UNIVERSIDAD PONTIFICIA COMILLAS

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA (ICAI)

OFFICIAL MASTER'S DEGREE IN THE
ELECTRIC POWER INDUSTRY

Master's Thesis

STOCHASTIC PRICING MODEL FOR THE IBERIAN POWER MARKET

Author: Vassiliki Boura
Supervisor: Antonio De Andrés
Co-Supervisor: Eva Porras

Madrid, September 2016

Master's Thesis Presentation Authorization

THE STUDENT:

Vassiliki Boura

……………………………………………………….

THE SUPERVISOR

Antonio de Andrés

Signed: ……………………… Date: 11/07/2016

THE CO-SUPERVISOR

Eva Porras

Signed: ……………………… Date: 11/07/2016

Authorization of the Master's Thesis Coordinator

Dr. Luis Olmos Camacho

Signed: ……………………… Date: ……/ ……/ ……

UNIVERSIDAD PONTIFICIA COMILLAS

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA (ICAI)

OFFICIAL MASTER'S DEGREE IN THE ELECTRIC POWER INDUSTRY

Master's Thesis

STOCHASTIC PRICING MODEL FOR THE IBERIAN POWER MARKET

Author: Vassiliki Boura
Supervisor: Antonio De Andrés
Co-Supervisor: Eva Porras

Madrid, September 2016

SUMMARY

The current thesis project studies the price forecasting procedures used nowadays by the electricity market and try to contribute by providing a stochastic pricing model that can be used for the Iberian power market as well as other power markets after certain adjustments. At first the project examines the past work that has been done in the field of pricing models, examining their different methodologies and assumptions and analyzing their strong as well as their less important points. Then the actual models used in this thesis projects are examined. The above mentioned study of the Iberian power market through a stochastic model was achieved by initially creating a simple deterministic pricing model, that level evolved into being able to examine multiple deterministic (independent) scenarios at the same time and finally reached its stochastic form, in which each scenario is accompanied by a corresponding probability and results were found examining different cases. Additionally, a simplified company representation was created in order to examine the effect of the above mentioned models in companies that own different thermal generation assets and operate at the same market. In that respect, three representative companies were created, owners of nuclear, coal and CCGT plants, and their incomes and productions were examined both for the multiple deterministic and the stochastic representation. Finally, ideas and suggestions for future improvements and expansions of the model are provided, in order to cover possible weaknesses and make it produce even more accurate and realistic results.

RESUMEN

La presente tesis estudia diferentes procedimientos para determinar previsiones de precios de mercados eléctricos y trata de ofrecer un modelo estocástico de determinación de precios con aplicación al mercado ibérico de electricidad así como a otros mercados tras la realización de determinados ajustes. En primer lugar el proyecto examina otros trabajos realizados con el fin de llevar a cabo previsiones de precio, examinando las diferentes metodologías y consideraciones y analizando su fortaleza así como los factores menos importantes. A continuación se examinarán los modelos actuales empleados en la presente tesis. El previamente mencionado modelo estocástico para la previsión de precios en el mercado ibérico de electricidad se ha conseguido tras haber desarrollado en primer lugar un modelo simple determinista que ha permitido examinar varios escenarios (independientes) al mismo tiempo hasta conseguir la forma estocástica, bajo la cual a cada escenario le acompaña una valor de probabilidad. Adicionalmente se desarrollo un modelo de compañía simplificado para examinar el efecto de los modelos previamente mencionados en compañías eléctricas que poseen plantas de generación térmicas y que operan en el mismo mercado. Se crearon tres compañías representativas dueñas de plantas de carbón, nucleares y ciclos combinados con el fin de analizar la energía producida e ingresos bajo una representación

determinística y estocástica. En ultimo lugar, se han propuesto ideas de mejoras y futuros trabajos del presente

ACKNOWLEDGEMENTS

I would first like to thank my thesis supervisors, Eva Porras and Antonio de Andrés of the Market Modeling department at Gas Natural Fenosa, were I did my thesis project. They have always been more than willing to help whenever I ran into a trouble spot or had a question about my research or writing. They consistently allowed this paper to be my own work, but steered me in the right the direction whenever they thought I needed it.

I would also like to acknowledge the professors of the IIT institute at Comillas Pontificia university as for their valuable contribution to my scientific background (especially regarding mathematical models), and I am gratefully indebted to their very valuable knowledge sharing that has been a great help for my project

Finally, I must express my very profound gratitude to my parents and to my brothers for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Author

Vassiliki Boura

CHAPTER ORGANISATION

The thesis project has been organized in the following chapters:

Chapter 1 – Introduction

- ➔ Purpose/Motivation of the thesis – Reference to preliminary works
- ➔ Thesis' objectives
- ➔ Structure of assignment

Chapter 2 - State of the art

- ➔ A survey of the literature (journals, conferences, book chapters) on the areas that is relevant to the research question.
- ➔ Original conclusions deriving from the above surveys.
- ➔ Gaps/problems to be overcome with this thesis.

Chapter 3: Problem setting, description, etc.

- ➔ Presentation of the problem
    - Description of system used
    - Data sources & initial assumptions

Chapter 4: Proposed Method (model, analysis, solution, etc.)

- ➔ Modeling : mathematical / review of optimization techniques
- ➔ Use of GAMS (General Algebraic Modeling System)
- ➔ Deterministic vs stochastic models
- ➔ Stochastic modeling of certain variables

Chapter 5: Results

- ➔ Scenarios to be examined
- ➔ Presentation-analysis of numerical results
- ➔ Comparison with theoretical expectations - comments

Chapter 6: Conclusions

- ➔ Summary of the problem, main findings and overview
- ➔ Feedback for further research & analysis

Contents

MASTER'S THESIS PRESENTATION AUTHORIZATION 1

SUMMARY 3

RESUMEN 3

ACKNOWLEDGEMENTS 5

CHAPTER ORGANISATION 6

CHAPTER 1 - INTRODUCTION 8

CHAPTER 2 – STATE OF THE ART 10

CHAPTER 3 – SYSTEM ANALYSIS 17

CHAPTER 4: METHODS & TECHNIQUES 26

CHAPTER 5 – RESULTS 40

CHAPTER 6 – CONCLUSIONS 55

BIBLIOGRAPHY 58

APENDIXES 59

CHAPTER 1 - INTRODUCTION

Since the early 1990s, two main events have been reshaping the landscape of the traditionally monopolistic and government-controlled power sectors – these are the process of deregulation and the introduction of competitive markets. In a number of countries worldwide, electricity is now traded under market rules using spot and derivative contracts. However, electricity is a very special commodity because of certain characteristics, not individually unique but quite uncommon when it comes to their combination: it is economically non-storable, and power system stability requires a constant balance between production and consumption. At the same time, electricity demand strongly depends on the weather conditions (temperature, wind speed, precipitation, etc.) as well as the intensity of business and everyday activities (on-peak vs. off-peak hours, weekdays vs. weekends, holidays and near-holidays, etc.). On the one hand, these unique and specific characteristics lead to price dynamics not observed in any other market, exhibiting seasonality at daily, weekly and annual levels, and also abrupt, short-lived and generally unanticipated price spikes. On the other hand all the above have encouraged researchers to intensify their efforts in the development of better forecasting techniques.

At the corporate level, electricity price forecasts have become a fundamental input to energy companies' decision-making mechanisms. As several worldwide events have shown, electric utilities are the most vulnerable, with the main reason being the fact that they generally have difficulty in transferring their costs on to the retail consumers. The costs of over-/under-contracting and then selling/buying power in the balancing (or real-time) market are typically so high that they can lead to huge financial losses or even bankruptcy. Also, extreme price volatility, which can be up to two orders of magnitude higher than that of any other commodity or financial asset, has forced market participants to hedge not only against volume risk but also against price movements. Price forecasts from a few hours to a few months ahead have become of particular interest to power portfolio managers and not only. A generator, utility company or large industrial consumer who is able to forecast the volatile wholesale prices with a reasonable level of accuracy can adjust its bidding strategy and its own production or consumption schedule in order to reduce the risk or maximize the profits in day-ahead trading.

A variety of methods and ideas have been tested regarding the *electricity price forecasting* (EPF), with varying degrees of success. This part of the thesis project aims to explain the complexity of the available solutions, with a special emphasis on the strengths and weaknesses of the individual methods.

Purpose / Motivation for thesis project

The importance and basic advantage of using stochastic models in the calculation of electricity pricing, compared to deterministic ones, is the ability to simultaneously examine different scenarios for the market, related to the variation of certain variables (stochastic variables) – this is a far more realistic representation of the market than the one offered by the deterministic model, as the real world is buffeted with stochasticity and cannot be described by a unique set of circumstances. What is more, and with regards to the electricity market, the variations that demand as well as non thermal generation present are nowadays even more important because of the increased penetration of RES technologies as well as a global economic crisis affecting recent demand patterns. Summing up, the pre-mentioned facts are indicative of the urgent need for further development of stochastic pricing models, through the use of which it will be possible to reach more realistic and significant results, to be used also as signals from generators and consumers.

The above constitute a major motivation for the current thesis project, whose practical applications are multiple and it can also be further evolved to include more complex cases, such as balancing markets , connected electricity markets or/and more technical constraints.

Objectives of the Master Thesis

The objectives of the thesis project are the followings:

1. Development of a stochastic model of electricity market prices in the medium – term horizon

   → Given the uncertainty of demand, fuel prices and renewable resources the goal is to develop a model using stochastic techniques to forecast price of electricity market.

The above will be achieved by performing the following steps:

- ✓ Development of an initial deterministic model, to be used as a basis for the development of the stochastic one. The system to be used in that stage will be simplified and based on the generation mix currently existing in the Spanish market.

- ✓ Review of the "state of the art" stochastic optimization techniques.

- ✓ Development of a stochastic model for forecasting electricity market prices in the medium - term in an environment of uncertainty.

After the completion of the above steps, the next objective is to:

2. Apply to an example of a simplified system with companies owners of different generation assets, with a data set that resembles the Spanish power market.

3. Commenting on the final results and proposals about future improvements and expansions of the model in order to depict even more accurately the evolution of prices and help them work as a signal for future uses.

CHAPTER 2 – STATE OF THE ART

The most significant change regarding electricity markets in the last decades has been the start of liberalization of the electric market; prices are now determined on the basis of contracts on regular markets and their behavior is mainly driven by usual supply and demand forces. Because of the importance of their evolution, a large body of literature has been developed in order to analyze and forecast them: it includes works with different aims and methodologies depending on the temporal horizon being studied (short, medium or long term).

As a conclusion, from the studying of the different kinds of models, six fundamental points arise: the peculiarities of electricity market, the complex statistical properties of prices, the lack of economic foundations of statistical models used for price analysis, the primacy of uni-equational approaches, the crucial role played by demand and supply in prices determination and the lack of clear-cut evidence in favor of a specific framework of analysis.

As it has been analyzed above, the importance of a stochastic pricing model is extremely big for various reasons. Specifically referring to the electricity market, whose volatility concerning demand as well as RES production can be significantly high, a stochastic approach is more than necessary. It allows to examine a number of different scenarios simultaneously, giving them the corresponding probability – the choice of that parameter is crucial, since it should be somehow matched to the real data of the market.

2.1. General Contributions

Except for a few isolated cases, EPF (Electricity Price Forecasting) publications did not appear in the literature before the year 2000. The next major breakthrough occurred in the years 2005 and 2006, when the number of publications first doubled and then tripled with respect to 2002–2004 figures. Initially, this increased inflow of EPF publications was due mostly to proceedings or conference papers; journal articles followed with a delay. The overall publication rate increased until 2009/2010 and then dropped to 2006–2008 levels because of a reduced number of conference papers. As of 2013, the topic seems to have saturated the research community, although the number of citations is still increasing. Possibly a new fundamental impulse–like the deregulation of the late 1990s or the increased volatility of electricity spot prices in the mid-2000s–is needed in order to propel electricity price forecasting to a new level of publication intensity.

EPF is mostly addressed in the three books presented as follows:

- **Shahidehpour et al.** : discuss the basics of electricity pricing and forecasting

  (price formation, volatility, exogenous variables), describe a price forecasting module based on neural networks, and comment on performance evaluation.

- **Weron** : provides an overview of modeling approaches, then concentrates on practical applications of statistical methods for day-ahead forecasting (ARMA-type, ARMAX, GARCH-type, regime-switching), discusses interval forecasts,

  and moves on to quantitative stochastic models for derivatives pricing (jump-diffusion models and Markov regime-switching).

- **Zareipour** : begins by reviewing linear time series models (ARIMA, ARX,

  ARMAX) and nonlinear models (regression splines, neural networks), then uses them for forecasting hourly prices in the Ontario power market.

There are a few more books which touch upon the topic of electricity price forecasting, but they generally concentrate on modeling the stochastic price dynamics for risk management and derivatives valuation, rather than on day-ahead price forecasting; some examples Benth et al., 2008, Bunn, 2004, Burger et al., 2007, Eydeland and Wolyniec, 2003, Fiorenzani, 2006, Huisman, 2009, Keppler et al., 2007 and Lewis, 2005, and Weber (2006). There is also a recent monograph by Yan and Chowdhury (2010a), based on the master's thesis of the first author, but it considers only mid-term electricity price forecasting, with a time frame of between one and six months. Although mid-term EPF is important for resource reallocation, maintenance scheduling, bilateral contracting, budgeting and planning purposes, it is beyond the few hours to few days ahead forecasting horizons that are typically considered in the EPF literature.

Regarding review and survey articles, the situation looks a little better: the first review papers were already being published in the early 2000s. In an invited paper that appeared in the *Proceedings of the IEEE,* Bunn (2000) reviews some of the main

In an *IEEE Power & Energy Magazine* discussion article on real-world market representation with agent-based models, Koritarov (2004) argues that the purpose of ABM is not necessarily to predict the outcome of a system; rather, it is to reveal and explain the complex and aggregate system behaviors that emerge from the interactions of the heterogeneous individual entities. At the same time, he concludes that the ABM approach is positioned well for performing short- and long-term electricity price forecasting, resource forecasting and asset valuation. Unfortunately, he does not provide any examples of EPF applications of ABM. Weidlich and Veit (2008) also fail to find any examples of EPF in a survey of agent-based wholesale electricity market models in *Energy Economics*.

In another *IEEE Power & Energy Magazine* discussion article, Amjady and Hemmati (2006) explain the need for short-term price forecasts, review problems related to EPF, and put forward proposals for such predictions. They argue that time series techniques (AR, ARIMA, GARCH) are generally only successful "in the areas where the frequency of the data is low, such as weekly patterns. Furthermore, they advocate the use of artificial (or computational) intelligence and hybrid approaches (neural networks, fuzzy regression, fuzzy neural networks, cascaded architecture of neural networks, and committee machines), which are "capable of tracking the hard nonlinear behaviors of hourly load and especially price signals". In a later publication, Amjady (2012, Chapter 4) briefly reviews EPF methods, then focuses again on artificial intelligence-based methods, and in particular feature selection techniques and hybrid forecast engines. He also discusses forecast error measures, the fine tuning of model parameters, and price spike predictions.

In the year 2009, two similar survey articles, co-authored by the same three researchers, appeared in parallel in the *International Journal of Electrical Power and Energy Systems* and the *International Journal of Energy Sector Management*. Aggarwal, Saini, and Kumar (2009a) review 47 'time series' and 'neural network' papers published between 1997 and 2006 in terms of the model type and architecture, forecast horizon(s), model input and output variables, preprocessing and datasets used. They conclude that "there is no systematic evidence of out-performance of one model over the other models on a consistent basis", which may be attributed to "the large differences in price developments (...) in different power markets". In a more recent–in terms of the publications reviewed–article, Aggarwal, Saini, and Kumar (2009b) also compare 'time series' and 'neural network' papers. They classify EPF models as falling into one of three categories (although differently from Aggarwal et al., 2009a): heuristics (naïve, moving average), simulations (production cost and game theoretical) and statistical models, where the last category–somewhat surprisingly–includes both time series (regression) and artificial intelligence models. They expand the analysis to include quantitative comparisons of **(i)** the forecasting accuracy and **(ii)** the computational speed of different forecasting techniques. In our opinion, the value of **(i)** is disputable. Even if the forecasting accuracy is reported for the same market and the same out-of-sample (forecasting) test period, the errors of the individual methods are not truly comparable if different in-sample (calibration) periods are used. Moreover, the implementation of the algorithms differs between software packages, and is generally very sensitive to the initial conditions in the case of nonlinear or multi-parameter models. It may be impossible to replicate the results, even given the exact model structure, as was reported by Weron (2006) for the case of the multi-parameter transfer function (ARMAX) model of Nogales, Contreras, Conejo, and Espinola (2002). On the other hand, a table with the computation speeds of different forecasting techniques is interesting. Unfortunately, though, it cannot be used to draw quantitative conclusions, due to the differences in processors used, software implementations, calibration periods, etc. Finally, Aggarwal et al. (2009b) conclude that "there is no hard evidence of out-performance of one model over all other models on a consistent basis" and that longer "test periods of one to two years should be used". We cannot argue with these conclusions.

