 2b show only two partitions (i.e. $L = 2$) but the formulation is valid for any desired number of partitions.

The main drawback of both modeling approaches is that the feasible solution space of losses is not bounded from above, potentially resulting in artificial losses. However, when artificial losses do not occur in the system, four or five steps should provide a reasonably accurate value of losses (see in the results section).

3.4.1.3. *Model 3a—Piecewise Linear Approximation*

This model, which is described in [42] and further used in a long-term TEP problem in a deregulated environment [65], is based on the piecewise linearization of the nonlinear losses term. It should be noted here that we have modified the originally developed piecewise linear models. Line flows—instead of angle differences—are discretized here when computing losses, for the reasons already mentioned above. In other words, we compute here a piecewise linear approximation of the quadratic term of the losses expression in equation (40). In order to do this, we represent the absolute value of the line flow variable by the sum of positive step-size flow variables $\Delta p_{k,l}$ associated with each partition of line losses computed using the corresponding linear expressions. This can be understood as a piecewise linear fitting (or first order approximation) of the quadratic losses function, as depicted in Figure 2b. Generally, the model includes constraints (81)–(84).

$$PL_k = r_k \sum_{l=1}^{L} \alpha_{k,l} \Delta p_{k,l} \; ; where \; \alpha_{k,l} = (2l - 1)\Delta p_k^{max} \tag{81}$$

$$0 \leq \Delta p_{k,l} \leq \Delta p_k^{max} \tag{82}$$

$$|P_k| = P_k^+ + P_k^- = \sum_{l=1}^{L} \Delta p_{k,l} \tag{83}$$

$$PL_k \geq 0 \tag{84}$$

Note that $\Delta p_{k,l}$ is a discrete flow variable associated to the lth linear constraint used to represent the losses curve. Equation (81) provides the expression of linearized losses, which are computed as the accumulated sum of step-size losses; equation (82) ensures that the step-size variables do not exceed a preset value. According to equation (83), the discrete flow variables should add up to the absolute value of flows in line k. Equation (84) ensures losses are non-negative.

Note that it is also possible to piecewise–linearize the losses curve by using secants (which allow positive and negative errors in losses representation) as reported in [43] instead of chords (which allow only positive errors) as in this model. Under normal conditions, the former may result in a slightly lower estimation error of overall losses than the latter provided that the expressions of the secants are properly optimized. This is because of the partial cancelation of the positive and the negative errors. However, this has to be weighed in the context of TEP, where losses computed on an individual line basis have more relevance than the overall system losses. Transmission investment decisions are especially sensitive to the underestimation of losses.
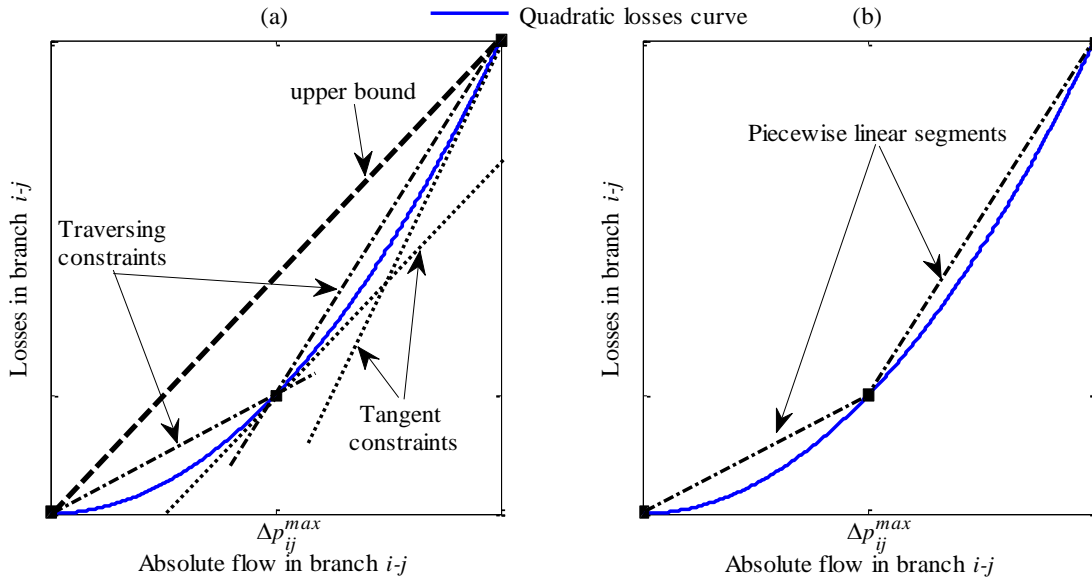
Fig. 3. 12 Method of linearizing losses by (a) tangent or traversing linear inequality constraints with or without an upper bound and (b) piecewise linear approximation.

The main drawback of this model is the large number of additional flow variables needed to represent losses. This model limits artificial losses, since the equality constraints in (8) and (10) guarantee that computed losses are bound to be less than or equal to $r_k(S_k^{max})^2$. The authors in [36] acknowledge that artificial losses computed with such a losses model can sometimes have a dramatic impact on the optimality of the transmission expansion solution. In order to avoid this effect, they reformulate the above model by introducing binary variables to ensure that the angle difference and the losses pair fall exactly on either of the linear segments. However, they conclude that introducing binary variables makes the problem highly complex to solve. The same model has been used in an ED problem [218], and with numerical results of the ED problem which embeds this

model, the authors have showed that "the optimization arrives at an infeasible solution from the physical point of view" [218]. They have demonstrated that the solution algorithm "tries to optimize losses" (referred to artificial losses here) in one of the lines to send more power through other links. And, computed losses in their study are 3.4 times higher than the actual ones. Note that even if Model 3a is based on equality linear constraints, artificial losses can still be computed under certain circumstances. Alternatively speaking, losses can be optimized, but not as high as the losses that would be computed by Model 2a. In Model 3a, losses cannot be higher than $r_k(S_k^{max})^2$.

3.4.4. Coping with Artificial Losses

We have already stated that, unless properly addressed in the TEP optimization model, artificial losses may negatively affect the optimality of a TEP solution. We use the results of the simple system shown in Figure 3.5 to illustrate how the models already described above deal with artificial losses.

We can see in Table 3.6 that Model 2a results in high artificial losses because of the reasons mentioned in Section 2. With Model 3a, losses in line 1-3 are reasonably limited, yet they are 50% higher than the actual ones. Model 1 avoids artificial losses, but it underestimates losses, in this case, producing a value of losses that is 35% lower than actual one. Such inaccuracy (which can even be higher in large-scale systems) is not acceptable in the TEP context.

When conducting a TEP optimization for this system, Model 2a does not result in the reinforcement of the network, since it is cheaper to assume high artificial losses than to reinforce line 2-3. With Model 3a, line 2-3 is chosen to be reinforced because the computed artificial losses in this illustrative example are too small to influence the result. But it should be noted that, when using Model 3a, the resulting expansion solution depends on the amount of artificial losses computed (the higher their level is, the fewer the network investments will most probably be). Model 1 also results in the reinforcement of line 2-3. Note that reinforcing corridor 2-3 relieves the congestion and allows the full use of the low-cost generator at node 1, reducing the overall system cost.

Table 3. 6 Economic Dispatch Results Considering Different Losses Models

Models	System losses (MW)	Generated power (MW)		Total cost (€)
		G1	G2	
Model 1	2.109	720.900	281.210	55,168,454.68
Model 2a[†]	100.689	833.370	267.319	48,286,036.20
Model 2a[§]	100.731	833.391	267.340	48,291,302.83
Model 3a, 2b[§], 3b	4.646	725.410	279.236	56,065,934.04

[†] With tangent linear constraints [§] With traversing linear constraints

These results show that some of the customary models provide artificial losses that may significantly influence expansion plans. Moreover, in long-term TEP problems stated for large-scale systems, the level of penetration of renewable sources is normally relevant.

Together with the computing time (due to the size of the system), the presence of large amounts of low-cost generation is of great concern because large artificial losses may make economic sense.

In conclusion, an appropriate linear model for losses is needed in large-scale TEP problems. This should be chosen considering:

- Computational efficiency
- Estimation accuracy of losses, and especially
- Effective limits to artificial losses

In the next subsection, we present some models that can represent losses more accurately (and reduce artificial losses) and more efficiently (from computational point of view) than the ones already described. A comparative analysis of all these models is included in Section 3.5.

3.4.5.   Proposed Linear Losses Models

3.4.5.1.   *Model 2b—Tangent or Traversing Linear Inequality Constraints with an Upper Bound*

We have already stated that when using Model 2a, artificial losses may appear in some lines for economic reasons. This can be partially avoided by including an additional linear constraint in the formulation of Model 2a that sets an upper limit to the feasible losses region, as shown in Figure 3.6a (using two constraints). Including such a constraint may also accelerate convergence because it further shrinks the feasible region.

The expression of the upper bound constraint is given by $PL_k \leq r_k S_k^{max}|P_k|$, where $|P_k| = P_k^+ + P_k^-$. Note that the inclusion of such a constraint does not fully avoid artificial losses, but limits their value (see Table 3.6) to the polygon area depicted in Figure 3.6a.

3.4.5.2.   *Model 3b—Piecewise Linear Approximation*

This is a modified version of Model 3a. When using Model 3a, it is desirable that the losses segments be "filled up" successively, i.e. in increasing order of the segment indices. Otherwise, under the circumstances that lead to artificial losses, upper segments (with larger slopes leading to higher losses) would be filled up first and to a greater extent than lower segments (with smaller slopes). Despite this fact, Model 3a lacks a constraint that enforces the right behavior in the filling of losses segments. For this reason, we add here the constraint (85) to guarantee that at least upper segments are not filled to a greater extent than lower ones.

$$\Delta p_{k,l} \geq \Delta p_{k,l+1} \tag{85}$$

Including (85) helps to limit artificial losses but does not fully eliminate them, since all step-size variables $\Delta p_{k,l}$ could be made equal, instead of forcing a given step variable to be at its maximum before the following step is allowed to be non-zero (see Table 3.6). As a

result, this development achieves a reduction in the level of artificial losses computed but does not manage to fully avoid them.

3.4.5.3. *Model 4—Traversing Linear Equality Constraints*

The problem of computing artificial losses (when using existing models) can be completely avoided if we can (i) convert the inequality constraints in Model 2a into equality constraints, and (ii) stipulate that no more than one constraint in Model 2 can be active simultaneously. The condition in (ii) can be met by introducing binary variables, as many as the number of linear constraints. Only one of these binary variables will have a value equal to 1, while the others will be set to 0. This can be expressed mathematically as in (86) or (87), where losses are made equal to the expression of the particular linear constraint whose binary variable is 1.

$$PL_k = r_k \left\{ \begin{array}{l} b_{k,1}\big(\alpha_{k,1}|P_k| + c_{k,1}\big) + b_{k,2}\big(\alpha_{k,2}|P_k| + c_{k,2}\big) + \cdots \\ + b_{k,l}\big(\alpha_{k,l}|P_k| + c_{k,l}\big) + \cdots \left(1 - \sum_{l=1}^{L-1} b_{k,l}\right)\big(\alpha_{k,L}|P_k| + c_{k,L}\big) \end{array} \right\} \tag{86}$$

$$PL_k = r_k \left\{ \sum_{l=1}^{L-1} b_{k,l}\big(\alpha_{k,l}|P_k| + c_{k,l}\big) + \left(1 - \sum_{l=1}^{L-1} b_{k,l}\right)\big(\alpha_{k,L}|P_k| + c_{k,L}\big) \right\} \tag{87}$$

$$\sum_{l=1}^{L-1} b_{k,l} \leq 1 \tag{88}$$

Equation (88) ensures that at most, one of the binary variables has a value equal to 1. Although this formulation is mathematically correct, it is non-linear because it includes products of binary and flow variables. These products can be easily expressed using an alternative linear expression by replacing the equations (86) and (87) above with their disjunctive equivalents as in (89) and (90). One of the drawbacks of the big-*M* formulation is the complication associated with the selection of the right value of the big-*M* parameter. Very large values may lead power flow matrices to be ill-conditioned, while low values may cause convergence and inaccuracy problems [55]. To avoid such problems, equations (89) and (90) are reformulated as in (91) and (92), respectively. Equations (93) and (94) are included as well to ensure that the line segment considered to represent losses corresponds to the one defined for an interval that includes the value of the flow variable.

$$\left| PL_k - r_k\big(\alpha_{k,l}|P_k| + c_{k,l}\big) \right| \leq M_k(1 - b_{k,l}) \tag{89}$$

$$\left| PL_k - r_k\big(\alpha_{k,L}|P_k| + c_{k,L}\big) \right| \leq M_k \left(1 - \left(1 - \sum_{l=1}^{L-1} b_{k,l}\right)\right) \tag{90}$$

$$-LB_k\big(1 - b_{k,l}\big) \leq PL_k - r_k\big(\alpha_{k,l}|P_k| + c_{k,l}\big) \leq UB_k\big(1 - b_{k,l}\big) \tag{91}$$

$$-LB_k\left(1-\left(1-\sum_{l=1}^{L-1}b_{k,l}\right)\right)\le PL_k - r_k\big(\alpha_{k,L}|P_k| + c_{k,L}\big)$$

$$\le UB_k\left(1-\left(1-\sum_{l=1}^{L-1}b_{k,l}\right)\right) \tag{92}$$

$$|P_k|\le \sum_{l=1}^{L-1}b_{k,l}\,l\,\Delta p_k^{max} + S_k^{max}\left(1-\sum_{l=1}^{L-1}b_{k,l}\right) \tag{93}$$

$$|P_k|\ge \sum_{l=1}^{L-2}b_{k,l+1}\,l\,\Delta p_k^{max} + (L-1)\Delta p_k^{max}\left(1-\sum_{l=1}^{L-1}b_{k,l}\right) \tag{94}$$

where

$\Delta p_k^{max} = S_k^{max}/L$ is the step-size of each linear constraint; $|P_k| = P_k^+ + P_k^-$; $\alpha_{k,l}$ and $c_{k,l}$ are the slope and the intercept of the lth constraint in Model 2, which are given by $(2l-1)\Delta p_k^{max}$ and $(l-l^2)(\Delta p_k^{max})^2$, respectively;

$b_{k,l}$ is the binary variable associated with the lth constraint;

M_k is a big-M parameter;

UB_k is an upper bound constant given by:

$$\max\left(PL_k - r_k\big(\alpha_{k,l}|P_k| + c_{k,l}\big)\right) \approx r_k\big(\alpha_{k,L}S_k^{max} + c_{k,L}\big);$$

LB_k is a lower bound constant whose value in this case is zero; and

L is the total number of linear losses constraints.

In this case, the model includes constraints (88) and (91)–(94). The disadvantage of this model is its mathematical complexity. It is easy to understand that the higher the number of constraints and binary variables, the larger the computational burden of the model is. Regarding artificial losses, test results obtained for the system in Figure 3.5 show they are effectively avoided.

3.4.5.4.   Model 5—SOS2 Approach

We have already mentioned earlier that the first three losses models have certain drawbacks, particularly in terms of their accuracy in estimating losses. The improvement achieved as a result of the additional constraints included in Models 2b and 3b may not be sufficient to properly limit artificial losses in some situations. Moreover, despite the fact that it provides a more accurate estimate of losses than the first three models, Model 4 cannot be suitably applied to large-scale TEP problems because of its computational complexity. Because of this reason, we propose a new modeling approach based on the use of Special Ordered Sets of type 2 (SOS2) [219] (also discussed in detail in [220]), which is explained in detail in the following paragraphs.

Piecewise linear functions, as in Figure 3.13, can be modeled by introducing a set of positive variables $\lambda_{k,l}$, where $l \in (0,1,...,L)$, that form an SOS2 (see [219]). These variables can be understood as weights associated to the points where the linear losses constraints cross the quadratic losses curve, here called intersection points. The main

property of SOS2 variables, as described in [219], is that at most, two consecutive variables among them can have non-zero values. As we shall explain below, this property leads to the fact that losses are computed by carrying out a linear interpolation between two consecutive intersection points.

The formulation of this losses model is as follows. First, the absolute flow in a line is expressed as the sum of the products of the flow values at the partitions defined on the flow axis in Figure 3.13 {i.e. $P_{k,l}$ where $l \in (0, 1, ..., L)$ } and the corresponding lambda variables, as in equation (95). Without loss of generality, the intersection points can be assumed to be equally spaced on the horizontal (flow) axis, where the distance between two consecutive points is given by Δp_k^{max}, as in the previous models. Thus, the flow at the lth intersection point becomes $l * \Delta p_k^{max}$. Substituting this in (40) gives the line losses at this point as $r_k(l * \Delta p_k^{max})^2$, where $l \in (0, 1, ..., L)$. Then, the flow expression in (96) is derived accordingly.

Similarly, the line losses can be expressed as in (97), from which (98) is derived by considering the quadratic expression of losses at each intersection point. Equation (15) is a general upper bound for the lambda variables. Note that $P_{k,0}$ and $PL_{k,0}$ are both zero since they correspond to the flow and the losses at the first intersection point (i.e. the zero coordinates as in Figure 3.13). Elsewhere, the SOS2 approach has been applied for dealing with nonlinear functions in a mixed integer programming gas network optimization [221]. The authors in [220] also extend this concept to linearization of a two-dimensional function.

Remember that in this model, it is additionally required that at most two consecutive lambda variables are non-zero. This requirement combined with (99) makes the lambda variables have the same properties as SOS2 variables, thoroughly described in [219]. Adding this condition ensures the values of the flow and the losses for each line are linked and correspond to a point that lies exactly on one of the linear segments between two consecutive intersection points (see Figure 3.13).

$$|P_k| = \sum_{l=0}^{L} P_{k,l} * \lambda_{k,l} \tag{95}$$

$$|P_k| = \Delta p_k^{max} \sum_{l=0}^{L} l * \lambda_{k,l} \tag{96}$$

$$PL_k = \sum_{l=0}^{L} PL_{k,l} * \lambda_{k,l} \tag{97}$$

$$PL_k = r_k(\Delta p_k^{max})^2 \sum_{l=0}^{L} l * \lambda_{k,l} \tag{98}$$

$$\sum_{l=0}^{L} \lambda_{k,l} = 1 \tag{99}$$

To further clarify how this linearization works, an example is provided next. Suppose the value of the line flow lies between p_3 and p_4 (see Figure 3.13). In this case, all but λ_3 and λ_4 variables would be zero, forcing the flow-losses pair to lie on the fourth segment. This

is because, due to constraints represented by equations (95) and (97), the line losses are computed by linearly interpolating the value of the losses function (the expression for PL_k for a value of the line flow, $p_3 \lambda_3 + p_4 \lambda_4$, lying between the two extreme flow values in the corresponding segment, p_3 and p_4. Then, the expression of losses is computed as $\varphi_3 \lambda_3 + \varphi_4 \lambda_4$. Since $\lambda_3 + \lambda_4 = 1$ from (15), $p_3 \lambda_3 + p_4 \lambda_4$ and $\varphi_3 \lambda_3 + \varphi_4 \lambda_4$ can be equivalently expressed as $p_3 \lambda_3 + p_4 (1 - \lambda_3)$ and $\varphi_3 \lambda_3 + \varphi_4 (1 - \lambda_3)$, respectively, which clearly implies that line losses are computed by linear interpolation.

As an example of the computation of the parameters of the expression used here to approximate actual losses, suppose the actual flow PL_k is 75 MW for a given line with a rated capacity and a resistance of 100 MW and 0.1 per-units, respectively. Note that the above formulations are based on per-units. But if one wants to instead work with megawatts, the per-unit flows and losses should be multiplied by the base power S_B (here, assumed to be 100 MVA). Based on this, multiplying the value of per unit losses PL_k in equation (40) by S_B gives the value of losses in megawatts ϕ_k i.e. $\phi_k = S_B * PL_k$ which is also equal to $S_B * r_k p_k^2$. We know by definition the per-unit flow p_k is obtained by dividing the MW flow P_k by the base power S_B i.e. P_k/S_B. Therefore, we can rewrite the losses expression as $S_B r_k P_k^2/S_B^2$ or equivalently as $\xi_k P_k^2$ where ξ_k is a coefficient given by r_k/S_B. For the example case, ξ_k is equal to 0.001/MW.

Taking five equally spaced partitions, the set of evenly distributed flow steps taken in the losses representation $\{p_0, p_1, p_2, p_3, p_4, p_5\}$ becomes $\{0, 20, 40, 60, 80, 100\}$. Clearly, we can see that the line flow, 75, lies in the fourth partition (i.e. between 60 and 80). Actual losses corresponding to the flows 60 and 80 MW, computed using the quadratic expression, are 3.6 and 6.4 MW, respectively. In the losses model presented here, only the lambda variables corresponding to the intersection points (60, 3.6) and (80, 6.4) should be different from zero. Thus, equations (13)—(15) become: $75 = 60 \lambda_3 + 80 \lambda_4$, $PL_k = 3.6\lambda_3 + 4.6\lambda_4$ and $\lambda_3 + \lambda_4 = 1$, respectively. Solving these equations simultaneously, we get $\lambda_3 = 0.25$, $\lambda_4 = 0.75$ and $PL_k = 5.7$ MW. In this case, the difference between the losses value computed with the proposed model (5.7 MW) and those that would be computed using the quadratic losses function (5.625 MW) is practically negligible, clearly showing the accuracy of this model. It is also interesting to observe that the point (75, 5.7) lies exactly on the linear segment which passes through the intersection points (60, 3.6) and (80, 6.4), thus resulting in the linear equation for losses $PL_k = 0.14P_k - 4.8$.

The number of additional variables required to represent losses using this model may create a considerable computational burden, but this is counterbalanced by the fact that this model is fully based on the use of equality constraints. Artificial losses are not a concern here since the lambda variables are SOS2. In other words, constraints (96) and (98), together with SOS2 properties, guarantee that losses are not oversized for economic reasons.

It should be noted, however, that even if the SOS2 stipulation is not explicitly included in the model, losses should be bounded to be lower than $r_k(S_k^{max})^2$. This has been experimentally proven by applying this losses model to the 3-node test system in Figure 3.5. The results of the economic dispatch for this system, when employing this SOS2-based losses model, confirm its ability to effectively eliminate artificial losses. Calculated losses coincide with the actual losses (3.222 MW), with generators G1 and G2 producing 721.456 and 281.766 MW, respectively. The total system operation cost in this case is € 56,217,891.45.
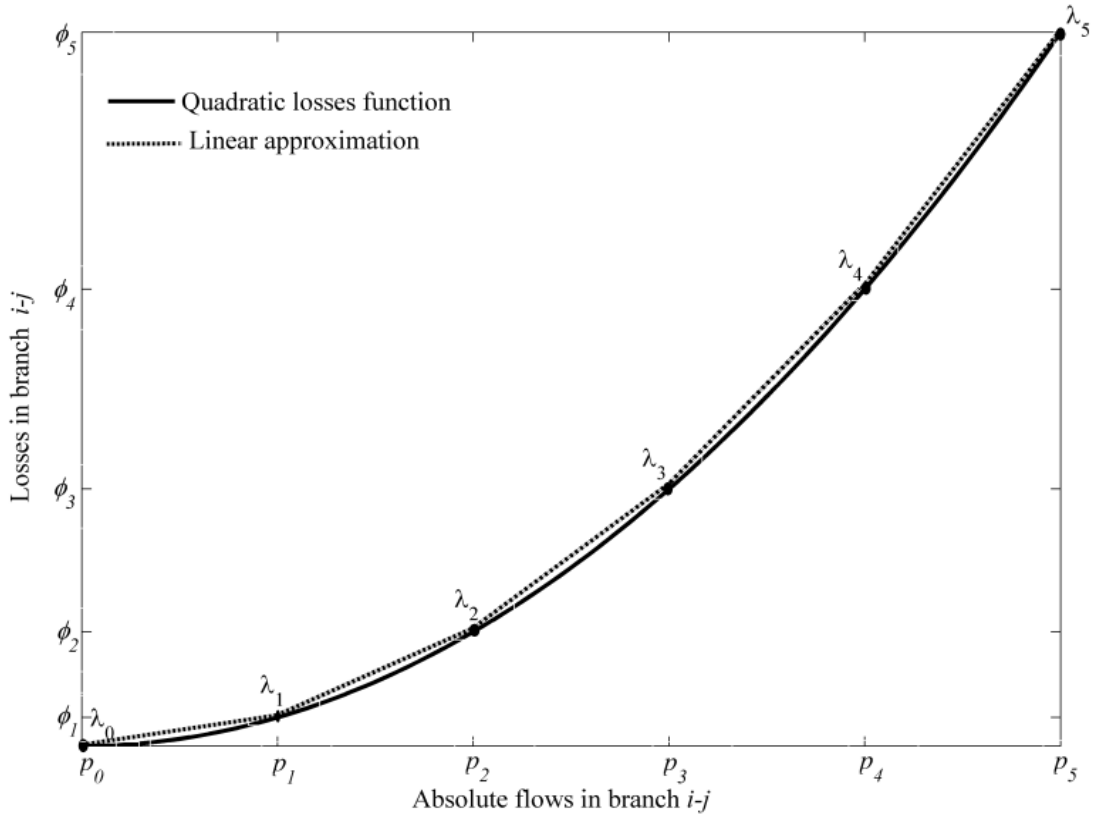
Fig. 3. 13. Piecewise linearization of losses in the SOS2 approach

3.5.   NUMERICAL COMPARISONS OF THE LOSSES MODELS

Case studies including small, medium and large-scale networks are used here to analyze the performance of the models considered in this work. For this purpose, a static version of the DC-OPF based TEP model, described in Section 3.3 and in [15], is considered. In addition, the hourly forecast of electricity demand at each node is assumed to be given for the whole target (planning) year and a load duration curve is used to aggregate the demand at each node into 5 load blocks by means of piecewise approximation. The demand level and number of hours in each load block are determined in such a way that the peak hours are modeled more precisely than the off-peak and shallow ones. All case studies have been solved using a computing machine Core 2 Duo SU7300 processor with 4 GB RAM clocking at 1.3 GHz.

3.5.1. Impact of Losses on TEP Results—Numerical Results

For the analyses here, a static version of the DC-OPF based TEP model in [15] is employed. The considered TEP model minimizes the sum of operation and transmission investment costs while simultaneously satisfying a number of customary technical constraints. The operation cost includes generation and reliability costs. The latter are simply modeled by including a factor in the objective function which penalizes unserved power computed at each node. The standard Garver's 6-bus [39], shown in Figure 3.14, and the IEEE 118-bus [222] test systems are used in the analyses.
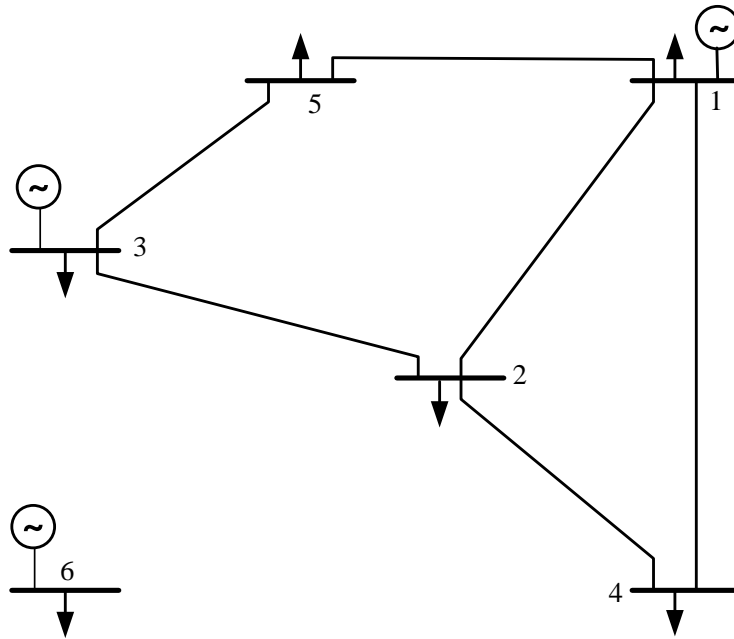
Fig. 3. 14 Garver's 6-bus test system

Table 3. 7 Impact of Network Losses on Expansion Results

Garver's 6-bus System	Lossless TEP	Lossy TEP
Total cost (investment + operation costs)	0.9157 [§]	1 [§]
Investment cost (as a fraction of total costs)	0.1468	0.2244
Losses (MW)	0	110.0089
#Corridors with investments	3	5
IEEE 118-bus system		
Total costs (investment +operation costs)	0.9047 [†]	1 [†]
Investment cost (as a fraction of total cost)	0	0.1024
Losses (MW)	0	501.1706
#Corridors with investments	0	11

[§] *Expressed as the ratio of the total costs of lossy TEP* in Garver's system

[†] *Expressed as the ratio of the total costs of lossy TEP in IEEE 118-bus system*

Table 3.7 summarizes the results of the analyses, showing the values of the most relevant output variables of the TEP problem with and without losses. Here, *lossy TEP* and *lossless TEP* refer to the TEP optimization problems with and without considering transmission losses, respectively. The total costs in Table 3.7 correspond to the optimal values of the objective function in the corresponding TEP optimization problems, which is given by the sum of the operation and the transmission investment costs. Comparing the total cost figures under lossy and lossless TEP (see in Table 3.7); the latter apparently results in a network expansion solution at a lower overall cost. However, this does not include real costs as the effect of losses is not accounted for in the lossless TEP problem. We shall explain next why this is the case.

Let us define *lossy ED* as the economic dispatch (ED) problem, where the total operation cost is minimized taking into account transmission losses and considering the network configuration computed in the lossless TEP problem (i.e. the network consisting of lines in the base-case system plus the network reinforcements computed by running the lossless TEP optimization).

In order to determine how good the network investment decisions obtained by lossless TEP are, it suffices to compare the following two cost figures: (i) the objective function value (operation cost) of the lossy ED plus the cost of network investments computed in the lossless TEP, and (ii) the total costs (i.e. network investment plus operation costs) computed in the lossy TEP problem. Note that the full operation costs, including transmission losses, resulting from the network configuration (investments) computed taking into account losses are the ones already computed in the lossy TEP problem. As expected, the total costs corresponding to the first and the second case studies, computed as in (i), are found out to be approximately 16% and 14% higher than those computed as in (ii), respectively. In other words, the total costs of a system—including operation and transmission investment—resulting from lossy TEP can be significantly lower than that of the system expanded according to lossless TEP. This is because lossless TEP underestimates the operation cost of the system (since it does not take into account the extra costs related to the existence of losses) and results in underinvestment which, in the end, turns out to be significantly more costly. Generally, incorporating network losses in TEP shifts the costs incurred from operation to line investments, resulting in different expansion results. We can also observe in Table 3.7 (especially for the second case study) that the reduction of losses achieved in lossy TEP may, by itself, justify network investments. This is because of the corresponding reduction achieved in operation costs. Related to this, the losses computed in the lossy ED case are nearly 5.3% higher than those computed in the lossy TEP for the Garver's case study. This figure even gets as high as 29% in the second case study. The costs corresponding to these extra losses amount to approximately 11% and 23% of the total costs computed in the lossy TEP problem, respectively.

The optimal network expansion strategy for the Garver's system in the lossless TEP case comprises investments in corridors (2,4), (3,5) and (4,6). However, when the network is expanded taking into account losses, two more lines are built in corridors (2,3) and (2,5).

It should be noted here that the network expansion results obtained by the lossy TEP agree with the full AC-OPF network expansion results reported in [36]. This shows that a DC-OPF TEP model with losses can result in a realistic and reasonably accurate TEP solution.

Analyzing the lossy DC-OPF results for the lossless TEP solution (i.e. lossy ED) clarifies the reasons for the higher costs compared to the lossy TEP solution. The results show that corridor 3-5 is congested in the first case, forcing the curtailment of about 30 MW of load at node 5. In addition, losses in the system are 5.3% higher than those in the system expanded according to lossy TEP. As expected, the increase in losses along with the increase in non-served energy causes an increase in the operation cost of the system which exceeds the savings in network investments.

The overall cost reduction achieved by considering losses in TEP can be obtained by subtracting the total system costs computed for the system expanded according to the lossy TEP from those of the system expanded according to the lossless TEP. In this case, the operation cost reduction achieved when considering losses is approximately 20% higher than the cost of the two extra investments in the lossy TEP. As a result, net savings achieved in this particular case are about 2.3% of the total system costs obtained for the lossy TEP.

Similarly, the results of the second case study, i.e. the IEEE 118-bus system [222], also highlight the undesirable consequences of ignoring losses. Given the data in [222], this system does not require investments regardless of which TEP model is used (lossy TEP or lossless TEP). However, in order to create a need for line investments, the base case electricity demand of 3733.07 MW has been increased by 90% to 7092.83 MW.

Even in this case, no line investment is deemed necessary when the lossless TEP model is used. But in the lossy TEP exercise, up to 11 network reinforcements are planned, mainly due to the substantial 29% reduction in losses they bring about. This is a good example of reinforcements solely justified by the reduction of losses.

3.5.2. Numerical Comparison of the Losses Models

The level of accuracy of the results provided by different losses models is evaluated by two criteria: (1) the relative error in the estimation of losses, and (2) the impact of such error on the intermediate and final results of the TEP problem, i.e. the differences occurring due to losses computation in the set of network investment decisions and in the overall system cost (investment + operation). To this end, three test systems including the Garver's 6-bus, IEEE 118-bus and a 425-node Spanish network systems are employed. The results for the first case study are presented here. Test results and further discussions can be found in [13].

The standard Garver's 6-bus test system (a complete description and data can be found in [39]), is used as a case study. This system comprises eleven candidate lines across different corridors, and is shown in Figure 3.14.

One of the goals of this work is to test the accuracy and computational burden of different linear losses models in a TEP context. Table 3.8 shows the accuracy in the computation of system losses and system costs for each model for various numbers of linear losses constraints, or segments. Note that the system costs in this table refer to the investment plus the operation costs computed after running the TEP optimization model which embeds each losses model.

The total resistive losses value computed using an AC-OPF of the expanded system (in this case, 109.19 MW) is taken as a reference when assessing the relative accuracy of each model. Given that artificial losses do not make economic sense in this small system, the results obtained by all models are found out to be very similar in terms of their accuracy in estimating losses, no matter how many linear constraints or segments, L, are considered. As it can be seen in Table 3.8, when L is larger than 5, the relative error induced when computing the overall system losses falls below 5%, which is practically negligible from a TEP perspective. Table 3.8 also shows that considering a single linear constraint (i.e. $L=1$) can result in greatly overestimated losses.

Table 3. 8 Effect of Number of Partitions in Losses Linearization on System Costs and Relative Error in the Estimation of Losses for the Garver's System

		Number of losses constraints (L)						
		1	2	3	5	10	15	20
Relative error of losses (%)	Model 2a[†]	26.740	5.712	3.322	1.270	0.354	0.089	0.052
	Model 2a[§]	46.549	10.866	6.102	2.212	0.538	0.209	0.110
	Model 2b[§]	46.549	10.883	6.102	2.212	0.538	0.209	0.110
	Model 3b	46.549	10.883	6.102	2.212	0.538	0.209	0.110
	Model 5	46.549	10.883	6.102	2.212	0.538	0.209	0.110
	Model 4	46.549	10.883	5.278	1.387	0.505	0.027	0.030
Cost, in M€	Model 2a[†]	286.29	291.11	291.67	292.21	292.41	292.48	292.49
	Model 2a[§]	305.21	295.60	294.21	293.13	292.66	292.56	292.54
	Model 2b[§]	305.21	295.60	294.21	293.13	292.66	292.56	292.54
	Model 3b	305.21	295.60	294.21	293.13	292.66	292.56	292.54
	Model 5	305.21	295.60	294.21	293.13	292.66	292.56	292.54
	Model 4	305.21	295.60	294.50	293.51	293.02	293.05	292.94

[†] *With tangent linear constraints* [§] *With traversing linear constraints*

As depicted in Table 3.8, losses have a considerable impact on the overall system cost (which includes the operation and the network investment costs). For small values of L, the system costs tend to be overestimated in all the models, except for Model 2a (the one based on tangent constraints), in which the total system costs are underestimated. As L increases beyond 5, the effect of the model choice on the total system costs becomes insignificant because losses are represented accurately.

The performance of models related to their computational requirements can be assessed based on the figures provided in Table 3.9. In this table, we show the time elapsed when running the TEP problem for the Garver's system using each of the losses models considered. As L increases, Model 4 becomes computationally very demanding compared to the others, because its formulation includes binary variables.

Table 3.9 also shows there are small differences among the computational performances of the considered models. Model 2b behaves very well despite the fact that it is mathematically more complex due to the additional constraint added as an upper limit to the feasible losses space. This suggests that shrinking the feasible space by adding an upper bound has a substantial contribution to speeding up the solution process.

Table 3. 9 Effect of Numbers of Partitions in Losses Linearization on TEP's Computation Time in the Garver's System

L	Computation times for each model (in seconds)						
	Model 2a[†]	Model 2a[§]	Model 2b[§]	Model 3b	Model 3a	Model 5	Model 4
1	0.875	0.953	0.813	0.829	0.833	0.823	0.906
2	0.900	0.984	0.875	0.925	0.834	0.838	1.188
5	0.930	1.010	0.899	1.078	0.855	0.872	1.688
10	1.010	1.050	0.954	1.110	0.870	0.928	2.186
15	1.050	1.080	0.985	1.172	0.901	0.963	3.016
20	1.091	1.130	1.020	1.192	0.923	0.981	6.226

[†] *With tangent linear constraints* [§] *With traversing linear constraints*

In general, it seems that Model 5 is attributed with the lowest computational requirements, with the exception of model 3a. In this particular case, the savings in computing time achieved by Model 5 with respect to the other models, apart from 3a, ranges from 3% to 24% when L is set to 5, as depicted in Table 4. Using Model 3a in a TEP optimization may result in a faster convergence of the algorithm than using Model 5. However, in order for the solution provided by model 3a to be acceptable, either of the following two conditions must be met: (i) there should not be any operational condition in the considered system that can lead to artificial losses, or (ii) the effects of artificial losses should be deliberately neglected. However, knowing a priori (i.e. before solving a TEP or an OPF problem) whether artificial losses make economic sense in a system is very difficult.

When planning the expansion of the grid in this test system for the set of data in [39] (i.e. in the base-case scenario), artificial losses are not computed regardless of which losses model is used in the TEP optimization. . This may be related to the fact that the original generation in this system does not include any renewable generation. For the purposes of assessment here, another scenario with some specific changes to the data in [39] has been defined. Changes made to the original data are as follows:

- A low-cost wind power generator is included at node 1, with a capacity five times greater than the original capacity of the generator at that node. This leads to a 45% penetration level of wind power in terms of installed capacity.
- The capacity of line 1-4 is upgraded from 100 MW to 500 MW
- The capacity of line 2-4 is derated by 50%
- Demand at node 2 is decreased by 50% and
- Demand at node 4 is increased by 400%

Note that apart from the above changes, the remaining data (including demand, generation and network parameters) are kept the same as in the base-case scenario. In the new scenario, artificial losses as high as 5 times the actual losses are computed for line 1-2 using Model 2a. Thus, the only network investment found optimal when using Model 2a is in corridor 3-5. In contrast, using Models 4 and 5, which avoid artificial losses, result in reinforcements in corridors 2-4, 3-5 and 4-6.

As expected from previous analyses, using Model 2b, 3a or 3b significantly reduces artificial losses to about 89, 74, and 72% of the actual value of losses, respectively. This indicates that the features added to Models 2b and 3b manage to effectively limit artificial losses. However, sometimes, artificial losses computed in these three models also have an impact on the optimal network expansion solution. Thus, in the case of this test system, the optimal expansion solution computed when using either of these losses models differs from that computed when using Models 4 and 5. Thus, if Models 2b, 3a, or 3b are used, network investments computed do not concern corridor 2-4. Reinforcing this corridor is avoided by artificially increasing losses.

Even though Model 1 avoids artificial losses as well, its use in TEP or OPF problems results in significant errors in losses estimation. In this case, losses computed with this model are about 37% lower than real ones, which is not acceptable. The impact of such a deviation on the network expansion solution computed can also be substantial. This occurs when using Model 1 in the new scenario defined above, in which reinforcing corridor 4-6 is not deemed optimal.

To support the conclusions drawn from the analyses of the previous case study, similar tests have been carried out on a system featuring a medium-scale network: the IEEE 118-bus system. Data used in this analysis can be found in [222]. Since the analysis conducted is basically the same as in the previous case study, the results are skipped in this document but can be found in [13]. In order to assess the accuracy of losses computed by each model, the benchmark level of losses is obtained by solving an AC-OPF problem. Note that the AC-OPF problem is formulated for the network configuration which includes the truly optimal network investments. This benchmark value amounts to 489.52 MW. The relative error made in the estimation of losses, with respect to this reference value, drops below 10% in all models for L greater than or equal to 5. Having a 10% error in losses estimates may be deemed acceptable in many cases because an error of this magnitude normally do not have a relevant impact on the network expansion solution computed. The results here support the choice of Model 5, or Model 3a, since the savings

achieved by Model 5 in the computation time with respect to other models, apart from model 3a, ranges from 17% to 40%, as shown in Table V. However, given the inability of Model 3a to sufficiently limit artificial losses, it should be considered as a reasonable option only in systems where artificial losses are not relevant.

To further validate the findings in this work, a real-life large-scale system featuring the Spanish system has been considered. The electricity network in this case-study comprises 425 nodes and 628 transmission lines. Both wind and solar power generation existing in the Spanish system are included. All in all, a 25% penetration level of power generation from RES is considered in the case study. Test results from this system generally shows that Model 3a demands the least computation effort while delivering similar results to other models in terms of accuracy. Model 5 appears to be, computationally speaking, the second best performing model. Differences in computation times between Model 5 and Model 3a are nevertheless insignificant. As pointed out earlier, using Model 3a in TEP studies makes sense only if one can anticipate that artificial losses computed by any model will be negligible, and/or the system considered in the studies is unlikely to result in artificial losses.

Talking about the accuracy in the computation of losses and the resulting overall system costs, it can be observed that all models provide quite similar values for losses and system costs when L is greater than or equal to 5. Regarding the computational time, Model 5 and Model 3a clearly outperform the rest of the models, achieving savings in computation time as high as 25% with respect to the third best performing model in this regard.

All this suggests that Model 5 is the most appropriate losses model for large-scale network expansion optimization problems, since it strikes a good trade-off between accuracy in losses representation (including avoidance of artificial losses), and computation time required to solve the TEP problem.

3.5.3.   Effects of Number of Partitions on TEP Solutions

Since the estimation accuracy of losses by the linear models depends on the value of L, it can also be expected that the choice of L affects network expansion planning outcomes. For the Garver's system, the optimal network investment plan involves the reinforcement of corridors 2-3, 2-5, 2-6, 3-5 and 4-6, in line with the AC network expansion solution reported in [36]. When artificial losses are not computed, using any of the models assessed here find the same optimal network expansion solution for any value of L greater than or equal to 3.

However, the threshold value for L in the two larger test systems is 5, beyond which the investment decisions do not change. When L is set to a value lower than the threshold (i.e. 3 in the Garver's system or 5 in the other two), investment decisions depend on which losses model is used. Generally, Model 2a (the one based on tangent constraints) underestimates both losses and line investments required. The remaining models overestimate losses and, consequently, result in overestimated network expansions.

