Swedish and Spanish electricity market
Comparison, improvements, price forecasting and a global future perspective

Edgar Bahilo Rodríguez
09-06-2017

Student thesis, Advanced level (Master degree, one year), 15 HE
Energy Systems
Master Programme in Energy Systems

Supervisor: Shahnaz Amiri
Examiner: Taghi Karimipanah
Abstract

This report aims to make a comparison between the Swedish and Spanish electricity market, the design of new improvements that could achieve a better operation for both markets as well as the price forecasting for both spot markets. These enhancements are oriented to decrease electricity prices, energy use and the system CO₂ emissions.

Also, the main organizations of the market and their roles has been characterized, clarifying the functions of the Market Operator and the System Operator. In addition, the different markets, the trading products and the price formation have been explained and the picture of the market structure has been achieved with enough depth.

Moreover, some of the most used methods in Time Series Analysis has been enumerated to understand which techniques are needed for forecast the electricity prices and the methodology used (Box-Jenkins Method) has been explained in detail. Later, all these methods have been implemented in an own code developed in Python 3.6 (TSAFTTools.py) with the help of different statistics libraries mentioned during the method chapter.

On the other hand, the description of the market situation has been carried out for both countries. Power installed capacity, electricity generation, average prices, main renewable technologies and policies to increase the renewable energy share has been analysed and corresponding described.

Then, to estimate the market’s future spot electricity prices, ARIMA models have been selected to analyse the evolution of the day-ahead price using the TSAFTTools.py. The final models show a proper performance in the two markets, especially in the Nordpool, achieving an RMSE: 37.68 and MAPE: 7.75 for the year in 2017 in Nordpool and a RMSE: 270.08 and MAPE: 20.24 in OMIE for 2017. Nordpool spot prices from 2015 to 2016 has been analysed too but obtaining a result not as good as the year 2017 with an RMSE: 49.01 and MAPE: 21.42.

After this analysis, the strengths and weaknesses of both markets are presented and the main problems of the Spanish electricity system (power overcapacity, fuel dependency, non-cost-efficient renewable energies policies, lack of interconnexion capacity etc.) and the Swedish electricity system (dependency for nuclear power, uncertainty for solar electricity Generation) are presented.

Finally, due to the quick development of the energy sector in the last years and the concern of the European Committee to reach a new design for the electricity market, different kinds of recommendations for the future have been considered.

Acknowledgments

Along the course of my studies I have grown as a person and as an engineer and this is mainly because all the wonderful friends and university colleagues that I have met in all this years. Everything that I have learned is in part because of them. I would like to take this opportunity, which is my final step in my studies, to thank all of them, especially to the following ones: Carlos Moya, David Martínez, Marcos Llamazares, Mario Blasco, Víctor Sevilla, Pepe Bonet, Andrea Gabaldón, Pilar Calatayud, Moisés Antón, Lucía Arcos, Silvia Alba, Leonie Schneh, Adrián Gómez, Lierni Arnaiz, David Gracia, Alvar Villegas and Óscar Almingol.

Secondly, I would like to express my gratitude to my mother who took care of me when my father did not have his best moments. Of course, I would like to say thank you to my father because he looked after me when my mother could not do it anymore.

Besides my friends and my family, I wish to thank the rest of the teachers that supported me in this thesis project: Dr. Nawzad Mardan, Dr. Saeid Homayoun, Prof. Shahnaz Amiri from Högskolan i Gävle. Then, I want to show my gratitude to Prof. Guillermo Escrivà from Polytechnic University of Valencia for their encouragement and insightful comments.