In a recent survey article published in the *IEEE Signal Processing Magazine*, Chan et al. (2012) review neural networks, support vector machines, time series models (ARMA, ARMAX, GARCH), and functional principal component analysis (FPCA) models for electricity prices/load, wind and solar forecasting. They advocate the use of multivariate factor models, and especially of the robust FPCA, which is shown to outperform both the standard FPCA and an AR model with a time varying mean in a limited forecasting study.

In a chapter in the *Wiley Encyclopedia of Electrical and Electronics Engineering*, Garcia-Martos and Conejo (2013) review short- and medium-term EPF, with a focus on time series models. Specifically, they consider ARIMA and seasonal ARIMA models calibrated to hourly prices for day-ahead predictions, and vector ARIMA (essentially VAR) and unobserved component (i.e., factor) models for medium-term horizons. Sadly, in the most novel part on factor models, the authors limit the discussion to their own approach (Garcia-Martos et al., 2011 and Garcia-Martos et al., 2012), and neither review nor compare other relevant publications. Interestingly, though, the chapter includes an introduction to the computation of prediction intervals, a topic which is addressed very rarely in the EPF literature.

In a short review article, Hong (2014) briefly discusses spatial load forecasting, short-term load forecasting, EPF, and two 'smart grid era' research areas: demand-response and renewable-generation forecasting. He classifies EPF models into three groups: simulation methods (which require a mathematical model of the electricity market, load forecasts, outage information, and bids from market participants), statistical methods, and AI methods. Perhaps the most important contribution of the paper is that the author emphasizes the need for rigorous out-of-sample testing of the different methods proposed in the literature.

In the most recent survey of structural models, published as a chapter in the book *Quantitative Energy Finance*, Carmona and Coulon (2014) present a detailed analysis of the structural approach for electricity modeling, emphasizing its merits relative to traditional reduced-form models. Building on several recent articles, they advocate a broad and flexible structural framework for spot prices, incorporating demand, capacity and fuel prices in several ways, while calculating closed-form forward prices throughout.

The above-mentioned articles, book chapters and Ph.D. thesis are complemented by a few survey conference papers of varying quality. Niimura (2006) studies over 100 papers and classifies them as either simulation models (production cost and game theoretical) or statistical models (which again include time series, regression, and artificial intelligence models). Haghi and Tafreshi (2007) construct a different classification in which they categorize 'time series' models as either 'stationary' (including ARIMA, ARIMA-Wavelet, ARX and ARMAX models) or 'non-stationary' (including neural networks, regime-switching models, GARCH, jump-diffusions and mean-reversion models). This is a very confusing classification, as some of the 'stationary' models are non-stationary in a statistical sense (for instance, ARIMA), while some of the 'non-stationary' models are stationary (for instance, mean-reversion models) Daneshi and Daneshi (2008) consider over 100 papers and classify them as time series models, neural networks, fuzzy set models, fuzzy neural networks and other techniques. Similar in scope are the papers of Hu, Taylor, Wan, and Irving (2009) and Negnevitsky, Mandal, and Srivastava (2009), together with the more recent survey of Cerjan, Krzelj, Vidak, and Delimar (2013).

To start with, the stochastic pricing of the electricity market has been studied by institutions like the MIT, in whose study an interconnected, multi-market system was taken into account. This is by far a more complex system than the one examined in the current thesis project and the conclusions deriving from this paper are mainly related to the effect that the connection among markets has on prices. In the "An Ambit Stochastic Approach to Pricing Electricity Forward**"**, of Luca Di Persio and Isacco Perin (2015), the writers acknowledge the need for including stochasticity in the pricing of the electricity market – starting from the market liberalization and moving on to the creation of new markets, e. g the NordPool, they state that "There is no doubt that such markets will play a vital role in the future given the constant expansion of global demand for energy. From the financial point of view, standard products traded on energy markets are spot prices, forward and futures contracts, and options written on them."

An overview of all the candidate models suitable to describe the features of the electricity market is provided by Misiorek, Trueck, and Weron (2006). The aim of their paper is to assess the short-term point and interval forecasting performance of different time series models of the electricity spot market during normal (calm), as well as extremely volatile, periods. Since the authors want to mimic a typical practitioner *praxis*, adopting a truly real time forecasting approach, they chose as test ground the California power market, that offers freely accessible high quality electricity price and load data; moreover this is a quite interesting market, since it provides the ideal framework for studying those behaviors typically leading to a market crash (really occurred in winter 2000/2001)

2.2. The State of the Art - Summarizing Major and Minor Issues

Six fundamental points arise from the analysis of the theoretical and empirical literature on electricity prices:

- The electricity market retains absolutely peculiar characteristics: it is an auction market that, although liberalized, is not strictly a spot one, but it requires both price and quantity of equilibrium to be defined one day in advance on the basis of expected supply and demand. This guarantees a good match among supply and demand that due to the non-storability of electricity, to unexpected peaks in demand and to congestions over the distribution network, could fail, causing jumps in prices and leading in extreme cases to the system blackout.

- The series of electricity prices have complex statistical properties that vary depending on spectral frequency to which data are measured and on sample size. Depending on the cases, it is possible to notice phenomena of seasonality at different frequencies, trends which are more or less linear at low frequencies, phenomena of auto-correlated volatility at high frequencies, and combinations of outliers apparently managed by non standard distributions.

- A wide range of models dedicated to the analysis of the properties of price series follow an approach that can be defined as being agnostic from the point of view of economic interpretation, meaning they do not foster the inference on (economic) factors that influence prices, but they limit the analysis to only their statistical properties.

- Despite the above, it seems evident that the evolution of prices over time is driven by the interaction between supply and demand of electricity, that is, from two phenomena not directly measurable and in some way latent. Therefore, in order to effectively model demand and supply it would be suitable to include in the model those factors that determine their trend: for example, climatic factors or the business cycle state that affect demand; productivity, size of the plant and costs of production concerning supply. It is a insidious approach, as these determinants play a role at different frequencies and usually statistical data on them are characterized by significant measurement errors, which makes more difficult the correct identification of the effects caused by each phenomenon on prices.

- Even for the hidden dangers previously mentioned, the econometric models dedicated to the analysis of electricity prices adopt very simplified specifications, often uni-equational, taking into account only a few aspects of the issue at a time.

- Among the models proposed by the literature, none of them seems to be characterized by a uniformly better capability of fitting the data and by an outperforming forecasting behavior; depending on the market taken as reference, on the sample of data being considered and on the measure of forecasting performance chosen, now prevail very simple autoregressive models, whereas other times Markov switching models with changing regimes.

2.3. Gaps / Problem to be overcome by this thesis project

By thorough examination and careful choice among the scenarios that were used so as to test both possible extreme situations and also the normal ones in terms of demand and also renewable production, this thesis project is a model that will provide the user with the ability to work with multiple scenarios of different probabilities (that can be adjusted depending on the kind of analysis the user wants to provide) and obtain results that correspond to the real life ones – this can be proven valuable for example for companies who wish to realize the impact of the unit commitment decisions to final results and adjust their bidding decisions accordingly.

As mentioned before, at the corporate level electricity price forecasting has become a fundamental input to energy companies' decision-making mechanisms. The high level of vulnerability that the companies face because of the variations in the prices is an important factor in determining their (quite often high) costs. Moreover, specifically referring to the Iberian electricity market, the great amount of CCGT plants, the lower cost of coal plants and also the increased availability of hydro and wind production at certain year periods lead to great diversification of the prices, increasing the level of difficulty for model developers to obtain realistic and interpretable results. In the current thesis project there has been an effort to capture this effect to the final results, and also to provide a deep analysis of them in correlation to extreme variations in the thermal demand estimation – by this we refer to the diversity of the scenarios used, that have been selected through clustering of an initial big amount of scenarios. Finally, for greater analysis purposes, model companies of different generation assets have been created and their contribution to the final prices has been examined and explained. All these will be further discussed at chapter 5.

CHAPTER 3 – SYSTEM ANALYSIS

Description of system used

In the current chapter, the thermal generation system used for the model is presented and analyzed – this includes description of the demand data, the power plants used, their selected characteristics and the inclusion of any other parameter important within the system settings (e. g fuel and plant cost).

The sub-chapters to be analyzed are:

- Demand
- Power plants (capacities, costs, maintenance)

DEMAND

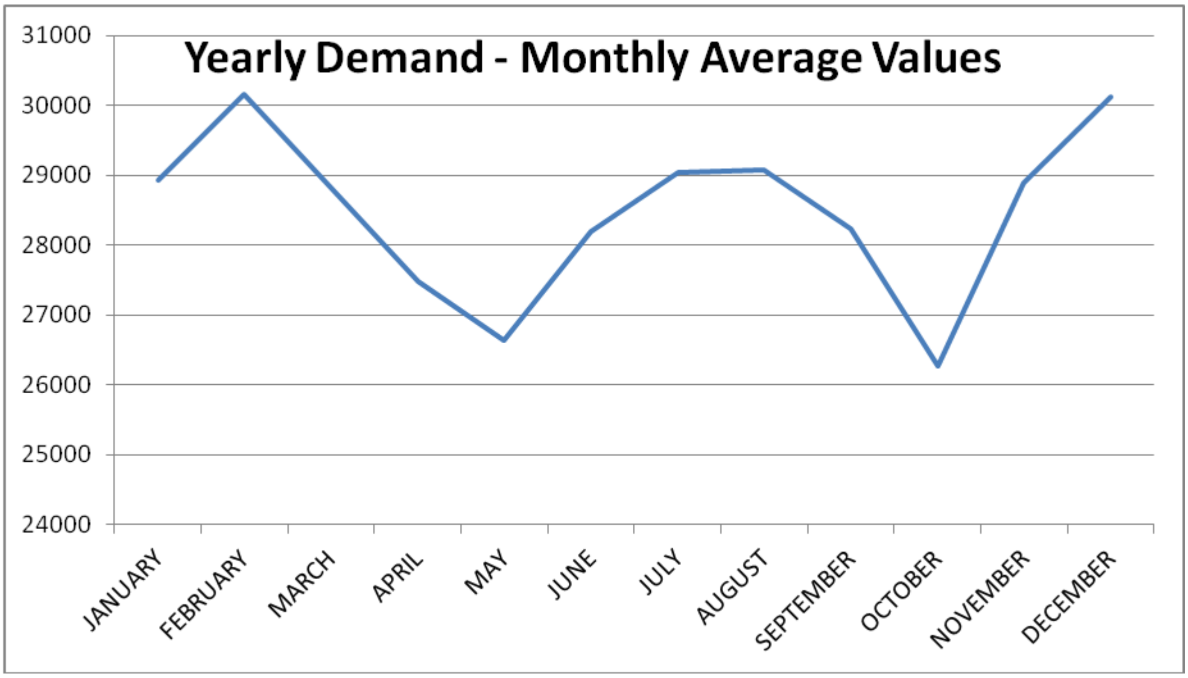
The final model that was created is formulated in 3 different stages – these include a simple deterministic model with a single scenario, the same model but running independently and simultaneously a number of scenarios and finally a probabilistic model that uses different probabilities for a number of scenarios and runs them all together. While the demand used in each one of these cases is a bit differentiated (mainly in terms of data presentation, for example periods instead of single hours used), the basis is the same and it includes hourly data for the time horizon between January 2017–December 2017. These data were provided by Gas Natural Fenosa, and are mainly predictions calculated from available historical data. The analytical demand data table is given at the end of the thesis project (Appendix).

In the next step certain characteristics of the demand will be examined such as its peak and lowest value, comparison with the installed system capacity and also a monthly representation that can provide greater insight to the variations that the final prices presented after running the scenarios.

Also, there will be a preliminary reference to the clustering procedure that took place regarding the demand as well wind and solar production scenarios from the corresponding technologies. This clustering was necessary in order to decrease the initial number of available scenarios from 200 down to 5 that could be used for the purposes of the model. More detail about this process is provided in chapter 5.

As it will be mentioned in the next sub-chapter the total installed system capacity (including nuclear, coal and CCGT plants) is around 36 GW, in comparison to the peak demand that is 30 GW. More details regarding the exact percentage each technology covers will be also given to the next sub-chapter.

In the diagram following a monthly representation of the demand is given, where its lowest and highest values can be seen. (values of the vertical axis are provided in MW). That is:

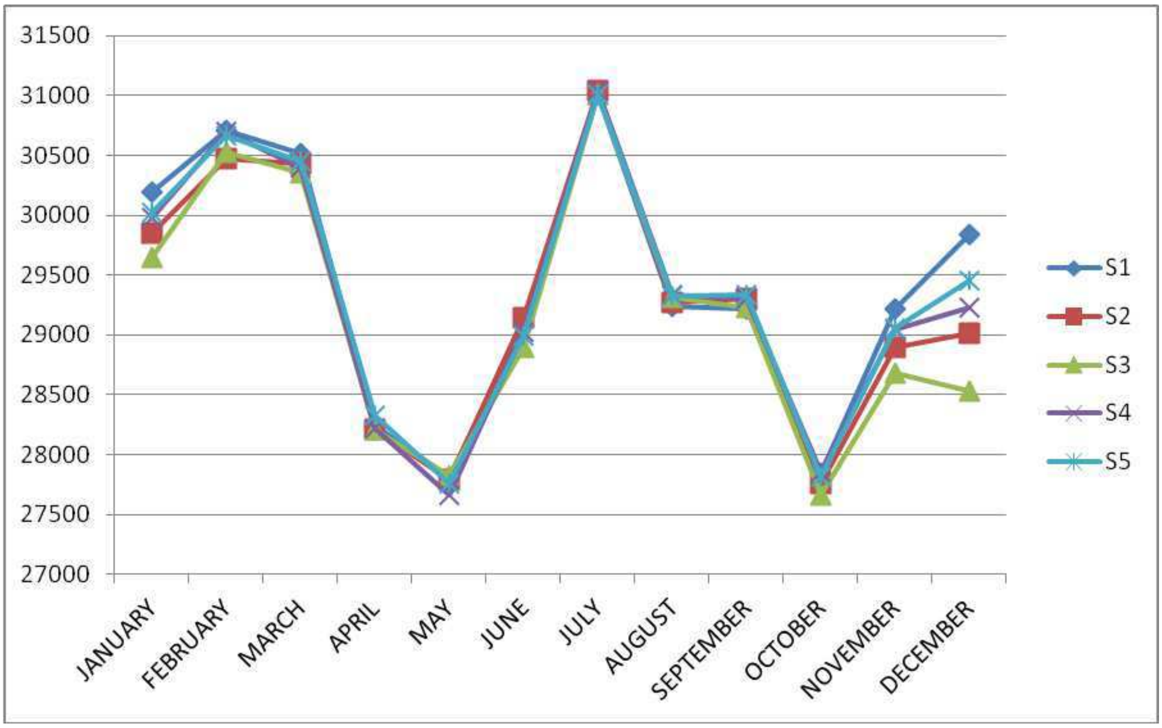


Monthly average representation of annual demand

As expected, demand presents its highest value around February, while it reaches its lowest at October. Summer months also present a relatively high demand value, while spring and autumn have with a significantly lowered demand, as is reflected in the diagram.

Regarding the clustering procedure that will be analyzed in detail in chapter number 5, this process took place in order to obtain a limited number of different final scenarios (in terms of demand as well as wind and solar production) that could be used, followed by their corresponding probabilities so as to test the stochastic model that has been the purpose of the thesis. Specifically, it was considered essential to take into account medium demand scenarios that follow the Norma of a usual system as well as extreme situations meaning extremely high or low demand values. The initial number of the 200 scenarios was obtained through a suitable software called ALEA, and was later clustered through use of XLSTAT (excel software application) in order to obtain five final scenarios – the clustering took account of similar values among scenarios and created groups of common characteristics, followed by their corresponding probabilities.