3.5.4. Concluding Remarks

The main motivation behind this study is the need to choose or develop an adequate losses model for large-scale TEP applications. Such a model should be computationally efficient, provide a reasonably accurate estimate of losses in every line and in particular, avoid the computation of artificial losses aimed at alleviating network congestion (a common drawback of many linear models).

The compliance with these requirements has been separately analyzed for several losses models. Besides, and most importantly, the impact of the use of each model on the outcome of network expansion planning has also been assessed. In particular, four alternative linearization methods have been evaluated together with other four variants of existing methods. The performance of these methods has been assessed by embedding each method in TEP problems pertaining to small, medium and large-size systems.

The results show how the accuracy of estimated losses increases with the number of linear partitions L considered in the linearization of nonlinear losses curve. However, increasing L beyond a certain threshold has no significant effect on losses estimates and TEP results. Then, it is not worth the extra computational burden. It seems that 5 partitions are sufficient to compute a reasonably accurate estimate of losses for medium and large-scale systems. Higher number of partitions (greater than or equal to 5) results in relative errors below 10% and 5% in the estimates of losses for the medium-scale system, respectively. This is acceptable from the TEP context since such small deviations in the estimation of losses are not likely to influence TEP results.

Regarding the computational results, the SOS2-based linear losses model is found out to be the most efficient, having computational advantage over the other assessed models with as high as 40% reduction in solution time. In contrast, models which involve regular binary variables are certainly the most computationally intensive, which makes them inappropriate for large-scale expansion planning problems.

The additional features included in Models 2b and 3b achieve some improvements in the accuracy and/or the computational efficiency of their original versions, Models 2a and 3a. For instance, the inclusion of an upper losses constraint in Model 2a makes the resulting Model 2b perform better than the former in terms of computational efficiency while yielding similar results. On the other hand, adding the logical precedence constraint in Model 3a when deriving Model 3b results in an increase in the computation time required to solve the TEP problem but achieves an increase in the level of accuracy of the representation of network losses.

The results also show the importance of avoiding, or at least limiting, artificial losses computed. This is especially true when the problem being dealt with is the computation of long-term network expansion of large-scale systems featuring large amounts of RES generation because, in such a problem, having artificial losses may make economic sense.

Models which make use of inequality constraints to represent losses define an unbounded feasible space of losses. As a result, such models fail to limit artificial losses. However,

those models that include an upper bound constraint for the feasible losses space, as well as the piecewise-linearized losses models, largely suppress artificial losses. The SOS2-based model, on the other hand, avoids them completely. Despite their complexities, models based on additional binary variables guarantee that artificial losses are effectively avoided.

All in all, the proposed SOS2 based approach balances accuracy very well with computational burden of the resulting TEP problem. For instance, in the large-scale system considered in the present work, using this model demands lower computational effort than using any other model considered in the study (with reductions in time achieved higher than 33%). This makes the SOS2 based approach the best candidate for modeling losses in very large-scale TEP problems.

3.6. SUMMARY

Both the tractability of a TEP problem and the accuracy of an expansion solution largely depend on the level of system details captured by the expansion model. This is associated with the characterization of physical network variables, in particular, flows and losses. From this angle, this chapter has presented an extensive review of the most commonly used TEP optimization models with different mathematical complexity levels, theoretical and numerical comparisons of these models from the viewpoint of expansion solution accuracy and computational requirements. Contributions from this chapter include the systematic comparisons of various existing TEP models, and some improvements and proposed changes to the mathematical modeling of existing TEP models that can speed up the computational process. Some of these include two variants of the DC expansion planning model and flow-based losses representations in all TEP models. Instead of the angle-based losses representation commonly used in TEP studies, this work proposes a flow-based losses model which has a significant computational advantage over the angle-based equivalent losses model.

The comparative analyses of linear TEP models also includes the effect of network losses on the expansion outcome. Analysis results have showed that neglecting network losses can lead to underestimation of network investment needs. Hence, modeling losses should be an integral part of TEP models. The fact that network losses are a function of quadratic flow adds complexity to the TEP model because of its nonlinear and nonconvex nature. It should be linearized to keep the entire problem linear. In this regard, the provision of rigorous theoretical and technical analyses, and exhaustive performance comparisons, of several losses models has been presented in this chapter. Existing linear losses models are thoroughly assessed in terms of their accuracy in losses representation as well as the contribution in computational burden. Generally, this assessment has revealed that existing models are not adequate because of either their accuracy related issues or computational limitations. Motivated by this, this thesis has proposed two novel linear losses models as well as two modified versions of existing ones that address accuracy and computational issues inherent to the existing losses models.

4

IV. UNCERTAINTY AND OPERATIONAL VARIABILITY MANAGEMENT

This chapter introduces the novel method developed in this thesis for handling the uncertainty and variability in a very efficient manner. Numerical results as well as the computational implications from applying this method in TEP problems are presented and discussed at the end of this chapter.

4.1. INTRODUCTION

4.1.1. Description of Terminologies

The terminologies uncertainty and variability are sometimes incorrectly used interchangeably in the literature despite the fact that they are different. Variability, as defined in [223], refers to the natural variation in time of a specific uncertain parameter, whereas uncertainty refers to "*the degree of precision with which the parameter is measured*" or predicted. For example, wind power output is characterized by both phenomena; its hourly variation corresponds to the variability while its partial unpredictability (i.e. the error introduced in predicting the wind power output) is related to uncertainty. The schematic illustration in Figure 4.1 clearly distinguishes both terminologies. As demonstrated in this figure, the hourly differences in wind power outputs are due to the natural variability of primary energy source (wind speed); whereas, the likelihood of having different power outputs at a given hour is a result of uncertainty (partial unpredictability) in the wind speed.

Other terminologies used in this thesis are snapshot and scenario. A snapshot refers to an hourly operational situation. Alternatively, it can be understood as a demand—generation pattern at a given hour. A scenario, on the other hand, denotes the evolution of an uncertain parameter over a given time horizon (often yearly). For example, the hourly variations of wind power production and electricity consumption collectively form a group of snapshots; whereas, the annual demand growth (which is subject to uncertainty) and emission price uncertainty are represented by a number of possible storylines (scenarios).
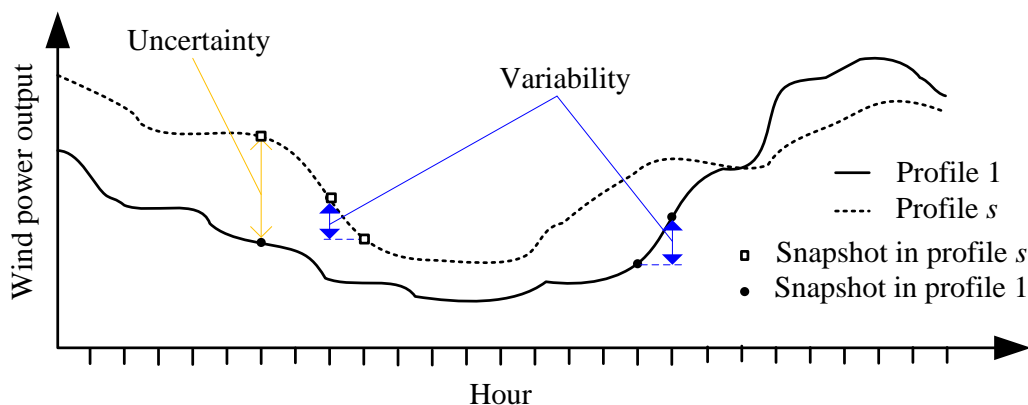
Fig. 4. 1 Illustration of variability and uncertainty in wind power output

4.1.1. Overview of the Chapter

The global drive for integration of renewable energy sources (RESs) means they will have an increasing role and a profound impact on power systems. On one hand, it is inevitable that such resources introduce more variability and uncertainty to the system operation because of their intermittent nature. On the other hand, achieving large shares of RES power production results in a more relevant the role of electricity networks since the variability of the power production from such energy sources involves the need to develop larger amounts of interconnection capacities among zones to ensure security of supply at zones where RESs (wind and sun, for example) are scarce or not available. Hence, Transmission Expansion Planning (TEP) becomes a more relevant issue since the variability and uncertainty of RES power production significantly increase the amount of operational situations to be considered.

TEP involves solving an optimization problem subject to multiple sources of complexity, such as the use of discrete variables, its non-linear behavior, and the existence of several levels of uncertainty. As aforementioned, this problem is especially hard to solve when the goal is the long-term expansion of a large network in a power system featuring large amounts of generation from RESs, since in this case the size of the problem increases very substantially. Moreover, the addition of new transportation load such as electric vehicles, railways, etc. also brings in more operational uncertainty to the system. It is therefore mandatory that long-term TEP tools consider the operational impact of such uncertainties and variability in system conditions, since additional investments may be required to expand the network. In principle, such objective can be met by considering a large number of operational states but this leads to a computationally intractable TEP problem. Improving the management of such kind of uncertainty in TEP problems is one of the main focus areas of this thesis, contributing therefore, to a more cost efficient penetration of RES energy in power systems.

The different sources of uncertainties in long-term TEP that are related to the variability and unpredictability of situations are usually classified as random and nonrandom [12]. The random ones are also known as high-frequency uncertainties because they correspond to situations that occur repeatedly. Hence, they can be characterized by probability distribution functions (PDFs), estimated by fitting historical data. Such uncertainties have a profound impact on the operation of power systems. Demand variability is one example of random uncertainty. On the other hand, nonrandom uncertainties do not occur repeatedly; so they cannot be estimated by PDFs. A good example here is generation expansion.

In light of this, an appropriate long-term TEP tool should account for both types of uncertainties. Because of their aforementioned differences, different methodologies are employed to effectively deal with each type of uncertainty. Nonrandom uncertainties are often modeled by a set of possible future scenarios, each with a certain probability of realization. This will not, however, be the subject of this thesis; instead, the work here focuses on the art of dealing with the variability of operational states and the associated random uncertainties (i.e. sources of operational uncertainty). Therefore, it should be noted here that the literature review is also limited to this subject area.

As introduced above, this thesis focuses on a particular aspect of the global TEP problem, namely the operational variability and uncertainty of the system which is introduced by the so-called random uncertainties. This is the level of uncertainty that remains when one considers known and constant factors of nonrandom uncertainties such as generation investments, costs and prices, economic growth (average demand growth level) and policy-related parameters. Operational variability and uncertainty include, for instance, component outages or availability, demand variability, and wind and solar power output variability. If such uncertainties are not properly managed, the quality of network expansion planning solutions could be significantly jeopardized.

In spite of the aforementioned facts, the network expansion planning of power systems has often been solved using a deterministic approach, where the effect of operational uncertainty and variability is not accounted for, or represented in an overly simplified way, such as planning for the worst-case state —traditionally peak demand (i.e. the most stressful state from network point of view), or a proxy of this (for instance, dealing with a very limited number of operational states). However, this is not valid in current power system planning, especially in long-term TEP problems because of the large variability in operation conditions that can result in added stress to the system. The variation of operation conditions throughout the planning horizon, which cannot be predicted appropriately, is the main source of the operational uncertainty. Even if planning the network expansion for the most-stressful (worst-case) operation situation were an elegant approach, it would be very difficult to identify the "worst-case", since it would be unrealistic to expect it to happen at the peak load time. In relation to this, it has been particularly reported in [224] that transmission investment decisions made under uncertainty are more robust than their deterministic counterparts. Authors in the former work highlight the benefits of including uncertainty in TEP studies. However, mainly because of computational reasons, many sources of uncertainty and variability are frequently ignored, partially addressed, or represented by few predefined operational states in TEP models. It is obvious that handling large instances of operational situations is not computationally feasible and/or efficient in power systems planning. On the other hand, inadequate consideration of operational situations could adversely affect decision-making. Therefore, an operational variability and uncertainty management tool that balances accuracy with computational burden is needed in TEP studies. Given that operational uncertainty and variability resulting in the need to consider a multiplicity of operational states in TEP studies are very much linked, and largely related to the existence of RES generation, we shall deal with both of them jointly in the remainder of this chapter under the name of operational uncertainty management.

4.2. PROPOSED METHOD OF OPERATIONAL UNCERTAINTY MANAGEMENT

Despite the vast literature on TEP, current modeling and planning practices have some limitations with regards to handling operational uncertainty because: (1) they tend to incorporate only a few sources of operational uncertainty (often one or two) while many sources of uncertainty are unaccounted for; and (2) spatial and temporal correlations among the uncertain parameters are largely neglected. In general, currently available

network expansion planning methods are not adequate to handle large-scale systems while appropriately taking account of operational uncertainty. Therefore, there is still a need to develop a scheme to accurately represent uncertainty in the context of TEP applied to large scale systems. The scheme adopted for uncertainty treatment should be able to capture the variability of relevant uncertain parameters and correlations existing among them, especially for long-term TEP with high penetration levels of renewable generation. The work in this thesis may be deemed a probabilistic method as explained above in detail. As it shall be explained in the following paragraphs and sections, differences with existing approaches are related to the criteria employed to select the set of operational states considered in the TEP problem, and the level of detail considered in representing the system.

As mentioned above, operational uncertainty can be handled as the variation of stochastic parameters which are repeatable in time (often hourly) and exhibit a random behavior with known approximate probability distributions. It can broadly be represented by a set of operational states, here referred to also as "snapshots", each containing a generation—demand pattern (i.e. with different levels of demand at each node and generator outputs). Each operational state can be considered as a generation—demand pattern of the power system, which leads to an OPF pattern in the network. A large set of snapshots, each one with an estimated probability of occurrence, is assumed to be already available to evaluate and optimize the network expansion. In particular, hourly generation—demand data for a given target planning year (8760 snapshots, in total) are considered in this work. All snapshots are assumed to have the same probability of occurrence, therefore given by 1/8760. It should be noted here that this can be scaled to any number of snapshots.

A common practice to handle such uncertainty is to perform a clustering process over the multi-dimensional stochastic input dataset (i.e. generation—demand patterns) [165]. Clearly, the overall accuracy of the TEP solutions in this regard depends on the selection of the clustering variables. The accuracy is indeed determined by how representative the clusters are with respect to the original set of operational situations or snapshots. Usually, a large number of clusters are required to achieve a reasonable level of accuracy.

This thesis shows how to reduce the number of clusters, corresponding to operation snapshots considered in the TEP problem, without a relevant loss of accuracy in the TEP results, using an adequate selection of the classification variables in the clustering process. The proposed method relies on two ideas. First, the snapshots are characterized by their OPF patterns (the effects) instead of generation—demand patterns (the causes). This is because the network expansion planning is the target problem, and losses and congestions (resulting from the OPF) are the drivers of network investments. Second, OPF patterns, after some processing to represent their relevant features as "fingerprints", are classified using a "moments" technique, a well-known approach to address Optical Pattern Recognition problems. To the best knowledge of the authors, this is the first time this technique has been applied in a TEP problem.

83

The proposed clustering method is conceptually illustrated in Figure 4.2. This Figure illustrates the process to follow in order to compute, for each snapshot t taken from the generation—demand dataset, the values of classification variables, the relevant moments, to be considered in the clustering analysis. A description of this process follows:

1) The OPF of the snapshot is computed neglecting transmission line capacities.
2) The transmission lines with more relevant congestions (overloads) and losses are selected.
3) The selected lines are represented as graphical objects, with properties such as location, orientation, thickness (overload or losses) and length. This arrangement of objects can be deemed the "fingerprint" of the snapshot.
4) The graphical pattern, or snapshot's fingerprint, is then coded into a reduced-dimension space defined by moments. This technique is common in Optical Pattern Recognition problems.
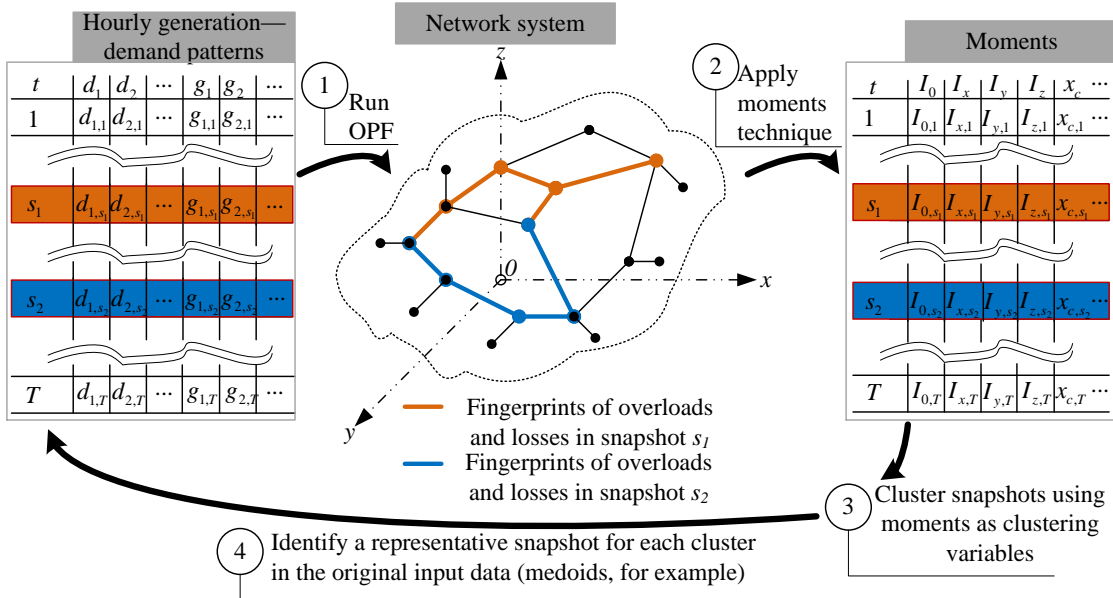
Fig. 4. 2 Conceptual illustration of the proposed clustering methodology

Snapshots where network investment needs are similar have similar fingerprints of network overloads caused by non-capacity constrained economic flows. Similar fingerprints should result in similar "moments". Thus, if the set of moments is properly selected, non-similar snapshots should result in different moments. This means that the snapshots can be effectively clustered using their moments as clustering variables, i.e., computing the distances in the moments space as measures of similarity.

Once the clustering process is completed, the last step involves choosing from the original dataset a representative snapshot for each cluster. This can, for example, be the medoid of the set of snapshots grouped together in the corresponding cluster. Note that $d_{i,t} (\forall t \in \Omega^{SBC} \quad and \ \forall i \in \Omega^{NS})$ and $g_{i,t} (\forall t \in \Omega^{SBC} \quad and \ \forall i \in \Omega^{NS})$ in Figure 4.2 represent the demand

84

and the generation levels at node i for snapshot t; whereas, $I_{m,t}(\forall t \in \Omega^{SBC} \ and \ \forall m \in \Omega^{MS})$, $x_{c,t}(\forall t \in \Omega^{SBC})$, etc. are the corresponding computed moments.

In general, the main contributions of this work include:

- The definition of a novel method for clustering operational states, including a detailed description of the process being followed, and the provision of formulations of each stage in the process. The main features of this clustering method are listed next:
  - o The method is a tailor-made approach for TEP problems;
  - o It involves systematic management of operational uncertainty in TEP problems, leading to an accurate representation of uncertainty, which makes this approach suitable to be applied in TEP of systems with significant share of generation from RESs;
  - o It allows compact representation of snapshots via a new set of clustering variables, and the compactness of the set of the clustering variables derived leads to a significant reduction in computational burden, which makes this method suitable for the TEP of large systems.
- The comparative analyses of results produced by our method and other snapshot clustering methods that, contrary to the former, are based on the causes of optimal power flows.

Other contributions in this chapter include the new quasi-linear losses model used in the capacity unconstrained economic dispatch problem, and the nonlinear optimization approach developed to estimate the geographical coordinates of a test system.

4.3. NETWORK CAPACITY UNCONSTRAINED ECONOMIC DISPATCH

In order to characterize the snapshots by their OPF patterns (the effects or results of system operation), a Network Capacity Unconstrained Economic Dispatch (NCUED) model is used. This model is similar to the "copper sheet" TEP model described in Chapter 3. In this model, transmission capacity constraints are neglected (relaxed), leading to the assumption of having a flexible network. Technically speaking, this means that the constraints corresponding to the power transfer capacity limits of existing corridors are not active in the NCUED model. As a result, each existing corridor has the flexibility to accommodate any amount of flows that increase the overall system welfare as far as the flows respects the Kirchhoff's laws. For example, suppose a given line has a capacity of 100 MW. In an ordinary economic dispatch problem, this constraint has to be included, which means the line cannot carry more than 100 MW of power. However, in the NCUED problem, the capacity constraint is not imposed, allowing the line to transport more than 100 MW of power. This assumption makes sense in order to detect TEP investment needs, since the latter are closely related to the corridors of the system in which investments would have the largest impact on system operation by allowing the largest increase in flows that are efficient from an economic point of view. In this way, the aim here is to consider the relevance of each

snapshot on prospective expansion needs. Snapshots that result in similar patterns of overflows in lines may then be grouped together, because this means that similar network investments will be needed to increase the efficiency of the system operation.

The NCUED model minimizes the total operation cost in (100), which includes the costs of generation (I), unserved power (II), and emissions (III), subject to the set of DC-OPF based constraints in (101) and the losses model provided by the set of constraints in (102). Issues related to the formulation of these cost terms are discussed in Chapter 3.

$$
\min_{PG_{g,b}, p_{i,b}} Z = \underbrace{\sum_{t\in\Omega^t}\sum_{s\in\Omega^s}\rho_s\sum_{w\in\Omega^w}\pi_w\sum_{g\in\Omega^g}(1+\sigma)^{-t}PG_{g,s,w,t}\lambda_{g,s,t}}_{I}
$$
$$
+ \underbrace{\sum_{t\in\Omega^t}\sum_{s\in\Omega^s}\rho_s\sum_{w\in\Omega^w}\pi_w\sum_{i\in\Omega^i}(1+\sigma)^{-t}p_{i,s,w,t}\Lambda}_{II} \tag{100}
$$
$$
+ \underbrace{\sum_{t\in\Omega^t}\sum_{s\in\Omega^s}\rho_s\sum_{w\in\Omega^w}\pi_w\sum_{g\in\Omega^g}(1+\sigma)^{-t}\lambda_{s,t}^{CO_2e}ER_gPG_{g,s,w,t}}_{III}
$$

Subject to:

$$
\left.\begin{array}{c}
P_{k,s,w,t} + V_{nom}^2 b_k\theta_{k,s,w,t} = 0; \ where - b_k = 1/x_k \\
\left|P_{k,s,w,t}\right| + 0.5PL_{k,s,w,t} \le u_{k,s,w,t}S_{k,max} \\
\sum_{k\in i}P_{k,s,w,t} + \sum_{g\in i}PG_{g,s,w,t} + p_{i,s,w,t} - \sum_{d\in i}PD_{d,s,w,t} + 0.5\sum_{k\in i}PL_{k,s,w,t} = 0 \\
u_{g,s,w,t}PG_{g,s,w,t.min} \le PG_{g,s,w,t} \le u_{g,s,w,t}PG_{g,s,w,tmax} \\
\theta_{min} \le \theta_{i,s,w,t} \le \theta_{max}; \ \theta_{ref} = 0 \\
P_{k,s,w,t} = P_{k,s,w,t}^+ - P_{k,s,w,t}^- \to \left|P_{k,s,w,t}\right| = P_{k,s,w,t}^+ + P_{k,s,w,t}^- \\
u_{k,s,w,t} \in \mathbb{R} \ and \ 0 \le u_{k,s,w,t} \le \infty
\end{array}\right\} \tag{101}
$$

$$
\left.\begin{array}{ll}
L1: & PL_{k,,s,w,t} \ge \dfrac{0.5r_k S_k^{max}\left|P_{k,s,w,t}\right|}{S_B} \\[3mm]
L2: PL_{k,s,w,t} \ge \dfrac{r_k\left(0.96S_k^{max}\left|P_{k,s,w,t}\right| - 0.24(S_k^{max})^2\right)}{S_B} \\[3mm]
L3: & PL_{k,s,w,t} \le \dfrac{r_k S_k^{max}\left|P_{k,s,w,t}\right|}{S_B} \\[3mm]
L4: & PL_{k,s,w,t} \ge 0
\end{array}\right\} \tag{102}
$$

If required, an additional term may be added to Equation (100) to factor in the investment cost of the "would-be" lines capable of accommodating the extra flows in a corridor (i.e. for flows beyond $S_{k,max}$) by multiplying the net extra MW needed in each line by a fixed capital cost per MW. This should be then weighted by the capital recovery factor and amortized in fixed installments during the lifetime of the line. Should this be adopted, Equation (100) need to be modified to account for the

associated costs during and after the planning horizon. This is extensively discussed in Chapters 3 and 5.

In equation (102), $L1$, $L2$, $L3$ and $L4$ represent the linear constraints of the losses model considered in the NCUED model. Note that the set each scenario in a given planning stage contains a certain number of operational samples (or at least 8760 hourly demand-generation samples corresponding to the total number of hours in a given year). Note that the constraints in (101) in the NCUED model, are applicable only for existing lines. In order to obtain proper estimates of the losses in overloaded lines, the losses model formulation for each individual line in Chapter 3 is replaced here with the quasi-linear losses model in (102). This is based on the following plausible assumption: in a congested corridor, there "exist" parallel lines (able to transport all flow in the corridor) whose capacity limits are identical to that of the existing line in that corridor. This is based on the nature of the NCUED model (also known as the "copper sheet" model). Note that the coefficients in $L2$ are obtained by minimizing the mean squared error as a result of representing losses by a linear curve. The quasi-linear losses model used here is illustrated in Figure 4.3. This figure shows the losses model considering the installation of four parallel lines in a given corridor. It is straightforward to extend this to a higher number of parallel lines.
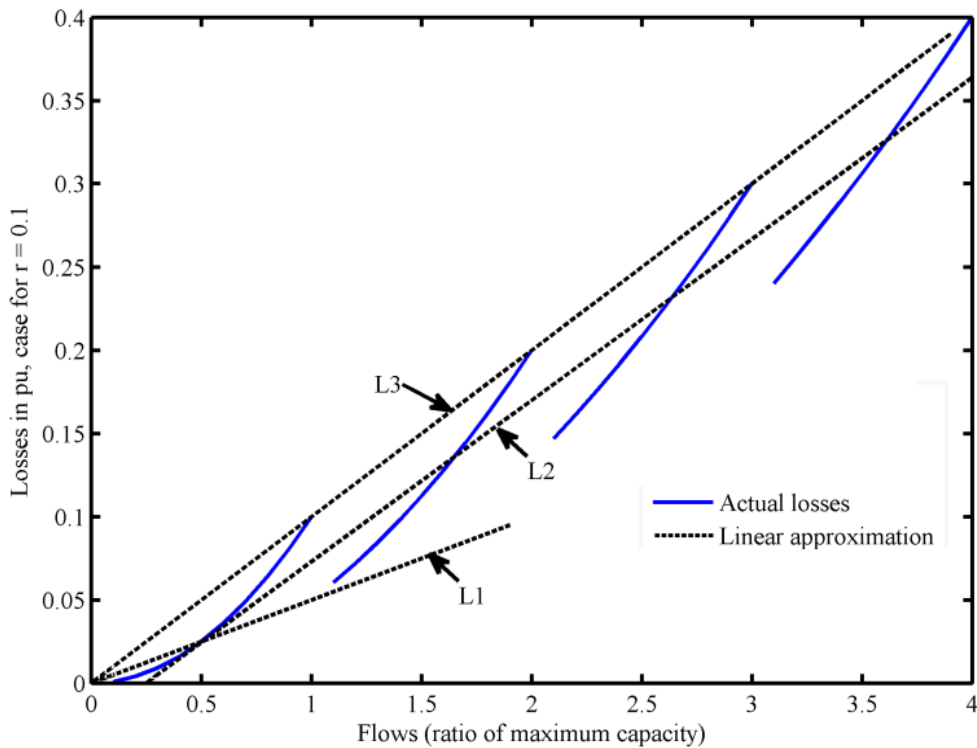
Fig. 4. 3 Losses model for the NCUED model (with potentially 4 parallel lines)

4.4. DEFINITION OF CLUSTERING VARIABLES

This thesis proposes the use of new classification variables in the clustering process of operation states (snapshots), which is especially devised for TEP analyses. These variables may then be used in any clustering algorithm, and the standard *k-means*

algorithm has been applied in the test cases. This section describes the criteria applied for the selection of clustering variables.

4.4.1.   Selection of Operation Variables for Network Expansion Planning

Power production and demand patterns are used as clustering variables in many power system planning applications such as contingency and reliability analysis [165], electricity supply analysis [166], TEP [130] and medium-term thermal scheduling [225] problems. However, such an approach (hereinafter, clustering based on causes, CbC), is not appropriate for TEP because some snapshots, apparently different, may result in the same transmission investment needs.

Instead of considering the production/demand patterns (the causes), the clustering process proposed here is based on the effects of such patterns on the transmission grid, because the effects (congestions and losses) are more closely related to network investments needs.

For the sake of simplicity, we use the two node system in Figure 4.4 to illustrate the proposed clustering methodology. Let us assume that we have two intermittent generation sources connected at each node. The electricity demand at each node is assumed to be 100 and this remains the same for the seven snapshots which we will consider here (see Table 4.1). Assume further that the capacity of the transmission line is 50, and 40 considering a 20% security margin.
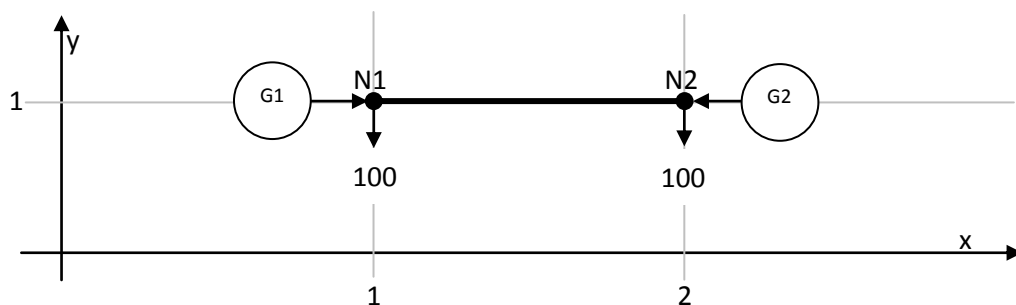
Fig. 4. 4 A system for illustrating the methodology

Given the snapshots in Table 4.1 for this system, we want to obtain four clusters using the generation patterns as clustering variables (the conventional approach) and the proposed method, and compare the results. In this regard, when we apply the conventional clustering method, we obtain the clusters in column 7 of Table 4.1. Statistically, these clustering results make sense. However, one can observe some inconsistency in the clusters when measured in terms of expansion needs. They are not generally representative because the effect of each snapshot on the system is lost. For instance, the first two snapshots have the same effect in terms of expansion needs because both result in congestion in the line with an overload of 60 in either direction. Yet, they are grouped into two different clusters. Snapshots 3 and 4 also have the same effect in terms of TEP, both creating an overload of 10 MW in the line, but these

snapshots again fall into different groups. This means that the first four snapshots result in an overload in the line (noting the difference in magnitudes of the overloads). The last three snapshots do not overload the line; hence, they are non-overloading snapshots.

Table 4. 1 Illustrative example

Snapshots	Classification Variables			Unbalances		Clustering Results	
	G1	G2	Absolute overflow	D1-G1	D2-G2	Clustering Based on G1 and G2	Proposed Clustering Method
Snapshot 1	200	0	60	NA	NA	1	1
Snapshot 2	0	200	60	NA	NA	3	1
Snapshot 3	150	50	10	NA	NA	2	2
Snapshot 4	50	150	10	NA	NA	3	2
Snapshot 5	100	100	0	0	0	4	3
Snapshot 6	125	75	0	-25	+25	4	4
Snapshot 7	75	125	0	+25	-25	3	4

NA: Not Applicable

If we cluster the snapshots by taking into account their effects instead of the causes, we obtain very realistic clusters. Note that we can determine the moments of overloads about any axis as per the proposal and the clustering results do not change. However, it is not necessary to do so here because we only have one line. As far as the non-overloading snapshots are concerned, we can see that the last two snapshots have the same effect when it comes to losses in the line. As a result, it makes sense from TEP point of view that they should be grouped together. Figures 4.5 and 4.6 compare the clustering results obtained by classical and the proposed method, respectively.
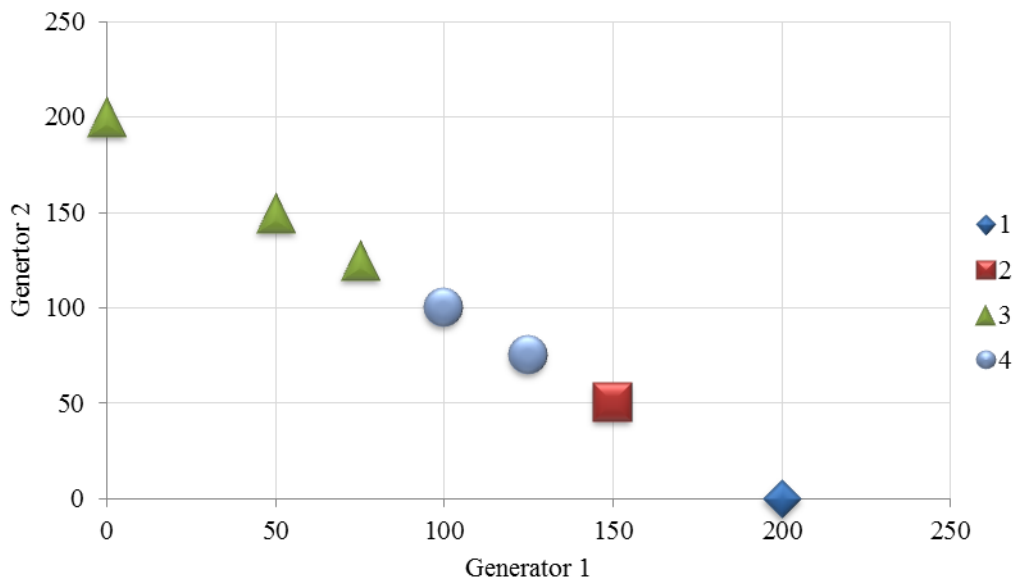
Fig. 4. 5 Clustering results using conventional approach
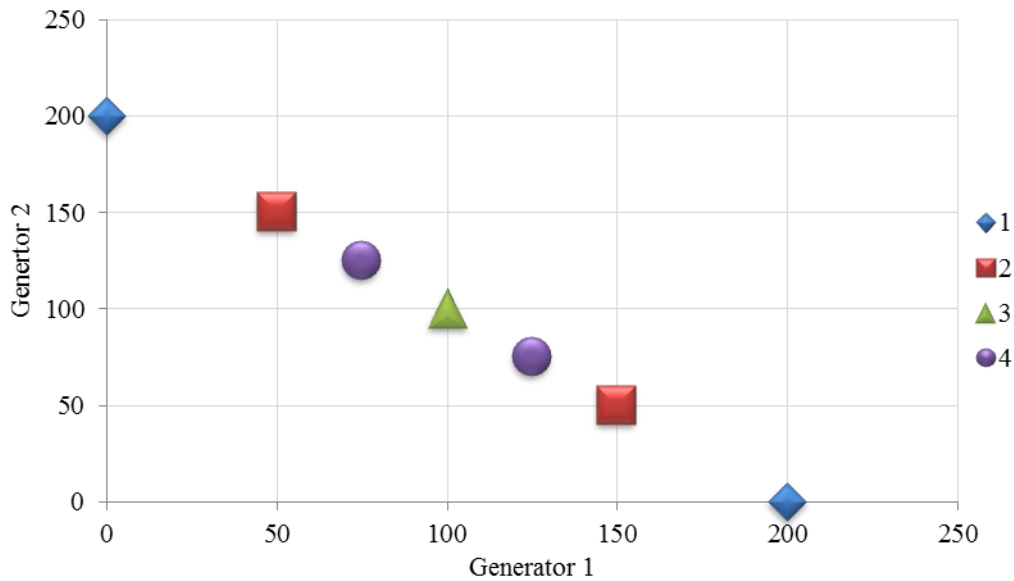
Fig. 4. 6 Clustering results using proposed method

Generally, comparison results of this simple example demonstrate how effective the proposed clustering methodology is in obtaining representative clusters in terms of network expansion needs.

In network expansion planning, the benefits of investing in a certain corridor can be measured in terms of reduction of network congestion, either under normal conditions or in a contingency situation, and/or the reduction of overall losses. Since both congestion and losses are directly related to power flow patterns, these patterns are the subjects of the proposed classification method.

Power flow patterns should identify the areas of the network where reinforcements have the largest potential to reduce operation costs. They may also show the estimated size of the reinforcements to be made. Thus, snapshots with similar power flow patterns should also lead to similar investment solutions, and when clustered together (with accumulated probability), a reduced set of clusters representing a reduced set of flow patterns may be successfully used in a TEP problem, instead of the large original set of operational states.

According to the features of flow patterns leading to investments, snapshots may be grouped into two big categories: overloading snapshots (those which lead to relevant network congestion), and non-overloading ones.

In overloading snapshots, where network congestion is relevant, making flows compatible with existing line capacities probably causes a significant increase in operation costs. Then, one should expect that network reinforcements are largely related to the need to reduce congestion. On the other hand, in non-overloading snapshots, only those reinforcements that are able to significantly reduce network losses can make economic sense.

90

Line overloads in the NCUED model, whose formulation has been provided in the previous section, reflect the size of the extra flow of power in each line, beyond its capacity, that would make economic sense given the current topology of the grid. Hence, network investment needs associated with overloading snapshots are closely related to the size and location in the network of overloads resulting from the NCUED. Therefore, in overloading snapshots, only the flows in lines that are close to congestion are taken into account in the clustering process.

On the other hand, the pattern of losses in the transmission allows characterizing potential network investments related to non-overloading snapshots. Losses in the network are the result of flows created by unbalances of power production and demand in the network. Therefore, different patterns of transmission losses should be the result of different patterns of unbalances of power production and demand in the system. Given that the location of conventional generation available to produce power is relatively stable across operation snapshots, non-overloading snapshots can reasonably be clustered using the pattern of demand and available RES power production.

It can, therefore, be concluded that the size and the location of *line overloads* caused by economic power flows in the NCUED of the system are probably the most relevant to cluster those snapshots that result in congestion-relieving investment needs. For the rest of the snapshots, the pattern of unbalances of demand and available renewable electricity production can be used as clustering criteria since their related investment are aimed at reducing losses.

4.4.2. The Use of Moments of Relevant Network Expansion Drivers

In overloading snapshots, considering line overloads and their location as clustering variables would define a number of clustering variables equal to three times the number of overloaded lines (N_{ol}), which may be a relevant fraction of total lines.

In non-overloading snapshots, considering net demands in system nodes (demand minus net available RES power)–and their location—as clustering variables, would define a number of clustering variables equal to three times the number of system nodes (N_N). Then, grouping all the hourly snapshots of a year into clusters would require managing a matrix of samples of $8760 * N_{ol} * 3$ and $8760 * N_N * 3$, respectively. In order to overcome this dimensionality problem, methods such as principal component analysis may be applied.

Along with these methods, the theory of moments provides a powerful tool to represent information, both for overload patterns and net demand patterns (unbalances between demand and gross RES power production).

The theory of moments, widely used in statistics and mechanics, describes the geometrical properties of physical objects. The basic two-dimensional Cartesian moment, m_{pq}, of order $p + q$ and with a density function of $f(x, y)$, is given by (103) [226].

91

$$\int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} x^p\, y^q f(x,y)dxdy \tag{103}$$

where (x,y) stand for the Cartesian coordinates.

A reduced set of Cartesian moments can be used to characterize the pattern followed by a much larger set of variables distributed throughout a certain space. For this reason, pattern recognition and classification techniques based on moments are widely used. The low-order moments starting from the zeroth to the fourth orders are often employed for such purposes. A review of the method of moments and significant research works on this issue are reported in [226]. For the clustering purpose considered in this work, the first and the second order moments are sufficient.

Both line overloads in the system and net demands can be represented as masses of a size proportional to their actual values, placed in those locations where these overloads and power unbalances occur. By making use of the moments technique in [226], the pattern (location and size) of these masses can be accurately characterized using a reduced set of moments. The first and the second order moments of these masses are used for obtaining the clusters of operation snapshots in a year.

Overloads in lines are represented as bars, with a distributed mass proportional to the overload, while net demands are represented as punctual masses (positive or negative).

The first order moments determine the center of mass of equivalent objects representing the relevant network expansion drivers (overloads and net demands); whereas, second order moments describe the "inertia" of these equivalent objects to rotate about a given axis.

4.5. Details of the Process of Defining Clustering Variables

Practical implementation details of the definition of clustering variables employed to choose operation snapshots in TEP problems are described here. Since the classification variables considered for overloading and non-overloading snapshots are different, the definition of both sets of variables is discussed separately.

4.5.1. Overloading Snapshots

The direction of flows in overloaded lines does not have any influence on expansion needs. Hence, the absolute value of excess flows is considered when computing clusters.

Moments are computed considering normalized distances among nodes and normalized levels of overloads in lines, so that magnitudes are comparable (i.e. in "per unit" quantities).

For instance, in the case example considered here, coordinates of nodes are all divided by the maximum length of a line between two nodes in the system, while the overloads

are divided by the system base power used in power flow computations. In this way, one can make sure that variables representing overloads and distances range between similar values.

Given that network expansion needs should also be computed taking contingency conditions into account, overloads have been defined as the excess of flows in lines over 80% of their rated capacity. This is a common technique used to consider contingencies through some safety margin in the absence of a detailed model to represent *N-1* operation conditions.

The next paragraphs describe the computation of moments to be chosen as classification variables, which, as already mentioned, are first and second-order ones. The description and derivation of these moments can be found in [226]. In this case, the mass density ψ_k in per unit values is given by:

$$\psi_k = \begin{cases} |P_k| - 0.8 S_k^{max}; & if\ |P_k| > 0.8 S_k^{max} \\ 0 & ;\ otherwise \end{cases} ; \forall k \in \Omega^{EL} \tag{104}$$

Thus, the total mass is given by the product of mass density and length ℓ_k i.e. $\psi_k \ell_k$. Based on this, the centroid, or the first order moment (FOM), of a group of masses can be determined by equations (105) and (106).

$$x_{c,w} = \frac{\sum \psi_k \ell_k x_{k,c}}{\sum \psi_k \ell_k}; \forall k \in \Omega^{OL_s}; \forall w \in \Omega^{OS}; OL_s \in \Omega^{EL} \tag{105}$$

$$y_{c,w} = \frac{\sum \psi_k \ell_k x_{k,c}}{\sum \psi_k \ell_k}; \forall k \in \Omega^{OL_s}; \forall w \in \Omega^{OS}; OL_s \in \Omega^{EL} \tag{106}$$

The second order moments (SOM), i.e. moments of inertia about different axes can be derived similarly from the general moment expression. For instance, the SOM about a given vertical, horizontal and perpendicular axes can be determined using (107)—(109), respectively.

$$I_{x=x',w} = \sum \psi_k \ell_k \left(\frac{\ell_k^2}{12} sin^2 \varphi_k + d_{x,k}^2\right); \forall k \in \Omega^{OL_s}; \forall w \in \Omega^{OS}; OL_s \in \Omega^{EL} \tag{107}$$

$$I_{y=y',w} = \sum \psi_k \ell_k \left(\frac{\ell_k^2}{12} cos^2 \varphi_k + d_{y,k}^2\right); \forall k \in \Omega^{OL_s}; \forall w \in \Omega^{OS}; OL_s \in \Omega^{EL} \tag{108}$$

$$I_{z=z',w} = I_{x=x',w} + I_{y=y',w} = \sum \psi_k \ell_k \left(\frac{\ell_k^2}{12} + d_{z,k}^2\right); \forall k \in \Omega^{OL_s}; \forall w \in \Omega^{OS}; OL_s \in \Omega^{EL} \tag{109}$$

where $I_{a,w}$ is the moment of inertia of a set of overloads about a given axis a, for snapshot w, whereas d_x, d_y and d_z represent the distances from each line to the particular axis of rotation, in this case, $d_z = \sqrt{d_x^2 + d_y^2}$; whereas, φ_k denotes the angle in which a particular line forms with the vertical axis.