Last but not the least, I would like to show my deepest love and gratitude to my couple Rocío, who has given to me all the support that I needed these last years.
Acronyms

AC: Alternating Current.
ACF: Autocorrelation Function.
ADF: Augmented Dickey–Fuller test.
ANN: Artificial Neural Networks.
AR: Autoregressive Model.
ARIMA: Autoregressive Integrated Moving Average.
ARMA: Autoregressive Moving Average.
ARMAX: Autoregressive Moving Average Model with eXogenous inputs.
ASM: Ancillary Services Markets (Adjustment Markets).
DSO: Distribution System Operator.
EC: European Committee.
ELF: Electricity Load Forecasting.
EM: Electricity Market.
EPF: Electricity Price Forecasting.
EUPHEMIA: Pan-European Hybrid Electricity Market Integration Algorithm.
ETD: Exchange-traded derivative.
EU: European Union.
GARCH: Generalized Autoregressive Conditional Heteroscedasticity.
GDP: Gross Domestic Product.
MA: Moving Average Model.
MAPE: Mean Absolute Percentage Error.
MCP: Market Clearing Price.
MLE: Maximum Likelihood Estimation.
MO: Market Operator.
OTC: Over the Counter.
PACF: Partial Autocorrelation function.
P/C: Generation / Capacity
PCR: Price Coupling of Regions.
**RMSE**: Root Mean Square Error.

**SsES**: Spanish Electricity System.

**SES**: Swedish Electricity System.

**SARIMAX**: Seasonal Autoregressive Moving Average Model with eXogenous inputs.

**TNREG**: Total Net Renewable Electricity Generation.

**TS**: Time Series.

**TSA**: Time Series Analysis.

**TSO**: Transport System Operator.
Table of Contents

1. Introduction
   1.1. Background
   1.2. Aims
   1.3. Limitations
   1.4. Literature review
   1.5. Approach

2. Theory
   2.1. Electricity Market Operation:
      2.1.1. System Operator
         2.1.1.1. Spanish Operator
         2.1.1.2. Swedish Operator
      2.1.2. Market Operator
         2.1.2.1. OMIE and OMIP
         2.1.2.2. Nordpool Spot and Nasdaq OMX
         2.1.2.3. Strategic Bidding and Market Clearing Price (EUPHEMIA)
         2.1.2.4. Market Splitting
   2.2. Time Series Analysis
      2.2.1. Stationarity
      2.2.2. Classical Method (Decomposition)
      2.2.3. Statistical Models: Box Jenkins Method
         2.2.3.1. Model Identification
         2.2.3.2. Model Estimation
         2.2.3.3. Model Checking
      2.2.4. AR Models
      2.2.5. MA Models
      2.2.6. ARIMA Models

3. Method
   3.1. Comparison between the two markets.
   3.2. Forecasting method (TAFTools.py and FNRVTTool.m)

4. Results
   4.1. Electricity use per capita
   4.2. Average Price of the kWh and Price components by sector
   4.3. Electricity Generation and Power Capacity Evolution
   4.4. Main Renewable Generation Technologies
      4.4.1. Wind Power
      4.4.2. Hydropower
      4.4.3. Biomass for Electricity
      4.4.4. Solar Power
   4.5. Summary of subsidies
      4.5.1. Swedish Case: Green Certificates
      4.5.2. Spanish Case: Special and ordinary system, fixed tariff and premium systems.
4.6. ARIMA Results
   4.6.1. Nordpool Spot Prices
   4.6.2. OMIE Spot Prices

5. Conclusions
   5.1. Study results
   5.2. Outlook & Perspectives

6. References

7. Appendix.
   7.1. Comparison Graphs
   7.2. Python Code: TSAFTools.py (Time Series Analysis and Forecasting Tools)
   7.3. FNRVTol.m (Fixing NaN and Repetitive Values Tool)
   7.4. ARIMA (1,1,0) Results Nordpool Spot Prices 2015-2016
   7.5. ARIMA (1,1,0) Results Nordpool Spot Prices 2017
   7.6. ARIMA (1,1,0) Results OMIE Spot Prices 2017
1 Introduction

1.1 Background

It is being known that the energy use has a tremendous importance nowadays. Society as we know it would be very different without the electricity or heating use, energy is in practically all the fields in the world and due to its importance, different management methods have been created for having the best possible energy supply at every moment.

Since the introduction of the electric lighting in 1882 [1], the use of electricity has increased with a speed never seen before. Moreover, the introduction of the AC systems as transformers, power conversion stations and the power lines caused that the electricity use was scattered for the whole world in a few years. However, it was not until the 1980-1990 that the main topic of this research, the electricity markets, was developed.