The next diagram presents monthly values for the final five (demand) scenarios that were used in our model. It is important to mention that for the initial deterministic model(s) the demand shown in diagram () was used.



Final demand scenarios – probabilistic model

The main differences among these scenarios are significantly obvious in high demand months such as January and December. For example, in December scenarios 1,3 present a difference of 5% in their corresponding values. The fact of extreme variations appearing mainly in high and low demand months is logical because the clustering procedure has to deal with realistic scenarios, meaning that not extreme variations exist in the non peak-bottom values.

POWER PLANTS

Capacities

Three different kinds of thermal power plants compose the system under examination - nuclear plants, coal plants and CCGT plants. For these three categories, fixed values that correspond to real ones were decided and are shown in the following matrix:

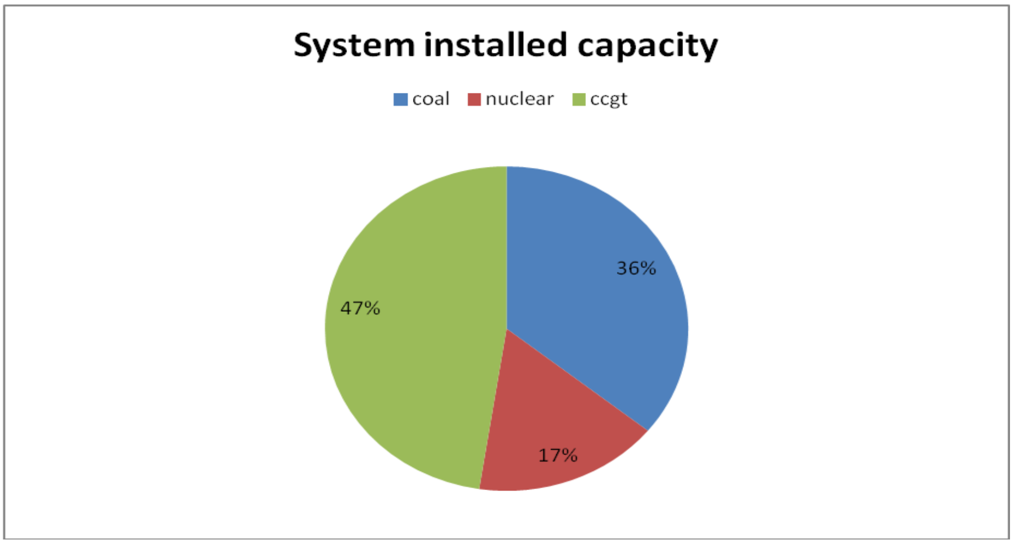
Plants – Capacities	KIND N.1	KIND N.2	KIND N.3
Nuclear	1000	-	-
Coal	200	350	550
CCGT	400	800	-

As one can notice, all selected nuclear plants have a fixed capacity of 1000 MW, which is also the actual value of nuclear plants in a thermal system. On the other hand, for coal and CCGT plants it was decided to use a certain variety over their capacities, and for this reason we chose to include plants of 200, 350 & 550 MW in the case of coal plants and 400 & 800 MW for CCGT ones. This gave certain flexibility to our system, since demand could be satisfied by different kinds and bigger number of plants, which provides support and stability to the system. Apart from this fact, this variation also affects the final prices because of different marginal plants within the year representation of the model, which also accounts for different final monthly average values.

As it is known, thermal power plants operate within a certain limit, where the maximum is usually set as their capacity (or a small percentage above it, e. g 105% of their nominal capacity), and the minimum is an important thermal limit, below which the operation of plants presents a lot of problems and can prove harmful to the system. Because of this, it is extremely important to set the minimum technical for the operation of the plants, and use this limitation as a constraint to the mathematical formulation of the operation of the system, to ensure a stability at all times. The technical limits, as set for each one of the above kinds of plants, are shown in the following matrix:

PLANTS-TECHNICAL MINIMUM	**KIND N.1**	**KIND N.2**	**KIND N.3**
Nuclear	800	-	-
Coal	100	150	170
CCGT	175	350	-

Regarding general characteristics of the system, the total number of thermal plants it includes is 70, with a total installed thermal capacity equal to 36,2 GW. As it is mentioned above, the maximum thermal demand appearing in the system is approximately 30 GW, so this installed capacity responds well to covering the demand plus a security limit (almost 20% above maximum demand). Out of this installed capacity, 13 GW are produced by coal plants, 6 GW come from nuclear plants and the remaining almost 17 GW from CCGT plants. Graphically, these are shown as follows:



Finally, regarding the characterisation of the plants the following coding was applied:

- Nuclear plants are all adressed as NUC, following by a number (NUC1, NUC2 etc). They are 6 in total.

- For coal plants, the three different kinds are named as CA, CB, CC, followed by their corresponding number as in the case of nuclears. In total they are 29, out of which 4 belong to the 'CA' category, 10 to the 'CB' one and the remaining 15 are of 'CC' type.

- Lastly, CCGT plants are under the coding CGA & CGB, again followed by an indicative number. Out of the total 35 CCGT plants 8 belong to the second category and 27 to the first one.

A matrix containing all the thermal plants of the system, with their corresponding capacities and technical limits can be found at the Annex at the end of the project.

Costs

Moving on to one of the most important parameters of the system that is the thermal plants cost, in this subchapter the detail calculation of cost for all plants is given and any assumptions made within it are thoroughly explained. The estimation of the plants´ cost is crucial since through them the model is able to provide us with the system price per period, equal to the price of the last plant participating in covering the system demand (marginal plant). Each one of the three categories of the thermal plants the system contains will be analyzed.

1. NUCLEAR PLANTS

These kind of plants only present a fixed cost (that is as a general rule quite high compared to other kinds of plants), but on the contrary a variable cost almost equal to zero. For simplification reasons, it was considered equal to zero for all the hours, throughout the examined yearly horizon, and for all the plants.

For both coal & CCGT plants, the estimation of their cost was done through calculating the fuel cost, and changing certain parameters to diversify among the different plants and provide a variety to the system as far as final price is concerned. Specifically:

2. COAL PLANTS

The calculation of fuel cost is provided for the 12-month period starting at January 2016 and finishing at December 2016. The results for the different parameters of fuel cost as well as total cost, **without including any premiums**, are presented right below:

COAL PLANTS	Jan	Feb	Mrh	Ap	May	Jn	Jl	Aug	Sep	Oct	Nov	Dec
API2 Index ($/t)	45,4	43,5	45,95	45,3	44,15	44,15	41,9	41,9	41,9	41,1	41,1	41,1
Cambio	1,09	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,09	1,08	1,08	1,09
Efficiency	0,35	0,35	0,35	0,35	0,35	0,35	0,35	0,35	0,35	0,35	0,35	0,35
Electrical tax	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,07
O&M	2,5	2,5	2,5	2,5	2,5	2,5	2,5	2,5	2,5	2,5	2,5	2,5
Transport cost	0	0	0	0	0	0	0	0	0	0	0	0
Emission factor	0,97	0,97	0,97	0,97	0,97	0,97	0,97	0,97	0,97	0,97	0,97	0,97
Green tax	0,65	0,65	0,65	0,65	0,65	0,65	0,65	0,65	0,65	0,65	0,65	0,65
EUA	6,8	5,15	5,14	5,07	5,07	5,07	5,07	5,07	5,07	5,07	5,07	5,12
OMIP	43,65	39,25	41,94	40,29	39,24	42,81	46,19	43,73	43,73	44,72	39,53	43,41
Clean Dark Spread	19,98	18,05	19,83	18,49	17,87	21,49	25,7	23,24	23,06	24,21	19,02	22,99
Dark Spread	26,6	23,06	24,84	23,43	22,81	26,42	30,64	28,18	27,99	29,14	23,95	27,97
CARBON COST	17,05	16,19	17,1	16,86	16,43	16,39	15,55	15,55	15,74	15,58	15,58	15,44
CARBON COST & EM.	23,67	21,2	22,11	21,8	21,37	21,32	20,49	20,49	20,67	20,51	20,51	20,42
TOTAL COST	35,15	32,51	33,48	33,15	32,69	32,64	31,75	31,75	31,95	31,78	31,78	31,68
MARGINAL COST	8,5	6,74	8,46	7,14	6,55	10,17	14,44	11,98	11,78	12,94	7,75	11,73

Each one of the above parameters in the cost calculation process is used is analyzed right above:

API2 Index ($/t) → The API 2 index is the benchmark price reference for coal imported into northwest Europe. It is calculated as an average of the Argus cif ARA assessment and the IHS McCloskey NW Europe Steam Coal marker. The Argus component of the price is published each day in Argus Coal Daily International. The

Argus/McCloskey's Coal Price Index report is published every Friday and includes daily, weekly and monthly average prices. Daily prices are also available as a data feed. The API 2 price is the primary price reference for physical and over-the-counter (OTC) coal contracts in northwest Europe. Some 90% of the world's coal derivatives are priced against the Argus/ IHS McCloskey API 2 and API 4 indexes. The Argus/McCloskey's Coal Price Index service also includes the API 4 index, which is the benchmark for coal exported out of Richards Bay in South Africa. For the purposes of the thesis, the API2 index prices estimations for the year 2016 were found at (??)

Exchange (cambio) → Change is the concept that denotes the transition occurs from one state to another, for example: the concept of change of state of matter in physics (solid, liquid and gas) or persons in marital status (single, married , divorced or widowed); or crises, or revolutions in any field studied by the social sciences, particularly history, which can be defined as the science of change.

Efficiency → For the case of coal plants a level of efficiency equal to 35% is considered, which reflects the level of efficiency of the fuel used in regards to the final outcome

Electrical tax → Set accordingly to current market values equal to 7%, fixed throughout the year

O&M → Operation and maintenance costs. Set as 2,5 () for every coal and CCGT plant, fixed within the yearly horizon

Emissions factor → It corresponds to the level of CO_2 emissions of coal plants, which is generally very high, in our case a value of 97% is assumed

Green Tax → Tax paid by consumers for products or services that are not environmentally friendly. Intended purpose of the green tax is to offset the negative impact resulting from the use of non-green products and services.

EUA → EU allowance (EUA) refers to the carbon credits traded under EU emission trading scheme. One EUA represents one ton of CO2 that the holder is allowed to emit. Allowances are freely allocated to firms which can be traded in carbon market. The firms should surrender EUAs equivalent to their emissions at the end of each compliance period. Those companies that emit more than their permitted allowance has to buy the extra allowances from the open market, while those firms that emit less can sell the balance allowance units to those firms that are in need of the same.

Dard Spread → The spark spread is the theoretical gross margin of a gas-fired power plant from selling a unit of electricity, having bought the fuel required to produce this unit of electricity. All other costs (operation and maintenance, capital and other financial costs) must be covered from the spark spread. Dark spread is the spark spread for coal. Instead of betting on the difference between the prices of gas and electricity, dark spread bets on the difference of coal and electricity prices.

Clean dark spread or **dark green spread** → is dark spread adjusted for the cost of carbon emissions

The formula used for the calculation of the total in the case of coal plants is the following:

T.C = (*carbon cost&emissions* + (green tax/efficiency)*3,6 + O&M) * (1 + electrical tax)

The carbon cost, including emissions, is calculated based on the next formula :

C.C = (((API$_2$ /Cambio+Transport Cost)/6000)*(1+ef./4,1868*3,6))*1000 + (EUA*emiss factor)

For the final formation of the plants' costs a premium was considered additionally to the above formula. For the case of coal plants the three different premiums considered were 0, 15 and 30, and the final results after including the premium are shown to the next matrix :

	Jan	Feb	Mrh	Apr	May	Jn	Jl	Aug	Sep	Oct	Nov	Dec
Premium 1	0	0	0	0	0	0	0	0	0	0	0	0
CARBON COST	17,05	16,19	17,1	16,86	16,43	16,39	15,55	15,55	15,74	15,58	15,58	15,44
plus emissions	**23,67**	**21,2**	**22,11**	**21,8**	**21,37**	**21,32**	**20,49**	**20,49**	**20,67**	**20,51**	**20,51**	**20,42**
Premium 2	15	15	15	15	15	15	15	15	15	15	15	15
CARBON COST	22,69	21,78	22,69	22,45	22,02	21,96	21,12	21,12	21,37	21,27	21,27	21,07
plus emissions	**29,3**	**26,78**	**27,69**	**27,38**	**26,95**	**26,89**	**26,06**	**26,06**	**26,31**	**26,2**	**26,2**	**26,06**
Premium 3	30	30	30	30	30	30	30	30	30	30	30	30
CARBON COST	28,32	27,36	28,27	28,03	27,6	27,53	26,69	26,69	27,01	26,96	26,96	26,71
plus emissions	**34,93**	**32,37**	**33,27**	**32,96**	**32,53**	**32,46**	**31,62**	**31,62**	**31,94**	**31,89**	**31,89**	**31,69**
Premium 4	30	30	30	30	30	30	30	30	30	30	30	30
CARBON COST	28,32	27,36	28,27	28,03	27,6	27,53	26,69	26,69	27,01	26,96	26,96	26,71
plus emissions	**34,93**	**32,37**	**33,27**	**32,96**	**32,53**	**32,46**	**31,62**	**31,62**	**31,94**	**31,89**	**31,89**	**31,69**
Premium 5	30	30	30	30	30	30	30	30	30	30	30	30
CARBON COST	28,32	27,36	28,27	28,03	27,6	27,53	26,69	26,69	27,01	26,96	26,96	26,71
plus emissions	**34,93**	**32,37**	**33,27**	**32,96**	**32,53**	**32,46**	**31,62**	**31,62**	**31,94**	**31,89**	**31,89**	**31,69**

3. *CCGT PLANTS*

The calculation of gas cost is provided for the 12-month period starting at January 2016 and finishing at December 2016. The results for the different parameters of gas cost as well as total cost, **without including any premiums**, are presented right below:

GAS PLANTS	Jan	Feb	Mrh	Ap	May	Jn	Jl	Aug	Sep	Oct	Nov	Dec
Efficiency	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5
Transport cost	2	2	2	2	2	2	2	2	2	2	2	2
O&M	2,5	2,5	2,5	2,5	2,5	2,5	2,5	2,5	2,5	2,5	2,5	2,5
ATR & Transport	3,4	3,4	3,4	3,4	3,4	3,4	3,4	3,4	3,4	3,4	3,4	3,4
Electrical tax	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,07
Green tax	0,65	0,65	0,65	0,65	0,65	0,65	0,65	0,65	0,65	0,65	0,65	0,65
Emission factor	0,36	0,36	0,36	0,36	0,36	0,36	0,36	0,36	0,36	0,36	0,36	0,36
EUA	6,8	5,15	5,14	5,07	5,07	5,07	5,07	5,07	5,07	5,07	5,07	5,12
Reference price gas	14,8	13,4	12,1	12,1	13,4	14,1	12,8	13,8	14	15,9	13,47	14,79
OMIP	43	39,2	41,9	40,2	39,2	42,8	46,1	43,7	43,7	44,7	39,5	43,4
Clean Spark Spread	4,64	3,75	8,94	7,37	3,68	5,91	11,8	7,4	7,03	4,15	3,97	5,18
Spark Spread	7,09	5,6	10,79	9,2	5,51	7,74	13,66	9,22	8,86	5,98	5,8	7,02
GAS COST	36,56	33,65	31,15	31,09	33,73	35,07	32,53	34,51	34,87	38,74	33,73	36,39
GAS COST & EM.	39,01	35,5	33	32,92	35,56	36,9	34,35	36,33	36,7	40,57	35,56	38,23
TOTAL COST	49,95	46,2	43,52	43,43	46,26	47,69	44,97	47,08	47,47	51,62	46,25	49,11
MARGINAL COST	-6,3	-6,95	-1,58	-3,14	-7,02	-4,88	1,22	-3,35	-3,74	-6,9	-6,72	-5,7

Following the same procedure as in the case of coal plants, we consider certain premiums for the CCGT plants, this time equal to 0, 10 and 15.