4.5.2. Non-overloading Snapshots

The density functions used to compute moments for non-overloading snapshots are the positive and the negative net demands of system nodes, i.e. the unbalances between demand and RES power production available at each node. The moments of negative and positive power unbalances are calculated separately to avoid the canceling out of net-demands of opposite signs in those nodes that are located symmetrically with respect to the axes considered in the computation process.

The demand and RES power production dispatched at each node should result from the NCUED, as for the case of overloading snapshots. After all, the amount of demand that can be served, and the RES power that can be used, will be conditioned by the expansion of the network, and should be as large as possible.

The equations (110)—(114), analogous to (105)—(109), are some of the expressions used here for computing the relevant moments.

$$x_{c,w'} = \frac{\sum (D_i - PG_{RES,i}) x_{i,c}}{\sum (D_i - PG_{RES,i})} ; \forall i \in \Omega^{NS}; \forall w' \in \Omega^{NOS} \tag{110}$$

$$y_{c,w'} = \frac{\sum (D_i - PG_{RES,i}) y_{i,c}}{\sum (D_i - PG_{RES,i})} ; \forall i \in \Omega^{NS}; \forall w' \in \Omega^{NOS} \tag{111}$$

$$I_{x=x',w'} = \sum (D_i - PG_{RES,i}) d_{x,i}^2 ; \forall i \in \Omega^{NS}; \forall w' \in \Omega^{NOS} \tag{112}$$

$$I_{y=y',w'} = \sum (D_i - PG_{RES,i}) d_{y,i}^2 ; \forall i \in \Omega^{NS}; \forall w' \in \Omega^{NOS} \tag{113}$$

$$I_{z=z',w'} = I_{x=x',w'} + I_{y=y',w'} = \sum (D_i - PG_{RES,i}) d_{z,i}^2$$
$$; \forall i \in \Omega^{NS}; \forall w' \in \Omega^{NOS} \tag{114}$$

In the above equations, Ω^{NOS} denotes the set of non-overloading snapshots, and D_i and $p_{RES,i}$ represent the demand and the total renewable power output at node i, respectively; whereas, $d_{x,i}$, $d_{y,i}$ and $d_{z,i}$ represent the distances from node i, whose Cartesian coordinate is $(x_{i,c}, y_{i,c})$, to a particular axis of rotation, and here, $d_{z,i}^2 = d_{x,i}^2 + d_{y,i}^2$. In the case study, 23 moments shown in Table 4.2 are calculated for each type of unbalances (positive or negative), resulting in a total of 46 moment variables.

4.6. NUMERICAL RESULTS AND DISCUSSIONS

4.6.1. Considered Moments

Moments considered in this thesis correspond to FOM and SOM about several axes, as tabulated in Table 4.2. The selected moments must correspond to features that altogether distinctly represent each overload pattern of the network under consideration.

The overall number of moments considered may vary with the power system analyzed. However, there is a threshold beyond which adding more moments only adds redundant information. Arbitrarily, a total of 23 moments are computed for each snapshot in our analysis though, as it shall be seen in the results section, not all of them are necessary to accurately represent the snapshots in the TEP problem. The selection of the appropriate number of moments for each power system is a separate problem by itself that needs to be addressed. However, since it depends on the particular system to be expanded, it can be determined off-line before the TEP process starts, and kept constant for all the snapshots.

Table 4. 2 Considered Moments

Information about moments	Considered moments (features)	# of moments
Center of masses (FOM)	x_c , y_c	2
About vertical axes (SOM)	$I_{x=x_c}$, $I_{x=x'}$ where $x' = -100, -60, 0, 40, 80$ *	6
About horizontal axes (SOM)	$I_{y=y_c}$, $I_{y=y'}$ where $y' = -10, 0, 70, 90$ **	5
About perpendicular-axes (SOM)	I_{z_c} and I_z where $z = (0,-10); (-60,0); (40,70);$ $(0,90); (0,0); (-100,-10); (80,-10);$ $(-100,90); (80,90)$ ***	10

*Vertical axis, ** horizontal axis*
**** axis perpendicular to the x-y plane at a given (x,y) coordinate*

4.6.2. Modeling System Operational Uncertainties

This thesis focuses on efficiently handling operational uncertainties in TEP by considering their expected influence in the final TEP solutions. The uncertainties considered here are discussed separately in the following subsections.

4.6.1.1. Demand Variability and Uncertainty

To capture demand variability, load aggregated models are often used, as in [227], based on the load duration curve. In real life, there exist spatial variations in demand which may significantly influence TEP solutions. Therefore, to account for this impact, demand correlations ranging from 0.7 to 1 are factored in to generate the demand series at different locations.

4.6.1.2. Conventional Generator Outages

A two–state model (online or offline) is used to represent the state of conventional power units based on their respective forced outage rates (FOR) , which range from 0.05 to 0.15 depending on the technology type of each generator. Then, a discrete random

binomial distribution is applied to generate availability patterns for different generators, obtained from their corresponding forced outage rates.

4.6.1.3. *RES Output Variability and Uncertainty*

The outputs of wind and solar power plants are subject to the wind speed and solar radiation regimes, respectively. A common approach to handle uncertainties in RES output is MCS, in which a number of samples are generated randomly from probability distributions. For the present analyses, historical hourly wind speed and solar irradiance data have been used. These are taken from publicly available meteorological websites (see[228], and[229], respectively).

Wind and solar power productions are correlated in space and time, and this effect is taken into account in the generation of input samples. In addition, the complementary nature of wind and solar power sources is also captured by taking correlations between them ranging from -0.3 to -0.1, which comply with the results in [230]. Note that wind and solar power outputs are determined by plugging in the hourly values of the primary renewable resource available in the wind [231] and solar [232] power output expressions (also known as power curves) . For instance, the hourly wind power output $P_{wnd,w}$ of each wind farm is determined by the nonlinear model of a typical wind turbine model as in (115).

$$
P_{wnd,w} = \begin{cases} 0 & ; 0 \leq v_w \leq v_{ci} \\ P_r(A + Bv_w^3) & ; v_{ci} \leq v_w \leq v_r \\ P_r & ; v_r \leq v_w \leq v_{co} \\ 0 & ; v_w \leq v_{co} \end{cases} \tag{115}
$$

In the above equation, *A* and *B* are parameters represented by the expressions in [233]. Similarly, the hourly solar power output $P_{sol,w}$ is determined by plugging in the hourly solar radiation levels in the solar power output expression given in (116), [234].

$$
P_{sol,w} = \begin{cases} \frac{P_r R_w^2}{R_{std}*R_c} & ; 0 \leq R_w \leq R_c \\ \frac{P_r R_w}{R_{std}} & ; R_c \leq R_w \leq R_{std} \\ P_r & ; R_w \geq R_{std} \end{cases} \tag{116}
$$

4.6.3.   Test Results and Discussion

The standard IEEE 24-bus Reliability Test System (RTS) [235] has been used to show the behavior of the proposed clustering approach and test its performance in the target TEP problem. The data used in this study can be found in [235].

The clustering method requires information about the location of each transmission line and node, but the information available in [235], as in many other standard test systems, does not include node coordinates. Because of this, estimates of the geographical coordinates of nodes in the system of the case study have been generated by computing

a geographical map of the network system. Distances among neighboring nodes are as close to the lengths of the lines linking these nodes as possible. And these distances, i.e. the lengths of the lines, are assumed to be proportional to their corresponding impedances. A non-linear optimization problem has been solved to generate the system network map including the required geographical information. Figure 4.7 shows the resulting map of the standard IEEE 24-bus system. Note that nodes 9 to 12 are in one substation which has 4 transformers, linking these nodes, and so are nodes 3 and 24 connected by a transformer (see Appendix C).
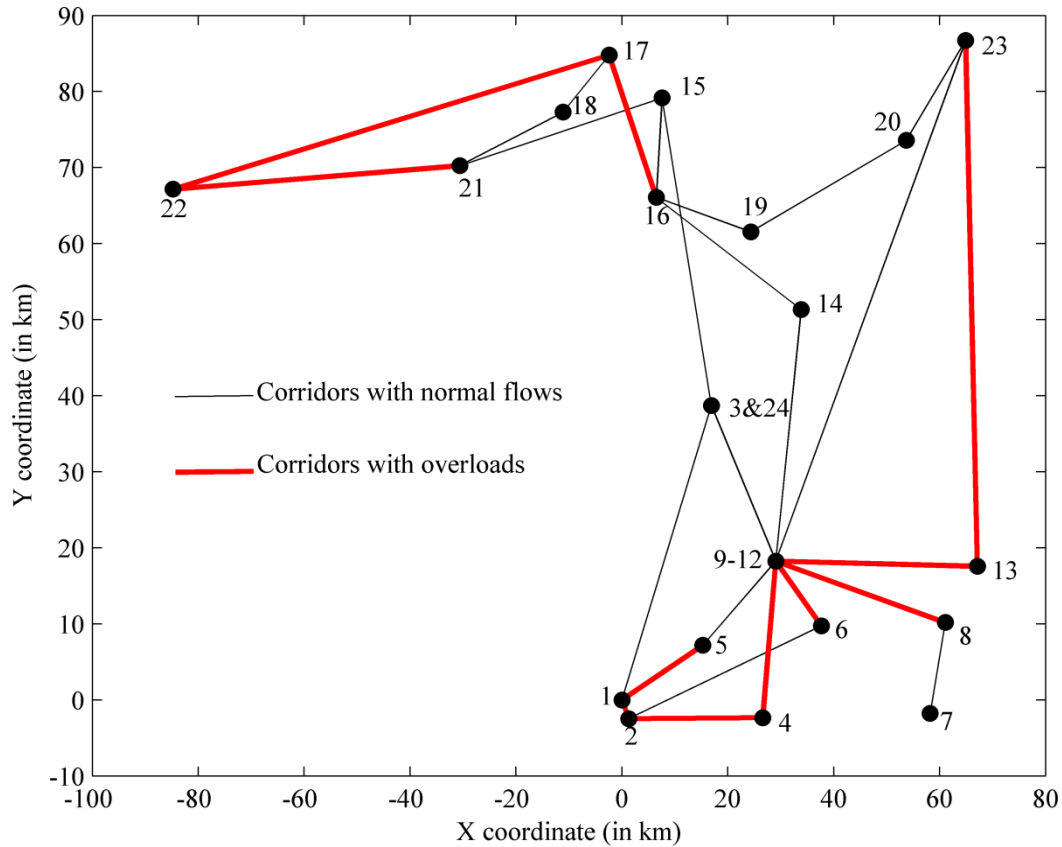
Fig. 4. 7 Generated map of IEEE 24-bus system

Our test system comprises 24 buses, 33 existing corridors, and 19 potential new ones, totaling 52 candidate corridors for potential investments. In addition to the existing generation capacity in the considered test system, three RES generators with a combined installed capacity of 3000 MW are added to the system, including a 500 MW solar farm connected to node 4 and two 1500 and 1000 MW wind farms connected to nodes 13 and 22, respectively. The hourly production time series of wind and solar farms for the planning year are determined as explained in section 2, as well as the hourly demand profile at each node and the availability profile of conventional generators. In total, 8760 samples, corresponding to hourly combinations of the regarded uncertain parameters, are subject to the clustering process. In particular, each sample includes the availability state of 11 conventional generators, the load level of 16 electricity consumers and the available power output of three RES generators, bringing the total dimension of the samples in the "uncertainty space" to 30. The dimension of samples

97

significantly increases with the network size, and the number and types of uncertain parameters being considered, leading to the curse of dimensionality and creating problems in the clustering process. The proposed clustering technique overcomes such problems by mapping high-dimensional samples to relatively lower-dimensional ones. Note that throughout this analysis, both terms–samples and snapshots–refer to operation states.

4.6.2.1. *Clustering Results for Overloading Snapshots*

The NCUED problem is solved for all snapshots to obtain the corresponding patterns of overloads. Figure 4.7 shows that there are a total of 13 overloaded lines, obtained by combining the sets of overloaded lines in all snapshots. In the considered case study, in a total of 4741 snapshots (out of the 8760 samples), there is at least one overloaded line that is congested (shown in Figure 4.7). This means that each overloading snapshot includes a subset of overloaded lines among those shown in Figure 4.7. In the remaining 4019 snapshots, there is no congestion in the system.

Once the fingerprint of each sample is obtained, the subsequent step is to compute the features of snapshots that are used as clustering variables in a TEP problem. In this case, the features considered are the moments of overloads, and groups of patterns are determined according to the set of moments. As mentioned before, the set of moments has to be adjusted to each power system under analysis, but only once and for all the further optimization processes to take place. The moments considered are selected here for the given case study using some performance metrics. One of these metrics is the similarity ratio, which is the ratio of average intra-cluster to average inter-cluster distances, in the space of moments. These are given by equations (117) and (118), respectively. The average intra-cluster distance measures the compactness of clusters; whereas, the average inter-cluster distance measures the cluster discrimination. The former should be as small as possible, while the latter should be as large as possible, resulting in a minimum value of the ratio.

$$x_{avg} = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{N_k} \sum_{X \in KS_k} \left\| X - \hat{X}_{KS_k} \right\|_2 \tag{117}$$

$$X_{avg} = \frac{2}{K(K-1)} \sum_{k=1}^{K-1} \sum_{l=k+1}^{K} \left\| \hat{X}_{KS_k} - \hat{X}_{KS_l} \right\|_2 \tag{118}$$

In equations (117) and (118), $\|\|_2$ represents the Euclidean distance.

This ratio has been calculated for several numbers of moments and clusters, as illustrated in Figure 4.8. For each set and number of clusters, the evolution of the considered ratio with the number of moments taken has been represented in a separate curve. It can be observed in Figure 4.8 that adding more moments beyond 15 seems to have little significance since the similarity ratio remains stable. This corresponds to a 50% reduction in the dimension of the clustering space. Another important conclusion is that adding more clusters does not improve the similarity ratio beyond some threshold. Here, the threshold seems to be close to 40 clusters.
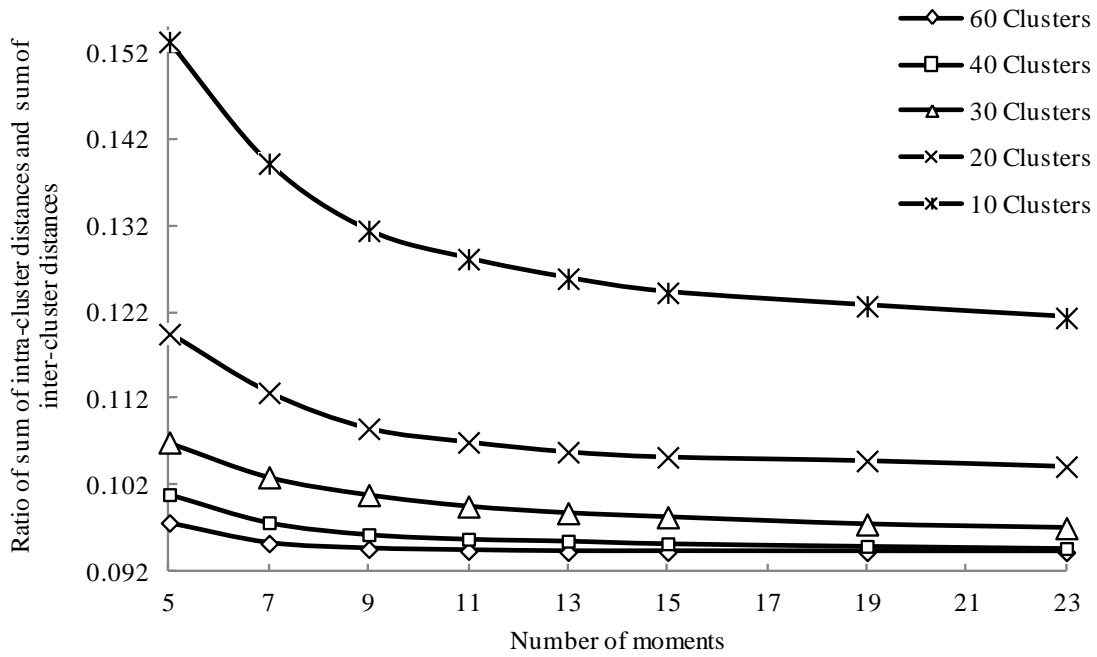
Fig. 4. 8 Variation of similarity ratio with number of moments and clusters (for overloading snapshots)

As stated earlier, the application of our clustering technique involves mapping the high-dimensional input dataset into the space of moments, which is quite convenient because it makes working with a relatively lower number of parameters possible when clustering the corresponding snapshots, thus reducing the dimension of the data set. The 4741x30 overloading dataset is, for example, clustered using the computed 4741 moment samples each including 15 moment values. Figure 4.9 displays the hourly time series of values for two of the moments considered. Each series comprises 4741 hourly values, and the hour for each value of the moment in the series is represented in the horizontal axis. The thick line represents the values of the "dominant" moment variable. The concept of dominance here should be understood in the following context. A moment variable about a given axis is dominant when the variance of its values is larger in magnitude than the variance of any of the other considered moment variables, which correspond to moments computed about different axes from that of the dominant moment variable. In the case study, the dominant moment variable is the moment about the corner point (-100, 90) of the network map in Figure 4.7. Here, it should be noted that the moment samples (snapshots) in Figure 4.8 are sorted by increasing index of the cluster they belong to.
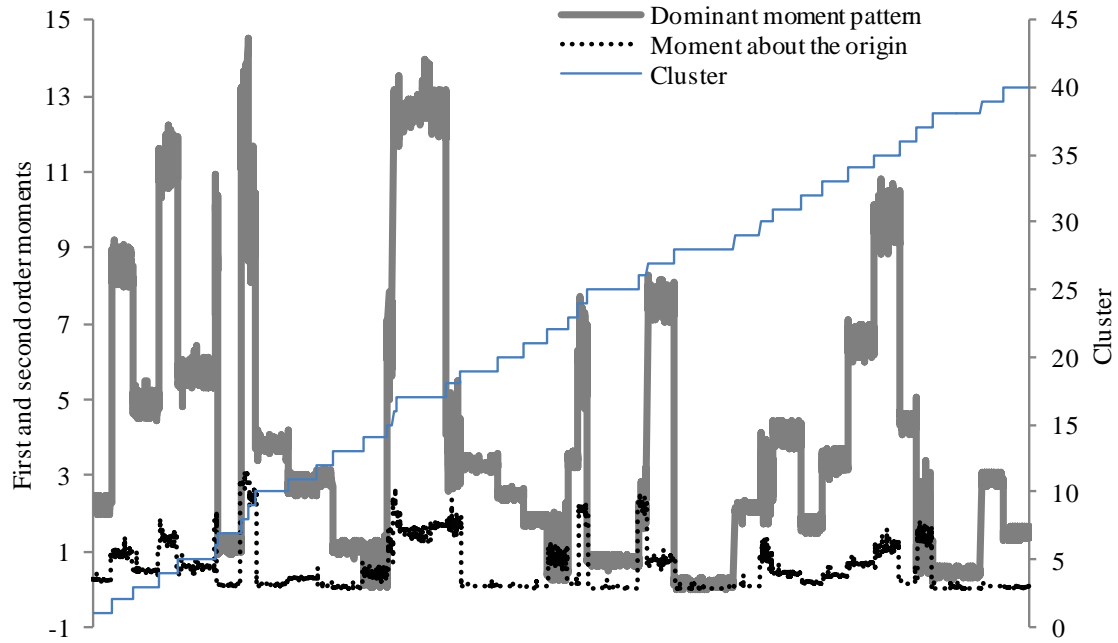
Fig. 4. 9 Patterns of moment values in the overloading snapshots sorted by increasing order of cluster indices (horizontal axis represents the number of samples)

The moment pattern about the origin is also shown in Figure 4.9 for comparison purposes. Basically, the remaining moments, which are not shown in Figure 4.9 for the sake of simplicity, follow similar patterns. The width of each discrete step in the cumulative clusters' curve is proportional to the number of snapshots grouped in that particular cluster. Generally, Figure 4.9 helps to observe whether the clustering results are accurate enough. Clearly, one can see that there is some discernible pattern in the plot i.e. some homogeneity in the values of the moments in each cluster and large differences among the values of the moments in different clusters, which validates the clustering approach.

4.6.2.2.   Clustering Results for Non-overloading Snapshots

As mentioned earlier, the non-overloading snapshots are clustered according to variables related to line losses and their locations. The selected variables are the net demand at each node along with its geographical location. As in the previous case, the standard *k-means* algorithm is used for clustering the moments of these variables. Moments considered here correspond to those listed in Table 4.2 except for the fact that the positive and the negative power unbalances (net demands) are treated separately, resulting in a total of 46 moments i.e. 23 moments for unbalances of each sign.

The number of clusters is decided based on the Elbow method, as in Figure 4.10, which allows balancing the accuracy of the clustering analysis (given by the objective value of the *k-means* algorithm) displayed on the primary vertical axis) and the number of clusters. The evolution of the objective value (minimized by the *k-means* clustering algorithm) with the number of clusters is shown in a curve. When plotting this curve, 46 moments have been considered. In this case, one can see that using 10 clusters seems a

100

reasonable trade-off. The evolution of similarity ratio (shown on the secondary vertical axis, in Figure 4.10) with the number of moments is depicted in another curve. For the analysis here, note that the moment variables have been taken from the set of moments of the positive and the negative power unbalances. One can see in Figure 4.10 that the changes in the similarity ratio are negligible when 15 or more moments are taken. Then, 15 is an appropriate number of moments.
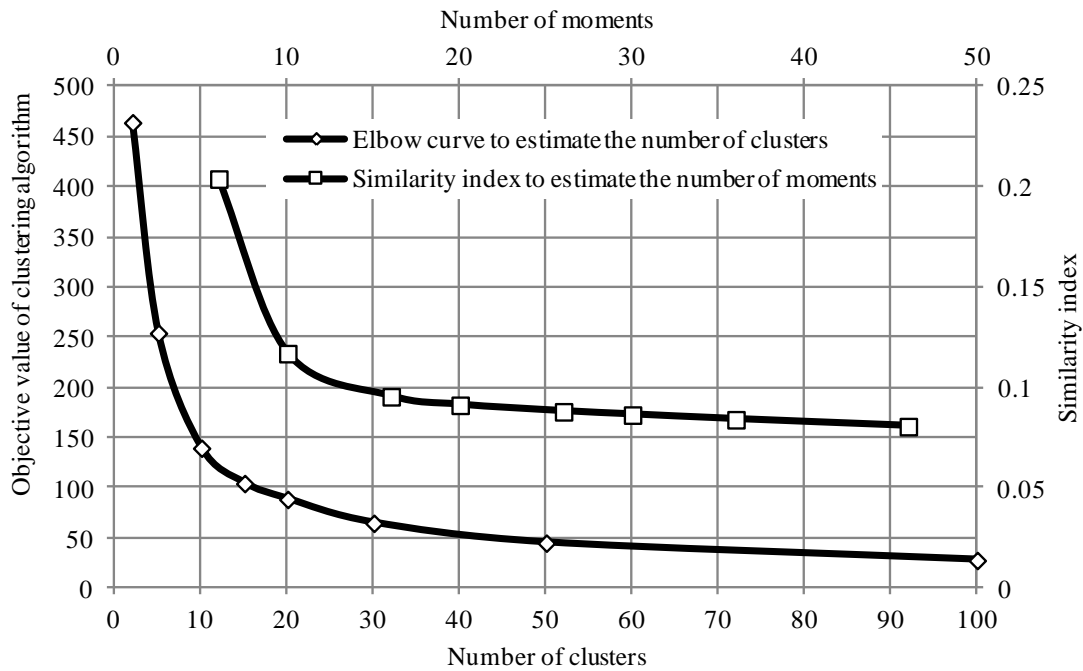
Fig. 4. 10 Estimating an appropriate number of moments and clusters for non-overloading clusters

4.6.2.3. *Clustering in the Principal Components Space*

As discussed in the previous subsection, moments of overloads and net-loads are used as clustering variables, and the results presented in this thesis are based on this. For the test system considered here, it has been already stated earlier that when clustering snapshots in the space of moments, a reasonably good balance between accuracy and computation burden is achieved using 15 to 20 moments, both in the case of overloading and non-overloading snapshots. However, this may not be the case for larger systems. Intuitively, the number moments required to distinguish properly the respective patterns may be higher for larger systems, potentially leading to a size problem. Therefore, additional ways may be required to reduce the number of clustering variables.
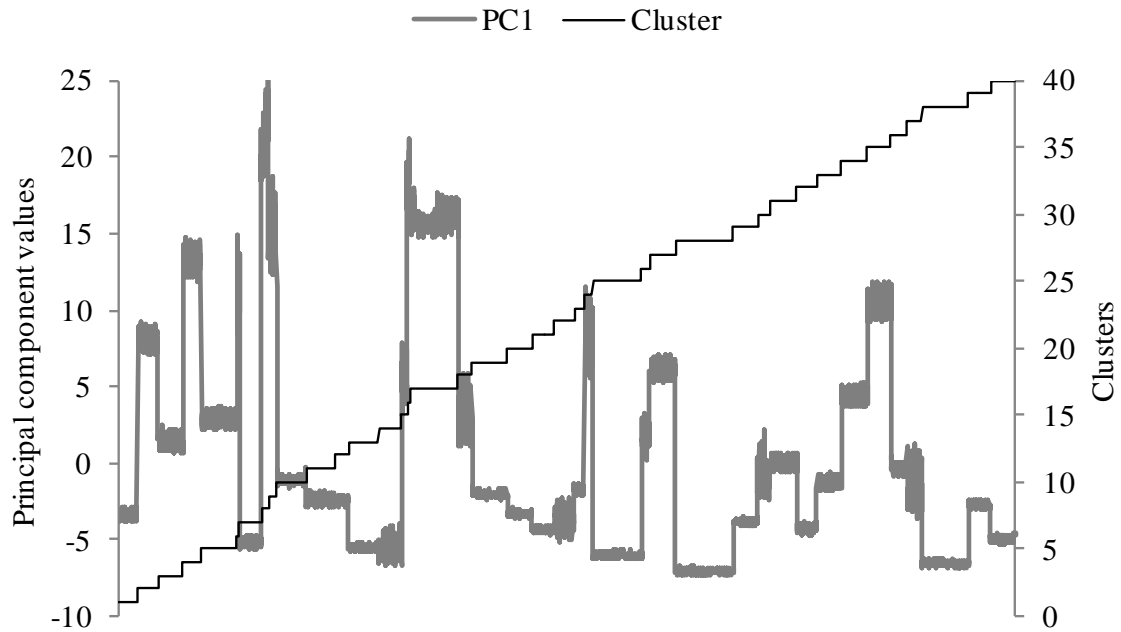
Fig. 4. 11 First principal component values sorted by increasing order of cluster indices (for overloading snapshots)

An interesting idea is to apply the Principal Component Analysis (PCA) to find the most relevant moments that represent most of the desired information regarding the variability of moment samples. For instance, in the considered system, using only 6 principal components as clustering variables of the overloading snapshots is enough to get the same results as with 15 moments (see Figure 4.8). It is even lower (4 principal components) in the case of non-overloading snapshots. This shows the capability of transforming the moment space into a principal component space in tackling the dimension problem.

Table 4. 3 Eigenvalues of Covariance Matrices of Moments

Principal component	Overloading snapshots		Non-overloading snapshots	
	Eigenvalues	Cumulative sum of eigenvalues	Eigenvalues	Cumulative sum of eigenvalues
1	58.854	85.02%	399.451	65%
2	9.002	96.59%	179.280	94%
3	1.839	99.35%	31.360	100%
4	0.510	99.93%	1.225	100%

Figure 4.11 shows a plot of the first principal component values (PC1), sorted by increasing order of cluster indices. One can see that the first principal component captures nearly 85% of the required information in terms of data variability in the principal components space. It is also interesting to note that this pattern closely resembles the pattern of the dominant moment variable in Figure 4.8. As shown in

Table 4.3, the first two principal components account for 97% and 94% of the variance of the principal component values of the moments taken in the case of overloading and non-overloading snapshots, respectively. In general, PCA can be a handy tool in reducing the dimension of the set of clustering variables without losing significant information.

4.6.2.4. *Comparisons in terms of TEP Results*

Since the clustering approach proposed here is to be applied to TEP, its efficiency should be verified in this context. This can be accomplished by running a DC-based TEP model (presented in the preceding Sections and Chapter 3) considering the set of snapshots identified as representatives of the clusters, and comparing TEP results with those of the full-scale (brute-force) problem that considers all the 8760 snapshots.

In this respect, investment decisions considering all the 8760 snapshots include new lines in corridors (2,4), (4,9), (9,11), (11,13), (13,23) and (21,22). Overall, investment costs in this brute-force problem amount to 82.8 M€. Now, one can check the evolution of network investment costs with the number of clusters, as shown in Figure 4.12. Investment costs with only 50 clusters, obtained using the moment-based clustering approach, are the same as those of the brute-force TEP solution. However, one can see in Figure 4.12 that selecting clusters according to generation—demand patterns, results in underinvestment even for higher number of clusters. This reveals a lack of representativeness of the snapshots selected according to this set of clustering variables.

In addition to investment decisions, one should also compare the total dispatch (operation) costs. Operation costs are computed by solving the economic dispatch problem for the whole target year considering investment decisions. Regarding clustering methods, it has been already stated from the outset that the method proposed here is denoted as clustering based on effects (CbE); while the traditional method based on generation—demand patterns is identified as clustering based on causes (CbC).

Figure 4.13 shows the evolution of the global dispatch costs with the number of clusters for both clustering approaches. Comparison of the total system dispatch costs in both cases (i.e. CbE and CbC) also strengthens the previous statement on the required number of clusters. In order to obtain the same results as in the brute-force problem in terms of investment solution and deviation in operation costs, CbC requires 310 or more clusters, while CbE only needs about 40 clusters of overloading snapshots along with 15 clusters of non-overloading ones.
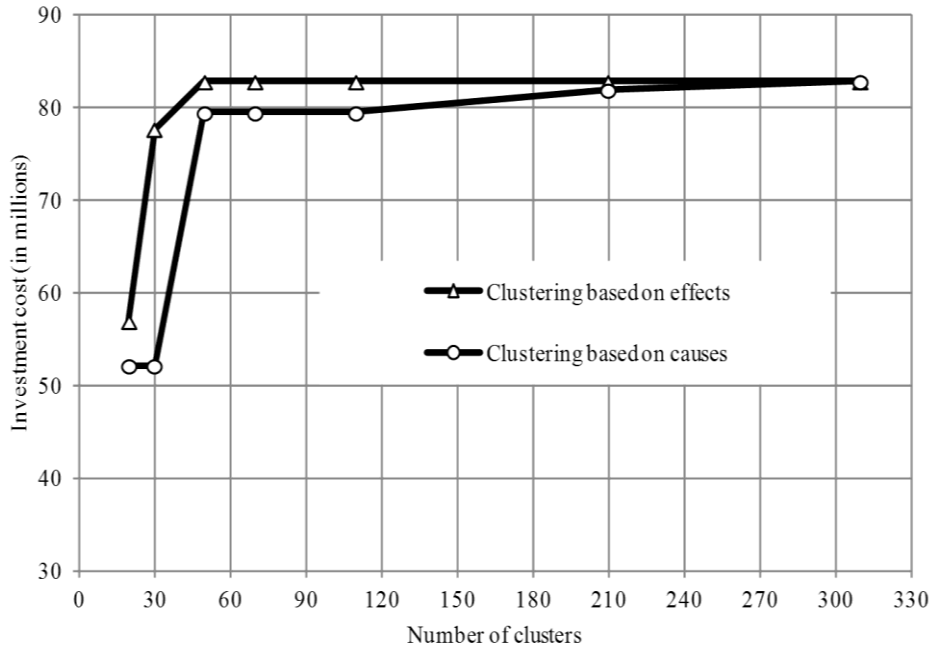
Fig. 4. 12 Evolution of investment costs with number of clusters
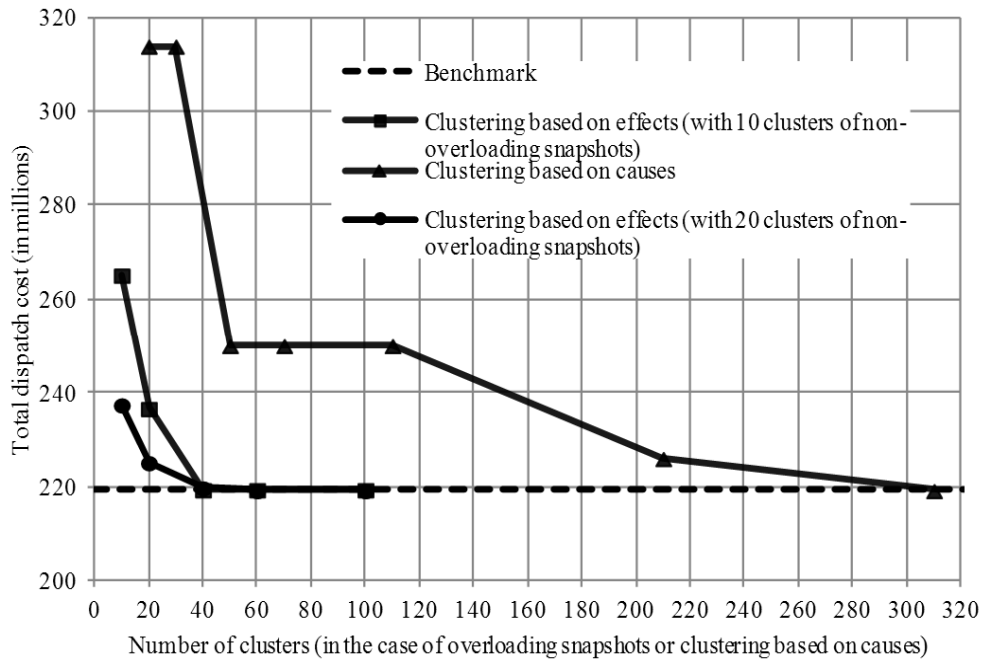
Fig. 4. 13 Evolution of total dispatch costs with number of clusters

In Figs. 12 and 13, the variations of costs with the number of clusters can be explained as follows. It is generally accepted in clustering theory that when increasing the number of clusters, the level of accuracy increases. When varying the number of clusters, two important parameters are affected: the representative snapshots and the cumulative probability of each cluster ρ_s. Note that the representative snapshot in a given cluster is selected among the snapshots, grouped to that same cluster, based on certain criteria (for example, being the closest to the centroid which means the medoid in the cluster). The

parameter ρ_s is the cumulative probability of all snapshots grouped together. Assuming the probability of occurrence of each snapshot is the same, ρ_s should be proportional to the number of snapshots in cluster s. This means ρ_s is the sum of all individual probabilities in the same cluster. Both these parameters define the accuracy level of the clustering outcome. The higher the number of clusters are, the more similar the snapshots grouped together will be (in terms of their effects on network expansion needs), and therefore the higher the clustering accuracy will be. On the other hand, a lower number of clusters increase the chance of clustering "dissimilar" snapshots together. Chances are also high that the representative snapshots selected for each cluster may not accurately represent their companions in their respective clusters. Therefore, when using a smaller number of clusters, the variability of operation situations is likely to be underestimated, potentially resulting in underinvestment, as shown in Figure 4.12. Obviously, the ultimate price of such inadequate network reinforcements in the system is an increase in operation cost due to the presence of congestion and unserved power. This is reflected in Figure 4.13, where one can easily see very high dispatch costs associated to smaller number of clusters. When the number of clusters is slowly increased, all curves gradually approach the benchmark one, showing an increasing trend of accuracy. Here, it is interesting to observe that the CbE-based method approaches the benchmark before the CbC-based one, showing the former's excellent performance and clear advantage in terms of computational burden —which is further increased by the compact representation of the snapshots. Another important result observed in Figure 4.13 is the decreasing trend in the dispatch costs achieved when one increases the number of clusters for non-overloading snapshots while keeping the number of clusters of overloading snapshots constant. This can be attributed to the better estimation accuracy of transmission losses achieved in this way, which may increase the accuracy of the computation of operation costs and justify some more line reinforcements. This is particularly shown in the case study, also depicted in Figure 4.13 by the two curves in the middle.

Therefore, the main conclusions from these analyses are two. First, it is much better to cluster snapshots based on relevant power flow effects (overloads and losses) than clustering them based on input system variables (generation—demand patterns). Second, the moments approach is an effective way to reduce the dimension of the clustering space.

These results are in line with those shown in Figs. 4.8 and 4.10, where the threshold for the number of clusters of overloading and non-overloading snapshots seems to be also 40 and 15, respectively, when the proposed clustering approach is applied.

4.7. COMPUTATIONAL IMPLICATIONS

Some computational implications of the proposed approach are discussed here. The efficacy of the proposed method has been already verified on the 24-bus test system. The results are very interesting in that the moment-based clustering using power flow variables results in a very compact optimization problem (because of the significant

reduction of snapshots) without considerable loss of accuracy. In addition, the results also show that using the generation—demand patterns (the causes, CbC) instead of power flow patterns (the effects, CbE) would require a far higher number of snapshots to achieve the same level of accuracy in the TEP context.

The method can generally be extended to large-scale TEP problems. As it is known, the main limiting factor in such problems is the computation burden. The computing time is directly related to the problem complexity. In this regard, given the DC-based TEP optimization problem [14], the marginal impact of reducing the number of snapshots on its computational burden can be quantified. This depends on complexity of the problem being considered, i.e., the number of equations, variables, non-zeros, etc. For example, in the TEP problem presented in [14], the total number of equations and continuous variables can be determined by (119) and (120), respectively.

$$1 + S * \left[2(N_N + N_G + N_{RES}) + (N_{CL} + N_{EL})(5 + 2L) \right] \tag{119}$$

$$S * \left[N_G + N_{RES} + 4(N_N + N_{CL} + N_{EL}) \right] \tag{120}$$

where L corresponds to the number of partitions in the losses modeling [15].

The above expressions, (119) and (120), clearly indicate that the impact of the number of snapshots on the problem size is linear i.e. a reduction in the number of snapshots by a certain fraction results in the same level of reduction in the number of equations and variables. One can observe in (119) and (120) that reducing the number of snapshots marginally (i.e. by one snapshot) leads to a reduction in the number of equations and continuous variables by an amount given by:

$\left[2(N_N + N_G + N_{RES}) + (N_{EL} + N_{CL})(5 + 2L) \right]$ and $\left[N_G + N_{RES} + 4(N_N + N_{EL} + N_{CL}) \right]$, respectively.

Such a reduction in complexity of the problem can indeed result in a huge difference in computing time. The impact can even be more noticeable in large-scale problems. For instance, for a 1000-node system, assume there are 2000 existing and 2000 candidate lines, five conventional generators of different technologies and five types of renewable energy sources at each node. Suppose the number of losses partitions, L, is set to 5. Under these assumptions, the reductions in the number of equations and continuous variables, with respect to the marginal reduction of one snapshot, amount to 82,000 and 30,000 respectively. Computationally speaking, such a huge reduction in complexity significantly enhances the tractability of the TEP problem.

Furthermore, it has already been stated that, when using the proposed method, the minimum number of clusters required for obtaining an optimal TEP solution in the test system is 50 (see Figure 4.13). However, up to 310 clusters are required to get the same solution using conventional clustering variables. This means that the resulting TEP optimization problem may have 6 times fewer equations and variables. Considering all the benefits, the proposed method seems to be very promising and can be extended to large-scale TEP problems which consider high uncertainty and a long temporal scope.

From a computational perspective, the implementation of the proposed clustering method is not burdensome. This is because it is formulated over a very fast NCUED model, and also because the OPF for each snapshot can be individually computed. This allows parallel computation, which further facilitates the computation process.

4.8. SUMMARY

This chapter has introduced a novel way of clustering operational states, or snapshots, based on classification variables that are closely related to TEP problems, instead of using the customary generation and demand variables.

In the proposed approach, snapshots are characterized according to their effects on the network, i.e. the congestions (overloads) and losses that will in fact create expansion needs. In the non-overloading snapshots, net power unbalances are instead used as significant variables. The effects on the network are then translated into a much more compact representation, namely a moments-based space of variables. Moments translate both the geographical location and the power-related parameters of potential investment needs into a reduced reference system.

The method has been tested comparing its results against both the original brute-force problem (using the whole original set of snapshots) and a clustering method based on generation and demand patterns.

For identical results of the TEP problem, the test results show that the proposed method reduces the number of required snapshots in almost 200 times with respect to the original problem, and in 6 times regarding the generation—demand pattern based clustering method. This work also estimates the savings in computing time related to the marginal reduction in the number of snapshots.

As a global conclusion, it can be stated that the proposed power-flow clustering criteria (the effects on the network), combined with the moments-based compact representation of those effects, seems to be an adequate and promising method to handle operation uncertainty in the context of TEP problems.

In addition to the new clustering method, contributions from this chapter include the losses model used in the NCUED model as well as the nonlinear optimization approach developed to generate a network map by making use of network parameters (impedances, in particular).

5

V. A Stochastic TEP Model Formulation and A Solution Strategy

This chapter presents the algebraic formulation of the stochastic TEP (STEP) model and a description of the proposed solution solution strategy.

5.1. THE TEP PROBLEM

5.1.1. Overview of the Multi-stage and Stochastic Programming Framework

A TEP problem is naturally dynamic because the solution has to explicitly provide necessary information regarding not only where and what but also when line investments are needed. Regarding the planning horizon and decision stages, on account of the dynamic nature of TEP, a more realistic approach would be to formulate the problem with multiple decision stages (i.e. multi-year decision framework). This modeling framework assumes that there are n probable future storylines (or scenarios) each associated with a probability of realization ρ_s that stochastically represent major long-term uncertainties. This modeling framework is, on one hand, the building block of complex dynamic models, and on the other hand, an appropriate model to combine short-term (first stage) and long-term strategic decisions (second stage). It also makes sense considering the nature of transmission planning practice, which often requires short- to medium-term decisions accompanied by long-term strategic plans for exploring future possible developments.