The key event for the birth of today’s electricity markets took place in 1990 [2], when the UK Government under the presidency of Margaret Thatcher privatised the UK electricity supply industry. After this event, the natural monopoly that ruled the electric power sector was gradually dissolved and the liberalisation of electricity markets started in the entire world. [3] [11]

The final purpose of the electricity markets, as it has been mentioned before, is to provide with the best possible energy supply at every moment. Nevertheless, it is important to understand that the electricity markets must connect two different realities. The first one is the physical grid, which rules the flow of electricity. This grid is formed by the electricity generators (power plants) and the electricity transmission system. The second one is the financial system, which rules the money exchange and consists of electricity suppliers, consumers, TSOs, DSOs, regulators and MOs. [4]

Due to the need of satisfying these two realities and the privatisation of the power sector, electricity has become a traded commodity using spot and derivative contracts (OTC and ETD). However, electricity is not a normal commodity because its storage is complex, currently expensive and not scalable to big sizes [5]. Moreover, the physical grid stability requires an immediate balance between supply and demand which adds another constrain to the energy trading.

EPF and ELF play a key role in the new electricity liberalised system, helping both sides of the financial market (suppliers and consumers) to achieve their optimal energy generation and use respectively to maximize their profitability. Also, the extreme volatility of the electricity market allows companies that are able to forecast the future prices and loads to have an incredible advantage over their competitors. In fact, the possible savings could be up to 2.600.000 SEK/y using both techniques. [6]

With the creation of the European Union and the Common European Market, the EC has been making diverse efforts in order to join the country’s electricity markets in a unique one. One of the most important work lines that the EC has developed is the PRC project. Basically, the PRC is a common algorithm which gives us a fair and clearly determination of the day-ahead electricity markets and the position of the bidding area across Europe.
This single solution has the purpose of increasing the efficiency, the liquidity and the social benefit of the electricity markets. [7]

But the EC has not only been focused in achieving the fair determination of the MCP for all the Eurozone. Another of the huge efforts performed by the EC is the Horizon 2020 project, which one of its main purposes is achieving secure, clean and efficient energy. Having a 5.931M€ budget, the priorities of this program are the increase of the energy efficiency, the development of low-carbon technologies and the creation of smart cities.

Under this economic political scenario, many different countries of the EU are trying to increase their renewable energy supply. These nations created a bonus to promote the renewable mix of the country. Nevertheless, these bonuses have had different effects and different implementations depending of each country. Of course, there is not a low carbon technology that is better than the other but the actual trend is perform this rise with wind and solar power, especially when their costs are decreasing faster than expected. [8]

As regards to the traditional renewable energies (hydropower and biomass), in 2016 Spain had a hydropower net electricity generation of 35.8 TWh and a biomass electricity generation of 4,615 TWh (including geothermal and marine hydropower). Meanwhile Sweden had a hydropower net electricity generation of 75,312 TWh a and attending to the biomass, the corresponding electricity generation was 13 TWh b.

When it comes to the “new” renewable energies, in Sweden the wind energy generation in 2016 was 15,496 TWh d meanwhile in Spain was 48.6 TWh e. In terms of solar energy generation, as it would be expected, the Spanish one was clearly higher than the Swedish with 12,96 e TWh versus 0.097 TWh f.

Performing a more global balance and including the other renewable sources (biogas, biofuel, marine power, geothermal and waste) the TNREG of Spain in 2016 was the 38.9% (107.46 TWh) g in front of the Swedish one that was 63% (94.5 TWh) d. Having described these general lines the question to answer is: if it is supposed that renewable energy is more expensive than the non-renewable and it needs subsidies, how can be possible that the Spanish electricity prices are more expensive than the Swedish ones? [9]

As stated above, wind and solar electricity generation are higher in Spain than in Sweden. Nonetheless, it should be pointed out that this situation increases the complexity of the grid system, due to the generation volatility of this two resources. Because of that fact, the SsES needs a “backup” power generation system which is done using CHP, Gas and Fuel Power Plants. Those requirements and the subsidies gave also to the CHP power plants, have led the SsES to a power overcapacity [10]. Moreover, the debt created by the huge amount of power installed in Spain has created a deficit in the SsES that still has effects on the country’s economy [11].