After including the considered premiums for each one of the five supposed plant categories, we obtain the following final results:

	Jan	Feb	Mrh	Apr	May	Jn	Jl	Aug	Sep	Oct	Nov	Dec
Premium 1	0	0	0	0	0	0	0	0	0	0	0	0
GAS COST	36,56	33,7	31,2	31,1	33,73	35,1	32,5	34,5	34,9	38,7	33,7	36,39
plus emissions	**39,01**	**35,5**	**33**	**32,9**	**35,56**	**36,9**	**34,4**	**36,3**	**36,7**	**40,6**	**35,6**	**38,23**
Premium 2	5	5	5	5	5	5	5	5	5	5	5	5
GAS COST	46,56	43,7	41,2	41,1	43,73	45,1	42,5	44,5	44,9	48,7	43,7	46,39
plus emissions	**49,01**	**45,5**	**43**	**42,9**	**45,56**	**46,9**	**44,4**	**46,3**	**46,7**	**50,6**	**45,6**	**48,23**
Premium 3	10	10	10	10	10	10	10	10	10	10	10	10
GAS COST	56,56	53,7	51,2	51,1	53,73	55,1	52,5	54,5	54,9	58,7	53,7	56,39
plus emissions	**59,01**	**55,5**	**53**	**52,9**	**55,56**	**56,9**	**54,4**	**56,3**	**56,7**	**60,6**	**55,6**	**58,23**
Premium 4	10	10	10	10	10	10	10	10	10	10	10	10
GAS COST	56,56	53,7	51,2	51,1	53,73	55,1	52,5	54,5	54,9	58,7	53,7	56,39
plus emissions	**59,01**	**55,5**	**53**	**52,9**	**55,56**	**56,9**	**54,4**	**56,3**	**56,7**	**60,6**	**55,6**	**58,23**

Premium 5	10	10	10	10	10	10	10	10	10	10	10	10
GAS COST	56,56	53,7	51,2	51,1	53,73	55,1	52,5	54,5	54,9	58,7	53,7	56,39
plus emissions	**59,01**	**55,5**	**53**	**52,9**	**55,56**	**56,9**	**54,4**	**56,3**	**56,7**	**60,6**	**55,6**	**58,23**

Maintenance

In order to have a realistic system representation, it was essential to provide maintenance schedule for the thermal plants, which was introduced to the system and was an important parameter affecting the final result.

The maintenance schedule corresponded, in terms of duration and frequency, to actual maintenance plans that can take place in a real life situation. The mathematical formulation of the maintenance used binary variables, meaning variables that can only take 0/1 values, indicating in this case whether the unit is connected or disconnected from the system for maintenance reasons. These kind of variables are the reason why the models created (both deterministic and its probabilistic evolution) are characterized as MIP (Mixed Integer Problems). Further details about the mathematical background of the model will be given in chapter ()

CHAPTER 4: METHODS & TECHNIQUES

Use of GAMS

INTRODUCTION

The General Algebraic Modeling System (GAMS) that is also used as part of the current thesis project is a high-modeling system for mathematical programming and optimization. Specifically, it is designed to be able to solve linear, non linear as well as mixed integer optimization problems. It can be proven to be really helpful when dealing with large, complex problems – it also allows the user to build them in such a way that can be adapted quickly to new situations. GAMS is available for use by personal computers, workstations, mainframes and supercomputers.

GAMS allows the user to concentrate on the modeling problem by making the setup simple. The system takes care of the time-consuming details of the specific machine and system software implementation.

As mentioned above, it is especially useful for handling large, complex, one-of-a-kind problems which may require many revisions to establish an accurate model. The system models problems in a highly compact and natural way. The user can change the formulation quickly and easily, by using for example different solvers, and can even make converts from linear to nonlinear with little trouble.

SYSTEM FEATURES

The important advantage of GAMS is that it allows the user to concentrate on modeling. By minimizing the need to think about purely technical machine-specific problems such as address calculations, storage assignments, subroutine linkage, and input-output and flow control, GAMS increases the time available for conceptualizing and running the model, and analyzing the results. GAMS structures good modeling habits itself by requiring concise and exact specification of entities and relationships. The GAMS language is formally similar to commonly used programming languages. It is therefore familiar to anyone with programming experience.

Using GAMS, data are entered only once in a list and table form. Models are described in concise algebraic statements which are easy for both humans and machines to read. Whole sets of closely related constraints are entered in one statement. GAMS automatically generates each constraint equation, and lets the user make exceptions in cases where generality is not desired. Statements in models can be reused without having to change the algebra when other instances of the same or related problems arise. The location and type of errors are pinpointed before a solution is attempted. GAMS handles dynamic models involving time sequences, lags and leads and treatment of temporal endpoints.

GAMS is flexible and powerful. Models are fully portable from one computer platform to another when GAMS is loaded to each platform. GAMS facilitates sensitivity analysis. The user can easily program a model to solve for different values of an element and then generate an output report listing the solution characteristics for each case. Models can be developed and documented simultaneously because GAMS allows the user to include explanatory text as part of the definition of any symbol or equation.

Deterministic vs. Stochastic Models

GENERAL INFORMATION

A simulation model is property used depending on the circumstances of the actual world taken as the subject of consideration. A deterministic model is used in that situation where the result is established straightforwardly from a series of conditions. In a situation where the cause and effect relationship is stochastically or randomly determined, the stochastic model is used.

A ***deterministic model*** has no stochastic elements and the entire input and output relation of the model is conclusively determined. A dynamic model and a static model are included in the deterministic model.

A ***stochastic model*** has one or more stochastic elements. The system having stochastic elements is generally not solved analytically and, moreover, there are several cases for which it is difficult to build an intuitive perspective. In the case of simulating a stochastic model, a random number is normally generated by some method or the other to execute trial. Such a simulation is called the Monte Carlo method or Monte Carlo simulation. In case the stochastic elements in the simulation are two or more persons and there is a competitive situation or some type of game being reproduced, this is specifically known as gaming simulation. Simulation by the deterministic model can be considered one of the specific instances of simulation by the stochastic model. In other words, since there are no random elements in the deterministic model, simulation can well be done just once. However, in case the initial conditions or boundary conditions are to be varied, simulation has to be repeated by changing the data. One the other hand, in Monte Carlo simulation, once the value has been decided by extracting a random number the simulation does not differ from deterministic simulation.

By gathering up the main characteristics of the two types of models, the following important points can be noted:

Deterministic Models:

- Model processes which are often described by differential equations, with a unique input leading to unique output for well-defined linear models and with multiple outputs possible for non-linear models;

- Equations can be solved by different numerical methods (after discretization: modification to run on a grid or a mesh, and parametrization: setting parameters to account for sub-grid processes):

    1. Finite difference
    2. Finite element
    3. Path simulation

- Models describe processes at various levels of temporal variation

    1. Steady state, with no temporal variations, often used for diagnostic applications
    2. Time series of steady state events, computed by running a steady state model with time series of input parameters, this approach is commonly used for estimation of long term average spatial distributions of modeled phenomena
    3. Dynamic, describing the spatial-temporal variations during a modeled event, used for prognostic applications and forecasting

Stochastic Models:

- Model spatial-temporal behavior of phenomena with random components

- Unique input leads to different output for each model run, due to the random component of the modeled process, single simulation gives only one possible result

- Multiple runs are used to estimate probability distributions

Conditional simulations combine stochastic modeling and geostatistics to improve characterization of geospatial phenomena

Behavior of dynamic stochastic systems can be described by different types of stochastic processes, such as Poisson and renewal, discrete-time and continuous- time Markov process, matrices of transition probabilities, Brownian processes and diffusion.

Deterministic Models "Vs" Stochastic Models:

A deterministic model assumes that its outcome is certain if the input to the model is fixed. No matter how many times one recalculates, one obtains exactly t he same result. It is arguable that the stochastic model is more informative than a deterministic model since the former accounts for the uncertainty due to varying behavioral characteristics.

In nature, a deterministic model is one where the model parameters are known or assumed. Deterministic models describe behavior on the basis of some physical law.

Deterministic models are usually developed by statistical techniques such as linear regression or non-linear curve fitting procedures which essentially model the average system behaviors of an equilibrium or steady/state relationship. In a live transportation system, a totally deterministic model is unlikely to include various dynamic random effects (or uncertainties). The uncertainty is commonly understood as factors related to imperfect knowledge of the system under concern, especially those being random in nature. It is closely related to heterogeneity, which denotes the state when entities within a given system are of non-uniform character. For example, when the heterogeneity is not faithfully recognized, the uncertainty increases. Conversely, a decrease in uncertainty, means that the system is better understood and thus the heterogeneity is better recognized.

All gathered:

Advantages

- Demand and generation are independent random variables
- Computation of units' output
- Computation of reliability measures
- Computation speed

Disadvantages

- Single loading order (heuristically obtained)
- No minimum load, no startup or shutdown
- No extensions for electricity markets

Stochastic Modeling

A quantitative description of a natural phenomenon is called a mathematical model of that phenomenon. Examples abound, from the simple equation S = Zgt2 describing the distance S traveled in time t by a falling object starting at rest to a complex computer program that simulates a biological population or a large industrial system. In the final analysis, a model is judged using a single, quite pragmatic, factor, the model's usefulness. Some models are useful as detailed quantitative prescriptions of behavior, as for example, an inventory model that is used to determine the optimal number of units to stock. Another model in a different context may provide only general qualitative information about the relationships among and relative importance of several factors influencing an event. Such a model is useful in an equally important but quite different way

Stochastic Processes

Stochastic processes are ways of quantifying the dynamic relationships of sequences of random events. Stochastic models play an important role in elucidating many areas of the natural and engineering sciences. They can be used to analyze the variability inherent in biological and medical processes, to deal with uncertainties affecting managerial decisions and with the complexities of psychological and social interactions, and to provide new perspectives, methodology, models, and intuition to aid in other mathematical and statistical studies.

Given a probability space $(\Omega, \mathcal{F}, P)$and a measurable space (S, Σ), an *S*-valued **stochastic process** is a collection of *S*-valued random variables on Ω, indexed by a totally ordered set *T* ("time"). That is, a stochastic process *X* is a collection

$$\{X_t : t \in T\}$$

where each X_t is an *S*-valued random variable on Ω. The space *S* is then called the **state space** of the process.

Finite-dimensional distributions

Let *X* be an *S*-valued stochastic process. For every finite sequence $T' = (t_1, \ldots, t_k) \in T^k$, the *k*-tuple $X_{T'} = (X_{t_1}, X_{t_2}, \ldots, X_{t_k})$ is a random variable taking values in S^k. The distribution $\mathbb{P}_{T'}(\cdot) = \mathbb{P}(X_{T'}^{-1}(\cdot))$ of this random variable is a probability measure on S^k. This is called a finite-dimensional distribution of *X*.

Under suitable topological restrictions, a suitably "consistent" collection of finite-dimensional distributions can be used to define a stochastic

Mixed integer programming (MIP)

Mathematically, the Mixed Integer Linear Programming (MIP) Problem looks like:

Maximize or Minimize c1t + c2u + c3v + c4w + c5x + c6y + c7z

subject to A1t + A2u + A3v + A4w + A5x + A6y + A7z

t	>= 0
u	>=0 & <=L2 & integer
v	∈(0,1)
w	∈ SOS1
x	∈ SOS2
y	=0 or >= L6
z	=0 or >=L7 & integer

where the

t variables are continuous real numbers
u variables can only take on integer values bounded above by L_2
v variables can only take on binary values

w variables fall into SOS1(*) sets exhibiting one nonzero
x variables fall into SOS2(**) sets exhibiting no more than two, adjacent non-zeros
y variables are semi-continuous being zero or in excess of L_6
z variables are semi-integer being zero or in excess of L_7 and integer
c1t + c2u + c3v + c4w + c5x + c6y + c7z is the objective function
A1t + A2u + A3v + A4w + A5x + A6y + A7z represents the set of constraints of various equality and inequality forms.

Specially ordered set variables of type 1 (SOS1) (*)

At most one variable within a specially ordered set of type 1 (SOS1) can have a non-zero value. This variable can take any positive value. SOS1 variables are declared as follows:

SOS1 Variable s1(i), t1(k,j), w1(i,k,j) ;
The members of the right-most index for each named item are defined as belonging to the SOS1 group or set of variables of which at most one of which can be non zero.

For example, in the SOS1 variables defined above:

- s1 forms one group of mutually exclusive SOS1 variables which contains elements for each member of the set i and thus only one variable for one of the cases of i can be nonzero with the rest being zero.

- t1 defines a separate SOS1 set for each element of k and within each of those sets the variables indexed by j are SOS1 or mutually exclusive.

- w1 a separate SOS1 set for each pair of elements in i and k and within each of those sets the variables indexed by j are SOS1 or mutually exclusive.

Notes

- By default each SOS1 variable can range from 0 to infinity. As with any other variable, the user may set these bounds to whatever is required.

- One is required to utilize a mixed integer (MIP) solver to solve any model containing SOS1 variables. However, the SOS1 variables do not have to take on integer solution levels.

- The MIP solver is required because the solution process needs to impose mutual exclusivity and to do this it implicitly defines an additional set of zero one integer variables, then solves the problem as a MIP.

The user can provide additional constraints say requiring the sum to the SOS1 variables in a set to be less than or equal to a quantity (often 1 for convexity).

In the following example :

```
SOS1 Variable s1(i) ;
Equation defsoss1 ;
defsoss1.. sum(i,s1(i)) =l= 3.5 ;
```

Here the equation defsoss1 defines the largest non-zero value that one of the elements of the SOS1 variable s1 can take.

- A special case of SOS1 variables is when exactly one of the elements of the set has to be nonzero and equal to a number. In this case, the defSoss1 equation will be

```
defSoss1.. sum(i,s1(i)) =e= 10 ;
```

A common use of the use of this type of restriction is for the case where the right hand side in the equation above is 1. In such cases, the SOS1 variable is effectively a binary variable. In such a case, the SOS1 variable could just have been binary and the solution provided by the solver would be indistinguishable from the SOS1 case.

- Not all MIP solvers allow SOS1 variables. Furthermore, among the solvers that allow their use, the precise definition can vary from solver to solver. A model that contains these variables may not be perfectly transferable among solvers.

Specially ordered set variables of type 2(SOS2) (**)

At most two variables within a specially ordered set of type 2 (SOS2) can take on nonzero values. The two non-zero values have to be for adjacent variables in that set.

Specially ordered sets of type 2 variables are declared as follows:

```
SOS2 Variable s2(i), t2(k,j), w2(i,j,k) ;
```

The members of the right-most index for each named item are defined as belonging to a special (SOS2) group or set of variables of which at most one of which can be non zero. For example, in the SOS1 variables defined above,

- s2 forms one group of SOS2 variables of which at most 2 can be non zero and they must be adjacent in terms of the set i. The adjacency means if the set i has elements /a,b,c,d,f,g/ that one could have any 2 variables like the ones associated with set elements a and b but never a and c since the set elements are not adjacent.

- t2 defines a separate SOS2 set for each element of k and within each of those sets no more than 2 variables can be non zero. Further, they they must be adjacent in terms of the set j. The adjacency means if the set j has elements /j1,j2,j3,j4,j5,j6/ that one could have any 2 variables like j3 and j4 but never j1 and j6 since the set elements are not adjacent.

- w2 defines a separate SOS2 set for each pair of elements in i and k. Within each of those sets no more than 2 variables can be non zero and they must be adjacent in terms of the set j. The adjacency means if the set j has elements /j1,j2,j3,j4,j5,j6/ that one could have any 2 variables like j3 and j4 but never j2 and j4 since the set elements are not adjacent.

Notes

- The most common use of SOS2 sets is to model piece-wise linear approximations to nonlinear functions using separable programming.
- One must use a mixed integer (MIP) solver to solve any model containing SOS2 variables. But, the SOS2 variables do not have to take on integer solution levels.
- The MIP solver is required because the solution process needs to impose both adjacency restrictions and the restrictions that no more than 2 nonzero level values can be present and to do this the solvers implicitly defines an additional set of zero one variables, then solves the problem as a MIP.
- The default bounds for SOS2 variables are 0 to plus infinity. As with any other variable, the user may set these bounds to whatever is required.
- Not all MIP solvers allow SOS2 variables. Furthermore, among the solvers that allow their use, the precise definition can vary from solver to solver. Thus a model that contains these variables may not be perfectly transferable among solvers.

Relaxed mixed integer programming (RMIP)

The relaxed mixed integer programming (RMIP) problem is the same as the mixed integer programming (MIP) problem in all respects except all the integer, SOS and semi restrictions are relaxed:

Maximize or Minimize c1t + c2u + c3v + c4w + c5x + c6y + c7z

Subject to	A1t + A2u + A3v + A4w + A5x + A6y + A7z						
	t						>= 0
		u					>=0 & <=L2
			v				>=0 & <=1
				w			>=0
					x		>=0
						y	>=0

z >=0

This problem type is sometimes helpful when one is having trouble attaining a feasible integer solution.