The length of the first period can be taken as 5 to 15 years because transmission expansions are planned well in advance (often within this range). Moreover, the construction permit process of lines is often accompanied by significant delays; most of the time, it takes several years (often in this range). Likewise, the length of the second period can be set in the range between 20 and 35 years long depending on the planner's choice. Overall this leads to a 50-years long planning horizon.

Figure 5.1 schematically illustrates the two-period TEP modeling framework and the form of its expansion solution. Each of these sub-horizons (periods) may have multiple planning stages. In the first period, we obtain a single and robust expansion strategy for each stage which is good enough for all scenarios. It should be noted here that the decisions made in the first stage are obtained considering the future scenarios. The long-term investment decisions made in the second-stage are adapted to each scenario, guaranteeing sufficient flexibility in the planning process by allowing the ability to postpone or alter decisions in the future.

As stated in the previous Chapters, the TEP problem is formulated considering its dynamic nature i.e. featuring multiple decision stages. In addition, in order to combine "here and now" and "wait and see" investment decisions, a two-period stochastic optimization framework is proposed in this work. This modeling framework assumes that there are n probable future storylines (or scenarios) each associated with a probability of realization ρ_s that stochastically represent relevant sources of uncertainties. The whole modeling scheme adopted in this thesis (i.e. the multi-stage and multi-scenario DGIP modeling framework and the expansion solution structure) is illustrated in Figure 5.1. The formulation is based on the assumption that there are two

investment pools, one for each period, from which the potential lines can be selected. Investments in the first period can be postponed to the second period if deemed necessary.
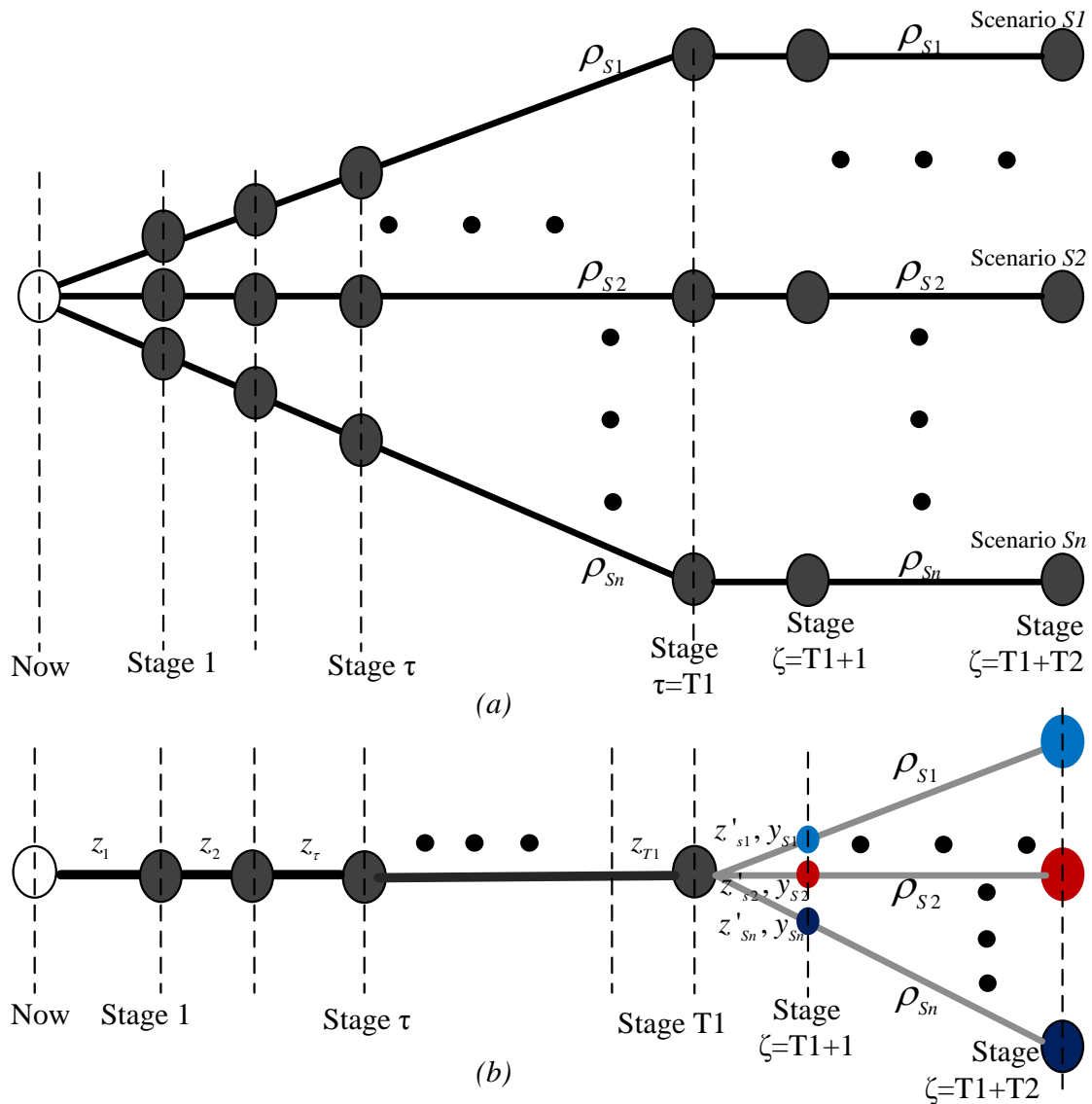
Fig. 5. 1 A schematic representation of (a) possible future scenario trajectories and (b) a decision structure

Figure 5.1 schematically represents the possible future scenario trajectories with multiple scenario spots throughout the planning horizon, along with the decision structure in the τth stage of the first planning period, showing a single investment decision z_τ (where $\tau = 1, 2, ..., T1$) in every stage of the first period which are common (or good enough) for all scenarios, and flexible or strategic decisions $z'_{s,\zeta}$ and $y_{s,\zeta}$ (where $\zeta = T1 + 1, T1 + 2, ..., T$) in every stage of the second one [15]. Note that the first-period decisions are more relevant than those made in the second period because they are implementable straightaway before uncertainties are uncovered i.e. "here and now" decisions. However, the second-period decisions can also be very useful if seen

from the flexibility/strategic planning perspective. In order to broaden the investment options, two investment pools are considered one for each period, in which it is possible to postpone the investments in the first z_τ to the second period $z'_{s,\zeta}$. The mathematical formulation of the model developed here is presented and explained in detail in the following sub-sections.

5.1.2.   Algebraic Formulation of the TEP Model

The stochastic TEP model developed in this thesis is described as follows.

5.1.2.1.   Objective Function

As mentioned earlier, this work develops a generalized optimization model that simultaneously determines the optimal location, time and size of transmission line investments under high penetration level of RESs. In other words, the objective is to expand the transmission network at a minimum cost possible from the system perspective. The resulting problem is formulated as a multi-objective stochastic MILP with an overall cost minimization.

The objective function in (121) is composed of NPV of six cost terms each weighted by a certain relevance factor $\alpha_j; \forall j \in \{1,2,\dots,5\}$. Note that, in this work, all cost terms are assumed to be equally important; hence, these factors are set to 1. However, depending on the relative importance of the considered costs, different weights can be adopted in the objective function. The first term in (121), $TInvC$, represents the total investment costs under the assumption of perpetual planning horizon [186]. In other words, the investment cost is amortized in annual installments throughout the lifetime of the installed component. Here, the total investment cost is the sum of investment costs of candidate lines as in (122).

The second term, TMC, in (121) denotes the total maintenance costs, which is given by the sum of individual maintenance costs of new and existing lines and generators, at each stage and the corresponding costs incurred after the last planning stage, as in (123). Note that the latter costs depend on the maintenance costs of the last planning stage. Here, a perpetual planning horizon is assumed. The third term TEC in (121) refers to the total cost of energy in the system, which is the sum of the cost of power produced by new and existing generators at each stage as in (124). Equation (124) also includes the total energy costs incurred after the last planning stage under a perpetual planning horizon. These depend on the energy costs of the last planning stage. The fourth term $TENSC$ represents the total cost of unserved power in the system and is calculated as in (125). The last term $TImiC$ gathers the total emission costs in the system, given by the sum of emission costs for the existing and new generators as well that of power purchased from the grid at the substations.

$$\min_{z,y,z',\dots} TC = \alpha_1 * TInvC + \alpha_2 * TMC + \alpha_3 * TEC + \alpha_4 * TENSC + \alpha_5 * TImiC \tag{121}$$

$$TInvC = \underbrace{\sum_{t\in\Omega^t} \frac{(1+\sigma)^{-t}}{\sigma} InvC_t^{LN}}_{NPV\ of\ investment\ cost} \tag{122}$$

$$TMC = \underbrace{\sum_{t\in\Omega^t} (1+\sigma)^{-t}\ (MntC_t^{NL} + MntC_t^{EL})}_{NPV\ of\ maintenance\ costs}$$
$$+ \underbrace{\frac{(1+\sigma)^{-T}}{\sigma}(MntC_T^{NL} + MntC_T^{EL})}_{NPV\ maintenance\ costs\ incured\ after\ stage\ T} \tag{123}$$

$$TEC = \underbrace{\sum_{t\in\Omega^t} (1+\sigma)^{-t}\ (EgyC_t^{NG} + EgyC_t^{EG})}_{NPV\ of\ operation\ costs}$$
$$+ \underbrace{\frac{(1+\sigma)^{-T}}{\sigma}(EgyC_T^{NG} + EgyC_T^{EG})}_{NPV\ operation\ costs\ incured\ after\ stage\ T} \tag{124}$$

$$TENSC = \underbrace{\sum_{t\in\Omega^t} (1+\sigma)^{-t}\ ENSC_t}_{NPV\ of\ reliability\ costs} + \underbrace{\frac{(1+\sigma)^{-T}}{\sigma}ENSC_T}_{NPV\ reliability\ costs\ incured\ after\ stage\ T} \tag{125}$$

$$TEmiC = \underbrace{\sum_{t\in\Omega^t} (1+\sigma)^{-t}\ (EmiC_t^{NG} + EmiC_t^{EG})}_{NPV\ emission\ costs}$$
$$+ \underbrace{\frac{(1+\sigma)^{-T}}{\sigma}(EmiC_T^{NG} + EmiC_T^{EG})}_{NPV\ emission\ costs\ incured\ after\ stage\ T} \tag{126}$$

The individual cost components in (122)—(126) are computed by the following expressions. Equation (127) represents the investment costs of lines. Notice that all investment costs are weighted by the capital recovery factor, $\frac{\sigma(1+\sigma)^{LT}}{(1+\sigma)^{LT}-1}$. The formulations in (127), along with the logical constraints which are described in the constraints section, ensure that the investment cost of each line added to the system is considered only once in the summation. For example, suppose an investment in a particular feeder k is made in the second year of a three-year planning horizon. This means that the feeder should be available for utilization after the second year. Hence, the binary variable associated to this feeder will be 1 after the second year while zero otherwise i.e. $z_{p,k,t} = \{0,1,1\}$. In this particular case, only the difference $(z_{p,k,2} - z_{p,k,1})$ equals 1, implying that the investment cost is considered only once. Equations (128) and (129) stand for the maintenance costs of new and existing lines at each time stage, respectively. The maintenance cost of a new/existing lines is included only when its corresponding investment/utilization variable is different from zero. The maintenance costs of new and existing generators at each stage can also be similarly formulated but this information is not often available for network planners.

112

$$InvC_t^{NL} = \sum_{k\in\Omega^k}\sum_{a\in\Omega^a}\frac{\sigma(1+\sigma)^{LT_{a,k}}}{(1+\sigma)^{LT_{a,k}}-1}IC_{p,k}^{N1}(z_{a,k,\tau}^{N1}-z_{a,k,\tau-1}^{N1})$$

$$+\sum_{s\in\Omega^s}\rho_s\sum_{k\in\Omega^k}\sum_{a\in\Omega^a}\frac{\sigma(1+\sigma)^{LT_{a,k}}}{(1+\sigma)^{LT_{a,k}}-1}IC_{a,k}^{N1}\big(z'^{N1}_{a,k,s,\zeta}$$

$$-z'^{N1}_{p,k,s,\zeta-1}\big)\tag{127}$$

$$+\sum_{s\in\Omega^s}\rho_s\sum_{k\in\Omega^k}\sum_{a\in\Omega^a}\frac{\sigma(1+\sigma)^{LT_{a,k}}}{(1+\sigma)^{LT_{a,k}}-1}IC_{p,k}^{N2}(y_{a,k,s,\zeta}^{N2}$$

$$-y_{a,k,s,\zeta-1}^{N2})\quad;\forall\tau\in\Omega^{T1};\forall\zeta\in\Omega^{T2};\ z_{p,k,0}$$

$$=0;\ y_{a,k,n,s,T1}^{N2}=0$$

$$MntC_t^{NL}=\sum_{k\in\Omega^k}\sum_{a\in\Omega^a}MC_{a,k}^{N1}z_{a,k,n,t}^{N1}+\sum_{s\in\Omega^s}\rho_s\sum_{k\in\Omega^k}\sum_{a\in\Omega^a}MC_{a,k}^{N1}z'^{N1}_{a,k,s,\zeta}$$

$$+\sum_{s\in\Omega^s}\rho_s\sum_{k\in\Omega^k}\sum_{a\in\Omega^a}MC_{a,k}^{N2}y_{a,k,s,\zeta}^{N2}\ ;\ \forall t\in\Omega^t;\forall\zeta\in\Omega^{T2}\tag{128}$$

$$MntC_t^{EL}=\sum_{k\in\Omega^{EL}}MC_k^{EL}u1_{k,\tau}^{EL}+\sum_{s\in\Omega^s}\rho_s\sum_{k\in\Omega^{EL}}MC_k^{EL}u2_{k,s,\zeta}^{EL}\ ;\ \forall\tau\in\Omega^{T1};\forall\zeta$$

$$\in\Omega^{T2}\tag{129}$$

$$EgyC_t^{NG}=\sum_{s\in\Omega^s}\rho_s\sum_{w\in\Omega^w}\pi_w\sum_{g\in\Omega^{NG}}OC_{g,s,w,\tau}^{NG}$$

$$+\sum_{s\in\Omega^s}\rho_s\sum_{w\in\Omega^w}\pi_w\sum_{g\in\Omega^{NG}}OC_{g,s,w,\zeta}^{NG}\ ;\ \forall\tau\in\Omega^{T1};\forall\zeta\in\Omega^{T2}\tag{130}$$

$$EgyC_t^{EG}=\sum_{s\in\Omega^s}\rho_s\sum_{w\in\Omega^w}\pi_w\sum_{g\in\Omega^{EG}}OC_{g,s,w,t}^{EG}\ ;\ \forall t\in\Omega^t\tag{131}$$

$$ENSC_t=\sum_{s\in\Omega^s}\rho_s\sum_{i\in\Omega^i}\pi_w\sum_{w\in\Omega^w}\pi_w\Lambda_{s,w,t}p_{i,s,w,t}\ ;\ \forall t\in\Omega^t\tag{132}$$

$$EmiC_t^G=EmiC_t^{NG}+EmiC_t^{EG}\ ;\ \forall t\in\Omega^t\tag{133}$$

$$EmiC_t^{NG}=\sum_{s\in\Omega^s}\rho_s\sum_{w\in\Omega^w}\pi_w\sum_{g\in\Omega^{NG}}\lambda_{s,w,t}^{CO_2e}ER_g^{NG}PG_{g,s,w,t}^{NG}\ ;\ \forall t\in\Omega^t\tag{134}$$

$$EmiC_t^{EG}=\sum_{s\in\Omega^s}\rho_s\sum_{w\in\Omega^w}\pi_w\sum_{g\in\Omega^{EG}}\lambda_{s,w,t}^{CO_2e}ER_g^{EG}PG_{g,s,w,t}^{EG}\ ;\ \forall t\in\Omega^t\tag{135}$$

The total operation costs given by (130) and (131) for new and existing generators, respectively, depend on the amount of power generated for each scenario, snapshot, stage and generator type. Therefore, these costs represent the expected costs of operation. Similarly, the penalty term for the unserved power, given by (132), is dependent on the scenarios, snapshots and time stages. Equation (132) therefore gives

the expected cost of unserved energy in the system. Equations (133) gathers the expected emission costs of power generated by new and existing generators, which are computed using (134) and (135), respectively. Note that, for the sake of simplicity, a linear emission cost function is assumed here. In reality, the emission cost function is highly nonlinear and nonconvex, as in [44]. Moreover, the cost of power generation $OC_{g,s,w,t}^G$; $\forall t \in \Omega^t$ is often modeled using a linear cost curve, where the marginal cost of power production is constant. Should there be a need to use more detailed generation cost curves (quadratic cost curves, for instance), nonlinear terms should be linearized using one of the linearization techniques, extensively discussed in Chapter 3. For the sake of simplicity, a linear cost curve is adopted throughout this thesis.

5.1.2.2. Constraints

Kirchhoff's Laws: Flows in AC systems are governed by Kirchhoff's voltage and current laws, abbreviated as KVL and KCL, respectively. The "DC" network model, described in Chapter 3, is reproduced here by extending the multi-load level equations to fit the proposed TEP framework. Inequality (136) and (137) represent the KVL constraints in existing lines in the first and the second investment sub-horizons, respectively. The corresponding constraints for candidate lines are given by (138)—(140), respectively.

$$\left| P_{k,s,w,t} + V_{nom}^2 b_k \theta_{k,s,w,t} \right| \le M_k (1 - u1_{k,t}); \ \ \forall t \in \Omega^\tau; \ s \in \Omega^s; \ w \in \Omega^w \tag{136}$$

$$\left| P_{k,s,w,t} + V_{nom}^2 b_k \theta_{k,s,w,t} \right| \le M_k (1 - u2_{k,s,t}); \ \ \forall t \in \Omega^\zeta; \ s \in \Omega^s; \ w \in \Omega^w \tag{137}$$

$$\left| P_{a,k,s,w,t} + V_{nom}^2 b_k \theta_{k,s,w,t} \right| \le M_{p,k} (1 - z_{a,k,t}) \ ; \ \forall t \in \Omega^\zeta; \ s \in \Omega^s; \ w \in \Omega^w \tag{138}$$

$$\left| P_{a,k,s,w,\zeta} + V_{nom}^2 b_k \theta_{k,s,w,\zeta} \right| \le M_{a,k} (1 - z'_{a,k,\zeta}) \ ; \ \forall \zeta \in \Omega^\zeta; \ s \in \Omega^s; \ w \in \Omega^w \tag{139}$$

$$\left| P_{a,k,s,w,\zeta} + V_{nom}^2 b_k \theta_{k,s,w,\zeta} \right| \le M_{a,k} (1 - y_{a,k,s,\zeta}^{N2}); \ \forall \zeta \in \Omega^\zeta; \ s \in \Omega^s; \ w \in \Omega^w \tag{140}$$

As mentioned in Chapter 3, the DC network model does not provide voltage magnitude information because one of the underlining assumptions in deriving this model is the consideration of flat voltage throughout the system. This can be somehow corrected by using the linearized active AC power flow equation, presented in Chapter 3, instead of the DC power flow equations described above, as in (141)—(144). Notice that these equations reduce to (136)—(140) if the voltage deviations (from the nominal value) at each node and line resistances are very small. These are among the simplifying assumptions in DC formulation.

$$\left| P_{k,s,w,t} - \{ V_{nom}(\Delta V_{i,s,w,t} - \Delta V_{j,s,w,t}) g_k - V_{nom}^2 b_k \theta_{k,s,w,t}) \} \right|$$
$$\le M_k (1 - u1_{k,t}); \ \ \forall t \in \Omega^\tau; \ s \in \Omega^s; \ w \in \Omega^w \tag{141}$$

$$\left| P_{k,s,w,t} - \{ V_{nom}(\Delta V_{i,s,w,t} - \Delta V_{j,s,w,t}) g_k - V_{nom}^2 b_k \theta_{k,s,w,t}) \} \right|$$
$$\le M_k (1 - u2_{k,s,t}); \ \ \forall t \in \Omega^\zeta; \ s \in \Omega^s; \ w \in \Omega^w \tag{142}$$

$$\left| P_{a,k,s,w,t} - \{ V_{nom} \left(\Delta V_{i,s,w,t} - \Delta V_{j,s,w,t} \right) g_k - V_{nom}^2 b_k \theta_{k,s,w,t} \) \} \right|$$
$$\leq M_{a,k} (1 - z_{a,k,t}) \, ; \, \forall t \in \Omega^\zeta; \, s \in \Omega^s; \, w \in \Omega^w \tag{143}$$

$$\left| P_{a,k,s,w,t} - \{ V_{nom} \left(\Delta V_{i,s,w,t} - \Delta V_{j,s,w,t} \right) g_k - V_{nom}^2 b_k \theta_{k,s,w,\zeta} \) \} \right|$$
$$\leq M_{a,k} (1 - z'_{a,k,\zeta}) \, ; \, \forall \zeta \in \Omega^\zeta; \, s \in \Omega^s; \, w \in \Omega^w \tag{144}$$

$$\left| P_{a,k,s,w,t} - \{ V_{nom} \left(\Delta V_{i,s,w,t} - \Delta V_{j,s,w,t} \right) g_k - V_{nom}^2 b_k \theta_{k,s,w,\zeta} \) \} \right|$$
$$\leq M_{a,k} (1 - y_{a,k,s,\zeta}^{N2}); \, \forall \zeta \in \Omega^\zeta; \, s \in \Omega^s; \, w \in \Omega^w \tag{145}$$

KCL constraints dictate that the load balance at each node should be respected all the time i.e. the sum of all injections should be equal to the sum of all withdrawals. This is enforced by adding the following constraints:

$$\sum_{(a,k)\in\Omega^{NL}} \sum_{k\in i} P_{a,k,s,w,t} + \sum_{k\in i; k\in\Omega^{EL}} P_{k,s,w,t} + \sum_{g\in i} PG_{g,s,w,t} + p_{i,s,w,t}$$
$$- \sum_{d\in i} PD_{d,s,w,t} + \sum_{(a,k)\in\Omega^{NL}} \sum_{k\in i} 0.5 * PL_{a,k,s,w,t} \tag{146}$$
$$+ \sum_{k\in i; k\in\Omega^{EL}} 0.5 * PL_{k,s,w,t} = 0; \, \forall t \in \Omega^t; \, s \in \Omega^s; \, w \in \Omega^w$$

Constraints Related to Network Losses: The real power losses in line k can be approximated as follows:

$$PL_{k,s,w,t} \approx V_{nom}^2 g_k \theta_{k,s,w,t}^2 \tag{147}$$

Clearly, Equation (147) is nonlinear and nonconvex function. Since keeping the linearity of the TEP problem is critical for computational reasons, Equation (147) need to be linearized. The most common linearization approach in the literature piecewise-linearizing the quadratic angular difference. However, instead of doing this, the expression in (147) can be expressed in terms of active power flows as in (148), as thoroughly described in Chapter 3. Issues related to network losses and linearization are extensively discussed in Chapter 3 and also in our published work [13].

$$PL_{k,s,w,t} = r_k P_{k,s,w,t}^2 / V_{nom}^2 \tag{148}$$

The quadratic expressions of active power flow in (148) can then be easily linearized using piecewise linearization, considering a sufficiently large number of partitions, *L*. There are a number of ways of linearizing such functions such as incremental, multiple choice, convex combination and other approaches in the literature [13], [183]. Here, the convex combination approach, which is implemented making use of special ordered sets of type 2 (SOS2). This losses modeling technique is described in Chapter 3. Further details can also be found in our published work [13]. For the sake of completeness, the model is reproduced here. For the linearization, two non-negative auxiliary variables are introduced for each of the flows $P_{k,s,w,t}$ such that $P_{k,s,w,t} = P_{k,s,w,t}^+ - P_{k,s,w,t}^-$. This implies $\left| P_{k,s,w,t} \right| = P_{k,s,w,t}^+ + P_{k,s,w,t}^-$. Note that these auxiliary variables (i.e. $P_{k,s,w,t}^+$ and $P_{k,s,w,t}^-$ represent the positive and negative flows of $P_{k,s,w,t}$, respectively. This helps one

to consider only the positive quadrant of the nonlinear curve, resulting in a significant reduction in mathematical complexity, and by implication the computational burden. In this case, the associated linear constraints are:

$$P_{k,s,w,t}^2 = \sum_{l=0}^{L} \lambda_{k,s,w,t}(l)\left[P_{k,s,w,t}(l)\right]^2 \quad ; \forall t \in \Omega^t; \ s \in \Omega^s; \ w \in \Omega^w; k \in \{\Omega^{NL} \cup \Omega^{EL}\} \tag{149}$$

$$\left|P_{k,b}\right| = P_{k,s,w,t}^+ + P_{k,s,w,t}^- = \sum_{l=1}^{L} \lambda_{k,s,w,t}(l)P_{k,s,w,t}(l) \quad ; \forall t \in \Omega^t; \ s \in \Omega^s; \ w \in \Omega^w; k \in \{\Omega^{NL} \cup \Omega^{EL}\} \tag{150}$$

$$\sum_{l=1}^{L} \lambda_{k,s,w,t}(l) = 1 \quad ; \forall t \in \Omega^t; \ s \in \Omega^s; \ w \in \Omega^w; k \in \{\Omega^{NL} \cup \Omega^{EL}\} \tag{151}$$

$$\lambda_{k,s,w,t}(l) \in SOS2 \quad ; \forall t \in \Omega^t; \ s \in \Omega^s; \ w \in \Omega^w; k \in \{\Omega^{NL} \cup \Omega^{EL}\} \tag{152}$$

where $P_{k,s,w,t}(l) = l * \frac{S_k^{max}}{L}$. Note that this has to be done for both existing and candidate lines. Further details about this model can be found in [13]. The losses in candidate lines are also linearized in a similar way.

Note that expressing the losses as a function of flows has a clear advantage over the angle-based losses. It avoids unnecessary constraints on the angle differences when a line between two nodes is not connected or remains not selected for investment. In the linearization of losses based on Equation (147), such problem is avoided by introducing additional binary variables and using a so-called big-*M* formulation [131]. However, this adds extra complexity to the problem.

Line Flow Limits: Flows in any line should lie within the permissible range i.e. within its thermal capacity limits. In existing lines, these constraints are enforced by (153) and (154) in the first and the second sub-horizons, respectively. The corresponding constraints in the case of candidates are given by (155)—(157).

$$\left|P_{k,s,w,t}\right| + 0.5PL_{k,s,w,t} \le u1_{k,t}^{EL}S_{k,\max} \quad ; \forall t \in \Omega^\tau; \ s \in \Omega^s; \ w \in \Omega^w; k \in \Omega^{EL} \tag{153}$$

$$\left|P_{k,s,w,t}\right| + 0.5PL_{k,s,w,t} \le u2_{k,s,t}^{EL}S_{k,\max} \quad ; \forall t \in \Omega^\zeta; \ s \in \Omega^s; \ w \in \Omega^w; k \in \Omega^{EL} \tag{154}$$

$$\left|P_{a,k,s,w,t}\right| + 0.5PL_{a,k,s,w,t} \le z_{a,k,t}S_{a,k,\max} \quad ; \forall t \in \Omega^\tau; \ s \in \Omega^s; \ w \in \Omega^w; (a,k) \in \Omega^{N1} \tag{155}$$

$$\left|P_{a,k,s,w,t}\right| + 0.5PL_{a,k,s,w,t} \le z'_{a,k,t}S_{a,k,\max} \quad ; \forall t \in \Omega^\zeta; \ s \in \Omega^s; \ w \in \Omega^w; (a,k) \in \Omega^{N1} \tag{156}$$

$$\left|P_{a,k,s,w,t}\right| + 0.5PL_{a,k,s,w,t} \le y_{a,k,s,t}^{N2}S_{a,k,\max} \quad ; \forall t \in \Omega^\zeta; \ s \in \Omega^s; \ w \in \Omega^w; (a,k) \in \Omega^{N2} \tag{157}$$

Active Power Limits of generators: The generation capacity limits of existing and candidate generators are given by (158) and (159), respectively. In the case of candidate generators, the corresponding constraints are (60). Note that the binary variables $u_{g,s,w,t}$

are required to indicate whether an existing generator is available or not. This makes sure that the power generation variable is zero when the generator is not being used.

$$PG_{g,s,w,t}^{EG,min}u_{g,s,w,t} \leq PG_{g,s,w,t}^{EG} \leq PG_{g,s,w,t}^{EG,max}u_{g,s,w,t} \; ; \; \forall t \in \Omega^t; \; s\epsilon\Omega^s; \; w\epsilon\Omega^w; \; g \in \Omega^{EG} \tag{158}$$

$$PG_{g,s,w,t}^{NG,min}x_{g,s,w,t} \leq PG_{g,s,w,t}^{NG} \leq PG_{g,s,w,t}^{NG,max}u_{g,s,w,t} \; \; ; \; \forall t \in \Omega^t; \; s\epsilon\Omega^s; \; w\epsilon\Omega^w; \; g \in \Omega^{NG} \tag{159}$$

It should be noted that, in the case of intermittent power source, the lower generation limits $PG_{g,s,w,t}^{EG,min}$ and $PG_{g,s,w,t}^{NG,min}$ are often set to 0 while the corresponding upper limits are set equal to the actual power output of the variable generation source corresponding to the level of primary energy source (wind speed and solar radiation, for instance). Hence, the upper bound in this case depends on the operational state (i.e. the snapshot) and the scenario.

Logical Constraints: Investment-related logical constraints (160)—(165) are included. These set of constraints ensure that an investment made at time stage t cannot be reversed or divested in the subsequent time stages; instead, the asset should be available for utilization immediately.

$$z_{a,k,\tau}^{N1} \geq z_{a,k,\tau-1}^{N1} \quad ; \; \forall\tau \in \Omega^{T1}; \; k \in \Omega^k; \; a \in \Omega^a \tag{160}$$

$$z_{a,k,\zeta}^{N1} = z_{a,k,T1}^{N1} \quad ; \; \forall\zeta \in \Omega^{T2}; \; k \in \Omega^k; \; a \in \Omega^a \tag{161}$$

$$z_{a,k,s,\zeta}'^{N1} \geq z_{a,k,s,\zeta-1}'^{N1} \; ; \; \forall\zeta \in \Omega^{T2}; \; s \in \Omega^s; \; k \in \Omega^k; \; a \in \Omega^a \tag{162}$$

$$z_{a,k,s,T1}'^{N1} = z_{a,k,T1}^{N1} \quad ; \; s \in \Omega^s; \; k \in \Omega^k; \; a \in \Omega^a \tag{163}$$

$$y_{a,k,s,\zeta}^{N2} \geq y_{a,k,s,\zeta-1}^{N2} \; ; \; \forall\zeta \in \Omega^{T2}; \; s \in \Omega^s; \; k \in \Omega^k; \; a \in \Omega^a \tag{164}$$

$$y_{a,k,s,T1}^{N2} = 0 \quad ; \; s \in \Omega^s; \; k \in \Omega^k; \; a \in \Omega^a \tag{165}$$

Budget Constraints: A budget constraint for line invests is enforced by adding constraint (166) for the first period and (167) for the second one.

$$\sum_{k\in\Omega^k}\sum_{a\in\Omega^a} IC_{a,k}^{N1}(z_{a,k,\tau}^{N1} - z_{a,k,\tau-1}^{N1}) \leq InvLim_\tau \; ; \; \forall\tau \in \Omega^{T1} \tag{166}$$

$$\sum_{k\in\Omega^k}\sum_{a\in\Omega^a} IC_{a,k}^{N1}(y_{a,k,s,\zeta}^{N2} - y_{a,k,s,\zeta-1}^{N2}) + \sum_{k\in\Omega^k}\sum_{a\in\Omega^a} IC_{a,k}^{N1}(z_{a,k,s,\zeta}'^{N1} - z_{a,k,s,\zeta-1}'^{N1})$$
$$\leq InvLim_{s,\zeta} \quad ; \; \forall\zeta \in \Omega^{T2}; \; s \in \Omega^s \tag{167}$$

Unserved Power Limits: The unserved power at any given node cannot exceed the demand at that node, and this is enforced by:

$$0 \leq p_{i,s,w,t} \leq d_{i,s,w,t} \quad ; \forall i \in \Omega^i; \; \forall t \in \Omega^t; \; s \in \Omega^s; \; w \in \Omega^w \tag{168}$$

Emission Related Constraints: Emission reduction targets can be achieved by imposing strict control, check and balance emission regulations. In this regard, constraint (169) is

added here which caps the expected emission levels at a given year t below a preset target.

$$\sum_{s \in \Omega^s} \rho_s \sum_{w \in \Omega^w} \pi_w \sum_{g \in \Omega^{NG}} (ER_g^{NG} PG_{g,s,w,t}^{NG} + ER_g^{EG} PG_{g,s,w,t}^{EG}) \leq TEEL_t \; ; \; \forall t \in \Omega^t \quad (169)$$

Angle and Voltage Related Constraints: For stability and power quality reasons, the voltage magnitude and its angle at each bus are bounded by:

$$\theta_{min} \leq \theta_{i,s,w,t} \leq \theta_{max}; \; \theta_{ref} = 0 \; ; \forall i \in \Omega^i; \; \forall t \in \Omega^t; \; s \in \Omega^s; \; w \in \Omega^w \quad (170)$$

$$\Delta V^{min} \leq \Delta V_{i,s,w,t} \leq \Delta V^{max} \; ; \forall i \in \Omega^i; \; \forall t \in \Omega^t; \; s \in \Omega^s; \; w \in \Omega^w \quad (171)$$

HVDC Line Constraints: DC lines are modeled as "flow" networks, which means that power flows in such lines are not goverened by Kirchhoff's voltage law. Unlike flows in AC lines, flows in DC ones are independent of the voltages and angles at the nodes where the DC lines are connected to. DC lines only respect load balance and capacity constraints and; hence, share the flow constraints, node balance as well as the losses constraints in (146)—(157) with their AC counterparts.

5.2. TEP MODEL REVISITED

From a computational standpoint, the TEP model presented Section 5.1, which is based on a yearly temporal planning scope, may not be sometimes affordable when applied to extra-large systems of the European network scale. Given the sheer size of such network systems, computing optimal power flow calculations for each time stage and scenario over a long planning horizon (often 30 to 50 years) renders significant computational challenge. To overcome this, the TEP problem can be re-formulated based on a reduced number of intermediate planning stages in each sub-horizon (planning period). In a two-period planning framework, assume the first period has two decision stages, one intermediate and one final stages, which are denoted as τ and $T1$, respectively, and, the second period is represented by one stage at the final planning horizon, as shown in Figure 5.2. This leads us to a three-stage problem. Note that an intermediate stage is intentionally added to the first period to somehow account for the investment lag inherent to TEP projects induced by the often lengthy permission process.

One way to formulate the objective function of such a three-stage problem with three stages is to minimize the total NPV sum of costs in each of the considered years (i.e. the three planning stages), as in (121a). The composition of these costs the same as the original model in Section 5.1, and are computed using Equations (122a)—(135a). The cost terms here differ from those described before in that they do not reflect the operation, maintenance, emission and reliability costs incurred outside these stages. In other words, these costs are not spread throughout and beyond the planning horizon to capture the short to long term impacts of expansion decisions on the levels of these costs.
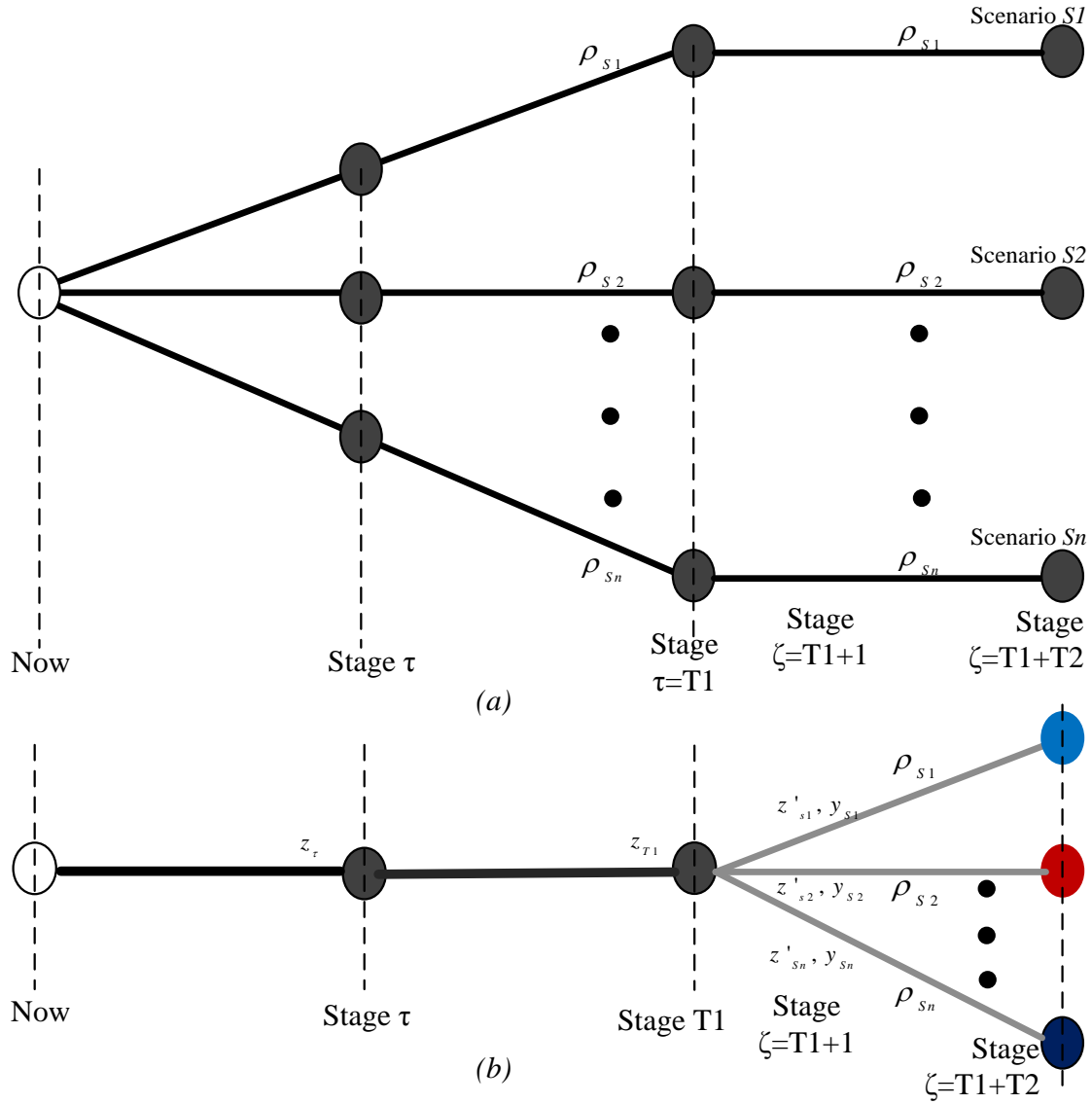
Fig. 5. 2 A schematic representation of (a) possible future scenario trajectories and (b) a decision structure

$$\min_{z,y,z',\dots} TC = \alpha_1 * TInvC + \alpha_2 * TMC + \alpha_3 * TEC + \alpha_4 * TENSC + \alpha_5 * TImiC \tag{121a}$$

$$TInvC = \underbrace{\sum_{t\epsilon\Omega^{t'}} (1+\sigma)^{-t} \, InvC_t^{LN}}_{NPV \ of \ investment \ cost} \tag{122a}$$

$$TMC = \underbrace{\sum_{t\epsilon\Omega^{t'}} (1+\sigma)^{-t} \, (MntC_t^{NL} + MntC_t^{EL})}_{NPV \ of \ maintenance \ costs} \tag{123a}$$

$$TEC = \underbrace{\sum_{t\epsilon\Omega^{t'}} (1+\sigma)^{-t} \, (EgyC_t^{NG} + EgyC_t^{EG})}_{NPV \ of \ operation \ costs} \tag{124a}$$

$$TENSC = \underbrace{\sum_{t\epsilon\Omega^{t'}} (1+\sigma)^{-t} \ ENSC_t}_{NPV \ of \ reliability \ costs} \tag{125a}$$

$$TEmiC = \underbrace{\sum_{t\epsilon\Omega^{t'}} (1+\sigma)^{-t} \ (EmiC_t^{NG} + EmiC_t^{EG})}_{NPV \ emission \ costs} \tag{126a}$$

$$\left.\begin{aligned} InvC_t^{NL} &= \sum_{k\in\Omega^k} \sum_{a\in\Omega^a} \frac{\sigma(1+\sigma)^{LT_{a,k}}}{(1+\sigma)^{LT_{a,k}}-1} IC_{a,k}^{N1}(z_{a,k,t}^{N1} - z_{a,k,t-1}^{N1}) \ ; \forall t \in \{\tau, T1\} \\ InvC_T^{NL} &= \sum_{s\in\Omega^s} \rho_s \sum_{k\in\Omega^k} \sum_{a\in\Omega^a} \frac{\sigma(1+\sigma)^{LT_{a,k}}}{(1+\sigma)^{LT_{a,k}}-1} IC_{a,k}^{N1}\big(z'^{N1}_{a,k,s,T} - z'^{N1}_{a,k,s,T1}\big) \\ &+ \sum_{s\in\Omega^s} \rho_s \sum_{k\in\Omega^k} \sum_{a\in\Omega^a} \frac{\sigma(1+\sigma)^{LT_{a,k}}}{(1+\sigma)^{LT_{a,k}}-1} IC_{p,k}^{N2}(y_{a,k,s,T}^{N2} - y_{a,k,s,T1}^{N2}) \\ & \qquad\qquad z_{p,k,0} = 0 \ ; \ y_{p,k,n,s,T1}^{N2} = 0 \end{aligned}\right\} \tag{127a}$$

$$\left.\begin{aligned} MntC_t^{NL} &= \sum_{k\in\Omega^k} \sum_{p\in\Omega^p} MC_{p,k}^{N1} z_{p,k,n,t}^{N1} \ ; \forall t \in \{\tau, T1\} \\ MntC_t^{NL} &= \sum_{s\in\Omega^s} \rho_s \sum_{k\in\Omega^k} \sum_{p\in\Omega^p} MC_{p,k}^{N1} z'^{N1}_{p,k,s,T} + \sum_{s\in\Omega^s} \rho_s \sum_{k\in\Omega^k} \sum_{p\in\Omega^p} MC_{p,k}^{N2} y_{p,k,s,T}^{N2} \end{aligned}\right\} \tag{128a}$$

$$\left.\begin{aligned} MntC_t^{EL} &= \sum_{k\in\Omega^{EL}} MC_k^{EL} u1_{k,t}^{EL} \ ; \forall t \in \{\tau, T1\} \\ MntC_T^{EL} &= \sum_{s\in\Omega^s} \rho_s \sum_{k\in\Omega^{EL}} MC_k^{EL} u2_{k,s,T}^{EL} \end{aligned}\right\} \tag{129a}$$

$$EgyC_t^{NG} = \sum_{s\in\Omega^s} \rho_s \sum_{w\in\Omega^w} \pi_w \sum_{g\in\Omega^{NG}} OC_{g,s,w,t}^{NG} ; \ \forall t \in \{\tau, T1, T\} \tag{130a}$$

$$EgyC_t^{EG} = \sum_{s\in\Omega^s} \rho_s \sum_{w\in\Omega^w} \pi_w \sum_{g\in\Omega^{EG}} OC_{g,s,w,t}^{EG} \ ; \forall t \in \{\tau, T1, T\} \tag{131a}$$

$$ENSC_t = \sum_{s\in\Omega^s} \rho_s \sum_{i\in\Omega^i} \pi_w \sum_{w\in\Omega^w} \pi_w \Lambda_{s,w,t} p_{i,s,w,t} ; \ \forall t \in \{\tau, T1, T\} \tag{132a}$$

$$EmiC_t^G = EmiC_t^{NG} + EmiC_t^{EG} ; \ \forall t \in \{\tau, T1, T\} \tag{133a}$$

$$EmiC_t^{NG} = \sum_{s\in\Omega^s} \rho_s \sum_{w\in\Omega^w} \pi_w \sum_{g\in\Omega^{NG}} \lambda_{s,w,t}^{CO_2e} ER_g^{NG} PG_{g,s,w,t}^{NG} \ ; \ \forall t \in \{\tau, T1, T\} \tag{134a}$$

$$EmiC_t^{EG} = \sum_{s\in\Omega^s} \rho_s \sum_{w\in\Omega^w} \pi_w \sum_{g\in\Omega^{EG}} \lambda_{s,w,t}^{CO_2e} ER_g^{EG} PG_{g,s,w,t}^{EG} \ ; \ \forall t \in \{\tau, T1, T\} \tag{135a}$$

The constraints to this optimization problem are the same as the ones in the original formulation except for the domain of variables and parameters related to the stage $t = \{\tau, T1, T2\}$.

As stated above, the major disadvantage of this formulation is the misbalancing of investment and other costs. Because the real costs throughout the planning horizon are not properly captured by this modeling, the resulting expansion solution can be suboptimal.

The above problem can be corrected by finding an estimate to the costs incurred on the "missing" years i.e. the years apart from the target years. By making use of the annuity concept [186], the operation, maintenance, emission and reliability costs incurred between the intermediate stages considered in the formulation can be approximated from the corresponding known quantities at these stages. This is further demonstrated in Figure 5.3. In this figure, suppose the first sub-horizon (period) spans over 15 years and the duration of the second one is 25 years. Furthermore, let us assume we have two target years for making investments in the first period i.e. one intermediate stage (5th year) and the last planning stage (15th year). The second sub-horizon (i.e. from the 16th year to the 40th one) have only one planning stage, which in this case is considered to be the last one. In effect, instead of having 40 yearly stages (15 in the first period and 25 in the second one), the whole planning horizon has now 3 decision stages. For this planning framework, the operation costs corresponding to these years are explicitly known. Now, the issue is to approximate the costs incurred in the years other than those explicitly considered i.e. the costs corresponding to Part I—IV in Figure 5.3.

Without loss of generality, the fixed payments in the years leading to stage 5 can be assumed to be the same the costs at this stage OC_5. Similarly, the annualized costs between the six and the 15th years can be assumed to be equal to those at the 15th year OC_{15}, while the annual costs in each year of the second period can be regarded to be equal to OC_{40}. Given all this, the concept of annuity [186] can be applied. Hence, the costs in Part I are assumed to be accrued and paid in full at the end of the fifth year, those in Part II at the fifth year and those in Part III at the last stage of the planning horizon. the total operation costs in each range (part) can be estimated by the difference of the perpetuity of the corresponding two known operation costs, updated by the NPV factor. Note that the present value of perpetuity, which is the sum of the net worth of infinite annual fixed payments, is determined by the ratio of the fixed payment at a given time by the interest rate σ. For the illustrative example, the total costs for the "missing" years can be estimated using: $OC_I = \frac{OC_5}{\sigma} - \frac{OC_5}{\sigma(1+\sigma)^5}$, $OC_{II} = \frac{OC_{15}}{\sigma(1+\sigma)^5} - \frac{OC_{15}}{\sigma(1+\sigma)^{15}}$, and $OC_{III} = \frac{OC_{40}}{\sigma(1+\sigma)^{15}} - \frac{OC_{40}}{\sigma(1+\sigma)^{40}}$. It is rather straightforward to express the costs incurred after the planning horizon, which depend on the costs in the final planning stage. Assuming perpetual planning horizon, OC_{IV} can be expressed as $\frac{OC_{40}}{\sigma(1+\sigma)^{40}}$. The remaining costs can be formulated in a similar manner. The objective function then minimizes the sum of all cost cost terms formulated in this way.
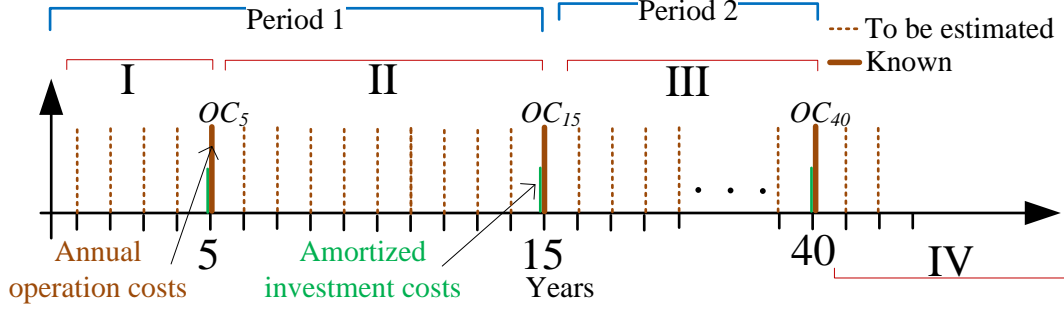
Fig. 5. 3 Illustration of cost components in the formulation

The complete formulation of the objective function of the TEP model with reduced number of stages is presented in (121b)—(135b). The constraints are for this optimization are the same as those in the original formulation with the exception of time stage domain, which in this case is $\{\tau, T1, T2\}$.

$$\min_{z,y,z',\dots} TC = \alpha_1 * TInvC + \alpha_2 * TMC + \alpha_3 * TEC + \alpha_4 * TENSC + \alpha_5 * TImiC \quad (121b)$$

$$TInvC = \underbrace{\sum_{t\epsilon\Omega^{t'}} \frac{(1+\sigma)^{-t}}{\sigma} InvC_t^{LN}}_{NPV\ of\ amortized\ investment\ cost} \quad (122b)$$

$$TMC = \underbrace{\left(\frac{1}{\sigma} - \frac{(1+\sigma)^{-\tau}}{\sigma}\right)(MntC_\tau^{NL} + MntC_\tau^{EL})}_{I}$$
$$+ \underbrace{\left(\frac{(1+\sigma)^{-\tau}}{\sigma} - \frac{(1+\sigma)^{-T1}}{\sigma}\right)(MntC_{T1}^{NL} + MntC_{T1}^{EL})}_{II} \quad (123b)$$
$$+ \underbrace{\frac{(1+\sigma)^{-T}}{\sigma}(MntC_T^{NL} + MntC_T^{EL})}_{III+IV}$$

$$TEC = \underbrace{\left(\frac{1}{\sigma} - \frac{(1+\sigma)^{-\tau}}{\sigma}\right)(EgyC_\tau^{NG} + EgyC_\tau^{EG})}_{I}$$
$$+ \underbrace{\left(\frac{(1+\sigma)^{-\tau}}{\sigma} - \frac{(1+\sigma)^{-T1}}{\sigma}\right)(EgyC_{T1}^{NG} + EgyC_{T1}^{EG})}_{II} \quad (124b)$$
$$+ \underbrace{\frac{(1+\sigma)^{-T}}{\sigma}(EgyC_T^{NG} + EgyC_T^{EG})}_{III+IV}$$

$$TENSC = \underbrace{\left(\frac{1}{\sigma} - \frac{(1+\sigma)^{-\tau}}{\sigma}\right)ENSC_\tau}_{I} + \underbrace{\left(\frac{(1+\sigma)^{-\tau}}{\sigma} - \frac{(1+\sigma)^{-T1}}{\sigma}\right)ENSC_{T1}}_{II}$$
$$+ \underbrace{\frac{(1+\sigma)^{-T}}{\sigma}ENSC_T}_{III+IV} \tag{125b}$$

$$TEmiC = \underbrace{\left(\frac{1}{\sigma} - \frac{(1+\sigma)^{-\tau}}{\sigma}\right)(EmiC_\tau^{NG} + EmiC_\tau^{EG})}_{I}$$
$$+ \underbrace{\left(\frac{(1+\sigma)^{-\tau}}{\sigma} - \frac{(1+\sigma)^{-T1}}{\sigma}\right)(EmiC_{T1}^{NG} + EmiC_{T1}^{EG})}_{II} \tag{126b}$$
$$+ \underbrace{\frac{(1+\sigma)^{-T}}{\sigma}(EmiC_T^{NG} + EmiC_T^{EG})}_{III+IV}$$

$$\left.\begin{array}{l} InvC_t^{NL} = \sum_{k\in\Omega^k}\sum_{a\in\Omega^a} \frac{\sigma(1+\sigma)^{LT_{a,k}}}{(1+\sigma)^{LT_{a,k}}-1} IC_{a,k}^{N1}(z_{a,k,t}^{N1} - z_{a,k,t-1}^{N1}) \; ; \forall t \in \{\tau, T1\} \\[2em] InvC_T^{NL} = \sum_{s\in\Omega^s}\rho_s \sum_{k\in\Omega^k}\sum_{a\in\Omega^a} \frac{\sigma(1+\sigma)^{LT_{a,k}}}{(1+\sigma)^{LT_{a,k}}-1} IC_{a,k}^{N1}(z'^{N1}_{a,k,s,T} - z'^{N1}_{a,k,s,T1}) \\[2em] + \sum_{s\in\Omega^s}\rho_s \sum_{k\in\Omega^k}\sum_{a\in\Omega^a} \frac{\sigma(1+\sigma)^{LT_{a,k}}}{(1+\sigma)^{LT_{a,k}}-1} IC_{p,k}^{N2}(y_{a,k,s,T}^{N2} - y_{a,k,s,T1}^{N2}) \end{array}\right\} \tag{127b}$$
$$z_{p,k,0} = 0 \; ; \; y_{p,k,n,s,T1}^{N2} = 0$$

$$\left.\begin{array}{l} MntC_\tau^{NL} = \sum_{k\in\Omega^k}\sum_{a\in\Omega^a} MC_{a,k}^{N1} z_{a,k,n,\tau}^{N1} \; ; \; \forall\tau \in \{\tau, T1\} \\[2em] MntC_T^{NL} = \sum_{s\in\Omega^s}\rho_s \sum_{k\in\Omega^k}\sum_{a\in\Omega^a} MC_{a,k}^{N1} z'^{N1}_{a,k,s,T} + \sum_{s\in\Omega^s}\rho_s \sum_{k\in\Omega^k}\sum_{a\in\Omega^a} MC_{a,k}^{N2} y_{a,k,s,T}^{N2} \end{array}\right\} \tag{128b}$$

$$\left.\begin{array}{l} MntC_t^{EL} = \sum_{k\in\Omega^{EL}} MC_k^{EL} u1_{k,t}^{EL} \; ; \; \forall t \in \{\tau, T1\} \\[2em] MntC_T^{EL} = \sum_{s\in\Omega^s}\rho_s \sum_{k\in\Omega^{EL}} MC_k^{EL} u2_{k,s,T}^{EL} \end{array}\right\} \tag{129b}$$

$$EgyC_t^{NG} = \sum_{s\in\Omega^s}\rho_s \sum_{w\in\Omega^w}\pi_w \sum_{g\in\Omega^{NG}} OC_{g,s,w,t}^{NG} \; ; \; \forall t \in \{\tau, T1, T\} \tag{130b}$$

$$EgyC_t^{EG} = \sum_{s\in\Omega^s}\rho_s \sum_{w\in\Omega^w}\pi_w \sum_{g\in\Omega^{EG}} OC_{g,s,w,t}^{EG} \; ; \; \forall t \in \{\tau, T1, T\} \tag{131b}$$

$$ENSC_t = \sum_{s\in\Omega^s}\rho_s \sum_{i\in\Omega^i}\pi_w \sum_{w\in\Omega^w} \pi_w \Lambda_{s,w,t} p_{i,s,w,t} \; ; \; \forall t \in \{\tau, T1, T\} \tag{132b}$$

$$EmiC_t^G = EmiC_t^{NG} + EmiC_t^{EG} \quad ; \forall t \in \{\tau, T1, T\} \tag{133b}$$

$$EmiC_t^{NG} = \sum_{s \in \Omega^S} \rho_s \sum_{w \in \Omega^W} \pi_w \sum_{g \in \Omega^{NG}} \lambda_{s,w,t}^{CO_2e} ER_g^{NG} PG_{g,s,w,t}^{NG} \quad ; \forall t \in \{\tau, T1, T\} \tag{134b}$$

$$EmiC_t^{EG} = \sum_{s \in \Omega^S} \rho_s \sum_{w \in \Omega^W} \pi_w \sum_{g \in \Omega^{EG}} \lambda_{s,w,t}^{CO_2e} ER_g^{EG} PG_{g,s,w,t}^{EG} \quad ; \forall t \in \{\tau, T1, T\} \tag{135b}$$

5.3. ROLLING WINDOW OF PLANNING

Power systems is subject to continuous changes and high level uncertainty. Because of this, it is almost impossible to exhaustively characterize its possible evolutions only in the form of a predefined storylines (scenarios). Moreover, the number of scenarios should be limited to ensure tractability. Yet, expansion decisions have to cope with the envitable changes in system evolutions. To adapt the decisions to a changing environment, the concept of a rolling window of planning is introduced in this thesis. This lays a quasi-dynamic planning framework which is attractive in intuitive terms because it recognizes the fact that the plan will be effectively readjusted as new data becomes available and tries to accommodate the effect of uncertainty by constantly readjusting the probabilities of realization of the scenarios. This is demonstrated in Figure 5.4. This planning framework uses the three-stage planning model developed in the preceding section. Figure 5.4 (a) shows the possible future scenario trajectories $\{s_1, s_2, \dots, s_n\}$ with three scenario spots along the planning horizon, in the three-stage and two-period planning framework. Whereas, Figure 5.4 (b) illustrates the decision structure in each stage, showing a single investment decision z_i common for all scenarios in the first period, and scenario-dependent decisions $\{y_1, y_2, \dots, y_n\}$ in the second period. Figure 5.4 (c) depicts new possible future scenario trajectories $\{s_1^*, s_2^*, \dots, s_n^*\}$ after new information is unveiled or made available. A new TEP optimization is carried out accounting for these changes, and as illustrated in Figure 5.4 (d), a new set of decisions are obtained. Note that Figures 5.4 (c) and (d) both demonstrate the moving window of planning.

It is understood that as time passes by, the scenarios unfold or new information becomes available that changes the probabilities of realizations of the scenarios under consideration. Either way, the planning can be repeated by rolling the planning window and new investment decisions are obtained. This process can be repeated as many times as desired. In doing so, there will be some overlaps in the planning windows, and as in the decisions. Of particular interest in this case are the decisions made in the first stage.
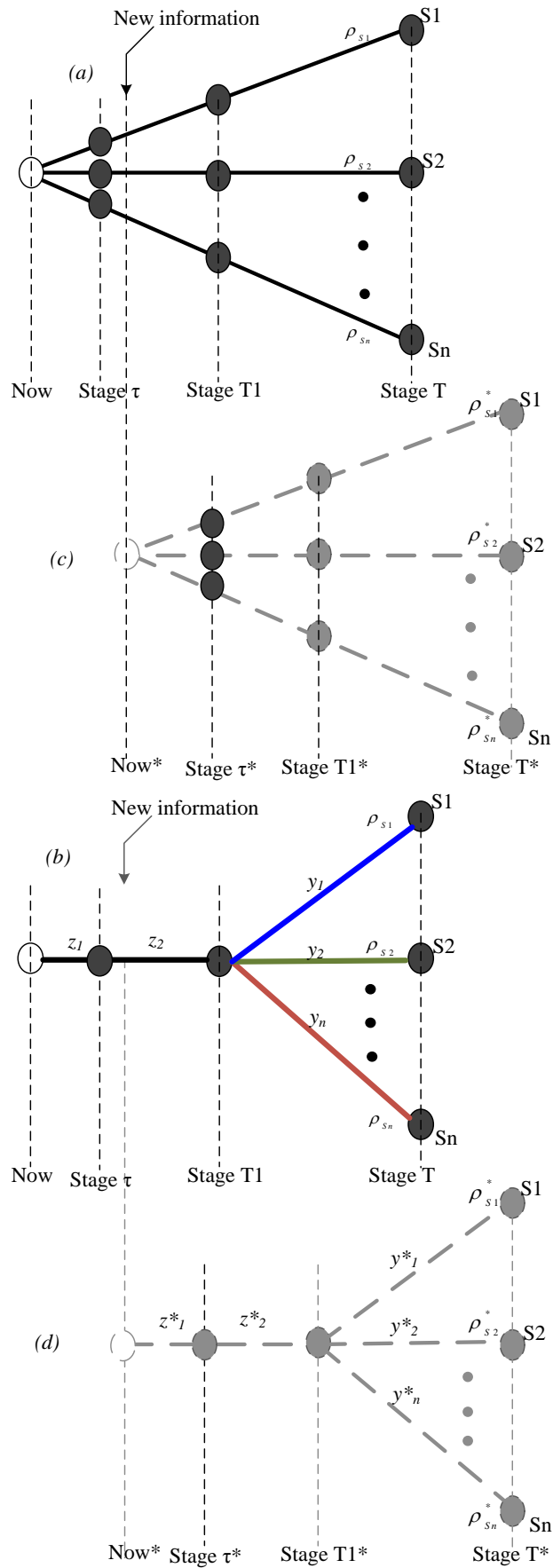
Fig. 5. 4 A schematic representation of the quasi-dynamic planning framework

In general, the quasi-dynamic planning framework helps to explore new expansion solutions as a result of dramatic changes in system evolution and/or obtain possible adjustments to a previously made expansion strategy for minor changes in the storylines. Experience shows that the first-stage expansion decisions are not built all overnight (i.e. at the same time). Considering the current practice in construction of lines, the permit process can be shorter for some, longer for others or even indefinite for some "unlucky" ones. This gives the planner an opportunity to revise the decisions taken in the first stage of the preceding planning window by comparing them with the decisions in the current window. Based on this, he can make some adjustments to the lines planned in the previous window. For instance, the lines in common can be understood as robust and retained in the current planning process. On the other hand, part of the lines of the first stage decisions of the previous window may not appear in the current planning window which may be taken as a reason to cancel them. This process somehow emulates the dynamism involved in TEP.

5.4. DESCRIPTION OF THE SOLUTION STRATEGY

TEP is a naturally combinatorial optimization problem because it includes several discrete (binary or integer) investment decision variables, which often pose significant computational burden. When the size of the system is not large, available solvers can explore the combinatorial search space and find the best expansion topology within a reasonable simulation time. However, for large-scale network systems, this is not possible. Suppose a given system has 2000 candidate corridors for line investments. Assume we have 5 transmission technologies to select from for the investment in each corridor. This would result in a combinatorial search space of $2^{(5*2000)}$ possible combinations, which is "maddeningly" huge. Currently available MILP solvers may not be able to efficiently handle a problem of this magnitude; or else, this may take unacceptably long simulation times even if the simplest TEP model is used. Because of this, the resolution of such a complex combinatorial problem needs to be supported by heuristic methods, which is one of the main aims of this work. This thesis proposes an effective solution strategy involving a gradual reduction of the combinatorial solution search (CSS) space, and parallel computation, largely discussed in [15]. The main of this approach is to significantly enhance the tractability of the TEP problem. Further details and descriptions of the proposed solution strategy is presented as follows.

The computational complexity of TEP is especially pronounced when the considered network is of a continental size, as this work aims to address, where one has to consider thousands of candidate lines in the expansion planning model.

Potential candidate lines for an expansion strategy have been traditionally identified/selected based on expert knowledge. Thus, a short list containing this information has been often made available for carrying out TEP studies. Sometimes, the candidate list by experts is complemented using some heuristic procedures such as the copper sheet method [237] or economic indicators such as locational marginal prices [237], [238] and [239], etc. However, given the huge network size (continental), such

information is unfortunately not available. This means that one cannot rely on expert knowledge; instead, consider a lot of candidates to complete this missing information. In addition, using a huge list of candidates intorudces sufficient flexibility in the search for the most economical expansion strategy.

Unfortunately, as mentioned earlier, increasing the number of candidate lines increases the CSS space, rendering significant burden to the solution process. In other words, given the network size and the huge number of candidates needed in the planning, the size of the optimization problem quickly increases and its computational complexity becomes beyond acceptable level. Unless the CSS space is sufficiently reduced, the resulting optimization can be intractable or demand an exceptionally huge computational effort. This makes it important to reduce the size of the problem without significantly compromising the quality (accuracy) of the solution. In light of this, the present work uses a successive decomposition technique to reduce the CSS space. The proposed solution strategy is schematically illustrated in Figure 5.5, summarizing the procedures followed.

The technique works by decomposing the problem into a number of successive optimization phases as illustrated in Figure 5.5. Each phase uses the results of the previous one to reduce the search space. This reduction in complexity allows each phase to use more complex models with a similar computational load. Moreover, each optimization phase could be defined and solved as an independent problem; thus, allowing the use of specific decomposition techniques, or parallel computation whenever possible.

Generally, as shown in Figure 5.5, this solution strategy can be understood as an approach that refines the large size of initial candidate list (ICL) by employing a mathematically simplified optimization model (in this case, MODELS I and II) before applying a more accurate and advanced optimization model (which in this case is MODEL III) to produce the final optimal investment decisions from the reduced candidate list (RCL), obtained by MODELS I and II. In effect, this approach signinificantly reduces the CSS space, facilitating the computational process.
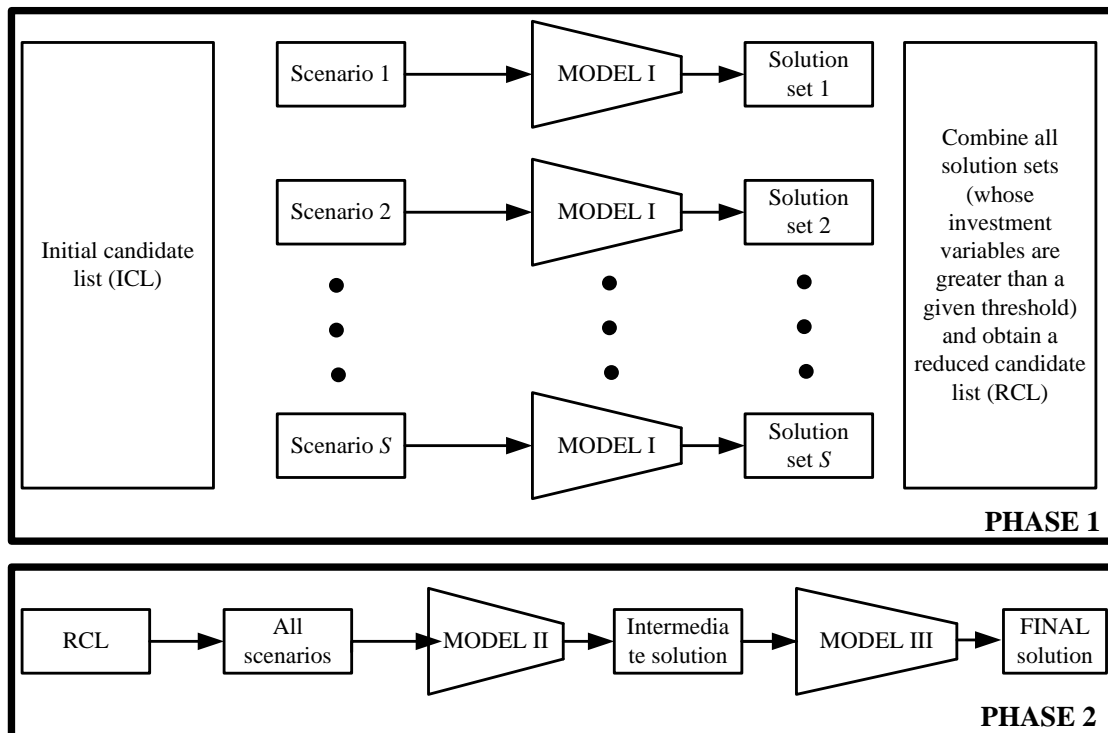
Fig. 5. 5 An illustration of the search space reduction approach and parallel implementation

In general, a significant computational gain is achieved by using this technique. This is because, on one hand, a relatively simplified optimization model (designated by MODEL I in Figure 5.5 and whose formulation can be found in Chapter 3 and [15]), is employed in the early phases before gradually switching to a more detailed and complex optimization models (designated as MODEL II and MODEL III). On the other hand, parallel computation is implemented wherever it is possible. In addition, MODEL I makes use of continuous investment variables which helps to further reduce the computation burden. Converting the naturally discrete investment variables into continuous ones might seem a coarse assumption but numerical results in Chapter 3 and [15] have demonstrated the effectiveness of this methodology.

In the context of the developed TEP problem, the first phase involves deterministic problems as many as the number of scenarios that can be independently solved. In the second phase, the reduced candidate list is further refined by MODEL II which involves a fully stochastic optimization model with continuous investment decision variables. This is followed by a final stochastic optimization, based entirely on an improved DC network model, but this time, considering only the lines selected after Model II of the second phase.

Since the foremost optimization phases assumes continuous investment variables, a set of investment decisions with fractional values is thus obtained for each scenario. A threshold is, therefore, set to limit the list of candidates that would be passed on to the second phase. The present work considers all candidates whose investment variables are different from zero after the optimization. And, these would be selected to eventually

form a reduced candidate list (RCL). This may rather seem very conservative assumption. Setting this threshold a little bit higher than zero as a selection criterion will hardly change the solution, rather it can help to further disregard lines with small investments, that are unlikely to appear in the final and optimal solution.

In the final stage of the first stage, the solution sets of each scenario (i.e. those lines whose investment decision variables are greater than zero) are combined to obtain reduced set of candidates i.e. RCL. In other words, for a candidate line to be considered in the RCL, the line should be selected at least in one of the scenarios i.e. its investment variable should be greater than zero; otherwise, it is rejected.

The second phase has two stochastic optimization processes sequenced one after another (i.e. MODELS II and III). The difference between the two models is that MODEL II is formulated based on the hybrid or the relaxed DC TEP model (see in Chapter 3), which allow continuous investment decisions, while MODEL III is fully based on an improved DC TEP model, in which only discrete investment variables are feasible. The optimization process in MODEL II is carried out considering all the scenarios together but only taking into account the lines in RCL. This optimization results in an intermediate solution comprising some of the lines in RCL. MODEL II further reduces the search space because not all the lines in the RCL are selected for an expansion plan. Therefore, we can also use here the same threshold to get rid of the lines which do not appear in the solution after running MODEL II of the second phase.

As a final step, the second optimization process is run with the intermediate solution as an input. This process finally obtains the required TEP solution.

It should be noted that MODELS I and II can be based on the hybrid or relaxed DC TEP model described in Chapter 3 including network losses. The main property of the hybrid model is that it exempts candidate lines from obeying the second Kirchhoff's law while the rest is the same as the DC network model in [240]. This property makes the hybrid model fit for the CSS space reduction process because it allows the use of continuous transmission investment variables. The relaxed DC TEP model (R-DCTEP), proposed in Chapter 3, also permits the use of continuous variables. This model fares better than the hybrid TEP model (HTEP) because flows in R-DCTEP are forced to obey the law of physics unlike in HTEP where reverse flows can occur in candidate lines. These issues, including numerical results (Tables 3.1—3.4 and Appendix D) are discussed in detail in Chapter 3.

5.5. SUMMARY

This chapter has presented the algebraic formulation of the stochastic TEP model in a multi-stage planning framework and considering multiple objectives including cost of operation and maintenance, emission, energy production, load shedding and line investments. The model is formulated in such a way that it combines mandatory short- to medium-term network expansion decisions with long-term (strategic) decisions both determined in the face of uncertainty. Another salient feature of the proposed model is

its account for the long-term impact of line investments on the overall system costs by means of econometric concepts. Since, a long-term TEP problem spans over 30 or more years, performing a yearly evaluation of the system operation and investment decisions throughout the planning horizon may render significant computational burden. Because of these reasons, a compact formulation, with fewer number of decision stages, is developed for large-scale TEP applications. The concept of rolling window of planning is also introduced to emulate the continuously changing evolution of the system. To address the combinatorial nature of such a problem, an effective solution strategy is described in full. The method works by decomposing the original problem into successive optimization phases, which use TEP models with increasing fidelity levels. This strategy dramatically reduces the combinatorial solution search space, which has a considerable influence on the solution process.

6

VI. CASE STUDIES

This chapter presents numerical results from case study on a reduced 1060-node European network system. The proposed methods and solution strategy are tested on this system.

6.1. CHAPTER OVERVIEW

Solving a TEP model of extra-large network systems (the European network, for example) under high level temporal and uncertainty scope is prohibitively expensive or it can even be impossible. In other words, network planning has to be carried out considering the enormous variability of expected system conditions (described as operation states or snapshots in the preceding chapters) and the high level uncertainty about the evolution of the system in the future (referred to as scenarios or storylines), demanding a new dimension of thinking to solve the resulting huge problem.

Here, the techniques proposed in the previous chapters are employed to reduce the complexity of the problem and enhance tractability. First, the moment-based technique that has been intorudced in Chapter 4 and dissemintated to the research community in [15] to cluster the operational states based on their effects on expansion needs. Second, the uncertainty regarding the evolution of the system is represented by a number of scenarios (or storylines) unfolding as time passes by. The number of storylines is limited, often defined according to expert knowledge. For the sake of brevity, three storylines are defined and used in this case study as the aim of the analysis in this chapters is to demonstrate the versatility of the proposed models, methods and solution strategy.

6.2. A 1060-NODE EUROPEAN SYSTEM

The TEP model, uncertainty and variability management methods and the solution strategy developed in this thesis have been tested on a reduced European transmission system. The analysis of the test results is presented as follows.

6.2.1. Data Preparation and Assumptions

In order to run a TEP on a continental scale, a great deal of data is required. For instance, hourly series of demand and generator output for each technology should be available for each node. In addition, network parameters (including transfer capacity and electrical parameters) of both existing and candidate lines should be known. However, most of this information is not publicly available for obvious reasons. We explain in the following subsections how we have extracted the information and data needed for the case study from various sources, and the corresponding assumptions that we have made to complement some missing information.

6.2.1.1. Base-case Network

Electricity network is a backbone for any TEP optimization process. Apparently, the required information about the existing European network is not readily available. For this reason, we relied on Enipedia database (which is developed by TUDelft, accessible

via http://enipedia.tudelft.nl) to generate a European network model, used here in the case study. The database contains plenty of information, yet incomplete when it comes to electricity networks (especially in countries of the Northern and Eastern Europe, where we had to almost generate the networks from scratch). Also, we have observed that a lot of details are missing especially at lower resolutions. In other words, those lines which carry power over relatively longer distances seem to be sufficiently available, but some lines that connect local electricity demand or generator are largely missing in the database. Because of this, we have decided to aggregate the demand and generation capacities by the smallest socio-economic regions of Europe (officially referred to as NUTS-3). And, we have developed a network model considering only the interconnections among these regions (with voltage levels higher or equal to 220 and 150 kV for AC and DC lines, respectively). The network connections within a NUTS-3 region are simply disregarded regardless of the type or voltage level; instead, represented by a single node located at the geographical center of the considered region.

Fig. 6. 1 Network model aggregated by NUTS-3 regions

The network model, developed based on the data extracted from Enipedia database, is not yet complete because a large number of important links are missing (especially in the Balkan and Nordic countries). The missing network links are recovered by visually inspecting the network extracted from Enipedia database, and painstakingly comparing it with the ENTSO-E's paper map of European networks.

Once we know the number and the type of lines interconnecting the NUTS-3 regions, we represent those lines with a corresponding equivalent line, whose transfer capacity and electrical parameters are approximated as follows. First, standard values of transmission line parameters [241] are used for each type and voltage level. In some cases, lines connecting two specific areas can be of the same type and voltage level but their lengths can be generally different. Therefore, the standard values need to be readjusted to account for the effect of distance. For instance, the transfer capacity of a line gets substantially lower when the distance increases. On the other hand, the impedance of a line increases with distance. With this in mind, the maximum transfer capacity of each of the lines connecting two areas is determined, the sum of which gives the total (maximum) transfer capacity between the two given areas. This can be understood as the transfer capacity of an artificial line connecting the two areas. However, because of N-1 security criterion, the actual transfer capacity is often far less than the arithmetic sum. As a proxy to this criterion, we deduct the maximum transfer capacity of a line among those connecting the two areas, and obtain the effective (net) transfer capacity between those areas. And, the corresponding electrical parameters (resistance and reactance, in particular) are determined by fitting a curve with known transfer capacities and electrical parameters. Figure 3 shows the final European network model developed this way.

6.2.1.2. *Generation Capacity by Technology*

The Enipedia database contains a huge list of generators of different technologies associated with their geographical coordinates and relevant tags such as generation capacity (in MW), annual MWh-production and emission intensity among others. Unfortunately, the database is not complete. Only a fraction of the generators have the generation capacity information, prompting us to devise other ways to recover the required data. For instance, the technology type (if missing in the database) of a generator is identified by its emission intensity because each technology has a comparatively unique carbon footprint. In addition, we have mapped the annual production values to capacities by using regression models to recover the generation capacities of the generators. The regression models (see Figure 6.2, for example) are technology-specific, and in some cases, are even different for the same technology situated in different countries. They are obtained from already known quantities.

$$MW = 0.0001 * MWh + 83.379$$
$$R^2 = 0.9502$$
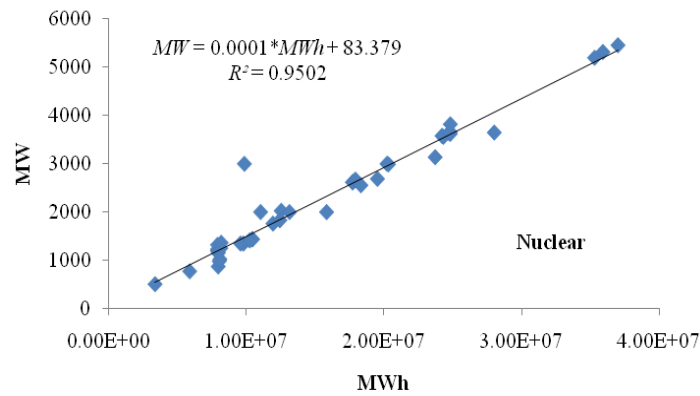
Fig. 6. 2 An example of a linear relationship between MWh-production and MW-generation capacity in nuclear technology.

When the generation capacities of all generators is determined, obtaining the total generation capacity corresponding to each technology in each NUTS-3 areas (nodes) is straightforward. Since we know the location of each generator, we add the capacities of the generators of the same technology which are located within the same region. This way, we get the total generation capacity for each technology and region.

6.2.1.3.   Electricity Demand

The total demand per country is available in the ENTSO-E website. We redistribute this aggregate demand among all nodes in the country in proportion to their respective population sizes. For instance, suppose country X has a total electricity demand of 100 MW and four NUTS-3 regions, with its population distributed across the regions in the following proportion {40%, 30%, 20%, 10%}. For this country, the corresponding electricity demand consumed by the population in each region would be {40, 30, 20, 10} MW, respectively.

ENTSO-E also regularly publishes records of hourly electricity demand aggregated at country level. We use this information in order to generate the demand series at all the nodes in each country. This is needed because electricity consumption varies with geographical locations and weather patterns. For instance, geographically dispersed demand regimes, particularly those in different time horizons, are likely to be less correlated. Therefore, spatial demand correlations ranging from 0.9 to 1 are factored in to account for such spatial variations of electricity demand within each country. This can be achieved by generating different time-lagged demand series or using Cholesky factorization to create different demand series with a given correlation matrix, whose entries depend on the distance among the nodes.

Electricity demand is assumed to grow by 1% annually, and this is kept the same for all scenarios. Accordingly, the demand growth at the end of first and second stages is 10% and 30%, respectively.

6.2.1.4.   Time-Series of Wind and Solar Power Sources

135

Solar irradiance and wind speed data are the most useful components in determining wind and solar power outputs. Because of this, historical records of solar irradiances with different time resolutions are collected from multiple sources such as http://project.mesor.net/web/guest/solemi-free and http://www.soda-is.com/eng/index.html. Similarly, wind speed data series have been collected from various sources. Majority of the meteorological websites in Europe have historical records of wind speed, spanning over several years, publicly available. Some missing information is complemented from the daily data provided by the European Climate Assessment and Dataset (ECA&D), available online on http://eca.knmi.nl/. In some cases, wind speed series for different years are used because the inter-annual wind speed variations are often very small (less than six percent of the mean [242]).

It should be noted here that whenever wind speed or solar irradiance information is not available for a specific place or country, the corresponding series are either generated from approximate probability distributions (given that the corresponding mean values are known) or simply assumed to be the same as that of neighboring nodes, where this information is already known. This has been the case for some nodes in the Balkan and Baltic countries.

Once the hourly series of wind speed and solar irradiance are known for each node in the test system, the corresponding power outputs are determined by plugging in these values in the respective power output expressions in [243] and [244].

6.2.1.5. Time-Series of Conventional Power Sources

The conventional power sources considered here are nuclear, gas- and coal-fired power plants. In order to generate the time-series for these technologies, we use a two–state model (online or offline) to represent the state of conventional power units based on their respective forced outage rates (FOR), which range from 0.05 to 0.15 depending on the type of generator. This way, a discrete random binomial distribution is employed to generate availability patterns for different generators, obtained from their corresponding forced outage rates.

6.2.1.6. Other Power Sources

The time series of hydro power plants are generated based on the assumption that hydro power outputs are closely related with rainfall pattern (which can be found in national meteorological sites). In this way, the highest power output from hydropower plants is assumed to occur at the same time with the highest rainfall, and for lower rainfalls, production is reduced proportionally and kept at its minimum during dry seasons (i.e. when there is no rainfall).

Power plants that generate electricity from municipal solid waste, biomass and geothermal at each node (if any) are assumed to be available year-round.

6.2.2. Scenario Definitions

The process of defining scenarios is in itself a very challenging task. A large number of scenarios are often required to fully explore the plausible future states. For the sake of simplicity, we only use three scenarios in the EU-1060 node system. These scenarios are characterized by large-scale power production from either wind in the Northern Europe, where a large portion comes from the North Sea area, distributed renewable energy (mainly, wind and solar), or solar resources in the MENA and Southern Europe. From now on, we refer the above scenarios as North-Wind, Distributed-RES, and South-Solar, respectively.

6.2.2.1. North-Wind Scenario

This scenario can be considered as a pro-wind scenario. By the end of the first stage, a 35 GW of wind power is assumed to be generated from the North Sea, West Coast and Baltic Sea areas. This amount is injected at 88 strongly connected nodes bordering these areas. It is distributed among the nodes in proportion to their corresponding average wind speeds. Another 15 GW of wind power is distributed among all the nodes in the system proportionally to the primary energy source (i.e. average wind speed) and total area suitable for wind turbine installations in each region. Hence, the total wind capacity in the first stage is 50 GW.

In the second stage, wind power with a total capacity of 200 GW is expected to be installed, 140 GW of which comes from the North Sea, West Coast and Baltic Sea areas. As in the first stage, this is assumed to be injected at the same nodes (i.e. the 88 strongly connected nodes bordering these areas), distributed among the nodes in the same manner. The remaining 60 GW balance is distributed among all the nodes in the system in proportion to the average wind speed and total area suitable for wind turbine installations in each region.

6.2.2.2. Distributed-RES Scenario

The amount of generation capacities is assumed to be added in the first and the second stages is the same as in the North-Wind scenario. However, in this scenario, large-scale wind or solar installations is limited; instead, distributed generation of wind and solar power is favored. It is assumed that 30% (i.e. 15 GW and 60 GW in the first and the second stages, respectively) of the total power comes from a total of 53 sunny and 125 windy regions identified across Europe, as in Figure 6.3. Equal amount of the 30% power is generated the wind and the solar sources. Again, distribution of the installed wind or solar among each set of nodes is made in proportion to the primary energy resource (either wind speed or solar radiation), and suitable areas for wind turbine or solar PV installations. The remaining balance (i.e. 70%) is redistributed among all nodes proportionally to the existing installed wind or solar power capacity at each node.

Fig. 6. 3 Hotspots for distributed solar (orange circle) and wind (blue circle) generation

6.2.2.3. *South-Solar Scenario*

This scenario is mainly characterized by large-scale solar power imports from MENA. The total amount of installed capacity is the same as in the above two scenarios (i.e. 50 and 200 GW in the first and the second stages), of which 70% is to be imported from MENA via 10 nodes, selected based on proximity and connectivity strength criteria. The remaining balance is redistributed among all nodes according to existing installed solar power at each node.

6.2.3. Candidate Lines For Expansion

The candidate selection involves selecting corridors to be possibly reinforced, technology and its cost structure.

6.2.3.1. *Identifying Corridors to Be Reinforced*

138

It is a daunting challenge to create an initial candidate pool for expansion at continental or inter-continental level. This is because, on one hand, from the transmission investment perspective, there are several technological options, which make the problem even more difficult. On the other hand, there is generally a lack of expert knowledge on the set of potential corridors to be investigated for future investments in such a big system. Because of this, some heuristic candidate selection methods have been used such as the copper sheet method in [237] and methods in [237] and [239] which makes use of marginal prices as economic indicators for the selection process of candidates. An extended and automatic version of the marginal prices –based candidate selection methods in [237] and [239] is reported in [238]. Yet, it is likely that such methods eventually end up with a huge list, because, in such a big system, there can be many possibilities which satisfy the conditions for the selection. In any case, a sufficiently large set of candidates (potentially encompassing some existing and new corridors) are required for TEP optimization. For the sake of simplicity, we consider each corridor as a candidate for reinforcement, resulting in a total of 1654 candidate lines in existing corridors (which comprise AC and DC connections). In addition, ten new HVDC submarine connections (in new corridors) are included in the initial candidate list, bringing the total number in the initial candidate list to 1664.

6.2.3.2. *Cable-Overhead Proportion, Selecting Technology and Construction Cost*

Recent study shows that underground cables with 315 kV higher voltage levels constitute less than 5% of the total circuit length in Europe [245]. However, there is a general consensus that this will significantly change in the future (mainly caused by the lack of right of ways for overhead lines and increasing urbanization). Because of this, we assume that one fifth of a given line being added to the network in the first stage will be underground, and a 50% is assumed by the end of the second stage. The total installation cost of a line is calculated by taking these assumptions into account. It should be noted here that these assumptions do not take effect on undersea power transmissions, where it is assumed that only HVDC cables are the only viable options (particularly, for distances higher than 50 km).