The final model to be formed is the stochastic one, examining simultaneously a number of multiple scenarios. In our case, the number of scenarios we chose to examine is 15, and it was chosen for multiple reasons. Theoretically, an increased number of scenarios could improve the quality of the results, in the sense that it could become more obvious how small or bigger changes affect the final outcome. This is in general very important but it has been proven hard to apply it in the specific case – the most important problem to be faced is the limited system memory that significantly reduced the number of scenarios that could be tested at the same time. Moreover, a further examination among the different possible scenarios indicated that picking “characteristic ones” (referring to the ones that can be seen as the mean scenarios of others with small differences), can be proven a better option than having a big number of scenarios that result into similar prices (the final outcome of our scenarios).

In the matrixes following below, a detailed representation of the scenarios is given, including explanations on the different parameters as well as on the results obtained by them. In the context of the stochasticity of the model, each one of these scenarios is regarded with a different probability. While the initial test considered an equal probability for each one of the scenarios, for a better understanding and control of the results, the probabilities were matched to the “actual probability” of the scenario – by this, it is assumed that scenarios presenting a higher probability have an actual higher chance of happening in a real situation. With this kind of match, it is expected to have a as-much-as-possible realistic final picture for the system price in the different time periods in the annual horizon.

To start with, the first matrix is regarding the three different hydro scenarios that were taken into consideration. As in a real system situation, these scenarios are about a wet, dry or normal year. Presented in an hourly basis, the three scenarios in this case are :

System Description (summary)

The system created for the deterministic model is simplistic and resembles the Spanish market in a quite general way. It includes:

1. The thermal generation plants include nuclear, CCGT & coal plants. Also, renewable technology is included in the system and is divided among the companies in a random way. For the initial, deterministic model, the renewable generation was not taken into account. Also, the final representation of the generation does not include the concept of companies at any point – this was only applied for means of simplification at the early stages of the problem formation.

2. One important element is that the generation mix was created in order to exceed the maximum demand by approximately 20%, while the real data of the Iberian market present a higher differentiation (more than 100%!). This assumption was made to achieve greater differences among prices coming from applying different demand and generation scenarios.

MAIN DATA FILES TO BE USED TO THE SYSTEM

The models use a great amount of data, that are passed to them through excel files. As follows we will refer to each one of them and the kind of information they contain.

Plants_Capacities_Coal

A file containing all the capacities for the different coal plants of the system. The total number of coal plants in the system is equal to 71 and they appear into three different capacity categories, as will be shown below.

Plants_Capacities_Others

A file containing all the capacities for the different nuclear and CCGT plants of the system. The nuclear plants have a fixed capacity of 1000 MW, while the CCGT plants come into two types – plants of either 400 or 800 MW. The system accounts for 7 nuclear plants, while the total number of CCGT plants is 78.

ThermalDemand_Final

This file contains the total information about the thermal demand of the system. As far as the calculation of this demand is concerned, historical data were used.

Plants_min

For each one of the plants of the system, this file contains their technical minimum , to be used as a constraint in the model regarding the plants operation.

CostEstimation_plants

In the following matrix, the plants of our system, as well as their capacities are presented. The plants are divided into two categories – coal plants and CCGT/nuclear plants. This is because of their different technical characteristics that made it easier for the formation of the model equations, to have them in separate categories. The exact formation of the costs and the analysis of their separate elements will be analyzed further below.

In the matrix provided in the Appendix we present all the thermal plants used in our system, by code name and corresponding capacity.

As one can notice, all nuclear plants have a fixed capacity of 1000 MW, the CCGTs appear in two different capacities, 400 or 800 MW, and for the coal plants we have considered plants of 200, 350 or 550 MW.

MAIN CHARACTERISTICS OF THE MODEL

<u>Information about the sets</u>

For the yearly horizon, all 12 months from January to December were used. Each month was considered with its normal days (e. g 28 days for February) and there was no discrimination between working and labor days. Regarding the hour periods, we separated the 24 hours of the day into six 4-hour period, so as to have an easier analysis and also to help the model run in that point by reducing the amount of used data.

Data for thermal demand, plant names – capacities – technical minimum – cost, are taken from excel files. The formation of the data within the excel files matches the corresponding equations' formation within the model so as to automate the process of retrieving data from the files and using them in the calculation processes.

At this point, we will write down and analyze the equations that were used for the deterministic model, and remained quite the same at the stochastic one.

These are:

E_DMND (mt,da,p)	Meeting the demand
E_FOBJ	Objective cost Function
E_QMAXG1(mt,da,coal)	Maximum power of generators that are coal
E_QMAXG2(mt,da,p,others)	Maximum power of nuclear/CCGT generators
E_QMING1(mt,da,coal)	Minimum power of generators that are coal
E_QMING2(mt,da,p,others)	Minimum power of nuclear/CCGT generation

As it can be seen, these equations do not include the probability within their identifiers. When formulating the stochastic problem this was altered in the needed equations. By this we mean that a new parameter pr(i) was included in the model, that is the probability that each scenario presents (the calculation of the probabilities is analyzed in a next chapter). After including this new parameter, all equations now added i within their identifiers, since they are all dependent on the probability of each scenario examined. This also applies to the objective function that is now formulated as follows:

E_FOBJ

fobj=E=
SUM[**i**,SUM[coal,SUM[(mt,da),SUM[p,$cost_1$(coal,mt)*qa(mt,da,coal,**i**)***pr(i)**]]] + SUM[others,SUM[(mt,da),SUM[p,$cost_2$(others,mt)*qb(mt,da,p,others,**i**)***pr(i)**]]]];

In this point, it is important to address two matters:

As one can notice, there is a separation in all kinds of calculations for coal plants (that form the one group), and nuclear / CCGT plants (that form the other). The reason for that can be made clear when the domains for the equations (which are also the domains of the variables used in the equations), are seen – when we refer to coal plants, no hourly period is taken into account, but only months and days, whereas that is not the case for nuclear and CCGT plants, where everything is taken into account. The reason for that is because for the operation of the coal plants we do not consider that changes in their operative status (on/off) can be made on consecutive hourly periods but only in consecutive days. Also, nuclear plants because of their zero variable cost operate as base plants for the whole time period apart from the hours contained into their maintenance schedule, that are set by proper binary variables. (analyzed further down).

For the stochastic approach to the problem, the first step was to change the demand excel sheet, to consider stochasticity and examine a number of different scenarios. The thermal demand, used in our model, will be calculated now as follows:

1. Obtain scenarios´ data (initial number of demand scenarios equal to 5, combined with the 3 basic scenarios for hydro, will give us a total of 15 scenarios to work with).

2. An assumption to be made at these first stages is that there are no imports – exports from our system to others and vice-versa. That is a reference to interconnections, which will be considered as zero, with a view to increase the complexity by including them, in the future.

3. For the renewable technology: we will consider 3 different types of renewable technology to be existent in our system, wind, sun and a category that is co-generation and others, and refers to independent producers of renewable energy. For the first two, also 5 scenarios will be examined, but this will not increase the number of scenarios in that stage, since these 5 scenarios will comply with

the 10 demand scenarios we mentioned before, so they will pair in that stage. For the co-generation, we will consider it constant among the different scenarios.

4. Considering all the above, thermal demand will be calculated by the formula :

THERMAL DEMAND = DEMAND +/- INTERCONNECTIONS – HYDRO – RENEWABLES

<u>More important things to be mentioned</u>

The data will be analyzed for the time period of a year, with hourly data, in the same format that we followed when processing the deterministic model. For making the calculations easier, as mentioned above time periods of 4 hours are considered Moving on to the probabilistic formation of the problem, the first aspect that was handled, was the demand. Demand includes a high level of stochasticity, and so multiple scenarios had to be taken into account.

<u>**Thermal demand**</u> in our model is formulated in the following way :

From the total amount of demand for the system, which is given in an 4-hour period form, for every day of the year (that makes 6 periods per day), we subtract the outputs of the hydro plants in our system, outputs of wind and sun and finally what we get from renewable. The final result is the thermal demand. Stochasticity appears in the consideration of different scenarios – specifically, we will consider 3 different kinds of hydro scenarios (wet, normal, dry) and 5 different scenarios for the initial demand, as well as for the outputs of wind and sun. A very important thing to be mentioned in that point is the fact that the 5 scenarios of the demand are **correlated** with the scenarios of the sun and wind production – by that way, we do not increase, in that stage of the problem, the complexity by having to consider numerous scenarios. The scenarios to be examined in that point are the result of a commercial program called ALEA, that is able to receive data and through correlation and matching procedures, to end up in a number of possible scenarios. As a final goal, the model will be examined for 200 scenarios. For the scenarios of the hydro units, in the driest year (or drier year series) demand has to be satisfied (deterministically)

<u>Type of Model Results</u>

The results to be obtained to this first stage of the problem are the following ones:

- For every month, day and period the production of each one of the plants of the system is presented, plus the total thermal demand of this period, and most importantly the marginal cost for this period that equals the cost of the last plant that entered in order for the demand to be covered.

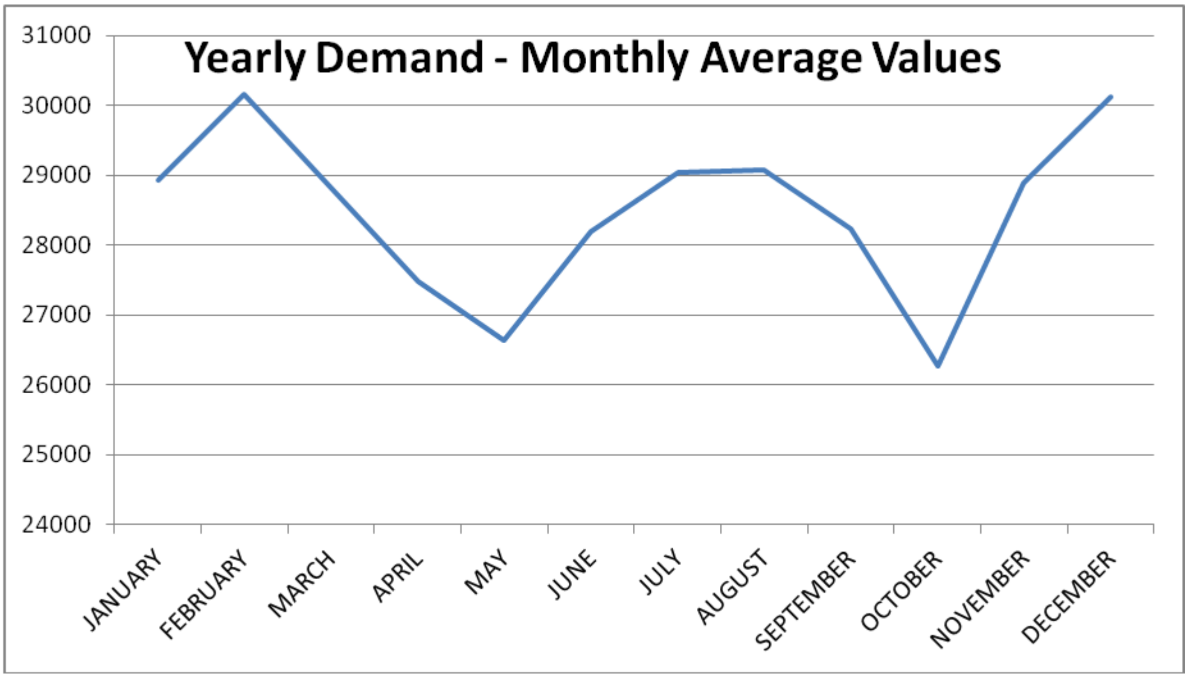
In this point, it is important to go back to chapter 2 where a thorough analysis of the cost of the power plants is given, since these costs are strongly connected to the final prices received by the model. Calculations took place independently from the GAMS environment and were passed to the program, through an excel file.

CHAPTER 5 – RESULTS

Different Cases Analysis

5.1 DETERMINISTIC SCENARIO

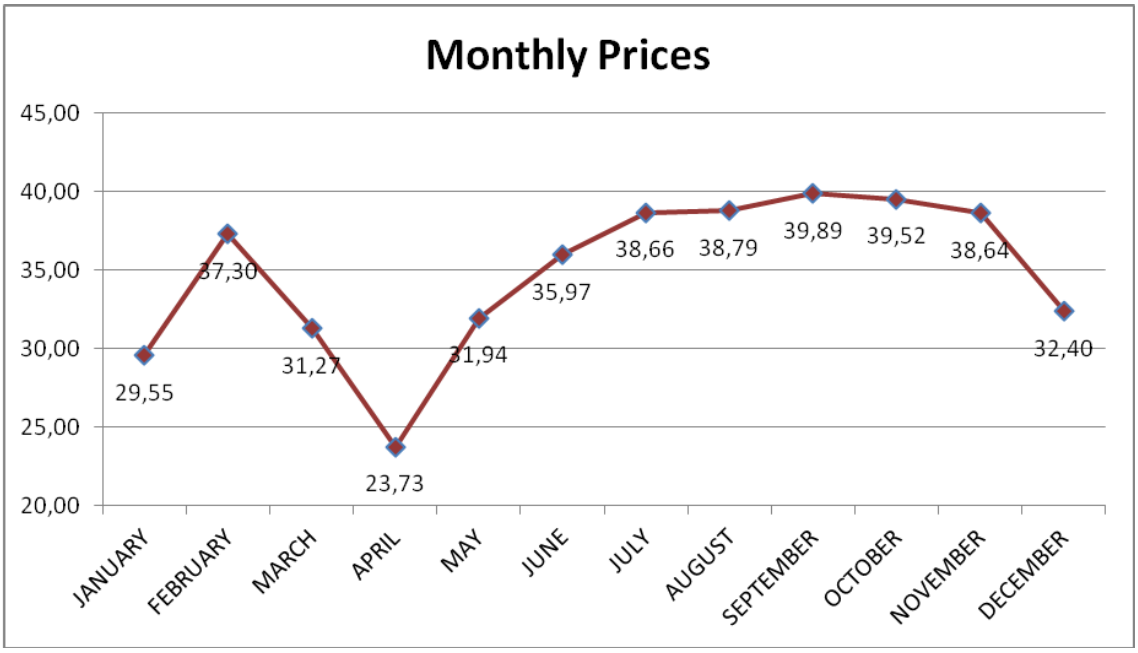
To start with, we will analyze the data used in this initial stage of the problem formation, where a simple deterministic model was created and only one scenario was considered. The system structure as well as any other planning such as the maintenance of the plants, is the same for the three stages of the problem, and the same accounts for costs, capacities etc. In the next diagram we represent, per month, the demand that was taken into account for the initial model.



As it can be seen, demand within the year presents a fair amount of variation, corresponding to real data in the sense that the months presenting lowest values (May, October) as well as the ones with the highest values (February, December) are the ones that in reality also present values like the above. This is a good indicator of the realistic representation of the system used in the thesis model.