Nowadays, there are a number of proven transmission technology options. In this regard, selecting the most economically viable option is in itself a separate problem. In this work, we only focus on selected technologies. Of a special importance here is the DC technology. The share of DC connections (in terms of length) in modern power system networks is very small. This will, however, significantly change in the future because DC technology alleviates some of the technical limitations of AC lines [241]. For instance, it is generally accepted that HVDC technology is more attractive for bulk power transmission over longer distances than HVAC one. Concerning this, different references report different break-even distance ranges between HVDC and HVAC connections, mostly beyond 100 km. But it also depends on the amount of power to be transmitted. It is usually about 50 km for submarine cables and 400 km (in some cases, as low as 200 km) for overhead lines. We have used the cost structure of different transmission technologies given in [241] for our calculations and analyses.

To select an appropriate transmission technology in a given corridor, we have set a criteria based on the value of an investment decision variable after the first phase of the solution process, size of excess flows in that corridor (i.e. higher than the maximum capacity), and its length. This is summarized in Table 6.1 below.

Table 6. 1 Transmission technology selection

	Investment decision variable after PHASE I	Extra flows in a corridor after PHASE I (MW)	Distance of corridor (km)	Transmission technology
1	≥ 4	≥ 3000	≥ 250	$\pm 600kV$ HVDC bipolar
2	≥ 1 and ≤ 4	≤ 3000 and ≥ 2000	≥ 250	$\pm 500kV$ HVDC bipolar
3	≥ 1	< 2000 and ≥ 1000	< 250 and ≥ 150	$500kV$ HVAC double circuit
4	≥ 1	< 1000 and > 500	< 150	$400kV$ HVAC double circuit
5	≥ 1	≥ 3000	< 150	$400kV$ HVAC Up to 4 circuits
5	≥ 1	≤ 500	< 150	$400kV$ HVAC double circuit

6.3. OPTIMIZATION RESULTS AND DISCUSSION

It should be noted that the optimization is carried out by a computing machine with Intel Xeon E5520 at 2.27 GHz frequency and with 32 GB RAM memory. First, using the moment-based technique clustering technique in [15], 8760 operation states (hourly snapshots) are reduced into 60 representative snapshots. Accordingly, we have obtained 60 representative snapshots for each spot in the scenario tree shown in Figure 6.1.

With these snapshots, the proposed solution methodology is tested on this system. The successive optimization process described earlier, is run starting with 1654 elements in the ICL. Consequently, in the first phase, which involves fully deterministic optimization based on the hybrid network model with continuous investment variables, the number of candidate lines is reduced from 1654 to 687. Moreover, this is further reduced to 640 using MODEL II of the second phase. With this reduced set of candidates and the ten proposed lines across new corridors as inputs (i.e. a total of 650 candidate lines), the two-stage stochastic discrete optimization (i.e. MODEL III) is run and the final investments are obtained. It should be noted here that the same number of candidates are used in both stages. The solution time (i.e. the total CPU time that the whole optimization process took) was about 25 hours, which is rather small for such a complex problem. An attempt to run the Brute-force TEP optimization problem (i.e.

without the implementation of the search space reduction methodology) was not successful due to computational limitation.

Regarding the resulting investment decisions, a total number of 331 are built in the first stage, approximately 11% of which are HVDC lines. This is shown in Figure 6.5. As mentioned earlier, these investments are considered to be good enough for all three scenarios. It is interesting to see that most of the biggest investments in the first stage are made across the borders of the European countries, where the main bottlenecks exist.

Fig. 6. 4 First stage expansion results (shown in bold)

Fig. 6. 5 Second stage expansions in North-Wind scenario (shown in bold)

Likewise, the investments corresponding to each scenario in the second stage are shown in Figures 6.5—6.7. The number of investments is 431, 349 and 423 in North-Wind, Distributed-RES and South-Solar scenarios, respectively. As expected, the number in each scenario here is a lot higher than in the first stage, especially in the North-Wind and South-Solar scenario. This is rather expected because large-scale renewable development prospects inevitably require huge network investments. On the other hand, we can observe that, with distributed generation, line investment requirements are much lower than in the case of highly dispersed large-scale renewable generations.

Fig. 6. 6 Second stage decisions in Distributed-RES scenario (shown in bold)

Fig. 6. 7 Second stage decisions in South-Solar scenario (in bold)

Another interesting observation here that the scenarios have a lot of investment decisions in common. The North-Wind and South-Solar scenarios, in particular, seem to use the same corridors in the central Europe for transporting power either south or north direction, respectively. This is contrary to the perception that different scenarios result completely different investment strategies. In fact, there are some differences in the investment decisions of both scenarios. Especially in the Southern and Northern Europe, where the power for each comes from, there seems to be shift in investments from north to south or vice versa.

Worth mentioning here is the substantial reduction in computation time. The reduction here is equivalent to a reduction of the combinatorial solution search space from 21654 to 2640. This is indeed significant from computational scale point of view.

6.4. SUMMARY

Transmission expansion planning at continental level is a very dimensionally huge and mathematically complex combinatorial problem which makes it difficult to solve by currently available computational machines. Obtaining optimal expansion solutions within a reasonable computation time is vital. To enhance tractability, we have

employed a solution strategy which effectively reduces the combinatorial solution search space, as a result, leading to a faster computation without significantly compromising the optimality of the solution. It is based on simple yet effective heuristic solution method which works by decomposing the complex problem into successive phases and making use of parallel implementation. It employs relatively fast optimization models whose formulations are based on the hybrid network model in order to refine a huge initial candidate list before switching on to a more accurate optimization model based on the DC network model. The results of the case study show the effectiveness of the proposed solution strategy in considerably reducing the computational complexity.

7

VII. CONTRIBUTIONS, CONCLUSIONS AND FUTURE WORK

This final chapter concludes the research carried out in this dissertation by summarizing its main contributions point-by-point, and drawing some conclusions from the case studies carried out throughout this thesis. In addition, this thesis points out the limitations of the developed approaches and suggests possible extensions as future work.

7.1. MAIN CONTRIBUTIONS

The contributions of this thesis include methodological and modeling aspects of the TEP problem. The main contributions are briefly summarized as bullet points below. It should however be noted that this summary does not include the contributions associated with the improvements and the modifications made to existing mathematical modeling techniques of TEP formulations. These are clearly stated in the body of this thesis.

- From a modeling perspective,
  - A new TEP model has been proposed for a long-term planning of transmission infrastructures under uncertainty with a multi-stage decision framework and considering a high level renewable integration. One of the salient features of the developed TEP model is its ability to capture the long-term impact of network investments on system costs. This has been partly published in [15].
  - Recognizing the significant impacts network losses have on TEP solutions (which are often neglected in most TEP studies because of computational limitations), new linear losses models have been proposed, some of which strike the right balance between accuracy and computational effort, particularly, in the context of medium to long-term TEP in large-scale power systems accommodating high level variable energy sources. An extensive analysis on this issue has been published in [13].
- From a methodological perspective,
  - A new clustering methodology is introduced to effectively and efficiently handle uncertainty and variability pertaining to the problem at hand. This contribution has been published in [14].
  - The entire TEP problem is formulated as a stochastic mixed-integer linear programming optimization, an exact solution method, for which efficient solvers are available and an optimal solution is guaranteed in a finite simulation time.
  - In order to significantly reduce the combinatorial solution search space and hence facilitate the computation, a new heuristic solution strategy has been devised. This approach works by primarily decomposing the problem into successive optimization phases.
- The extensive experimental and theoretical analysis made throughout the thesis.

7.2. CONCLUSIONS

147

A long-term expansion planning of large-scale transmission grids under high level renewable integrations has unprecedentedly huge uncertainty, temporal and geographical scope as well as the network size. Framed in this context, the main objective of this research has been to develop mathematical optimization models, uncertainty and variability management methods, and solution strategies that support the complex decision-making process of such a problem. To this end, first, a new TEP model has been developed which has the following salient features:

- It is a MILP optimization model, based on an improved "DC" network model, for which efficient off-the-shelf solvers are available and optimality is guaranteed within finite simulation time.
- It captures the uncertainty and/or variability of various uncertain parameters inherent to a long-term TEP problem with renewable generation via stochastic programming.
- It has a weighted sum of relevant costs such as emission, operation and maintenance, reliability and investment costs as its main objective.
- It provides a realistic measure of all cost terms during and after the planning horizon so that a proper comparison of the different costs is estabilished.
- Its formulation is based on a two-period planning framework which helps to combine/determine short- to medium-term decisions and long-term (strategic, adaptive) expansion decisions.

Second, a new methodology has been proposed in order to effectively manage the uncertainty and variability introduced by different uncertain parameters such as RES output and demand. A significant part of this uncertainty and variability is handled by a sufficiently large set of operational snapshots, which can be understood as generation-demand patterns of power systems that lead to OPF patterns in the transmission network. A large set of snapshots, each one with an estimated probability, is then used to evaluate and optimize the network expansion. In long-term TEP of large networks, the number of operational states must be reduced. Hence, the proposed methodology reduces these snapshots by means of clustering, without relevant loss of accuracy from the TEP solution perspective, by selecting classification variables is used in the clustering process. The proposed method relies on the following main ideas:

- The snapshots are first characterized by their OPF patterns (the effects) instead of the generation-demand patterns (the causes). This is simply because the network expansion is the target problem, and losses and congestions are the drivers to network investments.
- The OPF patterns are then classified using a "moments" technique, a well-known approach in Optical Pattern Recognition problems.

Third, to address the combinatorial nature of such a problem, an effective solution strategy has been proposed. This solution method works by decomposing the original

problem into successive optimization phases, which use TEP models with increasing fidelity levels. The proposed strategy dramatically reduces the combinatorial solution search space, which has a considerable influence on the solution process.

The developed models, methods and solution strategies has been tested on small-, medium- and large-scale network systems. In addition, to further validate the proposed TEP model, methods and solution strategies, an aggregated 1060-node European network system has been employed as a case study considering multiple RES development scenarios. Generally, numerical results show the versatility of the proposed TEP model. Moreover, the proposed methods and solution strategy are very effective in facilitating the solution process, and result in a significant reduction in computational effort while fairly maintaining optimality of the expansion solutions.

7.3. DIRECTIONS FOR FUTURE WORKS

The research work presented in this thesis has certain limitations, most of which can be translated into future directions of research. The methods, models and strategies developed in this thesis can be further extended or improved to support future works. Some of the shortcomings are listed below.

Transmission Technology Selection: There are several matured transmission technologies each having different physical and economic characterstics. Further technological advances in R&D will further add new transmission technologies that are expected to mature in due time. In addition, transmission lines have very long economic lifetimes, and the TEP problem is characterized by strong economies of scale. From the AC context, for instance, the cost per MW per km decreases with increasing voltage level. This, in the current work, is handled by associating a binary variable for each technology. However, this considerably increases the complexity of the problem, by implication increasing the computational requirement. From this perspective, devising a methodology (possibly heuristic) that can effectively determine which transmission technology to consider for investment and which to discard, at the same time reflecting the uncertainty in the maturity level of the transmission technology.

TEP from the Smart-grid Context: The present work focuses mainly on the development of models, methods and tools to handle wide-area and long-term grid expansion planning under high penetration level of variable energy sources. The effect of smart-grid technolgies large-scale deferrable loads, demand side management, energy storage technologies (centralized and/or distributed) and others on the network investment needs as well as on the system is not analyze. These technologies, along with substantial network expansions, are expected to be deployed in the system to support large-scale integration of variable energy sources, minimize the impact of high level variability and unpredictability of such energy sources, maintain system integrity, stability and power quality. Hence, this line of research can be very interesting for future works.

Integrated approach: The coordination of different sectors of energy infrastructure expansion and developments is becoming increasingly imprortant. Because of this, developing a multi-sectoral optimization problem is of paramount importance. It would be interesting to analyze this from the perspectives of coordinating different forms of energy consumption, improving overall system efficiency, enhancing energy security, optimally integrating and exploiting RESs, reducing GHGs, etc.

APPENDIX

APPENDIX A: DERIVATION OF THE FLOW-BASED ACTIVE AND REACTIVE POWER LOSSES

The derivations related to the losses equations in (40) and (41) are provided here. Squaring both sides of the flow equations in (23) and (24) and dividing each by V_{nom}^2, we get:

$$\frac{(P_k)^2}{V_{nom}^2} \approx \underbrace{[(\Delta V_i - \Delta V_j)g_k]^2}_{I} - \underbrace{2 * g_k V_{nom} b_k \theta_k * (\Delta V_i - \Delta V_j)}_{II} + (V_{nom} b_k \theta_k)^2 \quad \text{(A.1)}$$

$$\frac{(Q_k)^2}{V_{nom}^2} \approx \underbrace{[(\Delta V_i - \Delta V_j)b_k]^2}_{I} + \underbrace{2 * b_k V_{nom} g_k \theta_k * (\Delta V_i - \Delta V_j)}_{II} + (V_{nom} g_k \theta_k)^2 \quad \text{(A.2)}$$

Since the variables θ_k, ΔV_i and ΔV_j are very small, the second order terms (i.e. bilinear products of these variables) can be regarded to be close to zero. Hence, the first and the second terms in (A.1) and (A.2) can be neglected, leading to the following expressions, respectively.

$$\frac{(P_k)^2}{V_{nom}^2} \approx (V_{nom} b_k \theta_k)^2 \quad \text{(A.3)}$$

$$\frac{(Q_k)^2}{V_{nom}^2} \approx (V_{nom} g_k \theta_k)^2 \quad \text{(A.4)}$$

Multiplying both sides of (A.3) and (A.4) by r_k and adding both sides gives:

$$r_k \left(\frac{P_k}{V_{nom}}\right)^2 + r_k \left(\frac{Q_k}{V_{nom}}\right)^2 \approx r_k (V_{nom} b_k \theta_k)^2 + r_k (V_{nom} g_k \theta_k)^2 \quad \text{(A.5)}$$

After rearranging Equation (A.5), we get:

$$r_k (P_k^2 + Q_k^2)/V_{nom}^2 \approx g_k (V_{nom} \theta_k)^2 r_k \left(\frac{(b_k)^2}{g_k} + g_k\right) \quad \text{(A.6)}$$

One can easily verify that $r_k \left(\frac{(b_k)^2}{g_k} + g_k\right) = 1$, reducing Equation (A.6) to:

$$r_k (P_k^2 + Q_k^2)/V_{nom}^2 \approx g_k (V_{nom} \theta_k)^2 \quad \text{(A.7)}$$

Recall that the right hand side of (A.7) corresponds to the active power losses expression in (40), which proves the derivation. The flow-based reactive power losses in (41) are derived in a similar way. Multiplying both sides of Equations (A.3) and (A.4) by x_k instead of r_k, adding both sides and rearranging the resulting equation leads to:

$$r_k (P_k^2 + Q_k^2)/V_{nom}^2 \approx -b_k V_{nom}^2 \theta_k^2 x_k [-b_k + (g_k)^2/(-b_k)] \quad \text{(A.8)}$$

Note that, in Equation (A.8), $x_k [-b_k + (g_k)^2/(-b_k)] = 1$. Hence, the equation reduces to:

$$r_k(P_k^2 + Q_k^2)/V_{nom}^2 \approx -b_k V_{nom}^2 \theta_k^2 \tag{A.9}$$

And, observe that the right hand side of Equation (A.9) is equal to the reactive losses expression in (41).

For quick reference, this section presents compact forms of the lossy TEP models described in Chapter 3. The objective function in Equation (B.1) is common for all models.

$$
\begin{aligned}
\min_{z_{k,t},PG_{g,b,t},p_{i,b,t}} \quad & Z \\
= & \sum_t \sum_k (1+r)^{-t} \frac{r(1+r)^{LT_k}}{(1+i)^{LT_k}-1} z_{k,t} IC_k/r \\
& + \underbrace{\sum_t \sum_g \sum_b (1+r)^{-t}\Delta_b PG_{g,b,t}\lambda_g}_{I} \\
& + \underbrace{\sum_g \sum_b (1+r)^{-1}\Delta_b PG_{g,b,t}\lambda_g/r}_{II} \\
& + \underbrace{\sum_t \sum_g \sum_b (1+r)^{-t}\Delta_b p_{i,b,t}\Lambda}_{I} \\
& + \underbrace{\sum_t \sum_g \sum_b (1+r)^{-1}\Delta_b p_{i,b,t}\Lambda/r}_{II}
\end{aligned}
$$

(B.1)

$$
\begin{aligned}
& V_{i,b} = V_{nom} + \Delta V_{i,b}, \quad \text{where } \Delta V^{min} \leq \Delta V_{i,b} \leq \Delta V^{max} \\
& \left| P_{k,b} - \{V_{nom}(\Delta V_{i,b} - \Delta V_{j,b})g_k - V_{nom}^2 b_k\theta_{k,b})\} \right| \leq MP_k(1-u_k) \\
& \left| Q_{k,b} - \{-V_{nom}(\Delta V_{i,b} - \Delta V_{j,b})b_k - V_{nom}^2 g_k\theta_{k,b}\} \right| \leq MQ_k(1-u_k) \\
& \left| P_{k,b} - \{V_{nom}(\Delta V_{i,b} - \Delta V_{j,b})g_k - V_{nom}^2 b_k\theta_{k,b}\} \right| \leq MP_k(1-z_k) \\
& \left| Q_{k,b} - \{-V_{nom}(\Delta V_{i,b} - \Delta V_{j,b})b_k - V_{nom}^2 g_k\theta_{k,b}\} \right| \leq MQ_k(1-z_k) \\
& P_{k,b}^2 + Q_{k,b}^2 \leq u_k(S_{k,max})^2 \\
& P_{k,b}^2 + Q_{k,b}^2 \leq z_k(S_{k,max})^2 \\
& PL_{k,b} = r_k\{P_{k,b}^2 + Q_{k,b}^2\}/V_{nom}^2 \qquad \textit{LinACTEP} \\
& QL_{k,b} = x_k\{P_{k,b}^2 + Q_{k,b}^2\}/V_{nom}^2 \\
& P_{k,b} = P_{k,b}^+ - P_{k,b}^- \rightarrow |P_{k,b}| = P_{k,b}^+ + P_{k,b}^- \\
& Q_{k,b} = Q_{k,b}^+ - Q_{k,b}^- \rightarrow |Q_{k,b}| = Q_{k,b}^+ + Q_{k,b}^- \\
& P_{k,b}^2 = \sum_{l=1}^{L} \alpha_{k,b,l}\Delta p_{k,b,l} \\
& P_{k,b}^+ + P_{k,b}^- = \sum_{l=1}^{L} \Delta p_{k,b,l}
\end{aligned}
$$

$$Q_{k,b}^2 = \sum_{l=1}^{L} \beta_{k,b,l} \Delta q_{k,b,l}$$

$$Q_{k,b}^+ + Q_{k,b}^- = \sum_{l=1}^{L} \Delta q_{k,b,l}$$

$$\Delta p_{k,b,l} \leq P_k^{max}/L \; ; \; \Delta q_{k,b,l} \leq Q_k^{max}/L$$

$$\Delta p_{k,b,l+1} \leq \Delta p_{k,b,l} \; ; \; \Delta q_{k,b,l+1} \leq \Delta q_{k,b,l}$$

$$\sum_{k \in i} P_{k,b} + \sum_{g \in i} PG_{g,b} + p_{i,b} - \sum_{d \in i} PD_{d,b} + 0.5 \sum_{k \in i} PL_{k,b} = 0$$

$$\sum_{k \in i} Q_{k,b} + \sum_{g \in i} QG_{g,b} + q_{i,b} - \sum_{d \in i} QD_{d,b} + 0.5 \sum_{k \in i} QL_{k,b} = 0$$

$$u_g PG_{g,min} \leq PG_{g,b} \leq u_g PG_{g,max}$$

$$u_g QG_{g,min} \leq QG_{g,b} \leq u_g QG_{g,max}$$

$$\theta_{min} \leq \theta_{i,b} \leq \theta_{max}$$

$$V_{ref} = V_{nom} \; ; \; \theta_{ref} = 0$$

$$\alpha_{k,b,l} = \beta_{k,b,l} = (2l-1) S_k^{max}/L$$

$$\left| P_{k,b} + V_{nom}^2 b_k u_k \theta_{k,b} \right| \leq M_k(1 - u_k); \; where \; -b_k = 1/x_k$$

$$\left| P_{k,b} + V_{nom}^2 b_k z_k \theta_{k,b} \right| \leq M_k(1 - z_k) \; ; \; where \; -b_k = 1/x_k$$

$$\left| P_{k,b} \right| + 0.5 PL_{k,b} \leq u_k S_{k,\max}$$

$$\left| P_{k,b} \right| + 0.5 PL_{k,b} \leq z_k S_{k,\max}$$

$$\sum_{k \in i} P_{k,b} + \sum_{g \in i} PG_{g,b} + p_{i,b} - \sum_{d \in i} PD_{d,b} + 0.5 \sum_{k \in i} PL_{k,b} = 0$$

$$u_g PG_{g,min} \leq PG_{g,b} \leq u_g PG_{g,max}$$

$$\theta_{min} \leq \theta_{i,b} \leq \theta_{max}; \; \theta_{ref} = 0 \qquad\qquad \textit{DCTEP}$$

$$PL_{k,b} = r_k P_{k,b}^2/V_{nom}^2$$

$$P_{k,b} = P_{k,b}^+ - P_{k,b}^- \rightarrow \left| P_{k,b} \right| = P_{k,b}^+ + P_{k,b}^-$$

$$P_{k,b}^2 = \sum_{l=1}^{L} \alpha_{k,b,l} \Delta p_{k,b,l}$$

$$P_{k,b}^+ + P_{k,b}^- = \sum_{l=1}^{L} \Delta p_{k,b,l}$$

$$\Delta p_{k,b,l} \leq S_k^{max}/L \; ; \; \Delta p_{k,b,l+1} \leq \Delta p_{k,b,l} \; ; \; \alpha_{k,b,l} = (2l-1) S_k^{max}/L$$

$$V_{i,b} = V_{nom} + \Delta V_{i,b}, \qquad where \; \Delta V^{min} \leq \Delta V_{i,b} \leq \Delta V^{max}$$

$$\left| P_{k,b} - \{ V_{nom}(\Delta V_{i,b} - \Delta V_{j,b}) g_k - V_{nom}^2 b_k \theta_{k,b}) \} \right| \leq MP_k(1 - u_k) \qquad \textit{M-DCTEP}$$

$$\left|P_{k,b} - \{V_{nom}(\Delta V_{i,b} - \Delta V_{j,b})g_k - V_{nom}^2 b_k \theta_{k,b}\}\right| \le MP_k(1 - z_k)$$

$$\left|P_{k,b}\right| + 0.5PL_{k,b} \le u_k S_{k,\max}$$

$$\left|P_{k,b}\right| + 0.5PL_{k,b} \le z_k S_{k,\max}$$

$$\sum_{k \in i} P_{k,b} + \sum_{g \in i} PG_{g,b} + p_{i,b} - \sum_{d \in i} PD_{d,b} + 0.5 \sum_{k \in i} PL_{k,b} = 0$$

$$u_g PG_{g,min} \le PG_{g,b} \le u_g PG_{g,max}$$

$$\theta_{min} \le \theta_{i,b} \le \theta_{max}; \ \theta_{ref} = 0$$

$$PL_{k,b} = r_k P_{k,b}^2 / V_{nom}^2$$

$$P_{k,b} = P_{k,b}^+ - P_{k,b}^- \rightarrow \left|P_{k,b}\right| = P_{k,b}^+ + P_{k,b}^-$$

$$P_{k,b}^2 = \sum_{l=1}^{L} \alpha_{k,b,l} \Delta p_{k,b,l}$$

$$P_{k,b}^+ + P_{k,b}^- = \sum_{l=1}^{L} \Delta p_{k,b,l}$$

$$\Delta p_{k,b,l} \le S_k^{max}/L \ ; \ \ \Delta p_{k,b,l+1} \le \Delta p_{k,b,l} \ ; \ \ \alpha_{k,b,l} = (2l - 1) S_k^{max}/L$$

$$\left|P_{k,b} + V_{nom}^2 b_k u_k \theta_{k,b}\right| \le M_k(1 - u_k); \ where - b_k = 1/x_k$$

$$P_k = -V_{nom}^2 b_k \left(\phi_{1,k,b}^2 - \phi_{2,k,b}^2\right); \ where - b_k = \frac{1}{x_k}$$

$$\phi_{1,k,b} = \frac{z_k + \theta_{k,b}}{2} \quad ; \quad \phi_{2,k,b} = \frac{z_k - \theta_{k,b}}{2}$$

$$0 \le z_k \le z_{k,\max}$$

$$\frac{z_{k,\min} + \theta_{k,min}}{2} \le \phi_{1,k,b} \le \frac{z_{k,\max} + \theta_{k,max}}{2}$$

$$\frac{z_{k,\min} - \theta_{k,max}}{2} \le \phi_{2,k,b} \le \frac{z_{k,\max} - \theta_{k,min}}{2}$$

$$\left|P_{k,b}\right| + 0.5PL_{k,b} \le u_k S_{k,\max}$$

$$\left|P_{k,b}\right| + 0.5PL_{k,b} \le z_k S_{k,\max}$$

$$PL_{k,b} = r_k P_{k,b}^2 / V_{nom}^2$$

$$P_{k,b} = P_{k,b}^+ - P_{k,b}^- \rightarrow \left|P_{k,b}\right| = P_{k,b}^+ + P_{k,b}^-$$

$$P_{k,b}^2 = \sum_{l=1}^{L} \alpha_{k,b,l} \Delta p_{k,b,l}$$

$$P_{k,b}^+ + P_{k,b}^- = \sum_{l=1}^{L} \Delta p_{k,b,l}$$

R-DCTEP

$$\Delta p_{k,b,l} \le P_k^{max}/L \;\; ; \;\; \Delta p_{k,b,l+1} \le \Delta p_{k,b,l}$$

$$\sum_{k \in i} P_{k,b} + \sum_{g \in i} PG_{g,b} + p_{i,b} - \sum_{d \in i} PD_{d,b} + 0.5 \sum_{k \in i} PL_{k,b} = 0$$

$$u_g PG_{g,min} \le PG_{g,b} \le u_g PG_{g,max}$$

$$\theta_{min} \le \theta_{i,b} \le \theta_{max}; \; \theta_{ref} = 0$$

$$\phi_{1,k,b} = \frac{z_k + \theta_{k,b}}{2} \;\; ; \;\; \phi_{2,k,b} = \frac{z_k - \theta_{k,b}}{2}$$

$$\phi_{1,k,b} = \phi_{1k,b}^+ - \phi_{1,k,b}^- \rightarrow |\phi_{1,k,b}| = \phi_{1,k,b}^+ + \phi_{1,k,b}^-$$

$$\phi_{1,k,b}^2 = \sum_{l=1}^{L} \alpha_{k,b,l} \Delta\phi_{1,k,b,l}$$

$$\phi_{1,k,b}^+ + \phi_{1,k,b}^- = \sum_{l=1}^{L} \Delta\phi_{1,k,b,l}$$

$$\Delta\phi_{1,k,b,l} \le (z_k^{max} + \theta_k^{max})/(2L) \;\; ; \;\; \Delta\phi_{1,k,b,l} \ge \Delta\phi_{1,k,b,l+1}$$

$$\phi_{2,k,b} = \phi_{2k,b}^+ - \phi_{2,k,b}^- \rightarrow |\phi_{2,k,b}| = \phi_{2,k,b}^+ + \phi_{2,k,b}^-$$

$$\phi_{2,k,b}^2 = \sum_{l=1}^{L} \alpha_{k,b,l} \Delta\phi_{2,k,b,l}$$

$$\phi_{2,k,b}^+ + \phi_{2,k,b}^- = \sum_{l=1}^{L} \Delta\phi_{2,k,b,l}$$

$$\Delta\phi_{2,k,b,l} \le (z_k^{max} - \theta_k^{min})/(2L) \;\; ; \;\; \Delta\phi_{2,k,b,l} \ge \Delta\phi_{2,k,b,l+1}$$

$$|P_{k,b} + V_{nom}^2 b_k u_k \theta_{k,b}| \le M_k(1 - u_k); \; where \; -b_k = 1/x_k$$

$$|P_{k,b}| + 0.5 PL_{k,b} \le u_k S_{k,max}$$

$$|P_{k,b}| + 0.5 PL_{k,b} \le z_k S_{k,max}$$

$$\sum_{k \in i} P_{k,b} + \sum_{g \in i} PG_{g,b} + p_{i,b} - \sum_{d \in i} PD_{d,b} + 0.5 \sum_{k \in i} PL_{k,b} = 0$$

$$u_g PG_{g,min} \le PG_{g,b} \le u_g PG_{g,max}$$

$$\theta_{min} \le \theta_{i,b} \le \theta_{max}; \; \theta_{ref} = 0 \qquad\qquad HTEP^*$$

$$PL_{k,b} = r_k P_{k,b}^2 / V_{nom}^2$$

$$P_{k,b} = P_{k,b}^+ - P_{k,b}^- \rightarrow |P_{k,b}| = P_{k,b}^+ + P_{k,b}^-$$

$$P_{k,b}^2 = \sum_{l=1}^{L} \alpha_{k,b,l} \Delta p_{k,b,l}$$

$$P_{k,b}^+ + P_{k,b}^- = \sum_{l=1}^{L} \Delta p_{k,b,l}$$

$$\Delta p_{k,b,l} \leq S_k^{max}/L \;\; ; \;\; \Delta p_{k,b,l+1} \leq \Delta p_{k,b,l} \;\; ; \;\; \alpha_{k,b,l} = (2l-1)\, S_k^{max}/L$$

$$\left|P_{k,b}\right| + 0.5\, PL_{k,b} \leq u_k S_{k,\max}$$

$$\left|P_{k,b}\right| + 0.5\, PL_{k,b} \leq z_k S_{k,\max}$$

$$\sum_{k\in i} P_{k,b} + \sum_{g\in i} PG_{g,b} + p_{i,b} - \sum_{d\in i} PD_{d,b} + 0.5 \sum_{k\in i} PL_{k,b} = 0$$

$$u_g PG_{g,min} \leq PG_{g,b} \leq u_g PG_{g,max}$$

$$\theta_{min} \leq \theta_{i,b} \leq \theta_{max}; \; \theta_{ref} = 0$$

$$PL_{k,b} = r_k P_{k,b}^2 / V_{nom}^2 \qquad\qquad PTEP^*$$

$$P_{k,b} = P_{k,b}^+ - P_{k,b}^- \rightarrow \left|P_{k,b}\right| = P_{k,b}^+ + P_{k,b}^-$$

$$P_{k,b}^2 = \sum_{l=1}^{L} \alpha_{k,b,l} \Delta p_{k,b,l}$$

$$P_{k,b}^+ + P_{k,b}^- = \sum_{l=1}^{L} \Delta p_{k,b,l}$$

$$\Delta p_{k,b,l} \leq S_k^{max}/L \;\; ; \;\; \Delta p_{k,b,l+1} \leq \Delta p_{k,b,l} \;\; ; \;\; \alpha_{k,b,l} = (2l-1)\, S_k^{max}/L$$

$$P_{k,b} + V_{nom}^2 b_k u_k \theta_{k,b} = 0; \; where \; -b_k = 1/x_k$$

$$\sum_{k\in i} P_{k,b} + \sum_{g\in i} PG_{g,b} + p_{i,b} - \sum_{d\in i} PD_{d,b} + 0.5 \sum_{k\in i} PL_{k,b} = 0$$

$$u_g PG_{g,min} \leq PG_{g,b} \leq u_g PG_{g,max}$$

$$\theta_{min} \leq \theta_{i,b} \leq \theta_{max}; \; \theta_{ref} = 0$$

$$PL_{k,b} = r_k P_{k,b}^2 / V_{nom}^2$$

$$P_{k,b} = P_{k,b}^+ - P_{k,b}^- \rightarrow \left|P_{k,b}\right| = P_{k,b}^+ + P_{k,b}^- \qquad\qquad CSTEP^*$$

$$P_{k,b}^2 = \sum_{l=1}^{L} \alpha_{k,b,l} \Delta p_{k,b,l}$$

$$P_{k,b}^+ + P_{k,b}^- = \sum_{l=1}^{L} \Delta p_{k,b,l}$$

$$\Delta p_{k,b,l} \leq S_k^{max}/L \;\; ; \;\; \Delta p_{k,b,l+1} \leq \Delta p_{k,b,l} \;\; ; \;\; \alpha_{k,b,l} = (2l-1)\, S_k^{max}/L$$

* This model can also be formulated from models other than the DCTEP.

The data for the test systems used in the analysis throughout the thesis are provided here. Base power in all cases is 100 MVA.

C. 1.   Garver's 6-bus System

Table C. 1 Garver's 6-bus data

Generator data

Node	PG_{max} (MW)	PG_{min} (MW)	QG_{max} (MVAr)	QG_{min} (MVAr)	Marginal cost (€/MWh)
1	150	0	65	-65	30
3	360	0	150	-150	40
6	600	0	200	-200	5

Load data

Node	PD_{max} (MW)	QD_{max} (MW)
1	80	16
2	240	48
3	40	8
4	160	32
5	240	48
6	0	0

Existing lines data

From	To	r (pu)	x (pu)	S_{max} (MVA)
1	2	0.1	0.4	100
1	4	0.15	0.6	80
1	5	0.05	0.2	100
2	3	0.05	0.2	100
2	4	0.1	0.4	100
3	5	0.05	0.2	100

Candidate lines data

From	To	r (pu)	x (pu)	S_{max} (MVA)	IC (M€)
1	2	0.1	0.4	100	40
1	4	0.15	0.6	80	60
1	5	0.05	0.2	100	20
2	4	0.1	0.4	100	40
1	3	0.09	0.38	100	38
1	6	0.17	0.68	70	68
2	5	0.08	0.31	100	31
2	6	0.08	0.3	100	3

2	3	0.05	0.2	100	20
3	4	0.15	0.59	82	59
3	5	0.05	0.2	100	20
3	6	0.12	0.48	100	48
4	5	0.16	0.63	75	63
4	6	0.08	0.3	100	30
5	6	0.15	0.61	78	61

C. 2.   IEEE 24-bus System (Base power 100 MVA)
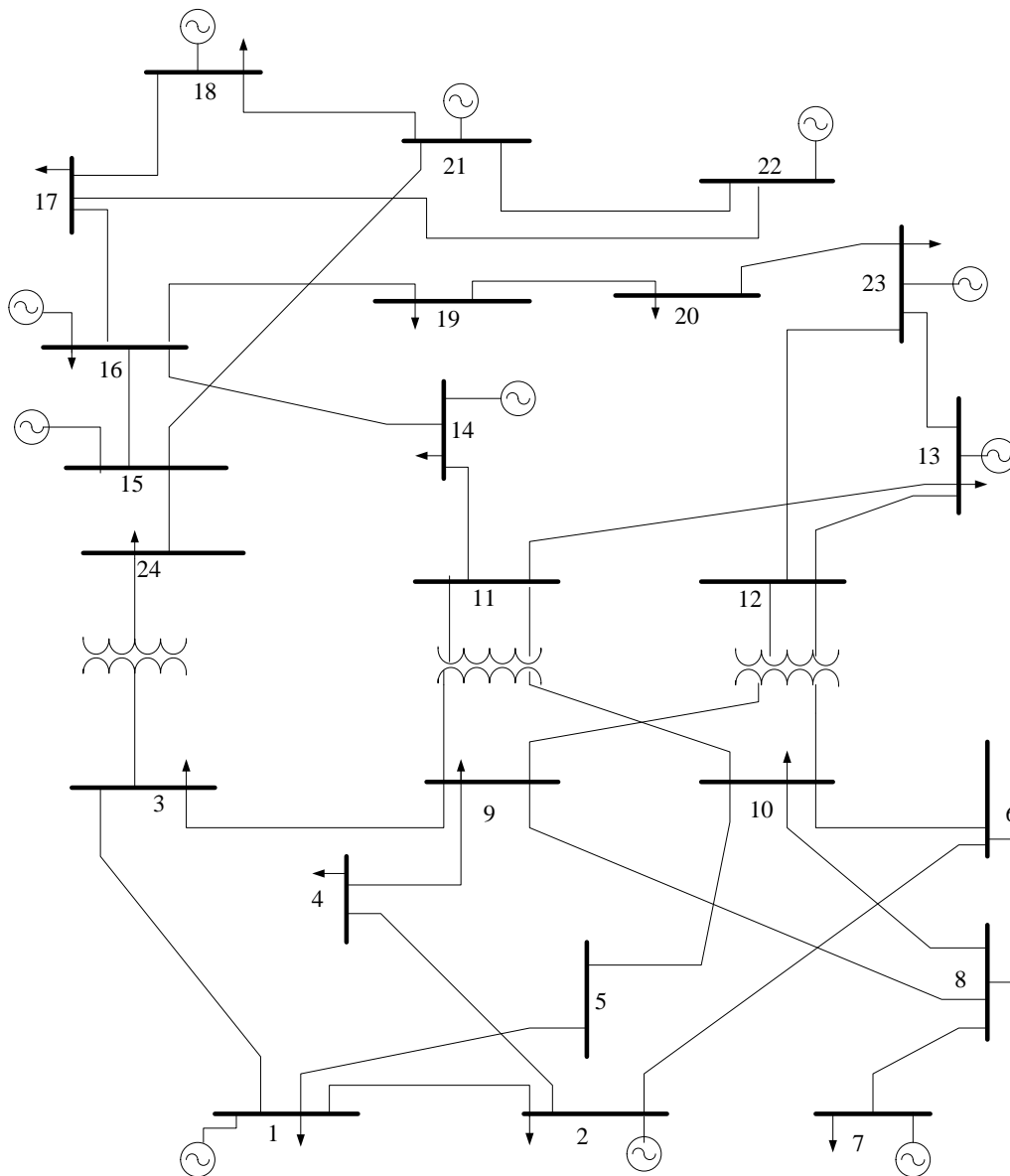
Fig. C. 1 Single line diagram of IEEE 24-bus test system

Table C. 2 IEEE 24-bus data

Generator data

Node	PG_{max} (MW)	PG_{min} (MW)	QG_{max} (MVAr)	QG_{min} (MVAr)	Marginal cost (€/MWh)
1	192	62.4	80	-80	16
2	192	62.4	80	-80	16
7	300	75	180	-180	43
13	591	207	240	-240	48
15	215	66.3	110	-110	58
16	155	54.3	80	-80	12
18	400	100	200	-200	4
21	400	100	200	-200	4
22	300	60	96	-96	0
23	660	248.6	310	-310	12
4	1500	0	0	0	0
17	1000	0	0	0	0

Demand data

Node	PD_{max} (MW)	QD_{max} (MW)
1	108	22
2	97	20
3	180	37
4	74	15
5	71	14
6	136	28
7	125	25
8	171	35
9	175	36
10	195	40
13	265	54
14	194	39
15	317	64
16	100	20
18	333	68
19	181	37
20	128	26

Existing lines data

From	To	r (pu)	x (pu)	S_{max} (MVA)
1	2	0.0026	0.0139	175
1	3	0.0546	0.2112	175
1	5	0.0218	0.0845	175

2	4	0.0328	0.1267	175	
2	6	0.0497	0.192	175	
3	9	0.0308	0.119	175	
3	24	0.0023	0.0839	400	
4	9	0.0268	0.1037	175	
5	10	0.0228	0.0883	175	
6	10	0.0139	0.0605	175	
7	8	0.0159	0.0614	175	
8	9	0.0427	0.1651	175	
8	10	0.0427	0.1651	175	
9	11	0.0023	0.0839	400	
9	12	0.0023	0.0839	400	
10	11	0.0023	0.0839	400	
10	12	0.0023	0.0839	400	
11	13	0.0061	0.0476	500	
11	14	0.0054	0.0418	500	
12	13	0.0061	0.0476	500	
12	23	0.0124	0.0966	500	
13	23	0.0111	0.0865	500	
14	16	0.005	0.0389	500	
15	16	0.0022	0.0173	500	
15	21	0.0027	0.029	1000	
15	24	0.0067	0.0519	500	
16	17	0.0033	0.0259	500	
16	19	0.003	0.0231	500	
17	18	0.0018	0.0144	500	
17	22	0.0135	0.1053	500	
18	21	0.0016	0.0129	1000	
19	20	0.0025	0.0198	1000	
20	23	0.0014	0.0108	1000	
21	22	0.0087	0.0678	500	

Candidate lines data

From	To	r (pu)	x (pu)	S_{max} (MVA)	IC (M€)
1	2	0.0026	0.0139	175	0.03
1	3	0.0546	0.2112	175	0.55
1	5	0.0218	0.0845	175	0.22
2	4	0.0328	0.1267	175	0.33
2	6	0.0497	0.192	175	0.5
3	9	0.0308	0.119	175	0.31
3	24	0.0023	0.0839	400	0.2
4	9	0.0268	0.1037	175	0.27
5	10	0.0228	0.0883	175	0.23

6	10	0.0139	0.0605	175	0.16
7	8	0.0159	0.0614	175	0.16
8	9	0.0427	0.1651	175	0.43
8	10	0.0427	0.1651	175	0.43
9	11	0.0023	0.0839	400	0.2
9	12	0.0023	0.0839	400	0.2
10	11	0.0023	0.0839	400	0.2
10	12	0.0023	0.0839	400	0.2
11	13	0.0061	0.0476	500	0.33
11	14	0.0054	0.0418	500	0.29
12	13	0.0061	0.0476	500	0.33
12	23	0.0124	0.0966	500	0.67
13	23	0.0111	0.0865	500	0.6
14	16	0.005	0.0389	500	0.27
15	16	0.0022	0.0173	500	0.12
15	21	0.0027	0.029	1000	0.68
15	24	0.0067	0.0519	500	0.36
16	17	0.0033	0.0259	500	0.18
16	19	0.003	0.0231	500	0.16
17	18	0.0018	0.0144	500	0.1
17	22	0.0135	0.1053	500	0.73
18	21	0.0016	0.0129	1000	0.36
19	20	0.0025	0.0198	1000	0.55
20	23	0.0014	0.0108	1000	0.3
21	22	0.0087	0.0678	500	0.47
1	8	0.0348	0.1344	175	0.35
2	8	0.0328	0.1267	175	0.33
6	7	0.0497	0.192	175	0.5
13	14	0.0057	0.0447	500	0.62
14	23	0.008	0.062	500	0.86
16	23	0.0105	0.0822	500	1.14
19	23	0.0078	0.0606	500	0.84