The above diagram shows the total demand, from which thermal demand was obtained by subtracting hydro, wind, sun and renewable production. The

In a previous chapter we have been analyzing the exact process followed for the formation of the final costs of the power plants used in the system. Also, the technical aspects of the power plants as well as procedures like the plant maintenance have been explained. All these taken into account, the final results for this model are shown in the following diagram:

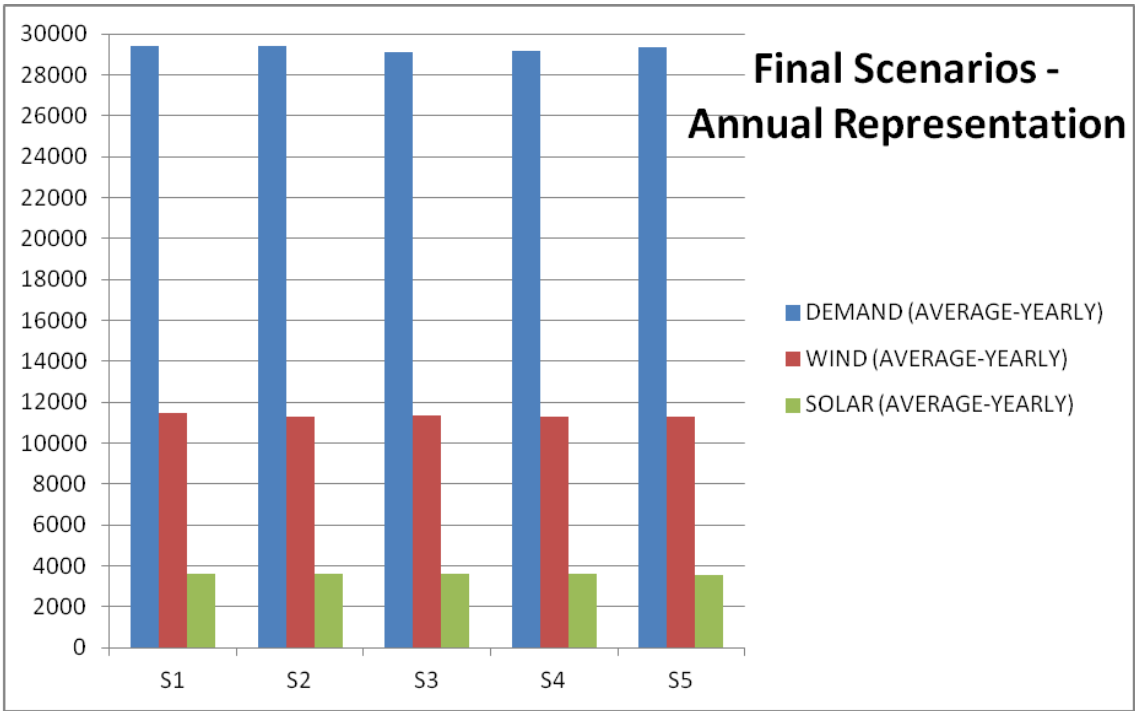


Prices present a variation of almost 40% (minimum to maximum price) throughout the year. As someone can notice, there is a kind of correlation between the demand values and the final prices of the corresponding month. Specifically, the spring season around May presents very low demand (bottom value) and the price also presents a somewhat common behavior, in the sense that it significantly decreased from its previous and future values.

5.2 MULTIPLE DETERMINISTIC SCENARIOS

As an evolution to the previous model, we rerun the model, considering a number of different scenarios, totally **independent** one from another (equals to a probability of 100% per scenario). This procedure was a middle step between the initial deterministic and the final probabilistic model, vital for gaining a better insight on the way prices are affected by variations in the demand, fuel cost and RES production and for building the data base for the probabilistic model.

The number of scenarios that was decided to be examined was 15, matching the combinations of the three available hydro scenarios (wet, mean, dry) with the five scenarios that were created through the clustering procedure also analyzed in a previous chapter. In the next matrix we present the yearly values per scenario corresponding to demand, wind and solar production.



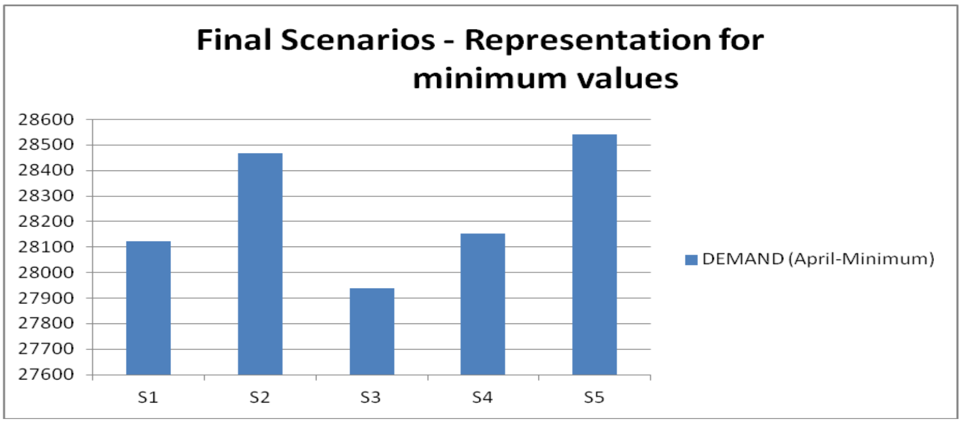
One very important conclusion to be drawn from the above diagram is the pretty small variation that the different scenarios present in all three categories. There are mainly two reasons for which this can be happening: to start with, an annual representation is not that much indicative of the real situation, meaning that it misses the actual variations that can be much more clearly shown in an hourly representation, or even daily. Because of lack of bigger computer memory, it was decided to do the clustering procedure taken monthly data into account – this probably resulted in a less accurate final result than if the clustering had used hourly data. It was in either way expected to have very small variations since the comparison was made in an annual horizon.

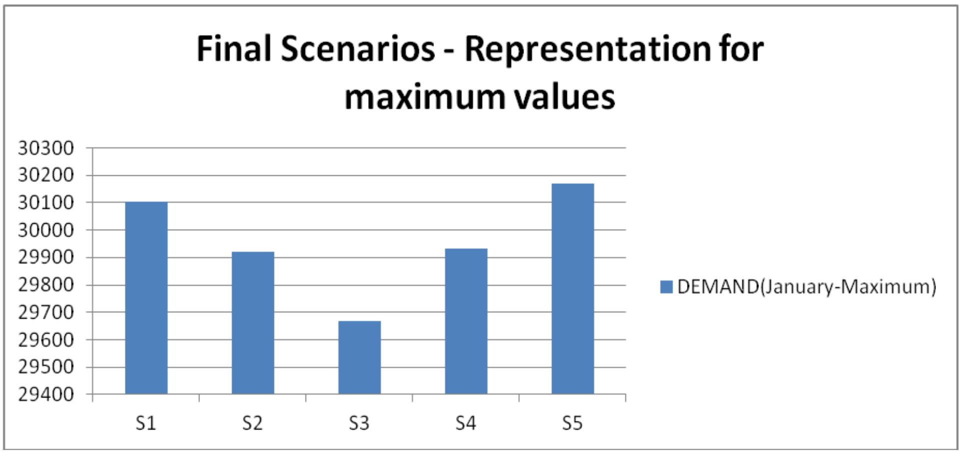
In order to obtain a better insight on the actual differences of the final scenarios that occurred from the clustering, we performed the same analysis as above but taking into account the extreme cases, meaning the months that demand, wind and solar production presented their maximum and minimum values.

These were:

	MAXIMUM	MINIMUM
DEMAND	January	April
SOLAR PRODUCTION	July	December
WIND PRODUCTION	February	September

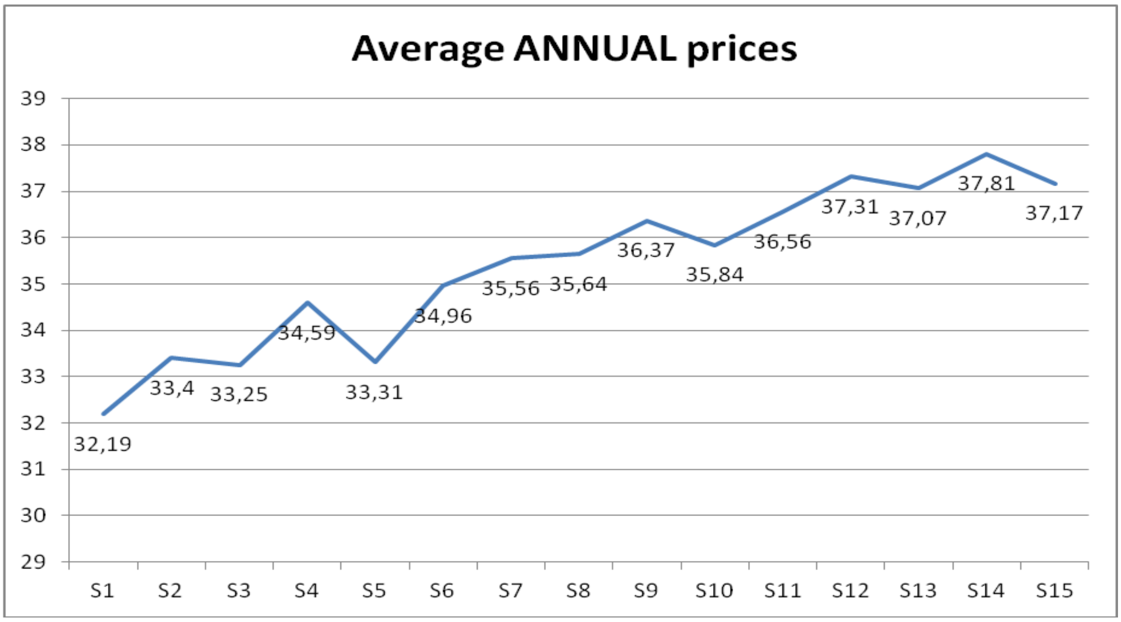
And the graphical representation is shown in the next diagrams:





Now, this representation reveals in a way the differentiations among the chosen clustered scenarios, which cannot be made obvious when we consider an annual horizon.

By running independently the scenarios (considering a 100% probability to each one of them), the results we receive are the ones following:



For means of simultaneous representation of all the scenarios, we chose to represent annual prices. As it can be noticed, the prices have a variation of almost 14%, comparing minimum and maximum prices. Comparing the above results with the annual price of the initial deterministic scenario (equal to 34,8), what is noticed is that is corresponds to a medium value with regards to the ones comi

5.3 PROBABILISTIC SCENARIO

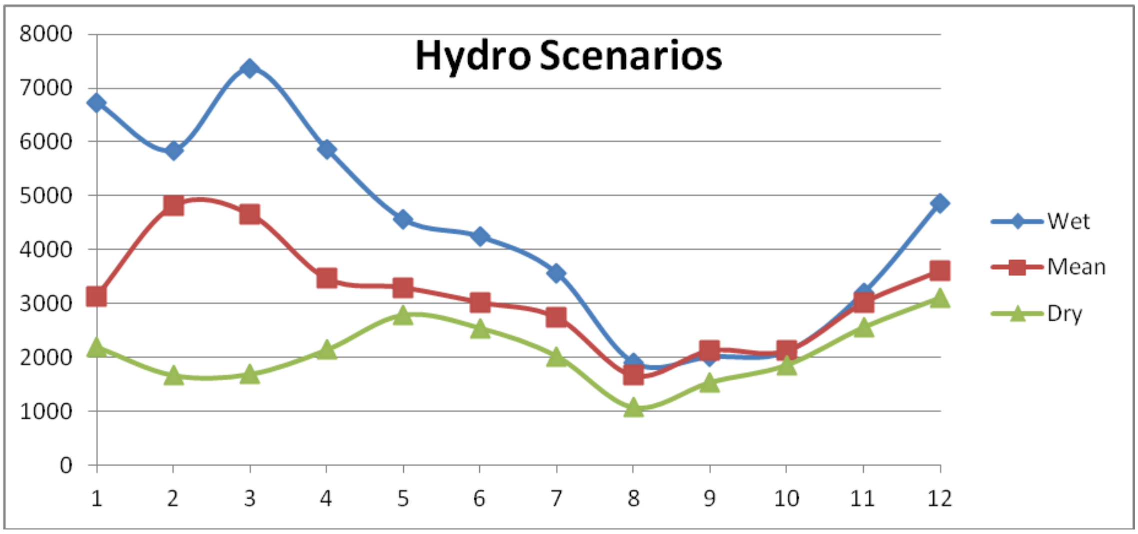
The final model to be formed is the stochastic one, examining simultaneously a number of multiple scenarios and resulting to one unique result. In our case, the number of scenarios we chose to examine is 15, and it was chosen for multiple reasons. Theoretically, an increased number of scenarios could improve the quality of the results, in the sense that it could become more obvious how small or bigger changes affect the final outcome. This is in general very important but it has been proven hard to apply it in the specific case – the most important problem to be faced is the limited system memory that significantly reduced the number of scenarios that could be tested at the same time. Moreover, a further examination among the different possible scenarios indicated that picking “characteristic ones” (referring to the ones that can be seen as the mean scenarios of others with small differences), can be proven a better option than having a big number of scenarios that result into similar prices (the final outcome of our scenarios).

In the matrixes following below, a detailed representation of the scenarios is given, including explanations on the different parameters as well as on the results obtained by them. In the context of the stochasticity of the model, each one of these scenarios is regarded with a different probability. While the initial test considered an equal probability for each one of the scenarios, for a better understanding and control of the results, the probabilities were matched to the "actual probability" of the scenario – by this, it is assumed that scenarios presenting a higher probability have an actual higher chance of happening in a real situation. With this kind of match, it is expected to have as-much-as-possible a realistic final picture for the system price in the different time periods in the annual horizon.

To start with, the first matrix is regarding the three different hydro scenarios that were taken into consideration. As in a real system situation, these scenarios are about a wet, dry or normal year. Presented in an hourly basis, the three scenarios in this case are:

	WET	**MEAN**	**DRY**
JANUARY	6728,50914	3133,294	2193,228
FEBRUARY	5843,570238	4810,693	1669,506
MARCH	7354,491398	4660,923	1689,179
APRIL	5862,945354	3467,741	2138,442
MAY	4557,814631	3298,007	2782,412
JUNE	4235,443611	3022,114	2540,351
JULY	3559,55745	2741,554	2022,774
AUGUST	1899,8	1660,9	1077,8
SEPTEMBER	2002,033773	2126,013	1538,182
OCTOBER	2100,627383	2122,231	1860,034
NOVEMBER	3193,397781	3013,099	2553,824
DECEMBER	4863,293289	3606,654	3099,468

Graphically, the above are represented as follows:



For means of simplification, the hydro scenarios are represented by their average monthly values. These three scenarios were generated by specific software called ALEA, that uses a combination of historical data and predictions available. The same software was used for the generating of demand and wind/solar scenarios that are presented as follows.

Next step was to obtain the scenarios for demand as well as solar and wind production. As already mentioned, for this purpose ALEA was again used, but some certain assumptions needed to be made in this stage. The reason for this is the huge amount of data and final scenarios that we would result into in case we kept all the scenarios generated for demand, for wind as well as solar. So, an important decision was taken at this level – scenarios generated for demand, and scenarios generated for solar and wind production is to be considered correlated. By this, it was meant that the first scenario corresponding to demand would ´match´ the first scenario for wind production and also the first one for solar production. This significantly reduces the number of combinations that can be made among scenarios for hydro and the ones for demand, wind and solar. This decision is not without a logical base – by analyzing the corresponding scenarios of demand, wind and solar, one can notice that there is an important level of correlation among them, so considering them correlated does not so highly alternates the final results in terms of quantity, but rather decreases their quantity.