C. 3. IEEE 118-bus System (Base power 100 MVA)

163
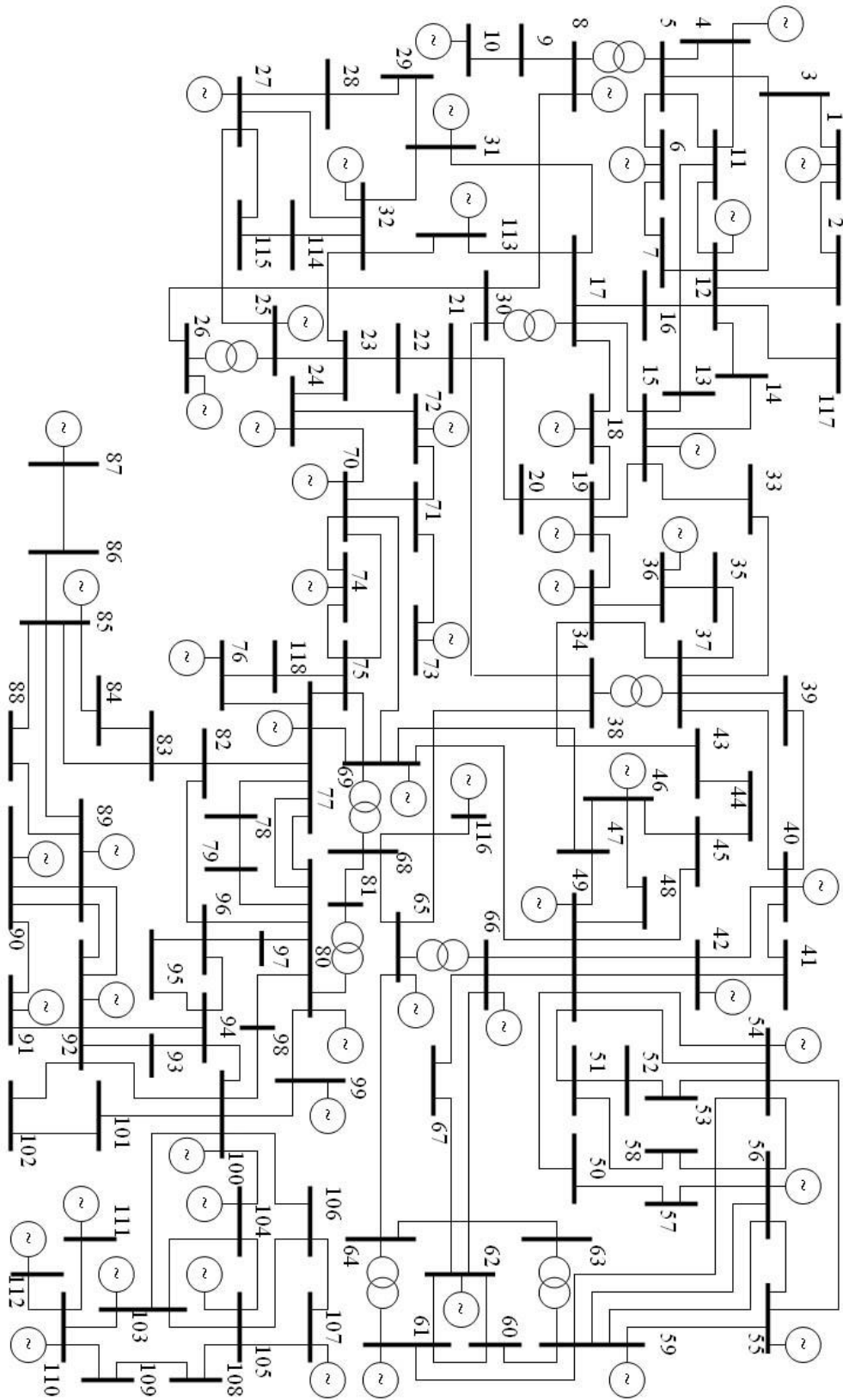
Fig. C. 2 Single line diagram of IEEE 118-bus test system

164

Table C. 3 IEEE 118-bus network data

Network data

From	To	r (pu)	x (pu)	S_{max} (MVA)	IC (M€)
1	2	0.0303	0.0999	115	18
1	3	0.0129	0.0424	115	7.6
4	5	0.0018	0.008	400	1.8
3	5	0.0241	0.108	115	16.2
5	6	0.0119	0.054	115	9.7
6	7	0.0046	0.0208	115	4.7
8	9	0.0024	0.0305	400	5.5
8	5	0	0.0267	400	6
9	10	0.0026	0.0322	400	5.8
4	11	0.0209	0.0688	115	12.4
5	11	0.0203	0.0682	115	12.3
11	12	0.006	0.0196	115	4.4
2	12	0.0187	0.0616	115	11.1
3	12	0.0484	0.16	115	24
7	12	0.0086	0.034	115	6.1
11	13	0.0223	0.0731	115	13.2
12	14	0.0215	0.0707	115	12.7
13	15	0.0744	0.2444	115	36.7
14	15	0.0595	0.195	115	29.3
12	16	0.0212	0.0834	115	15
15	17	0.0132	0.0437	400	7.9
16	17	0.0454	0.1801	115	27
17	18	0.0123	0.0505	115	9.1
18	19	0.0112	0.0493	115	8.9
19	20	0.0252	0.117	115	17.6
15	19	0.012	0.0394	115	7.1
20	21	0.0183	0.0849	115	15.3
21	22	0.0209	0.097	115	17.5
22	23	0.0342	0.159	115	23.9
23	24	0.0135	0.0492	115	8.9
23	25	0.0156	0.08	400	14.4
26	25	0	0.0382	400	6.9
25	27	0.0318	0.163	400	24.5
27	28	0.0191	0.0855	115	15.4
28	29	0.0237	0.0943	115	17
30	17	0	0.0388	400	7
8	30	0.0043	0.0504	115	9.1
26	30	0.008	0.086	400	15.5
17	31	0.0474	0.1563	115	23.4

29	31	0.0108	0.0331	115	6
23	32	0.0317	0.1153	115	17.3
31	32	0.0298	0.0985	115	17.7
27	32	0.0229	0.0755	115	13.6
15	33	0.038	0.1244	115	18.7
19	34	0.0752	0.247	115	37.1
35	36	0.0022	0.0102	115	2.3
35	37	0.011	0.0497	115	8.9
33	37	0.0415	0.142	115	21.3
34	36	0.0087	0.0268	115	6
34	37	0.0026	0.0094	400	2.1
38	37	0	0.0375	400	6.8
37	39	0.0321	0.106	115	15.9
37	40	0.0593	0.168	115	25.2
30	38	0.0046	0.054	115	9.7
39	40	0.0184	0.0605	115	10.9
40	41	0.0145	0.0487	115	8.8
40	42	0.0555	0.183	115	27.5
41	42	0.041	0.135	115	20.3
43	44	0.0608	0.2454	115	36.8
34	43	0.0413	0.1681	115	25.2
44	45	0.0224	0.0901	115	16.2
45	46	0.04	0.1356	115	20.3
46	47	0.038	0.127	115	19.1
46	48	0.0601	0.189	115	28.4
47	49	0.0191	0.0625	115	11.3
42	49	0.0358	0.1615	230	97
45	49	0.0684	0.186	115	27.9
48	49	0.0179	0.0505	115	9.1
49	50	0.0267	0.0752	115	13.5
49	51	0.0486	0.137	115	20.6
51	52	0.0203	0.0588	115	10.6
52	53	0.0405	0.1635	115	24.5
53	54	0.0263	0.122	115	18.3
49	54	0.0365	0.1445	230	87.1
54	55	0.0169	0.0707	115	12.7
54	56	0.0028	0.0096	115	2.1
55	56	0.0049	0.0151	115	3.4
56	57	0.0343	0.0966	115	17.4
50	57	0.0474	0.134	115	20.1
56	58	0.0343	0.0966	115	17.4
51	58	0.0255	0.0719	115	12.9
54	59	0.0503	0.2293	115	34.4

56	59	0.0413	0.1255	230	73.6
55	59	0.0474	0.2158	115	32.4
59	60	0.0317	0.145	115	21.8
59	61	0.0328	0.15	115	22.5
60	61	0.0026	0.0135	400	3
60	62	0.0123	0.0561	115	10.1
61	62	0.0082	0.0376	115	6.8
63	59	0	0.0386	400	6.9
63	64	0.0017	0.02	400	4.5
64	61	0	0.0268	400	6
38	65	0.009	0.0986	400	17.7
64	65	0.0027	0.0302	400	5.4
49	66	0.009	0.0456	800	33
62	66	0.0482	0.218	115	32.7
62	67	0.0258	0.117	115	17.6
65	66	0	0.037	400	6.7
66	67	0.0224	0.1015	115	15.2
65	68	0.0014	0.016	400	3.6
47	69	0.0844	0.2778	115	41.7
49	69	0.0985	0.324	115	48.6
68	69	0	0.037	400	6.7
69	70	0.03	0.127	400	19.1
24	70	0.0022	0.4115	115	61.7
70	71	0.0088	0.0355	115	6.4
24	72	0.0488	0.196	115	29.4
71	72	0.0446	0.18	115	27
71	73	0.0087	0.0454	115	8.2
70	74	0.0401	0.1323	115	19.8
70	75	0.0428	0.141	115	21.2
69	75	0.0405	0.122	400	18.3
74	75	0.0123	0.0406	115	7.3
76	77	0.0444	0.148	115	22.2
69	77	0.0309	0.101	115	15.2
75	77	0.0601	0.1999	115	30
77	78	0.0038	0.0124	115	2.8
78	79	0.0055	0.0244	115	5.5
77	80	0.0108	0.0332	800	24.5
79	80	0.0156	0.0704	115	12.7
68	81	0.0018	0.0202	400	4.5
81	80	0	0.037	400	6.7
77	82	0.0298	0.0853	115	15.4
82	83	0.0112	0.0367	115	6.6
83	84	0.0625	0.132	115	19.8
83	85	0.043	0.148	115	22.2

167

84	85	0.0302	0.0641	115	11.5
85	86	0.035	0.123	400	18.5
86	87	0.0283	0.2074	400	31.1
85	88	0.02	0.102	115	15.3
85	89	0.0239	0.173	115	26
88	89	0.0139	0.0712	400	12.8
89	90	0.0163	0.0651	800	46.1
90	91	0.0254	0.0836	115	15
89	92	0.0079	0.0383	800	32.8
91	92	0.0387	0.1272	115	19.1
92	93	0.0258	0.0848	115	15.3
92	94	0.0481	0.158	115	23.7
93	94	0.0223	0.0732	115	13.2
94	95	0.0132	0.0434	115	7.8
80	96	0.0356	0.182	115	27.3
82	96	0.0162	0.053	115	9.5
94	96	0.0269	0.0869	115	15.6
80	97	0.0183	0.0934	115	16.8
80	98	0.0238	0.108	115	16.2
80	99	0.0454	0.206	115	30.9
92	100	0.0648	0.295	115	44.3
94	100	0.0178	0.058	115	10.4
95	96	0.0171	0.0547	115	9.8
96	97	0.0173	0.0885	115	15.9
98	100	0.0397	0.179	115	26.9
99	100	0.018	0.0813	115	14.6
100	101	0.0277	0.1262	115	18.9
92	102	0.0123	0.0559	115	10.1
101	102	0.0246	0.112	115	16.8
100	103	0.016	0.0525	400	9.5
100	104	0.0451	0.204	115	30.6
103	104	0.0466	0.1584	115	23.8
103	105	0.0535	0.1625	115	24.4
100	106	0.0605	0.229	115	34.4
104	105	0.0099	0.0378	115	6.8
105	106	0.014	0.0547	115	9.8
105	107	0.053	0.183	115	27.5
105	108	0.0261	0.0703	115	12.7
106	107	0.053	0.183	115	27.5
108	109	0.0105	0.0288	115	6.5
103	110	0.0391	0.1813	115	27.2
109	110	0.0278	0.0762	115	13.7
110	111	0.022	0.0755	115	13.6
110	112	0.0247	0.064	115	11.5

168

17	113	0.0091	0.0301	115	5.4
32	113	0.0615	0.203	400	30.5
32	114	0.0135	0.0612	115	11
27	115	0.0164	0.0741	115	13.3
114	115	0.0023	0.0104	115	2.3
68	116	0.0003	0.0041	400	0.9
12	117	0.0329	0.14	115	21
75	118	0.0145	0.0481	115	8.7
76	118	0.0164	0.0544	115	9.8

This section presents simulation results corresponding to the IEEE 118-bus case study.

Table D. 1 Comparison of expansion decisions obtained by different TEP models – 118-bus case

Type	Model	Investment solution
Discrete	PTEP Lossy	34-43 (1), 43-44 (1), 44-45 (1), 82-83 (1), 85-86 (1), 86-87 (1)
	HTEP Lossy	34-43 (1), 43-44 (1), 44-45 (1), 77-78 (1), 82-83 (1), 83-85 (1), 85-86 (1), 86-87 (1)
	R-DCTEP Lossy	34-43 (1), 43-44 (1), 44-45 (1), 77-78 (1), 82-83 (1), 83-85 (1), 85-86 (1), 86-87 (1)
	DCTEP Lossy	34-43 (1), 43-44 (1), 44-45 (1), 77-78 (1), 82-83 (1), 83-85 (1), 85-86 (1), 86-87 (1)
	Lin ACTEP Lossy	1-3 (1), 2-12 (1), 34-43 (1), 43-44 (1), 44-45 (1), 77-78 (1), 82-83 (1), 83-85 (1), 85-86 (1), 86-87 (1)
Continuous	PTEP Lossy	15-17 (0.2), 23-25 (0.1), 25-27 (0.1), 26-30(0.3), 34-37 (0.1), 34-43 (0.7), 38-65 (0.1), 43-44 (1.0), 44-45 (1.0), 60-61 (0.1), 63-64 (0.1), 64-65 (0.2), 65-68 (0.1), 68-81 (0.06), 69-70 (0.1), 69-75 (0.1), 77-78 (0.2), 82-83 (0.4), 83-85 (0.2), 85-86 (0.6), 85-88 (0.2), 86-87 (0.6), 100-103 (0.07)
	HTEP Lossy	4-5 (0.02), 8-5 (0.07), 8-9 (0.07), 8-30 (0.2), 9-10 (0.07), 15-17 (0.15), 23-25 (0.1), 25-27 (0.1), 26-30 (0.3), 30-17 (0.09), 34-37 (0.1), 34-43 (0.8), 38-37 (0.4), 38-65 (0.1), 43-44 (1.0), 44-45 (1.0), 60-61 (0.1), 63-59 (0.3), 60-64 (0.2), 64-65 (0.2), 65-66 (0.1), 65-68 (0.1), 69-70 (0.1), 68-81 (0.1), 69-75 (0.1), 77-78 (0.2), 82-83 (0.6), 83-85 (0.2), 85-86 (0.6), 85-88 (0.07), 86-87 (0.6), 100-103 (0.07)
	R-DCTEP Lossy	4-5 (0.02), 8-5 (0.07), 8-9 (0.07), 8-30 (0.2), 9-10 (0.07), 15-17 (0.15), 23-25 (0.1), 25-27 (0.1), 26-27 (0.3), 30-17 (0.09), 34-37 (0.1), 34-43 (0.8), 38-37 (0.4), 38-65 (0.1), 43-44 (1.0), 44-45 (1.0), 60-61 (0.1), 63-59 (0.3), 60-64 (0.2), 64-65 (0.2), 65-66 (0.1), 65-68 (0.1), 69-70 (0.1), 68-81 (0.1), 69-75 (0.1), 77-78 (0.2), 82-83 (0.6), 83-85 (0.2), 85-86 (0.6), 85-88 (0.07), 86-87 (0.6), 100-103 (0.07)

Table D. 2 TEP model performances in terms of costs and simulation times—118-bus case

Type	Model	Investment cost (€)	Total cost (€)	CPU time (s)
Discrete	PTEP Lossy	166531673.8	13787863085	48.173
	HTEP Lossy	197508547.7	13835854730	1452.432
	R-DCTEP Lossy	197508547.7	13835854730	1542.35
	DCTEP Lossy	197508547.7	13843756116	1928.329
	LinACTEP Lossy	198679473.5	13895952854	78495.506
Continuous	PTEP Lossy	162317520.7	13729793181	5.647
	HTEP Lossy	174961933.2	13774538987	10.467

R-DCTEP Lossy	174961933.2	13774538987	23.883
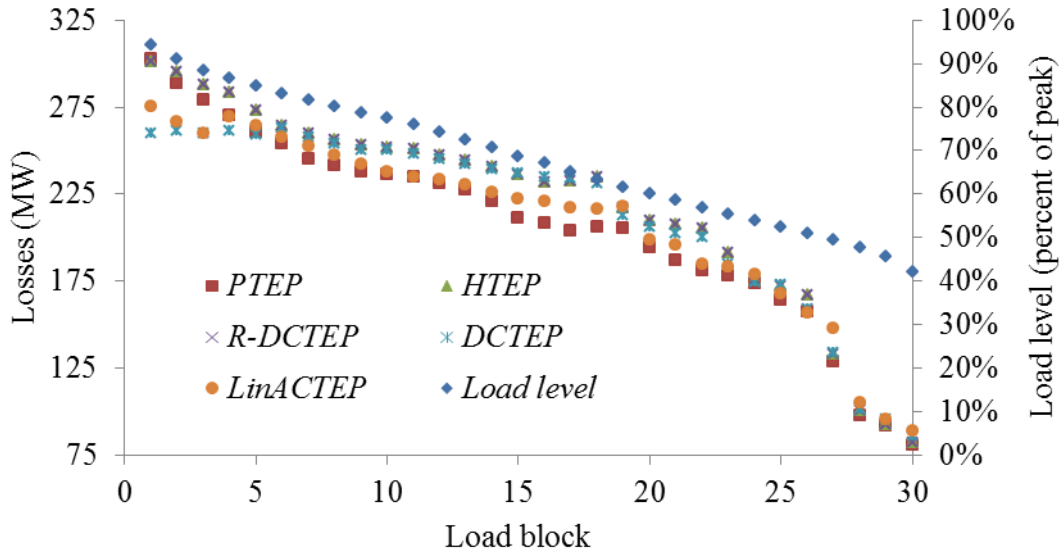
Fig. D. 1 Comparison of losses at each load level computed by different models 118-bus case
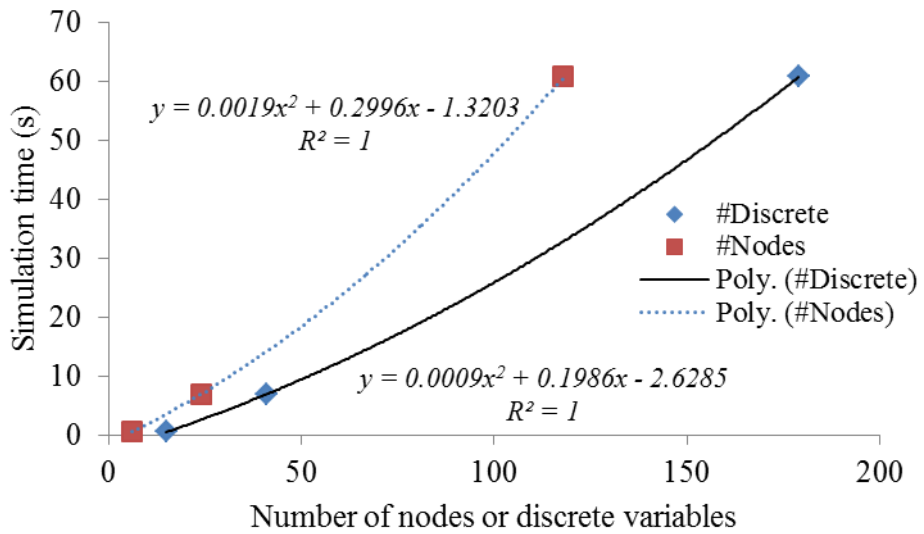

Fig. D. 2 Computational requirement of PTEP as a function of number of nodes and number of candidate lines
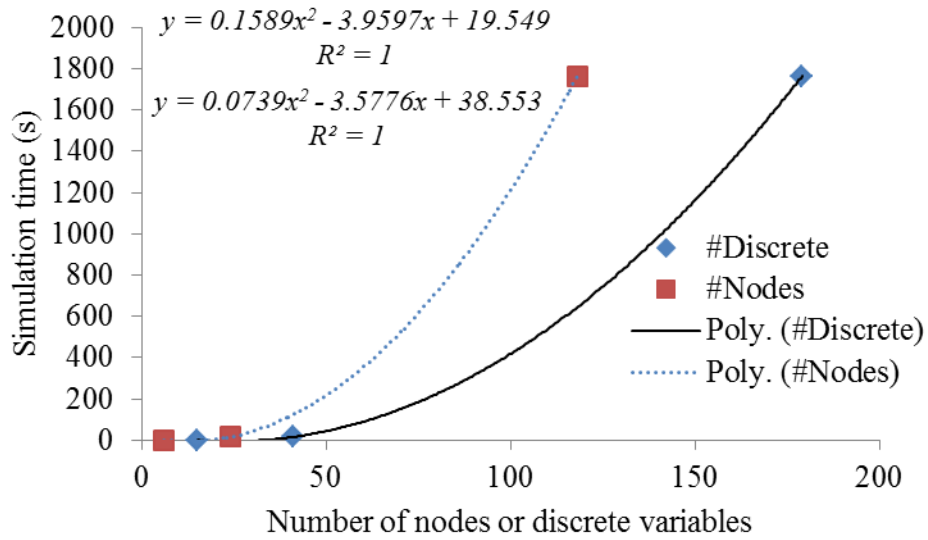
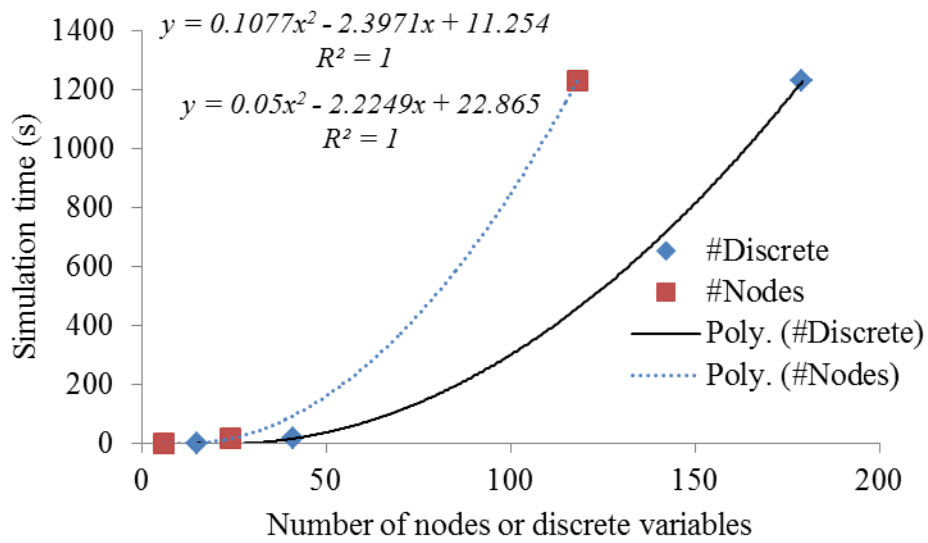Fig. D. 3 Computational requirement of HTEP as a function of number of nodes and number of candidate lines

Fig. D. 4 Computational requirement of R-DCTEP as a function of number of nodes and number of candidate lines

172

BIBLIOGRAPHY

REFERENCES

[1]   International Energy Agency (IEA), "World Energy Outlook Special Report - Energy and Climate Change," 2015.

[2]   International Energy Agency (IEA), "Energy Climate and Change - World Energy Outlook Special Briefing for COP21," 2015.

[3]   *Climate Change 2014: Synthesis Report*. Geneva, Switzerland: Intergovernmental Panel on Climate Change -IPCC.

[4]   *Renewables 2015 - Global Status Report*, Renewable Energy Policy Network for the 21st Century. .

[5]   N. Gupta, "A review on the inclusion of wind generation in power system studies," *Renew. Sustain. Energy Rev.*, vol. 59, pp. 530–543, Jun. 2016.

[6]   Eurostat, "Renewable energy statistics - Statistics Explained," 2015.

[7]   Eurostat, "Energy from renewable sources - Statistics Explained," 2016.

[8]   Eurostat, "Share of renewable energy in gross final energy consumption." [Online].   Available: http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode =t2020_31&plugin=1. [Accessed: 24-Jan-2016].

[9]   I. J. Pérez-Arriaga, "Security of electricity supply in Europe in a short, medium and long-term perspective," *Eur. Rev. Energy Mark.*, vol. 2, no. 2, 2007.

[10]   P. Favre-Perrod, "A vision of future energy networks," in *Power Engineering Society Inaugural Conference and Exposition in Africa, 2005 IEEE*, 2005, pp. 13–17.

[11]   Santiago Grijalva, Scott R. Dahman, Kollin J. Patten, and J. Anthony M. Visnesky, "Large-Scale Integration of Wind Generation Including Network Temporal Security Analysis," *Energy Convers. IEEE Trans. On*, vol. 22, no. 1, pp. 181–188, 2007.

[12]   M. O. Buygi, G. Balzer, H. M. Shanechi, and M. Shahidehpour, "Market-Based Transmission Expansion Planning," *IEEE Trans. Power Syst.*, vol. 19, no. 4, pp. 2060–2067, Nov. 2004.

[13]   D. Z. Fitiwi, L. Olmos, M. Rivier, F. de Cuadra, and I. J. Pérez-Arriaga, "Finding a representative network losses model for large-scale transmission expansion planning with renewable energy sources," *Energy*, vol. 101, pp. 343–358, Apr. 2016.

[14]   Desta Z. Fitiwi, F. de Cuadra, L. Olmos, and M. Rivier, "A new approach of clustering operational states for power network expansion planning problems dealing with RES (renewable energy source) generation operational variability and uncertainty," *Energy*, vol. 90, no. 2, pp. 1360–1376, Oct. 2015.

[15]   F. Desta Zahlay, F. de Cuadra, L. Olmos, M. Rivier, and I. J. Perez-Arriaga, "A formulation for large-scale transmission expansion planning problem and a solution strategy," in *10th International Conference on the European Energy Market (EEM)*, 2013, pp. 1–8.

174

[16] R. Nadira, R. R. Austria, C. A. Dortolina, and M. A. Avila, "Transmission Planning Today: A Challenging Undertaking," *Electr. J.*, vol. 17, no. 4, pp. 24–32, May 2004.

[17] G. Latorre, R. D. Cruz, J. M. Areiza, and A. Villegas, "Classification of publications and models on transmission expansion planning," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 938–946, May 2003.

[18] R. Hemmati, R.-A. Hooshmand, and A. Khodabakhshian, "State-of-the-art of transmission expansion planning: Comprehensive review," *Renew. Sustain. Energy Rev.*, vol. 23, pp. 312–319, Jul. 2013.

[19] R. Hemmati, R.-A. Hooshmand, and A. Khodabakhshian, "Comprehensive review of generation and transmission expansion planning," *IET Gener. Transm. Distrib.*, vol. 7, no. 9, pp. 955–964, Sep. 2013.

[20] M. J. Rider, A. V. Garcia, and R. Romero, "Power system transmission network expansion planning using AC model," *Gener. Transm. Distrib. IET*, vol. 1, no. 5, pp. 731–742, 2007.

[21] R.-A. Hooshmand, R. Hemmati, and M. Parastegari, "Combination of AC Transmission Expansion Planning and Reactive Power Planning in the restructured power system," *Energy Convers. Manag.*, vol. 55, pp. 26–35, Mar. 2012.

[22] J. Shu, L. Wu, B. Han, and L. Zhang, "Enhanced multi-dimensional power network planning based on ant colony optimization," *Int. Trans. Electr. Energy Syst.*, p. n/a-n/a, Feb. 2014.

[23] A. Arabali, M. Ghofrani, M. Etezadi-Amoli, M. S. Fadali, and M. Moeini-Aghtaie, "A Multi-Objective Transmission Expansion Planning Framework in Deregulated Power Systems With Wind Generation," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 3003–3011, Nov. 2014.

[24] T. Akbari, A. Rahimikian, and A. Kazemi, "A multi-stage stochastic transmission expansion planning method," *Energy Convers. Manag.*, vol. 52, no. 8–9, pp. 2844–2853, 2011.

[25] J. Aghaei, N. Amjady, A. Baharvandi, and M.-A. Akbari, "Generation and Transmission Expansion Planning: MILP - Based Probabilistic Model," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1592–1601, Jul. 2014.

[26] D. Pozo, J. Contreras, and E. Sauma, "If you build it, he will come: Anticipative power transmission planning," *Energy Econ.*, vol. 36, pp. 135–146, Mar. 2013.

[27] C. Muñoz, E. Sauma, J. Contreras, J. Aguado, and S. de La Torre, "Impact of high wind power penetration on transmission network expansion planning," *IET Gener. Transm. Distrib.*, vol. 6, no. 12, pp. 1281–1291, Dec. 2012.

[28] H. Park, R. Baldick, and D. P. Morton, "A Stochastic Transmission Planning Model With Dependent Load and Wind Forecasts," *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp. 1–9, 2015.

[29] S. Haffner, A. Monticelli, A. Garcia, J. Mantovani, and R. Romero, "Branch and bound algorithm for transmission system expansion planning using a transportation model," *Gener. Transm. Distrib. IEE Proc.-*, vol. 147, no. 3, pp. 149–156, 2000.

175

[30] G. Latorre-Bayona and I. J. Pérez-Arriaga, "Chopin, a heuristic model for long term transmission expansion planning," *IEEE Trans. Power Syst.*, vol. 9, no. 4, pp. 1886–1894, 1994.

[31] J. F. Alonso, A. Sáiz, L. Martín, G. Latorre, A. Ramos, and I. J. Pérez-Arriaga, "PERLA: An Optimization Model for Long Term Expansion Planning of Electric Power Transmission Networks," *Red Eléctrica Esp. SA*, 1991.

[32] R. Romero, A. Monticelli, A. Garcia, and S. Haffner, "Test systems and mathematical models for transmission network expansion planning," *Gener. Transm. Distrib. IEE Proc.-*, vol. 149, no. 1, pp. 27–36, 2002.

[33] G. C. Oliveira, S. Binato, L. Bahiense, L. Thomé, and M. V. Pereira, "Security-constrained transmission planning: A mixed-integer disjunctive approach," *Optim. Online August 2004*, 2004.

[34] H. Zhang, G. T. Heydt, V. Vittal, and H. D. Mittelmann, "Transmission expansion planning using an ac model: Formulations and possible relaxations," in *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1–8.

[35] J. A. Taylor and F. S. Hover, "Linear Relaxations for Transmission System Planning," *Power Syst. IEEE Trans. On*, vol. 26, no. 4, pp. 2533–2538, Nov. 2011.

[36] H. Zhang, G. T. Heydt, V. Vittal, and J. Quintero, "An Improved Network Model for Transmission Expansion Planning Considering Reactive Power and Network Losses," *IEEE Trans. Power Syst.*, vol. Early Access Online, 2013.

[37] R. A. Jabr, "Optimization of AC Transmission System Planning," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2779–2787, Aug. 2013.

[38] J. A. Taylor and F. S. Hover, "Conic AC transmission system planning," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 952–959, May 2013.

[39] L. L. Garver, "Transmission Network Estimation Using Linear Programming," *IEEE Trans. Power Appar. Syst.*, vol. PAS-89, no. 7, pp. 1688–1697, 1970.

[40] J.-C. Kaltenbach, J. Peschon, and E. H. Gehrig, "A Mathematical Optimization Technique for the Expansion of Electric Power Transmission Systems," *Power Appar. Syst. IEEE Trans. On*, vol. PAS-89, no. 1, pp. 113–119, 1970.

[41] H. M. D. R. H. Samarakoon, R. M. Shrestha, and O. Fujiwara, "A mixed integer linear programming model for transmission expansion planning with generation location selection," *Int. J. Electr. Power Energy Syst.*, vol. 23, no. 4, pp. 285–293, May 2001.

[42] N. Alguacil, A. L. Motto, and A. J. Conejo, "Transmission expansion planning: a mixed-integer LP approach," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1070–1077, Aug. 2003.

[43] F. Zhang, Z. Hu, and Y. Song, "Mixed-integer linear model for transmission expansion planning with line losses and energy storage systems," *IET Gener. Transm. Distrib.*, vol. 7, no. 8, pp. 919–928, Aug. 2013.

[44] R.-C. Leou, "A multi-year transmission planning under a deregulated market," *Int. J. Electr. Power Energy Syst.*, vol. 33, no. 3, pp. 708–714, Mar. 2011.

[45] R. S. Chanda and P. K. Bhattacharjee, "Application of computer software in transmission expansion planning using variable load structure," *Electr. Power Syst. Res.*, vol. 31, no. 1, pp. 13–20, Oct. 1994.

[46] S. H. M. Hashimoto, R. Romero, and J. R. S. Mantovani, "Efficient linear programming algorithm for the transmission network expansion planning problem," *Gener. Transm. Distrib. IEE Proc.-*, vol. 150, no. 5, pp. 536–542, 2003.

[47] J. F. Franco, M. J. Rider, and R. Romero, "A mixed-integer quadratically-constrained programming model for the distribution system expansion planning," *Int. J. Electr. Power Energy Syst.*, vol. 62, pp. 265–272, Nov. 2014.

[48] B. F. Hobbs, G. Drayton, E. Bartholomew Fisher, and W. Lise, "Improved Transmission Representations in Oligopolistic Market Models: Quadratic Losses, Phase Shifters, and DC Lines," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1018–1029, Aug. 2008.

[49] Z. M. Al-Hamouz and A. S. Al-Faraj, "Transmission expansion planning using nonlinear programming," in *Transmission and Distribution Conference and Exhibition 2002: Asia Pacific. IEEE/PES*, 2002, vol. 1, pp. 50–55 vols.1.

[50] H. K. Youssef and R. Hackam, "New transmission planning model," *Power Syst. IEEE Trans. On*, vol. 4, no. 1, pp. 9–18, 1989.

[51] I. G. Sanchez, R. Romero, J. R. S. Mantovani, and M. J. Rider, "Transmission-expansion planning using the DC model and nonlinear-programming technique," *Gener. Transm. Distrib. IEE Proc.-*, vol. 152, no. 6, pp. 763–769, 2005.

[52] J. Alvarez, K. Ponnambalam, and V. H. Quintana, "Transmission Expansion under Risk using Stochastic Programming," in *Probabilistic Methods Applied to Power Systems, 2006. PMAPS 2006. International Conference on*, 2006, pp. 1–7.

[53] N. Newham, "Power System Investment Planning using Stochastic Dual Dynamic Programming," 2008.

[54] Y. P. Dusonchet and A. El-Abiad, "Transmission Planning Using Discrete Dynamic Optimizing," *Power Appar. Syst. IEEE Trans. On*, vol. PAS-92, no. 4, pp. 1358–1371, 1973.

[55] S. Binato, M. V. . Pereira, and S. Granville, "A new Benders decomposition approach to solve power transmission network design problems," *IEEE Trans. Power Syst.*, vol. 16, no. 2, pp. 235–240, 2002.

[56] O. Alizadeh-Mousavi and M. Zima-Bočkarjova, "Efficient Benders cuts for transmission expansion planning," *Electr. Power Syst. Res.*, vol. 131, pp. 275–284, Feb. 2016.

[57] T. Norbiato dos Santos and A. L. Diniz, "A New Multiperiod Stage Definition for the Multistage Benders Decomposition Approach Applied to Hydrothermal Scheduling," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1383–1392, Aug. 2009.

[58] S. Haffner, A. Monticelli, A. Garcia, and R. Romero, "Specialised branch-and-bound algorithm for transmission network expansion planning," *Gener. Transm. Distrib. IEE Proc.-*, vol. 148, no. 5, pp. 482–488, 2001.

[59] M. J. Rider, L. A. Gallego, R. Romero, and A. V. Garcia, "Heuristic Algorithm to Solve the Short Term Transmission Network Expansion Planning," in *Power Engineering Society General Meeting, 2007. IEEE*, 2007, pp. 1–7.

[60] S. Binato, G. C. de Oliveira, and J. L. de Araujo, "A greedy randomized adaptive search procedure for transmission expansion planning," *Power Syst. IEEE Trans. On*, vol. 16, no. 2, pp. 247–253, 2001.

[61] M. V. F. Pereira, L. M. V. Pinto, S. H. F. Cunha, and G. C. Oliveira, "A Decomposition Approach To Automated Generation/Transmission Expansion Planning," *Power Appar. Syst. IEEE Trans. On*, vol. PAS-104, no. 11, pp. 3074–3083, 1985.

[62] Jaeseok Choi, A. A. El-Keib, and T. Tran, "A fuzzy branch and bound-based transmission system expansion planning for the highest satisfaction level of the decision maker," *Power Syst. IEEE Trans. On*, vol. 20, no. 1, pp. 476–484, 2005.

[63] V. A. Levi and M. S. Calovic, "A new decomposition based method for optimal expansion planning of large transmission networks," *Power Syst. IEEE Trans. On*, vol. 6, no. 3, pp. 937–943, 1991.

[64] L. P. Garces, A. J. Conejo, R. Garcia-Bertrand, and R. Romero, "A Bilevel Approach to Transmission Expansion Planning Within a Market Environment," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1513–1522, Aug. 2009.

[65] S. de la Torre, A. J. Conejo, and J. Contreras, "Transmission Expansion Planning in Electricity Markets," *IEEE Trans. Power Syst.*, vol. 23, no. 1, pp. 238–248, Feb. 2008.

[66] G. C. Oliveira, S. Binato, and M. V. F. Pereira, "Value-Based Transmission Expansion Planning of Hydrothermal Systems Under Uncertainty," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1429–1435, Nov. 2007.

[67] E. B. Fisher, R. P. O'Neill, and M. C. Ferris, "Optimal Transmission Switching," *Power Syst. IEEE Trans. On*, vol. 23, no. 3, pp. 1346–1355, 2008.

[68] G. A. Orfanos, I. I. Skoteinos, P. S. Georgilakis, and N. D. Hatziargyriou, "Transmission Expansion Planning Methodology Incorporating Contingency Analysis and Uncertainties of Load and Wind Generation," in *Energy Market (EEM), 2010 7th International Conference on the European*, 2010, pp. 1–7.

[69] L. P. Garcés, R. Romero, and J. M. Lopez-Lezama, "Market-Driven Security-Constrained Transmission Network Expansion Planning," vol. 7, no. 2, 2010.

[70] Risheng Fang and D. J. Hill, "A new strategy for transmission expansion in competitive electricity markets," *Power Syst. IEEE Trans. On*, vol. 18, no. 1, pp. 374–380, 2003.

[71] K. W. Hedman, Feng Gao, and G. B. Sheble, "Overview of transmission expansion planning using real options analysis," in *Power Symposium, 2005. Proceedings of the 37th Annual North American*, 2005, pp. 497–502.

[72] M. Mirhosseini and A. Gharaveisi, "Transmission Network Expansion Planning with a Heuristic Approach," vol. 2, no. 2, pp. 235 – 237, 2010.

[73] R. Romero, M. J. Rider, and I. d. J. Silva, "A Metaheuristic to Solve the Transmission Expansion Planning," *Power Syst. IEEE Trans. On*, vol. 22, no. 4, pp. 2289–2291, 2007.

[74] A. M. Leite da Silva, L. S. Rezende, L. A. da Fonseca Manso, and L. C. de Resende, "Reliability worth applied to transmission expansion planning based on ant colony system," *Int. J. Electr. Power Energy Syst.*, vol. 32, no. 10, pp. 1077–1084, Dec. 2010.

[75] Y.-X. Jin, H.-Z. Cheng, J. Yan, and L. Zhang, "New discrete method for particle swarm optimization and its application in transmission network expansion planning," *Electr. Power Syst. Res.*, vol. 77, no. 3–4, pp. 227–233, Mar. 2007.

[76] H. Faria Jr., S. Binato, M. G. C. Resende, and D. M. Falcao, "Power Transmission Network Design by Greedy Randomized Adaptive Path Relinking," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 43–49, Feb. 2005.

[77] R. J. Bennon, J. A. Juves, and A. P. Meliopoulos, "Use of Sensitivity Analysis in Automated Transmission Planning," *Power Appar. Syst. IEEE Trans. On*, vol. PAS-101, no. 1, pp. 53–59, 1982.

[78] R. Romero, C. Rocha, M. Mantovani, and J. R. S. Mantovani, "Analysis of heuristic algorithms for the transportation model in static and multistage planning in network expansion systems," *Gener. Transm. Distrib. IEE Proc.-*, vol. 150, no. 5, pp. 521–526, 2003.

[79] R. Romero, C. Rocha, J. R. S. Mantovani, and I. G. Sanchez, "Constructive heuristic algorithm for the DC model in network transmission expansion planning," *Gener. Transm. Distrib. IEE Proc.-*, vol. 152, no. 2, pp. 277–282, 2005.

[80] A. Monticelli, A. Santos, M. V. F. Pereira, S. H. Cunha, B. J. Parker, and J. C. G. Praca, "Interactive Transmission Network Planning Using a Least-Effort Criterion," *Power Appar. Syst. IEEE Trans. On*, vol. PAS-101, no. 10, pp. 3919–3925, 1982.

[81] A. O. Ekwue and B. J. Cory, "Transmission System Expansion Planning by Interactive Methods," *Power Appar. Syst. IEEE Trans. On*, vol. PAS-103, no. 7, pp. 1583–1591, 1984.

[82] M. V. F. Pereira and L. M. V. G. Pinto, "Application Of Sensitivity Analysis Of Load Supplying Capability To Interactive Transmission Expansion Planning," *Power Appar. Syst. IEEE Trans. On*, vol. PAS-104, no. 2, pp. 381–389, 1985.

[83] E. J. deOliveira, I. C. daSilva, J. L. R. Pereira, and S. Carneiro, "Transmission System Expansion Planning Using a Sigmoid Function to Handle Integer Investment Variables," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1616–1621, Aug. 2005.

[84] R. A. Gallego, A. Monticelli, and R. Romero, "Comparative studies on nonconvex optimization methods for transmission network expansion planning," in *Power Industry Computer Applications., 1997. 20th International Conference on*, 1997, pp. 24–30.

[85] K. Yoshimoto, K. Yasuda, and R. Yokoyama, "Transmission expansion planning using neuro-computing hybridized with genetic algorithm," in *Evolutionary Computation, 1995., IEEE International Conference on*, 1995, vol. 1, p. 126.

[86] Z. S. Temlyakova, V. Y. Lubchenko, and D. A. Pavlyuchenko, "Genetic algorithms in power engineering problems," in *Strategic Technology, 2007. IFOST 2007. International Forum on*, 2007, pp. 194–197.

[87] Y.-H. Li and J.-X. Wang, "Flexible Transmission Network Expansion Planning Considering Uncertain Renewable Generation and Load Demand Based on Hybrid Clustering Analysis," *Appl. Sci.*, vol. 6, no. 1, p. 3, Dec. 2015.

[88] R. Romero, R. A. Gallego, and A. Monticelli, "Transmission system expansion planning by simulated annealing," *Power Syst. IEEE Trans. On*, vol. 11, no. 1, pp. 364–369, 1996.

[89] P. Maghouli, S. H. Hosseini, M. O. Buygi, and M. Shahidehpour, "A Multi-Objective Framework for Transmission Expansion Planning in Deregulated Environments," *Power Syst. IEEE Trans. On*, vol. 24, no. 2, pp. 1051–1061, 2009.

[90] A. H. Escobar, R. A. Romero, and R. A. Gallego, "Transmission network expansion planning considering multiple generation scenarios," in *Transmission and Distribution Conference and Exposition: Latin America, 2008 IEEE/PES*, 2008, pp. 1–6.

[91] T. Sum-Im, "A novel differential evolution algorithmic approach to transmission expansion planning," 2009.