More specifically, the software generated a number of 200 scenarios for demand, 200 for wind and 200 for solar. In order to have a relative small number of final scenarios, which would cover all different market situations, the next step in the data analysis process described above, was to prepare a clustering of these 200 scenarios, in 5 final groups, with the **corresponding probabilities** – this was a very important stage, since the calculations of these probabilities is crucial for running the probabilistic model. The clustering procedure was achieved through the algorithm of the program XL stat that simultaneously generated the clusters and the corresponding probabilities. What is more, the initial data that were to be clustered were presented (again for means of decreasing volume) in a monthly basis, with values per month representing the sum of the values for the hours corresponding to each month. Also, because of the assumption mentioned above regarding the correlation among the demand, wind and solar scenarios, the above clustering procedure was only applied to demand, and its results were used for the clustering of wind and solar data

After running all the above, we receive the following results:

CLUSTERING RESULTS

Final scenarios	**1**	**2**	**3**	**4**	**5**
	SIM1	SIM2	SIM3	SIM5	SIM13
	SIM7	SIM11	SIM4	SIM6	SIM23

SIM9	SIM14	SIM8	SIM10	SIM28
SIM16	SIM15	SIM12	SIM17	SIM29
SIM24	SIM19	SIM21	SIM18	SIM30
SIM25	SIM20	SIM22	SIM41	SIM35
SIM32	SIM26	SIM27	SIM46	SIM38
SIM39	SIM33	SIM31	SIM47	SIM42
SIM43	SIM34	SIM36	SIM52	SIM49
SIM53	SIM40	SIM37	SIM57	SIM51
SIM61	SIM45	SIM44	SIM59	SIM64
SIM62	SIM50	SIM48	SIM60	SIM68
SIM69	SIM55	SIM54	SIM63	SIM71
SIM81	SIM65	SIM56	SIM70	SIM75
SIM86	SIM72	SIM58	SIM73	SIM76
SIM88	SIM77	SIM66	SIM74	SIM96
SIM89	SIM79	SIM67	SIM78	SIM98
SIM90	SIM85	SIM80	SIM84	SIM101
SIM93	SIM94	SIM82	SIM92	SIM102
SIM108	SIM97	SIM83	SIM99	SIM105
SIM120	SIM103	SIM87	SIM100	SIM106
SIM124	SIM107	SIM91	SIM116	SIM111
SIM126	SIM113	SIM95	SIM122	SIM117
SIM128	SIM123	SIM104	SIM125	SIM121
SIM131	SIM132	SIM109	SIM127	SIM136
SIM134	SIM139	SIM110	SIM130	SIM138
SIM135	SIM141	SIM112	SIM144	SIM140
SIM147	SIM143	SIM114	SIM145	SIM142
SIM151	SIM150	SIM115	SIM146	SIM154
SIM156	SIM152	SIM118	SIM159	SIM155
SIM160	SIM153	SIM119	SIM162	SIM157
SIM163	SIM168	SIM129	SIM167	SIM158
SIM165	SIM170	SIM133	SIM180	SIM164
SIM166	SIM171	SIM137	SIM187	SIM176
SIM169	SIM175	SIM148	SIM194	SIM181
SIM173	SIM179	SIM149	SIM197	SIM182
SIM186	SIM188	SIM161		SIM185
SIM192	SIM190	SIM172		SIM189
SIM193	SIM196	SIM174		SIM191
SIM195	SIM198	SIM177		
SIM200	SIM199	SIM178		
		SIM183		
		SIM184		

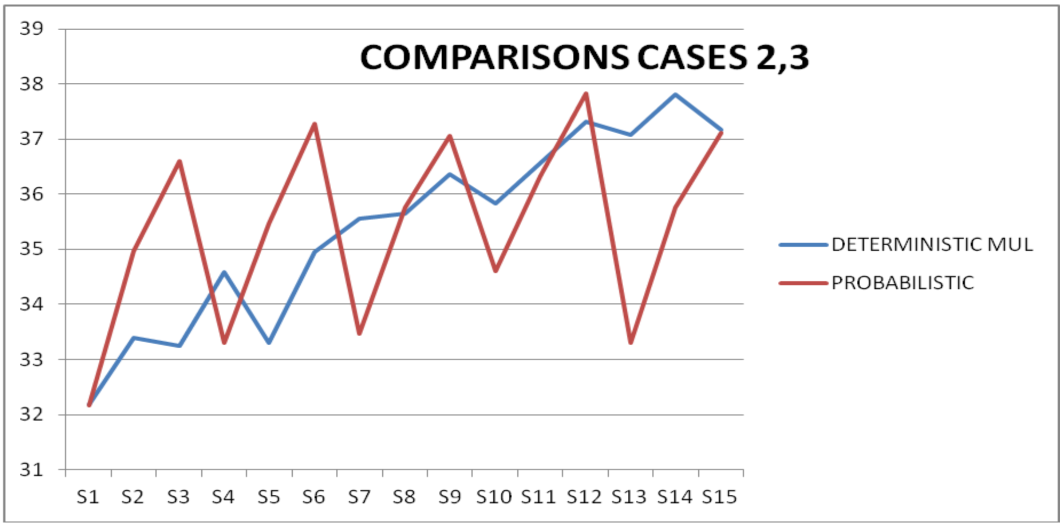
In the above matrix, the initial 200 scenarios (SIM1-SIM200), are clustered in 5 final scenarios. The exact same clustering was assumed for the wind and solar scenarios. In terms of probabilities used within the probabilistic model, they are shown in the following board:

PROBABILITIES MATRIX

Final Scenarios	Probability	s1_hydro	s2_hydro	s3_hdro	s1_others	s2_others	s3_others	s4_others	s5_others
S1	**0,0175**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S2	**0,0172**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S3	**0,0333**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S4	**0,0154**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S5	**0,0167**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S6	**0,1400**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S7	**0,1376**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S8	**0,2661**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S9	**0,1232**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S10	**0,1333**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S11	**0,0175**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S12	**0,0172**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S13	**0,0333**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S14	**0,0154**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167
S15	**0,0167**	0,1	0,8	0,1	0,175	0,172	0,333	0,154	0,167

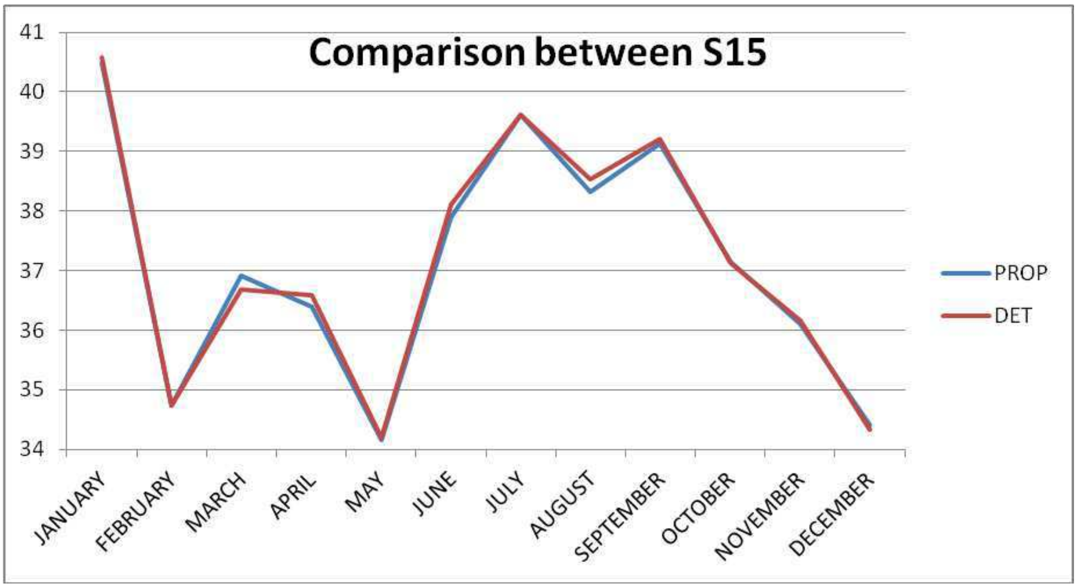
For these results, the probability of each scenario is calculated by multiplying the probability of each hydro scenario with the probability of demand (same for wind and solar), for each one of the three hydro scenarios – for example, the probability of scenario 13 is calculated by multiplying the probability of the s3 hydro scenario (dry), with the probability of the s3 scenario for demand/wind/solar. The final values are the ones marked in bold, for each one of the 15 final scenarios.

For purposes of better understanding of the final results, we chose to represent both the results of the multiple deterministic scenario case with the results of the probabilistic case, in a common diagram. This is what we obtain:



Again, the results are represented in an annual basis, with the values corresponding to average values. Among the 15 scenarios compared, some seem to have almost the same annual price value (e.g S1, S15), while others present quite a high variation (e.g S6, S13). While the differences in the first case are neglectable, we chose also to represent these 2 scenarios in new diagrams that contain the monthly average values, and can provide a better insight in any possible differences these two scenatrios can have within the year.

This is what we obtain :



It seems that also in the case of monthly representation, not a great deal of differentiation can be noticed – still, s15 has a certain variation between the two cases in spring and summer months. The most possible reason for this is that tha data for demand, hydro, wind and solar in these cases were very similar between the two cases, which lead to a very similar final result. One the other the scenarios 6,13 present a great deal of variation in the annual horizon considered – especially for the latter, the probabilistic models leads to a significantly decreased value comparing to the deterministic one, because of multiple factors that include the high probability of the specific scenario (0,0333 as can be seen from the probability matrix) but also lowered demand and perhaps increased penetration by renewable technologies.

5.1 DIFFERENT APPROACH

In order to examine the results under a different perspective and obtain more information regarding the kind of signal that these final prices could be in terms of companies, a different approach was also applied regarding the system under examination.

For this approach three companies of different thermal generation assets were considered. For these companies, whose specific assets will be analyzed as follows, we examined their final production and also incomes, taking into account the corresponding prices from the deterministic and stochastic model. After this, a comparison was made in order to determine the possible advantages of using a stochastic model in the calculation of prices.

More specifically, the total amount of thermal plants that have been mentioned above were divided into ten separate companies which for the purposes of the thesis were strictly considered as generation companies. There has been an effort to diversify these companies in terms of the kind of thermal plants used as well as regarding their final capacity. The model does not enter into great detail analyzing the bidding process of the companies into the daily and intra-day market, since this has not been the purpose of this thesis project. However, the results presented below can be used as a valuable price signal for companies with an alike generation mix, when it comes to bidding decision and in the long term investment ones.

The companies used for this reason were the following:

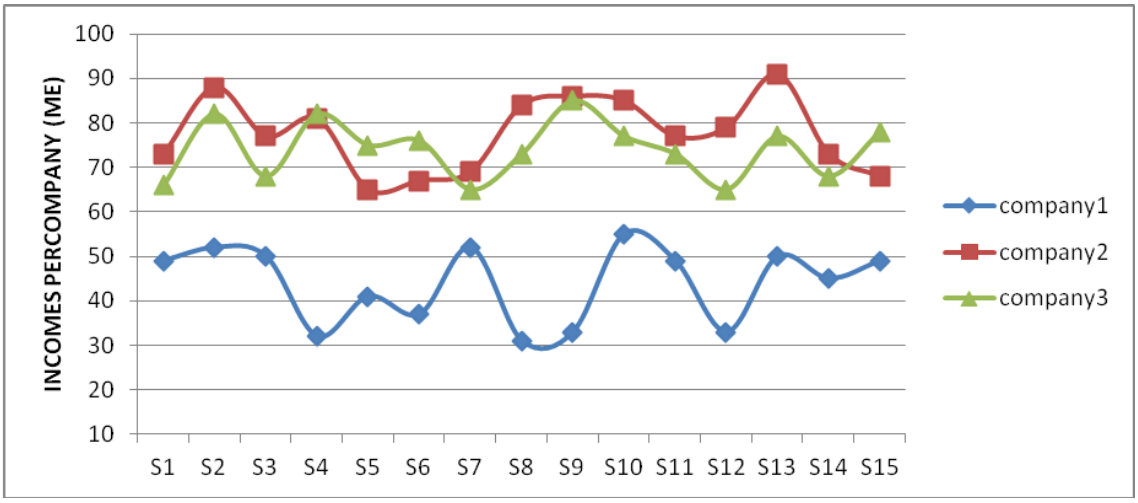
<u>COMPANY 1</u>			
COAL	*CCGT1*	*CCGT2*	*NUCLEAR*
CB8	CGA4		
	CGA5		
	CGA10		
	CGA11		
	CGA12		

<u>COMPANY 2</u>			
COAL	*CCGT1*	*CCGT2*	*NUCLEAR*
CC3		CGB3	
CA2			
CB14			
CB17			

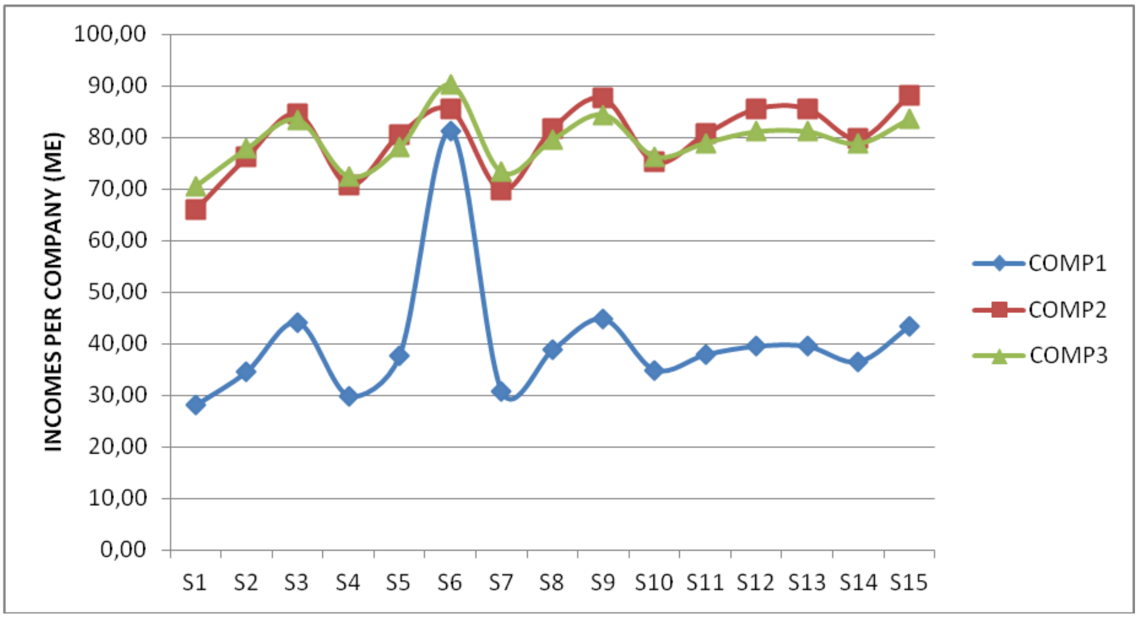
<u>COMPANY 3</u>			
COAL	*CCGT1*	*CCGT2*	*NUCLEAR*
CB18	CGA20		NUC3
CB20	CGA21		
	CGA22		

The reason for choosing to examine the results of the models for these three models is their differentiation in terms of the thermal generation mix – specifically, one company was chosen so as to contain all kinds of thermal plants (nuclear, coal and CCGT ones), another is focused on production from coal plants and the last one on thermal generation from CCGT plants. The different costs of these thermal plants are expected to lead to diversified incomes per scenario and some useful conclusion could be made regarding the stability of the incomes as well as regarding how including stochasticity can affect the accuracy of the predictions and consequently improve or not the decisions taken by the companies.

In order to perform the above, we will first examine the results from the second stage of the running of the model, meaning the results coming from the running of multiple deterministic models. For each one of the scenarios the daily production coming from the unit commitment results is multiplied by the corresponding price and final annual incomes are found and presented in the following graph.



At the next step we perform the same procedure for the stochastic model, for each one of the scenarios, following the steps described above. The results are as follows:



The above results show the annual income variation for the companies 1,2,3 for the 15 scenarios taken into account. Some important assumptions are to be made from the above graph:

❖ The company with the greatest variation and at the same time lowest incomes at all the examined scenarios is company 1, which is the company that owns an amount of CCGT plants. This kind of result was expected because of multiple reasons – this mainly refers to the high cost of CCGT plants, comparing to the one of coal and nuclear plants combined with a generally low demand in the majority of the scenarios. By this it is meant that CCGT are in general the last to be dispatched in the daily unit commitment and for this reason a company based mainly on them takes of risk of receiving lower incomes at periods of low demand or even increased hydro and RES generation.

❖ Companies 2 and 3 present a relative low variation between their results. There is a certain peculiarity at this point regarding the fact that company 3 is also the owner of a number of CCGT plants and it could be expected from it to have maybe lower incomes, because of the reason described above. However, the coal plants the same company owns as well the nuclear one, seem to provide importantly increased incomes comparing to company 1 and also a significant stability regarding variations among the different scenarios. This is also the case for company 2 – the combination of (in majority) coal plants with a CCGT one gives the company a high income for all scenarios and at the same time keeps the variation in relative low levels.

❖ A very important point to be mentioned in this stage is that when examining only the income of the company coming from its generation assets dispatch, not very clear conclusions can be drawn regarding the actual profit of the company and these kind of results cannot be used as the most efficient signal in order to organize a strategy and future investments. However, it still presents a good indicator of how a company of certain assets will perform in the market and of how variations in the demand as well as RES penetration should be predicted as accurately as possible in order to provide the company with the needed signals to adjust its behavior most efficient.

- When comparing the results from the multiple deterministic and the probabilistic model, a lot of remarks can be made regarding the importance of using a stochastic model in contrast to a stochastic one. One of the most important point is the accuracy of the results in the case of the stochastic model – as it can be noticed, the results provided by the deterministic one are more flat and this can be an indicator of the lack of accuracy and precision by the deterministic model. By taking into account the probabilities corresponding to variations in demand, hydro, wind and solar production it becomes obvious that a company could have a better insight on how its incomes are affecting by the multiple variable factors appearing in the volatile electricity market.

- For this kind of approach we chose to work with the companies´ income as an indicator of the good function of the stochastic model and its excellence against the deterministic one. It is however obvious that this goal could be achieved even more accurately if we worked with the companies´ margin instead – for that reason their corresponding costs should be taken into account. On the other hand, not extreme cases were expected even in the case of examining margin, in the sense that final patterns and changes among the companies remain more or less the same. It would in any case important to perform also this different kind of analysis as it is expected to give even better and more realistic results.