[92] P. S. Georgilakis, "Differential Evolution Solution to Transmission Expansion Planning Problem," *Handb. Power Syst. I*, pp. 409–427, 2010.

[93] R. A. Gallego, R. Romero, and A. J. Monticelli, "Tabu search algorithm for network synthesis," *Power Syst. IEEE Trans. On*, vol. 15, no. 2, pp. 490–495, 2000.

[94] E. L. Da Silva, J. M. . Ortiz, G. C. De Oliveira, and S. Binato, "Transmission network expansion planning under a tabu search approach," *Power Syst. IEEE Trans. On*, vol. 16, no. 1, pp. 62–68, 2002.

[95] A. S. Braga and J. T. Saraiva, "Transmission expansion planning and long term marginal prices calculation using simulated annealing," in *Power Tech Conference Proceedings, 2003 IEEE Bologna*, 2003, vol. 2, p. 7 pp. Vol.2.

[96] N. Leeprechanon, P. Limsakul, and S. Pothiya, "Optimal Transmission Expansion Planning Using Ant Colony Optimization," *J. Sustain. Energy Environ.*, vol. 1, pp. 71–76, 2010.

[97] Dong-Xiao Niu, Yun-Peng Ling, Qi Zhao, and Qing-Ying Zhao, "An Improved Particle Swarm Optimization Method Based on Borderline Search Strategy for Transmission Network Expansion Planning," in *Machine Learning and Cybernetics, 2006 International Conference on*, 2006, pp. 2846–2850.

[98] Yi-Xiong Jin, Hao-Zhong Cheng, Jian-Yong Yan, and Li Zhang, "Local Optimum Embranchment Based Convergence Guarantee Particle Swarm Optimization and Its Application in Transmission Network Planning," in *Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES*, 2005, pp. 1–6.

[99] T. Kuo, Z. Ming, Y. Fan, and L. Na, "Chance Constrained Transmission System Expansion Planning Method Based on Chaos Quantum Honey Bee Algorithm,"

in *Power and Energy Engineering Conference (APPEEC), 2010 Asia-Pacific*, 2010, pp. 1–5.

[100] R. C. G. Teive, E. L. Silva, and L. G. S. Fonseca, "A cooperative expert system for transmission expansion planning of electrical power systems," *Power Syst. IEEE Trans. On*, vol. 13, no. 2, pp. 636–642, 1998.

[101] H. Mori and K. Shimomugi, "Transmission network expansion planning with Scatter Search," in *Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on*, 2007, pp. 3749–3754.

[102] E. Handschin, M. Heine, D. Konig, T. Nikodem, T. Seibt, and R. Palma, "Object-oriented software engineering for transmission planning in open access schemes," *Power Syst. IEEE Trans. On*, vol. 13, no. 1, pp. 94–100, 1998.

[103] J. L. Ceciliano and R. Nieva, "Transmission network planning using evolutionary programming," in *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on*, 1999, vol. 3, p. 1803 Vol. 3.

[104] A. R. Abdelaziz, "Genetic algorithm-based power transmission expansion planning," in *Electronics, Circuits and Systems, 2000. ICECS 2000. The 7th IEEE International Conference on*, 2000, vol. 2, pp. 642–645 vols.2.

[105] H. Yu, C. Y. Chung, and K. P. Wong, "Robust Transmission Network Expansion Planning Method With Taguchi's Orthogonal Array Testing," *Power Syst. IEEE Trans. On*, vol. PP, no. 99, p. 1, 2010.

[106] H. Fan, H. Cheng, Z. Ying, F. Jiang, and F. Shi, "Transmission system expansion planning based on stochastic chance constrained programming with security constraints," in *Electric Utility Deregulation and Restructuring and Power Technologies, 2008. DRPT 2008. Third International Conference on*, 2008, pp. 909–914.

[107] S. Dehghan, H. Saboori, A. Kazemi, and S. Jadid, "Transmission network expansion planning using a DEA-based benders decomposition," in *Electrical Engineering (ICEE), 2010 18th Iranian Conference on*, 2010, pp. 955–960.

[108] "Transmission Network Expansion Planning with Simulation Optimization.pdf." .

[109] S. A. Khaparde, "Transmission Expansion Planning considering Contingency Criteria and Network Utilization," pp. 390–396, 2008.

[110] C. Cagigas and M. Madrigal, "Centralized vs. competitive transmission expansion planning: the need for new tools," in *Power Engineering Society General Meeting, 2003, IEEE*, 2003, vol. 2, p. 1017 Vol. 2.

[111] T. De la Torre, J. W. Feltes, T. Gomez San Roman, and H. M. Merrill, "Deregulation, privatization, and competition: transmission planning under uncertainty," *Power Syst. IEEE Trans. On*, vol. 14, no. 2, pp. 460–465, 1999.

[112] M. O. Buygi, H. M. Shanechi, G. Balzer, and M. Shahidehpour, "Transmission planning approaches in restructured power systems," in *Power Tech Conference Proceedings, 2003 IEEE Bologna*, 2003, vol. 2, p. 7 pp. Vol.2.

[113] R. Villasana, L. L. Garver, and S. J. Salon, "Transmission Network Planning Using Linear Programming," *Power Appar. Syst. IEEE Trans. On*, vol. PAS-104, no. 2, pp. 349–356, 1985.

[114] A. Seifu, S. Salon, and G. List, "Optimization of transmission line planning including security constraints," *Power Syst. IEEE Trans. On*, vol. 4, no. 4, pp. 1507–1513, 1989.

[115] A. dos Santos, P. M. Franca, and A. Said, "An optimization model for long-range transmission expansion planning," *Power Syst. IEEE Trans. On*, vol. 4, no. 1, pp. 94–101, 1989.

[116] E. L. Da Silva, H. A. Gil, and J. M. Areiza, "Transmission network expansion planning under an improved genetic algorithm," *Power Syst. IEEE Trans. On*, vol. 15, no. 3, pp. 1168–1174, 2000.

[117] R. Romero and A. Monticelli, "A hierarchical decomposition approach for transmission network expansion planning," *Power Syst. IEEE Trans. On*, vol. 9, no. 1, pp. 373–380, 1994.

[118] A. H. Escobar, R. A. Gallego, and R. Romero, "Multistage and Coordinated Planning of the Expansion of Transmission Systems," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 735–744, May 2004.

[119] J. Contreras and F. F. Wu, "A kernel-oriented algorithm for transmission expansion planning," *Power Syst. IEEE Trans. On*, vol. 15, no. 4, pp. 1434–1440, 2002.

[120] G. B. Shrestha and P. A. J. Fonseka, "Congestion-driven transmission expansion in competitive power markets," *Power Syst. IEEE Trans. On*, vol. 19, no. 3, pp. 1658–1665, 2004.

[121] P. Sanchez-Martin, A. Ramos, and J. F. Alonso, "Probabilistic midterm transmission planning in a liberalized market," *Power Syst. IEEE Trans. On*, vol. 20, no. 4, pp. 2135–2142, 2005.

[122] G. R. Kamyab, M. Fotuhi-Friuzabad, and M. Rashidinejad, "Transmission Expansion Planning in Restructured Power Systems Considering Investment Cost and n-1 Reliability," *J. Appl. Sci.*, vol. 8, no. 23, pp. 4312–4320, Dec. 2008.

[123] S. Sevin, *Transmission Expansion Planning to Alleviate Congestion in Deregulated Power Markets*. Auburn University, 2006.

[124] S. Sozer, C. S. Park, and J. Valenzuela, "Considering Uncertainty in Transmission Expansion Planning for Restructured Markets," *Power Syst. IEEE Trans. Spec. Sect. Transm. Invest. Pricing Constr. Submitt. Publ.*, 2006.

[125] A. S. D. Braga and J. T. Saraiva, "A Multiyear Dynamic Approach for Transmission Expansion Planning and Long-Term Marginal Costs Computation," *Power Syst. IEEE Trans. On*, vol. 20, no. 3, pp. 1631–1639, 2005.

[126] N. Yang and F. Wen, "A chance constrained programming approach to transmission system expansion planning," *Electr. Power Syst. Res.*, vol. 75, no. 2–3, pp. 171–177, 2005.

[127] T. Sum-Im, G. A. Taylor, M. R. Irving, and Y. H. Song, "A differential evolution algorithm for multistage transmission expansion planning," in *Universities Power Engineering Conference, 2007. UPEC 2007. 42nd International*, 2007, pp. 357–364.

[128] R. Bent, A. Berscheid, and G. L. Toole, "Transmission network expansion planning with simulation optimization," in *Los Alamos National Lab Report LA-*

UR-10-00165 submitted to the 24th AAAI Conference on Artificial Intelligence, 2009.

[129] A. K. Kazerooni and J. Mutale, "Transmission Network Planning Under Security and Environmental Constraints," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 1169–1178, May 2010.

[130] H. Zhang, "Transmission expansion planning for large power systems," Arizona State University, Tempe, AZ, 2013.

[131] H. Zhang, G. T. Heydt, V. Vittal, and J. Quintero, "An Improved Network Model for Transmission Expansion Planning Considering Reactive Power and Network Losses," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3471–3479, Aug. 2013.

[132] A. S. Braga and J. T. Saraiva, "Dealing with uncertainties in long term transmission expansion planning problems," in *Power Tech, 2005 IEEE Russia*, 2008, pp. 1–7.

[133] A. H. Seddighi and A. Ahmadi-Javid, "Integrated multiperiod power generation and transmission expansion planning with sustainability aspects in a stochastic environment," *Energy*, vol. 86, pp. 9–18, Jun. 2015.

[134] M. Moeini-Aghtaie, A. Abbaspour, and M. Fotuhi-Firuzabad, "Incorporating Large-Scale Distant Wind Farms in Probabilistic Transmission Expansion Planning - Part I: Theory and Algorithm," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1585–1593, Aug. 2012.

[135] M. Moeini-Aghtaie, A. Abbaspour, and M. Fotuhi-Firuzabad, "Incorporating Large-Scale Distant Wind Farms in Probabilistic Transmission Expansion Planning - Part II: Case Studies," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1594–1601, Aug. 2012.

[136] L. Bahiense, G. C. Oliveira, M. Pereira, and S. Granville, "A mixed integer disjunctive model for transmission network expansion," *IEEE Trans. Power Syst.*, vol. 16, no. 3, pp. 560–565, 2002.

[137] J. Choi, T. D. Mount, and R. J. Thomas, "Transmission Expansion Planning Using Contingency Criteria," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2249–2261, Nov. 2007.

[138] D. Mejia-Giraldo and J. McCalley, "Adjustable Decisions for Reducing the Price of Robustness of Capacity Expansion Planning," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1573–1582, Jul. 2014.

[139] A. Khodaei, M. Shahidehpour, L. Wu, and Z. Li, "Coordination of Short-Term Operation Constraints in Multi-Area Expansion Planning," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2242–2250, Nov. 2012.

[140] Min Xie, Jin Zhong, and F. F. Wu, "Multiyear Transmission Expansion Planning Using Ordinal Optimization," *Power Syst. IEEE Trans. On*, vol. 22, no. 4, pp. 1420–1428, 2007.

[141] A. Soroudi and T. Amraee, "Decision making under uncertainty in energy systems: State of the art," *Renew. Sustain. Energy Rev.*, vol. 28, pp. 376–384, Dec. 2013.

[142] J. M. Chamorro, L. M. Abadie, R. de Neufville, and M. Ilić, "Market-based valuation of transmission network expansion. A heuristic application in GB," *Energy*, vol. 44, no. 1, pp. 302–320, Aug. 2012.

[143] R. Hemmati, R.-A. Hooshmand, and A. Khodabakhshian, "Market based transmission expansion and reactive power planning with consideration of wind and load uncertainties," *Renew. Sustain. Energy Rev.*, vol. 29, pp. 1–10, Jan. 2014.

[144] G. A. Orfanos, P. S. Georgilakis, and N. D. Hatziargyriou, "Transmission Expansion Planning of Systems With Increasing Wind Power Integration," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1355–1362, May 2013.

[145] T. Akbari, M. Heidarizadeh, M. A. Siab, and M. Abroshan, "Towards integrated planning: Simultaneous transmission and substation expansion planning," *Electr. Power Syst. Res.*, vol. 86, pp. 131–139, May 2012.

[146] J. H. Zhao, J. Foster, Z. Y. Dong, and K. P. Wong, "Flexible Transmission Network Planning Considering Distributed Generation Impacts," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1434–1443, Aug. 2011.

[147] R. Billinton and W. Wangdee, "Reliability-Based Transmission Reinforcement Planning Associated With Large-Scale Wind Farms," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 34–41, Feb. 2007.

[148] Y. Gu, J. D. McCalley, and M. Ni, "Coordinating Large-Scale Wind Integration and Transmission Planning," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 652–659, Oct. 2012.

[149] Jae Hyung Roh, M. Shahidehpour, and Lei Wu, "Market-Based Generation and Transmission Planning With Uncertainties," *Power Syst. IEEE Trans. On*, vol. 24, no. 3, pp. 1587–1598, 2009.

[150] A. Lotfjou, Y. Fu, and M. Shahidehpour, "Hybrid AC/DC Transmission Expansion Planning," *IEEE Trans. Power Deliv.*, vol. 27, no. 3, pp. 1620–1628, Jul. 2012.

[151] N. E. Koltsaklis, P. Liu, and M. C. Georgiadis, "An integrated stochastic multi-regional long-term energy planning model incorporating autonomous power systems and demand response," *Energy*, vol. 82, pp. 865–888, Mar. 2015.

[152] B. Ji, X. Yuan, Z. Chen, and H. Tian, "Improved gravitational search algorithm for unit commitment considering uncertainty of wind power," *Energy*, vol. 67, pp. 52–62, Apr. 2014.

[153] Y.-K. Tung and B.-C. Yen, *Hydrosystems Engineering Uncertainty Analysis*, 1 edition. New York: McGraw-Hill Professional, 2005.

[154] C.-L. Su and C.-N. Lu, "Two-point estimate method for quantifying transfer capability uncertainty," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 573–579, May 2005.

[155] R. Azizipanah-Abarghooee, T. Niknam, A. Roosta, A. R. Malekpour, and M. Zare, "Probabilistic multiobjective wind-thermal economic emission dispatch based on point estimated method," *Energy*, vol. 37, no. 1, pp. 322–335, Jan. 2012.

[156] H. Yu, C. Y. Chung, K. P. Wong, and J. H. Zhang, "A Chance Constrained Transmission Network Expansion Planning Method With Consideration of Load and Wind Farm Uncertainties," *Power Syst. IEEE Trans. On*, vol. 24, no. 3, pp. 1568–1576, 2009.

[157] G. J. Osório, J. M. Lujano-Rojas, J. C. O. Matias, and J. P. S. Catalão, "A probabilistic approach to solve the economic dispatch problem with intermittent renewable energy sources," *Energy*, vol. 82, pp. 949–959, Mar. 2015.

[158] J. Dupačová, N. Gröwe-Kuska, and W. Römisch, *Scenario Reduction in Stochastic Programming: An Approach Using Probability Metrics*. 2003.

[159] N. Growe-Kuska, H. Heitsch, and W. Romisch, "Scenario reduction and scenario tree construction for power management problems," in *Power Tech Conference Proceedings, 2003 IEEE Bologna*, 2003, vol. 3.

[160] A. J. Conejo, F. J. Nogales, M. Carrión, and J. M. Morales, "Electricity pool prices: long-term uncertainty characterization for futures-market trading and risk management," *J. Oper. Res. Soc.*, vol. 61, no. 2, pp. 235–245, 2010.

[161] H. Park and R. Baldick, "Transmission Planning Under Uncertainties of Wind and Load: Sequential Approximation Approach," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2395–2402, Aug. 2013.

[162] B. Bahmani-Firouzi, E. Farjah, and R. Azizipanah-Abarghooee, "An efficient scenario-based and fuzzy self-adaptive learning particle swarm optimization approach for dynamic economic emission dispatch considering load and wind power uncertainties," *Energy*, vol. 50, pp. 232–244, Feb. 2013.

[163] A. Ghaderi, M. Parsa Moghaddam, and M. K. Sheikh-El-Eslami, "Energy efficiency resource modeling in generation expansion planning," *Energy*, vol. 68, pp. 529–537, Apr. 2014.

[164] J. Cristóbal, G. Guillén-Gosálbez, A. Kraslawski, and A. Irabien, "Stochastic MILP model for optimal timing of investments in CO2 capture technologies under uncertainty in prices," *Energy*, vol. 54, pp. 343–351, Jun. 2013.

[165] H. Kile and K. Uhlen, "Data reduction via clustering and averaging for contingency and reliability analysis," *Int. J. Electr. Power Energy Syst.*, vol. 43, no. 1, pp. 1435–1442, Dec. 2012.

[166] H. Kile, "Evaluation and Grouping of Power Market Scenarios in Security of Electricity Supply Analysis," Norwegian University of Science and Technology, 2014.

[167] Y. Ben-Haim, *Info-Gap Decision Theory, Second Edition: Decisions Under Severe Uncertainty*, 2 edition. Oxford: Academic Press, 2006.

[168] S. Dehghan, A. Kazemi, and N. Amjady, "Multi-objective robust transmission expansion planning using information-gap decision theory and augmented #x003F5;-constraint method," *IET Gener. Transm. Distrib.*, vol. 8, no. 5, pp. 828–840, May 2014.

[169] M. Sniedovich, "A critique of info-gap robustness model," *Saf. Reliab. Risk Anal. Theory Methods Appl. Taylor Francis Group Lond.*, pp. 2071–2079, 2009.

[170] A. Moreira, A. Street, and J. M. Arroyo, "An Adjustable Robust Optimization Approach for Contingency-Constrained Transmission Expansion Planning," *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp. 1–10, 2014.

[171] B. Chen, J. Wang, L. Wang, Y. He, and Z. Wang, "Robust Optimization for Transmission Expansion Planning: Minimax Cost vs. Minimax Regret," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 3069–3077, Nov. 2014.

[172] R. A. Jabr, "Robust Transmission Network Expansion Planning With Uncertain Renewable Generation and Loads," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4558–4567, Nov. 2013.

[173] R. Mínguez, R. García-Bertrand, and J. M. Arroyo, "Adaptive Robust Transmission Network Expansion Planning using Structural Reliability and Decomposition Techniques," *ArXiv Prepr. ArXiv150106613*, 2015.

[174] G.-R. Kamyab, M. Fotuhi-Firuzabad, and M. Rashidinejad, "Market-Based Transmission Expansion Planning Under Uncertainty in Bids by Fuzzy Assessment," *J. Electr. Eng. Technol.*, vol. 7, no. 4, pp. 468–479, Jul. 2012.

[175] S. M. Ryan, J. D. McCalley, and D. L. Woodruff, "Long Term Resource Planning for Electric Power Systems Under Uncertainty," 2011.

[176] T. J. Overbye, X. Cheng, and Y. Sun, "A comparison of the AC and DC power flow models for LMP calculations," in *Proceedings of the 37th Annual Hawaii International Conference on System Sciences, 2004*, 2004, p. 9 pp.-pp.

[177] K. Purchala, L. Meeus, D. V. Dommelen, and R. Belmans, "Usefulness of DC power flow for active power flow analysis," in *IEEE Power Engineering Society General Meeting, 2005*, 2005, p. 454–459 Vol. 1.

[178] B. Stott, J. Jardim, and O. Alsac, "DC Power Flow Revisited," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009.

[179] O. Ceylan, A. Ozdemir, and H. Dag, "Branch outage solution using particle swarm optimization," in *Power Engineering Conference, 2008. AUPEC '08. Australasian Universities*, 2008, pp. 1–5.

[180] M. R. Bussieck and A. Pruessner, "Mixed-integer nonlinear programming," *SIAGOPT Newsl. Views News*, vol. 14, no. 1, pp. 19–22, 2003.

[181] A. Wigderson, "P, NP and mathematics–a computational complexity perspective," in *Proc. of the 2006 International Congress of Mathematicians*, 2006.

[182] M. Rahmani, M. Rashidinejad, E. M. Carreno, and R. Romero, "Efficient method for AC transmission network expansion planning," *Electr. Power Syst. Res.*, vol. 80, no. 9, pp. 1056–1064, Sep. 2010.

[183] J. P. Vielma, S. Ahmed, and G. Nemhauser, "Mixed-Integer Models for Nonseparable Piecewise-Linear Optimization: Unifying Framework and Extensions," *Oper. Res.*, vol. 58, no. 2, pp. 303–315, Oct. 2009.

[184] L. S. Moulin, M. Poss, and C. Sagastizábal, "Transmission expansion planning with re-design," *Energy Syst.*, vol. 1, no. 2, pp. 113–139, Feb. 2010.

[185] J. Bisschop, *AIMMS optimization modeling*. Lulu. com, 2006.

[186] L. Blank and A. Tarquin, *Engineering Economy*, 7th ed. New York: McGraw-Hill Science/Engineering/Math, 2011.

[187] M. Ghasemi, S. Ghavidel, M. M. Ghanbarian, M. Gharibzadeh, and A. Azizi Vahed, "Multi-objective optimal power flow considering the cost, emission, voltage deviation and power losses using multi-objective modified imperialist competitive algorithm," *Energy*, vol. 78, pp. 276–289, Dec. 2014.

[188] ENTSO-E, "European Electricity Network." [Online]. Available: https://www.entsoe.eu/publications/grid-maps/. [Accessed: 15-Oct-2014].

[189] S.-Y. Lin and J.-F. Chen, "Distributed optimal power flow for smart grid transmission system with renewable energy sources," *Energy*, vol. 56, pp. 184–192, Jul. 2013.

[190] R. S. Salgado and E. L. Rangel Jr., "Optimal power flow solutions through multi-objective programming," *Energy*, vol. 42, no. 1, pp. 35–45, Jun. 2012.

[191] R. Arul, S. Velusami, and G. Ravi, "A new algorithm for combined dynamic economic emission dispatch with security constraints," *Energy*, vol. 79, pp. 496–511, Jan. 2015.

[192] Y. Z. Li, Q. H. Wu, M. S. Li, and J. P. Zhan, "Mean-variance model for power system economic dispatch with wind power integrated," *Energy*, vol. 72, pp. 510–520, Aug. 2014.

[193] M. Ghasemi, S. Ghavidel, E. Akbari, and A. A. Vahed, "Solving non-linear, non-smooth and non-convex optimal power flow problems using chaotic invasive weed optimization algorithms based on chaos," *Energy*, vol. 73, pp. 340–353, Aug. 2014.

[194] M. R. Narimani, R. Azizipanah-Abarghooee, B. Zoghdar-Moghadam-Shahrekohne, and K. Gholami, "A novel approach to multi-objective optimal power flow by a new hybrid optimization algorithm considering generator constraints and multi-fuel type," *Energy*, vol. 49, pp. 119–136, Jan. 2013.

[195] T. Niknam, M. rasoul Narimani, M. Jabbari, and A. R. Malekpour, "A modified shuffle frog leaping algorithm for multi-objective optimal power flow," *Energy*, vol. 36, no. 11, pp. 6420–6432, Nov. 2011.

[196] S. A.-H. Soliman and A.-A. H. Mantawy, *Modern Optimization Techniques with Applications in Electric Power Systems*. Springer Science & Business Media, 2011.

[197] T. N. dos Santos and A. L. Diniz, "A Dynamic Piecewise Linear Model for DC Transmission Losses in Optimal Scheduling Problems," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 508–519, 2011.

[198] A. Helseth, "A linear optimal power flow model considering nodal distribution of losses," in *European Energy Market (EEM), 2012 9th International Conference on the*, 2012, pp. 1–8.

[199] M. A. Farrag and M. M. El-Metwally, "New method for transmission planning using mixed-integer programming," *Gener. Transm. Distrib. IEE Proc. C*, vol. 135, no. 4, pp. 319–323, Jul. 1988.

[200] E. Litvinov, T. Zheng, G. Rosenwald, and P. Shamsollahi, "Marginal Loss Modeling in LMP Calculation," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 880–888, May 2004.

[201] F. Li and R. Bo, "DCOPF-Based LMP Simulation: Algorithm, Comparison With ACOPF, and Sensitivity," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1475–1485, Nov. 2007.

[202] V. Sarkar and S. A. Khaparde, "DCOPF-Based Marginal Loss Pricing With Enhanced Power Flow Accuracy by Using Matrix Loss Distribution," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1435–1445, Aug. 2009.

[203] Z. Hu, H. Cheng, Z. Yan, and F. Li, "An Iterative LMP Calculation Method Considering Loss Distributions," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1469–1477, Aug. 2010.

[204] A. J. C. Pereira and J. T. Saraiva, "Generation expansion planning (GEP) – A long-term approach using system dynamics and genetic algorithms (GAs)," *Energy*, vol. 36, no. 8, pp. 5180–5199, Aug. 2011.

[205] A. J. C. Pereira and J. T. Saraiva, "A long term generation expansion planning model using system dynamics – Case study using data from the Portuguese/Spanish generation system," *Electr. Power Syst. Res.*, vol. 97, pp. 41–50, Apr. 2013.

[206] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*. John Wiley & Sons, 2012.

[207] M. Modiri-Delshad and N. A. Rahim, "Solving non-convex economic dispatch problem via backtracking search algorithm," *Energy*, vol. 77, pp. 372–381, Dec. 2014.

[208] B. Bahmani-Firouzi, E. Farjah, and R. Azizipanah-Abarghooee, "An efficient scenario-based and fuzzy self-adaptive learning particle swarm optimization approach for dynamic economic emission dispatch considering load and wind power uncertainties," *Energy*, vol. 50, pp. 232–244, Feb. 2013.

[209] M. Younes, F. Khodja, and R. L. Kherfane, "Multi-objective economic emission dispatch solution using hybrid FFA (firefly algorithm) and considering wind power penetration," *Energy*, vol. 67, pp. 595–606, Apr. 2014.

[210] J. H. Zheng, J. J. Chen, Q. H. Wu, and Z. X. Jing, "Reliability constrained unit commitment with combined hydro and thermal generation embedded using self-learning group search optimizer," *Energy*, vol. 81, pp. 245–254, Mar. 2015.

[211] B. Mohammadi-ivatloo, A. Rabiee, A. Soroudi, and M. Ehsan, "Imperialist competitive algorithm for solving non-convex dynamic economic power dispatch," *Energy*, vol. 44, no. 1, pp. 228–240, Aug. 2012.

[212] S. Özyön, H. Temurtaş, B. Durmuş, and G. Kuvat, "Charged system search algorithm for emission constrained economic power dispatch problem," *Energy*, vol. 46, no. 1, pp. 420–430, Oct. 2012.

[213] R. Azizipanah-Abarghooee, T. Niknam, A. Roosta, A. R. Malekpour, and M. Zare, "Probabilistic multiobjective wind-thermal economic emission dispatch based on point estimated method," *Energy*, vol. 37, no. 1, pp. 322–335, Jan. 2012.

[214] G.-C. Liao, "A novel evolutionary algorithm for dynamic economic dispatch with energy saving and emission reduction in power system integrated wind power," *Energy*, vol. 36, no. 2, pp. 1018–1029, Feb. 2011.

[215] F. J. Heredia and N. Nabona, "Optimum short-term hydrothermal scheduling with spinning reserve through network flows," *IEEE Trans. Power Syst.*, vol. 10, no. 3, pp. 1642–1651, Aug. 1995.

[216] A. Sadegheih, "Optimization of network planning by the novel hybrid algorithms of intelligent optimization techniques," *Energy*, vol. 34, no. 10, pp. 1539–1551, Oct. 2009.

[217] S. M. Pedro and R. Andres, "Modeling Transmission Ohmic Losses in a Stochastic Bulk Production Cost Model." unpublished.

[218] R. Palma-Benhke, A. Philpott, A. Jofré, and M. Cortés-Carmona, "Modelling network constrained economic dispatch problems," *Optim. Eng.*, vol. 14, no. 3, pp. 417–430, Oct. 2012.

[219] H. P. Williams, *Model Building in Mathematical Programming*, 4th ed. Wiley, 1999.

[220] C. D'Ambrosio, A. Lodi, and S. Martello, "Piecewise linear approximation of functions of two variables in MILP models," *Oper. Res. Lett.*, vol. 38, no. 1, pp. 39–46, Jan. 2010.

[221] A. Martin, M. Möller, and S. Moritz, "Mixed Integer Models for the Stationary Case of Gas Network Optimization," *Math. Program.*, vol. 105, no. 2–3, pp. 563–582, Feb. 2006.

[222] -- --, "Data for Power Test Systems." [Online]. Available: http://motor.ece.iit.edu/Data/JEAS_IEEE118.doc. [Accessed: 10-Oct-2014].

[223] Gerald van Belle, *Statistical Rules of Thumb*, 2nd ed. Wiley, 2008.

[224] E. B. C. and S. Arora, "Performance comparison of Transmission Network Expansion Planning under deterministic and uncertain conditions," *Int. J. Electr. Power Energy Syst.*, vol. 33, no. 7, pp. 1288–1295, Sep. 2011.

[225] S. Wogrin, P. Duenas, A. Delgadillo, and J. Reneses, "A New Approach to Model Load Levels in Electric Power Systems With High Renewable Penetration," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2210–2218, Sep. 2014.

[226] R. J. Prokop and A. P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition," *CVGIP Graph. Models Image Process.*, vol. 54, no. 5, pp. 438–460, Sep. 1992.

[227] Y. Ding, P. Wang, L. Goel, P. C. Loh, and Q. Wu, "Long-Term Reserve Expansion of Power Systems With High Wind Power Penetration Using Universal Generating Function Methods," *Power Syst. IEEE Trans. On*, vol. PP, no. 99, p. 1, 2010.

[228] -- Swedish Meteorological and Hydrological Institute, "SMHI´s open data." .

[229] -- NREL, "National Solar Radiation Data Base." .

[230] G. Sinden, "Characteristics of the UK wind resource: Long-term patterns and relationship to electricity demand," *Energy Policy*, vol. 35, no. 1, pp. 112–127, Jan. 2007.

[231] Y. Z. Li, Q. H. Wu, M. S. Li, and J. P. Zhan, "Mean-variance model for power system economic dispatch with wind power integrated," *Energy*, vol. 72, pp. 510–520, Aug. 2014.

[232] M. R. Patel, *Wind and Solar Power Systems*. CRC Press, 1999.

[233] M. Zhao, Z. Chen, and F. Blaabjerg, "Probabilistic capacity of a grid connected wind farm based on optimization method," *Renew. Energy*, vol. 31, no. 13, pp. 2171–2187, Oct. 2006.

[234] M. Aien, M. Rashidinejad, and M. Fotuhi-Firuzabad, "On possibilistic and probabilistic uncertainty assessment of power flow problem: A review and a new approach," *Renew. Sustain. Energy Rev.*, vol. 37, pp. 883–895, Sep. 2014.

[235] C. Grigg, P. Wong, P. Albrecht, R. Allan, M. Bhavaraju, R. Billinton, Q. Chen, C. Fong, S. Haddad, S. Kuruganty, W. Li, R. Mukerji, D. Patton, N. Rau, D. Reppen, A. Schneider, M. Shahidehpour, and C. Singh, "The IEEE Reliability Test System-1996. A report prepared by the Reliability Test System Task Force of the Application of Probability Methods Subcommittee," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 1010–1020, Aug. 1999.

[236] P. Phonrattanasak, "Optimal placement of DG using multiobjective particle swarm optimization," in *2010 2nd International Conference on Mechanical and Electrical Technology (ICMET)*, 2010, pp. 342–346.

[237] Y. Gu, *Long-term power system capacity expansion planning considering reliability and economic criteria.* ProQuest, UMI Dissertation Publishing, 2012.

[238] S. Lumbreras, A. Ramos, and P. Sánchez, "Automatic selection of candidate investments for Transmission Expansion Planning," *Int. J. Electr. Power Energy Syst.*, vol. 59, pp. 130–140, Jul. 2014.

[239] S. Z. Moghaddam, H. Monsef, and M. Jafari, "A new heuristic method for transmission expansion planning using AHP," in *Environment and Electrical Engineering (EEEIC), 2011 10th International Conference on*, pp. 1–4.

[240] B. Stott, J. Jardim, and O. Alsac, "DC Power Flow Revisited," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009.

[241] G. Migliavacca, *Advanced Technologies for Future Transmission Grids*, vol. 1. Springer, 2013.

[242] G. N. Candlish, N. Law, P. Bett, R. Clark, H. Thornton, and C. Wilson, "Interannual wind variation from observations and numerical weather analyses." Met Office, Exeter, UK, EX1 3PB.

[243] P. Giorsetto and K. F. Utsurogi, "Development of a New Procedure for Reliability Modeling of Wind Turbine Generators," *IEEE Trans. Power Appar. Syst.*, vol. PAS-102, no. 1, pp. 134–143, Jan. 1983.

[244] R. Karki, P. Hu, and R. Billinton, "A simplified wind power generation model for reliability evaluation," *IEEE Trans. Energy Convers.*, vol. 21, no. 2, pp. 533–540, Jun. 2006.

[245] CIGRE Working Group B1.07, "Statistics of AC underground cable in power networks," CIGRE_B1.07, 2006.

The main set of computational tools employed in the presented research is the numerical decision support system jointly designated as STEP. These decision support models have been developed for this thesis and are documented mainly in Chapters 3-6, as well as the associated publications. Model code and all input parameters are freely available under CreativeCommons (CC) BY-SA 3.0 license, allowing free copies and redistribution of the material in any medium or format, as well as remixing, transforming, and building upon the material for any purpose, even commercially. The code can be requested directly from the author via email. STEP model is formulated and implemented in the General Algebraic Modeling System (GAMS©) BUILD 23.7-24.1.2. For handling input and output data, all calculations were performed using Microsoft Excel©. The optimization problems were for the most part solved with the CPLEX™ 12.5.1 solver for linear programming (LP) problems.

For more information about licensing and the public domain, please consider the CC homepage under: http://creativecommons.org/.

COMPLETE LIST OF RELEVANT PUBLICATIONS

Peer-reviewed Journal Articles

1. Desta Z. Fitiwi, L. Olmos, M. Rivier, and F. de Cuadra, "Finding a representative network losses model for large-scale transmission expansion planning with renewable energy sources," *Energy*, vol. 101C, pp. 343–358, Feb. 2016.
2. Desta Z. Fitiwi, F. de Cuadra, L. Olmos, and M. Rivier, "A new approach of clustering operational states for power network expansion planning problems dealing with RES (renewable energy source) generation operational variability and uncertainty," *Energy*, vol. 90, no. 2, pp. 1360–1376, Oct. 2015.
3. Desta Z. Fitiwi, L. Olmos, M. Rivier, F. de Cuadra and I. J. Pérez-Arriaga, "A Strategy for Combinatorial Solution Search Space Reduction under a Two-stage Modeling Framework of Large-scale Transmission Expansion Planning", *IEEE Trans. Power Syst.*, (*under review*).
4. Desta Z. Fitiwi, Sérgio F. Santos, Abebe W. Bizuayehu, M. Shafie-khah, J. P. S. Catalão, M. Asensio, J. Contreras, "Impact of Operational Variability and Uncertainty on DG Investment Planning: A Comprehensive Sensitivity Analysis", *IEEE Trans. Sust. Energy*, (*under review*).
5. Desta Z. Fitiwi Sérgio F. Santos, Abebe W. Bizuayehu, M. Shafie-khah, J. P. S. Catalão, M. Asensio, J. Contreras, "DG Investment Planning Considering Operational Variability and Uncertainty: A Multi-stage and Stochastic Approach", *IEEE Trans. Sust. Energy*, (*under review*).
6. Desta Z. Fitiwi, Sérgio F. Santos, Abebe W. Bizuayehu, M. Shafie-khah, J. P. S. Catalão, "A New Multi-Stage and Stochastic Mathematical Model for Maximizing RES Hosting Capacity—Part I: Problem Formulation", *IEEE Trans. Sust. Energy*, (*under review*).
7. Desta Z. Fitiwi, Sérgio F. Santos, Abebe W. Bizuayehu, M. Shafie-khah, J. P. S. Catalão, "A New Multi-Stage and Stochastic Mathematical Model for Maximizing RES Hosting Capacity—Part II: Numerical Results", *IEEE Trans. Sust. Energy*, (*under review*).
8. Desta Z. Fitiwi, Marco Cruz, Sérgio F. Santos, J.P.S. Catalão, "Impacts of Optimal Energy Storage Deployment and Network Reconfiguration on Renewable Integration Level in Distribution Systems", *Energies* (*under review*).

Peer-reviewed Conference Papers

9. Desta Z. Fitiwi, F. de Cuadra, L. Olmos, M. Rivier, and I. J. Perez-Arriaga, "A formulation for large-scale transmission expansion planning problem and a solution strategy," in *10th International Conference on the European Energy Market* (EEM), 2013, pp. 1–8.
10. Desta Z. Fitiwi, A. W. Bizuayehu, M. Shafie-khah, J. P. S. Catalão, M. Asenso, and J. Contreras, "DG Investment Planning Analysis with Renewable Integration and Considering Emission Costs," in *the 16th Int. Conf. on Computer as a Tool*, EuroCon2015, Salamanca, February 2015.

11. Desta Z. Fitiwi, Miadreza Shafie-khah, Abebe W. Bizuayehu, J. P. S. Catalão, Miguel Asensio, Javier Contreras, "A New Dynamic and Stochastic Distributed Generation Investment Planning Model with Recourse", IEEE PES General Meeting 2016, Boston (*accepted*).

12. S. F. Santos, Desta Z. Fitiwi, Miadreza Shafie-khah, Abebe W. Bizuayehu, J. P. S. Catalão, "Optimal Integration of RES-based DGs with Reactive Power Support Capabilities in Distribution Network Systems", in *13th European Energy Market Conference* (EEM), 2016 (*accepted*).

13. M. R. M. Cruz, Desta Z. Fitiwi, Sérgio F. Santos, J.P.S. Catalão, "Influence of Distributed Storage Systems and Network Switching/Reinforcement on RES-based DG Integration Level", in *13th European Energy Market Conference* (EEM), 2016 (*accepted*).

14. Abebe W. Bizuayehu, Desta Z. Fitiwi, J.P.S. Catalão, "Advantages of Optimal Storage Location and Size on the Economic Dispatch in Distribution Systems", *IEEE PES General Meeting*, 2016, Boston (*accepted*).

15. M. Shafie-khah, Desta Z. Fitiwi, J. P. S. Catalão, E. Heydarian-Forushani, M. E. H. Golshan, "Simultaneous Participation of Demand Response Aggregators in Ancillary Services and Demand Response Exchange Markets", *IEEE PES General Meeting* 2016, Boston (*accepted*).

16. M. Shafie-khah, M.H. Shoreh, P. Siano, Desta Z. Fitiwi, G.J. Osório, J. Lujano-Rojas, J.P.S. Catalão, "Optimal Demand Response Programs for Improving the Efficiency of Day-Ahead Electricity Markets using a Multi Attribute Decision Making Approach", in *IEEE Int. Energy Conf.* (EnergyCon), 2016, Leuven (*accepted*).

17. M. Shafie-khah, P. Siano, Desta Z. Fitiwi, J. P. S. Catalão, E. Heydarian-Forushani, "Regulatory Support of Wind Power Producers against Strategic and Collusive Behavior of Conventional Thermal Units", in 13th *European Energy Market Conference* (EEM), 2016 (*accepted*).

Book Chapters

18. Desta Z. Fitiwi, Sérgio F. S., M. Shafie-khah, Bizuayehu A. W., and J. P. S. Catalão, "Optimal Sizing and Placement of Smart Grid Enabling Technologies for Maximizing Renewable Integration," in *Smart Energy Grid Engineering*, Elsevier, *In press*, 2015.

19. Sérgio F. S., Desta Z. Fitiwi, M. Shafie-khah, Bizuayehu A. W., and J. P. S. Catalão, "Introduction to Renewable Energy Systems," in *Optimization in Renewable Energy Systems*, Elsevier, *In press*, 2016.

Publications from M.Sc. Studies

20. Desta Z. Fitiwi and K. S. Rama Rao, "Neuro-Prony and Taguchi's Methodology-Based Adaptive Autoreclosure Scheme for Electric Transmission Systems," Power Delivery, IEEE Transactions on, vol. 27, no. 2, pp. 575 –582, Apr. 2012.

21. Desta Z. Fitiwi, K. S. Rama Rao, and T. B. Ibrahim, "A New Intelligent Autoreclosing Scheme Using Artificial Neural Network and Taguchi's Methodology," IEEE Transactions on Industry Applications, vol. 47, no. 1, pp. 306–313, Feb. 2011.

22. Desta Z. Fitiwi, K. S. Rama Rao, and T. B. Ibrahim, "A new intelligent autoreclosing scheme using artificial neural network and Taguchi's methodology," in 2010 IEEE Industrial and Commercial Power Systems Technical Conference (I&CPS), 2010, pp. 1–8.

23. Desta Z. Fitiwi and K. S. Rama Rao, "Autoreclosure in Extra High Voltage Lines Using Taguchi's Method and Optimized Neural Networks," in International Conference on Computer Engineering and Technology, 2009. ICCET '09, 2009, vol. 2, pp. 151–155.

24. Desta Z. Fitiwi and K. S. Rama Rao, "Assessment of ANN-based auto-reclosing scheme developed on single machine-infinite bus model with IEEE 14-bus system model data," in TENCON 2009 - 2009 IEEE Region 10 Conference, 2009, pp. 1–6.

25. Desta Z. Fitiwi, K. S. Rama Rao, and T. M. Baloch, "Autoreclosure in extra high voltage lines using Taguchi's method and optimized neural networks," in Electric Power Conference, 2008. EPEC 2008. IEEE Canada, 2008, pp. 1–7.

CURRICULUM VITAE

Desta Zahlay Fitiwi was born on June 17th, 1980 in Tigray, Ethiopia. He completed his pre-university education in Gerebno, Zalambesa and Agazi Schools in 1998. In 1999, he began studying his Bachelor's Degree at Addis Ababa University and in the following year, he joined the technology faculty of Addis Ababa University, where he completed his first degree in Electrical and Computer Engineering in 2005. He worked as a Supervisory Control and Data Acquisition (SCADA) and Adaptation Engineer in the Ethiopian Load Dispatch Center Project from 2005 to 2007. He received the M.Sc. degree in electrical and electronics engineering from Universiti Teknologi PETRONAS, Tronoh, Malaysia, in 2009.

In 2010, he was selected as a PhD candidate of the Erasmus Mundus Joint Doctorate in Sustainable Energy Technologies and Strategies (SETS) and awarded an Erasmus Mundus Fellowship. This joint PhD program is offered by Comillas Pontifical University, KTH Royal Institute of Technology of Stockholm and Delft University of Technology. He was a visiting PhD student in KTH from August 2012 to September 2013 as part of the mandatory mobility plan of the SETS program.

Desta actively participated in the realization of four EU-funded projects. He is currently working as a researcher at University of Beira Interior, Covilhã, Portugal. His research interests include regulation and economics of the power industry, transmission expansion planning, sustainable energy modeling and strategic planning, artificial-intelligence applications and mathematical optimizations for power systems.

Desta Zahlay Fitiwi

fitiwi@kth.se

March 2016

Madrid, Spain

_____ The End _____