CHAPTER 6 – CONCLUSIONS

Assumptions & further research

6.1 Main findings

Summarizing the results and the conclusions drawn from the application of the three models to a market resembling a real one and up to a point specifically the Iberian one, we obtain the following:

- ✓ Pricing is an important feature of the market, and presents high level of correlation with the generation mix existing in the market. What is more, the kind of plants as well as the level of coverage of the demand is very important in terms of achieving efficiency and effectiveness regarding the unit commitment process.
- ✓ Regarding the specific market created for the current thesis project, its high resemblance to the Iberian one is also obvious by the fact that because of the over-capacity offered by the system, compared to the demand, not big variation of the price is achieved when applying the different demand/production scenarios.
- ✓ The above fact can also be a signal (for the short and medium term at least) of the relative stability that prices are going to present in the future for the Iberian power market.
- ✓ A stochastic approach to the calculation of prices within a market provides greater insight because of its ability to simultaneously examine different combinations of demand and renewable generation scenarios – specifically for the RES, taking into account the corresponding probability is vital, because of their level of uncertainty in real life situations.
- ✓ Despite of possible great changes in parameters such as the demand or the fuel cost (that was for example diversified among plants of the same kind as it has been described in a previous chapter), the achieved final diversification is not as big as might expected. This is equal to the fact that prices are in fact determined by a number of parameters and so it takes a simultaneous big change in all of them so as for a great variation to be achieved (15%-20%).

After applying the model that we created in the case of 10 companies of different generation assets (that we separated into two groups of similar characteristics), some further conclusions can be made.

- ✓ A company has to be able to adjust to variations of the market, that used to be uniquely linked to the variable demand but nowadays is mostly a matter of the increasing RES penetration and also a matter of inaccurate future demand prediction, which led to exaggerate investments.

- ✓ Companies owning nuclear plants operate more ‘safely’ comparing to the group assumed in the current thesis that only owns coal and CCGT plants. This refers mainly to the fact that nuclear plants are the first to be dispatched because of their zero variable cost, and so in cases of low demand or of high hydro availability coal and CCGT plants reduce their probability of being dispatched.

- ✓ In all cases it is very important for a company to be diversified in terms of its generation assets and try to interpret in a correct way the signals obtained for example by models as the one described in the thesis project.

6.2 Further research – expansion opportunities

With the current thesis project the goal was the creation of a stochastic pricing model to be used in an environment of uncertainty, within a market created in a realistic way (in terms of demand, generation mix, etc). This model was based on an initial deterministic model, later evolved into a deterministic one that could simultaneously run a number of scenarios of 100% each.

In terms of evolution of this project, certain steps can be taken in the direction of improving the model structure as well as the data that were used for running it. Specifically:

1. The final model could be changed in order to include a higher level of complexity, mainly when it comes to its constraints or to technical limits regarding the operation of the plants.

2. An important element that could be also differentiated so as to have a more detailed and realistic model, is to include the whole of the hydro operation process, which means formulating the needed equations for flows, reserves, different types of basins etc. By this procedure it will be able to have a greater control on the production variables and maybe achieve even more realistic results.

3. In terms of the data used, it is important to apply the model to markets with a higher level of differentiation in terms of generation so as to examine more thoroughly the effect of RES production to the final prices, and provide the market with insight of how to achieve a more successful operation. Also, it can be proven useful to give more detail to the market structure possibly by providing the model with further information on maintenance and operational issues of the plants.

4. One other aspect of the model that could change its final results and give a further realistic sense to them is increasing even more the penetration of renewable technologies that is one major future expectation by EE countries and if applied it shall transform the current market situation into a new one in which companies should be able to settle in as fast as possible.

5. In its current form the model is quite simplified in the sense that the interactions among agents operating in the market are not taken into account and decisions are only based to calculated thermal plants costs. By providing the model with this kind of interaction companies could also gain insight on how to adjust their particular strategies more than the simple unit commitment decisions or future investments.

6. One certain assumption made at this thesis project was the correlation among demand and RES production scenarios during the clustering procedure. While this was essential for simplification reasons in terms of the current thesis project, operating the model without taking the above assumption into consideration might lead to a more exact final result and a greater insight on the exact formation of final prices, which significantly improves the model function.

7. As mentioned also above regarding the companies approach, what could be a good idea for improving the results both in accuracy and also in terms of being signals for the future is to examine the effects of the different models in the companies´ final margin, instead of incomes. By this way companies are expected to be helped even more in taking decisions because costs is an important factor and vital to know how it affects decisions when in need to.

BIBLIOGRAPHY

- A General Stochastic Volatility Model for the Pricing of Interest Rate Derivatives, **Anders B. Trolle / Eduardo S. Schwartz**

- An introduction to Stochastic Modeling, **Howard M. Taylor / Samuel Karlin**

- Modeling Electricity Prices: From the State of the Art to a Draft of a New Proposal, **Massimiliano Serati / Matteo Manera / Michele Plotegher**

- Introduction to Stochastic Modeling, **Prof. Peter W. Glynn**

- Stochastic Behavioral Asset Pricing Models and the Stylized Facts, **Thomas Lux**

- Stochastic Optimization of Electricity Portfolios: Scenario Tree Modeling and Risk Management, **Andreas Eichhorn,/Holger Heitsch/ Werner Romisch**

- Stochastic Modeling of Electric Power Prices in a Multi-Market environment, **Petter Skantze / Marija Ilic / Jeffrey Chapman**

- A Medium-Term Integrated Risk Management Model for a Hydrothermal Generation Company, **J. Cabero Borrós, A. Baillo Moreno, S. Cerisola Lopez De Haro, M. Ventosa Rodríguez, A. García Alcalde , F. L. Perán Montero , G. Relaño Cobián**

- Pricing Energy and Reserves Using Stochastic Optimization in an Alternative Electricity Market, **Steven Wong, J. David Fuller**

- Practical stochastic modeling of electricity prices, **Steven Wong, J. David Fuller**

- *An introduction to stochastic modeling*, **Howard M. Taylor, Samuel Karlin**

- *Scenario tree modelling for multistage stochastic programs*, **Holger Heitsch, Werner Romisch**

APENDIXES

1. *Thermal power plants of the system – capacities (MW)*

NUCLEAR PLANTS	CAPACITY	COAL PLANTS	CAPACITY	CCGT PLANTS	CAPACITY
NUC1	1000	CA1	200	CGA3	400
NUC2	1000	CB1	350	CGA4	400
NUC3	1000	CB3	350	CGA5	400
NUC5	1000	CB6	350	CGA10	400
NUC6	1000	CB8	350	CGB1	800
NUC7	1000	CB9	350	CGB3	800
		CC3	550	CGA11	400
		CA2	200	CGA12	400
		CB14	350	CGA13	400
		CB17	350	CGA15	400
		CB18	350	CGA16	400
		CB20	350	CGA20	400
		CC8	550	CGB6	800
		CC11	550	CGB8	800
		CC12	550	CGA21	400
		CC13	550	CGA22	400
		CC14	550	CGB10	800
		CB22	350	CGA23	400
		CB23	350	CGB12	800
		CB25	350	CGA24	400
		CB27	350	CGA25	400
		CC16	550	CGA27	400
		CC21	550	CGA30	400
		CB28	350	CGA32	400
		CC24	550	CGB16	800
		CA4	200	CGA33	400
		CB30	350	CGA37	400

		CB31	350	CGA41	400
		CC31	550	CGA43	400
				CGB21	800
				CGA44	400
				CGA48	400
				CGA50	400
				CGA51	400
				CGA53	400

2. *Thermal power plants of the system – technical minimum capacities (MW)*

NUCLEAR PLANTS	CAPACITY	COAL PLANTS	CAPACITY	CCGT PLANTS	CAPACITY
NUC1	350	CA1	100	CGA3	175
NUC2	350	CB1	150	CGA4	175
NUC3	350	CB3	150	CGA5	175
NUC5	350	CB6	150	CGA10	175
NUC6	350	CB8	150	CGB1	350
NUC7	350	CB9	150	CGB3	350
		CC3	170	CGA11	175
		CA2	100	CGA12	175
		CB14	150	CGA13	175
		CB17	150	CGA15	175
		CB18	150	CGA16	175
		CB20	150	CGA20	175
		CC8	170	CGB6	350
		CC11	170	CGB8	350
		CC12	170	CGA21	175
		CC13	170	CGA22	175
		CC14	170	CGB10	350
		CB22	150	CGA23	175
		CB23	150	CGB12	350
		CB25	150	CGA24	175
		CB27	150	CGA25	175
		CC16	170	CGA27	175
		CC21	170	CGA30	175
		CB28	150	CGA32	175
		CC24	170	CGB16	350
		CA4	100	CGA33	175
		CB30	150	CGA37	175
		CB31	150	CGA41	175
		CC31	170	CGA43	175

	CGB21	350
	CGA44	175
	CGA48	175
	CGA50	175
	CGA51	175
	CGA53	175

3. *Code used for formulating the model*

```
$title Probabilistic Pricing Model

*Statement of sets to be used

SET

mt      Months of the year

/Jan, Febr, Mrh, Apr, May, Jn, Jl, Aug, Sep, Oct, Nov, Dec/

da      Days of month

/1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31/

p       periods of the day

/P1*P6/

i       different scenarios

/S1*S15/

*S16,S17,S18,S19,S20,S21,S22,S23,S24,S25,S26,S27,S28,S29,S30/

;
```

```
SET

coal/

$call =xls2gms r=BX1:BX71
i=C:\Users\90038718\Desktop\Vasiliki_project\Data\plants_avail_coal.xls
o=C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\coal.inc

$include C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\coal.inc

/;


SET

others/

$call =xls2gms r=CK1:CK85
i=C:\Users\90038718\Desktop\Vasiliki_project\Data\plants_avail_others.xls
o=C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\others.inc

$include C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\others.inc/;

*Statement of parameters



PARAMETERS

qmin1(coal)      Minimum gross power of coal generators [GW]

qmin2(others)    Minimum gross power of nuclear & ccgt generators [GW]

;


VARIABLES

fobj             Value of objective function

;


POSITIVE VARIABLES

qa(mt,da,coal,i)           Net power dispatched by coal generator  [GW]
```

qb(mt,da,p,others,i) Net power dispatched by other generators (nuclear and CCGT) [GW]

q1a(mt,da,coal,i) Net power dispatched by coal generator above minimum stable load [GW]

q1b(mt,da,p,others,i) Net power dispatched by other generator above minimum stable load [GW]

;

*System Demand

TABLE

d(mt,da,p,i)

$call =xls2gms r=B3:U2235 i=C:\Users\90038718\Desktop\Vasiliki_project\Data\Demand_prop_initial3.xls o=C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\d.inc

$include C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\d.inc

;

*Generators´ capacity

TABLE

cap1(mt,da,coal)

$call =xls2gms r=A1:BV366 i=C:\Users\90038718\Desktop\Vasiliki_project\Data\plants_avail_coal.xls o=C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\cap1.inc

$include C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\cap1.inc

;

TABLE

cap2(mt,da,others)

```
$call =xls2gms r=A1:CJ366
i=C:\Users\90038718\Desktop\Vasiliki_project\Data\plants_avail_others.xls
o=C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\cap2.inc

$include C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\cap2.inc

;


PARAMETER

min1(coal)/

$call =xls2gms r=A2:B72
i=C:\Users\90038718\Desktop\Vasiliki_project\Data\Plants_min.xls
o=C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\min1.inc

$include C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\min1.inc/;


PARAMETER

min2(others)/

$call =xls2gms r=D2:E86
i=C:\Users\90038718\Desktop\Vasiliki_project\Data\Plants_min.xls
o=C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\min2.inc

$include C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\min2.inc/;


PARAMETER

pr(i)/

$call =xls2gms r=A2:B16
i=C:\Users\90038718\Desktop\Vasiliki_project\Data\Scen_prop.xls
o=C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\prop.inc

$include C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\prop.inc/;


*Generators´ Costs
```

TABLE

cost1(coal,mt)

$call =xls2gms r=A2:M72
i=C:\Users\90038718\Desktop\Vasiliki_project\Data\plants_costs.xls
o=C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\cost1.inc

$include C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\cost1.inc;

TABLE

cost2(others,mt)

$call =xls2gms r=O2:AA87
i=C:\Users\90038718\Desktop\Vasiliki_project\Data\plants_costs.xls
o=C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\cost2.inc

$include C:\Users\90038718\Desktop\Vasiliki_project\Data\INC\cost2.inc;

BINARY VARIABLES

u(mt,da,p,others,i) Binary variable indicating whether unit gen is connected (1) or disconnected (0) in each period (where unit can be NUCLEAR or CCGT)

u1(mt,da,coal,i) Binary variable indicating whether COAL unit gen is connected (1) or disconnected (0) in each period;

EQUATIONS

E_DMND(mt,da,p,i) Meeting the demand

E_FOBJ Objective cost Function

E_QMAXG1(mt,da,coal,i) Maximum power of generators that are coal

E_QMAXG2(mt,da,p,others,i) Maximum power of generators that are nuclear or ccgt

E_QMING1(mt,da,coal,i) Minimum power of generators that are coal

E_QMING2(mt,da,p,others,i) Minimum power of generators that are nuclear or ccgt

;

*Formulation of equations :

E_DMND(mt,da,p,i) ..

SUM[coal, qa(mt,da,coal,i)] + SUM[others, qb(mt,da,p,others,i)] =E= d(mt,da,p,i);

E_QMAXG1(mt,da,coal,i) ..

q1a(mt,da,coal,i) =L= (cap1(mt,da,coal)-min1(coal))*u1(mt,da,coal,i);

E_QMAXG2(mt,da,p,others,i) ..

q1b(mt,da,p,others,i) =L= (cap2(mt,da,others)-min2(others))*u(mt,da,p,others,i);

E_QMING1(mt,da,coal,i) ..

qa(mt,da,coal,i) =E= u1(mt,da,coal,i)*min1(coal) + q1a(mt,da,coal,i);

E_QMING2(mt,da,p,others,i) ..

qb(mt,da,p,others,i) =E= u(mt,da,p,others,i)*min2(others) + q1b(mt,da,p,others,i);

```
E_FOBJ ..

          fobj=E=
SUM[i,SUM[coal,SUM[(mt,da),SUM[p,cost1(coal,mt)*qa(mt,da,coal,i)*pr(i)]]] +
SUM[others,SUM[(mt,da),SUM[p,cost2(others,mt)*qb(mt,da,p,others,i)*pr(i)]]]];

MODEL Char
/
E_FOBJ
E_DMND
E_QMAXG1
E_QMAXG2
E_QMING1
E_QMING2
/
;

OPTION solvelink=0;
Char.dictfile=0;
OPTION sysout = on;
Option RMIP = CPLEX;
$onecho > cplex.opt
names no
threads 1
memoryemphasis 1
lpmethod 1
nodefileind 2
```

```
nodesel 0
varsel 3
cuts -1
$offecho
Char.OptFile = 1;
Solve Char USING MIP MINIMIZING fobj;

Display qa.l, qa.m, qb.l, qb.m ;

FILE    Results/C:\Users\uf763286\Desktop\stochastic_model\Results.dat/;
Results.pw = 100000;
PUT Results;
PUT 'UNIT PRODUCTION[MW]'/
PUT '*':10;
PUT '*':10;
PUT '*':10;
PUT '*':10;
PUT '':5;
LOOP(others,PUT others.TL:10);
LOOP(coal,PUT coal.TL:10);
PUT 'Dem. Térm.':10;
PUT 'C.Marg.':10;
PUT /;
PUT '*':10;
PUT '*':10;
PUT '*':10;
```

```
PUT '*':10;
PUT '':5;


LOOP(others,PUT 'MW':10);
LOOP(coal,PUT 'MW':10);
PUT 'MW':20;
PUT '€/MWh':10;
PUT /;
loop(i,
    loop(mt,
          loop(da,
               loop(p,
               if(d(mt,da,p,i)>0,
                      put i.tl:10;
                      put mt.tl:10;
                      put da.tl:10;
                      put p.tl:10;
                      loop(others,
                        put qb.l(mt,da,p,others,i):10:0;
                      );
                      loop(coal,
                        put qa.l(mt,da,coal,i):10:0;
                      );
                      put d(mt,da,p,i):10:0;
                      put E_DMND.m(mt,da,p,i):10:2;
                      put/;
```

```
                    );
                );
            );
        );
put/;
put/;
);
```