Multiples investigations are required of how Sweden avoid the overcapacity and the debt problems and how the models created for EPF and ELF techniques could enclose the risk of increasing the sporadic electricity generation from solar, wind another renewable resource.

---

[a] Data extracted from the Statistical Centralbyran.
[d] Data calculated using the hourly generation profiles supplied by NordPool.
1.2 Aims

The aim of this research is describing the Swedish and Spanish electricity markets and the differences between them in terms of price, working way, technologies and developing some suggestions to improve the operation of both markets.

On the other hand, on account of the concern of the European Committee about having a unique market, some recommendations will be proposed for the future.

In addition, bearing in mind the financial problems mentioned in the background section that the Spanish market faced in the past, the author of this research will try to explain which were the main problems of the bonus tariff implemented in Spain and the advantages of the Swedish one.

Finally, with the help of a model developed in python language, it is expected to compare the price evolution in both market and the influence of the seasonality in both markets.

1.3 Limitations

Energy is not only about electricity, however due to the maximum extension that this thesis can have, the author of this report has preferred to delimitate this study only to the electricity market to get a more accurate perspective in only one topic. However, that do not mean that if is necessary, and if is going to clarify some aspects, it might be possible to talk about the primary energy or the district heating and fuels market.

Regarding the data validity in this document, only data from the following official and non-official organizations has been taken: NordPool AS (Market Operator of Sweden), OMIE (Market Operator of Spain), REE (Spanish Transmission System Operator), Svenska Krafträ (Swedish Transmission System Operator), Eurostat, INE (Statistical Spanish Office), Statistiska Centralbyrå (Statistical Swedish Office), MINETUR (Spanish Ministry of Energy, Tourism and Digital Agenda), UNESA (Spanish Association of the Electricity Industry), Energimyndigheten (Swedish Energy Agency), IEA (International Energy Agency), The World Data Bank, World Energy Council and ENTSOE (European Network of Transmission System Operators for Electricity). It has been observed differences depending of the used source, but they are not higher than 3%.

In addition, for making the comparison it has been considered the data from 1990 until 2014. It is true that the year of this publication is 2017 but, the electricity stats normally are revised several times and in fact the data about 2016 is not available yet and the data from 2015 is being checked right now.

Moreover, the forecasting method and model created are only be able to predict the day-ahead prices due to the complexity of integrate the derivative contracts and future exchanges.

On the other hand, ARIMA has been selected due to the problems appeared using SARIMAX models with python in Windows. The main library using for forecasting the values (StatModels 8.0) does not support any more the parameter statsmodels.tsa.statespace.sarimax.SARIMAX, and the only possibility for using that kind of models is installing an older version (specifically the 6.0) which has not been
possible. In addition, the maximum d order admitted for stats models is d=2 so the maximum times that the TS can be differenced (read chapter 2.2.3) is 2.

The author of this report does not have a huge knowledge of Swedish. Some information has been difficult to find, and sometimes difficult to understand, especially when was related with legislation. Probably some aspects about the Swedish market are missing due to this fact.

1.4 Literature review

1.4.1. Energy in the World

The importance of carrying out this research is based on the increase of the energy use in all Europe and developing countries. In the past years, several attempts have been done to find a correlation between the economic growth and the energy use. However, there were contradictory results until the 90s when the time series methods made significant advances. Different kind of models with several variables have been developed since that, and nowadays the trend is considering this relation as a dynamic one.

Dividing the countries in different groups depending of their income (low income, lower and upper income groups and high-income groups) the relationship between the energy demand growth (average) and the GDP growth is fairly different. Strong positive relationship appears in the low income and middle-income groups, but there is not the same relation in the high-income ones. It is understandable if we take care that after the financial crisis the high-income groups tried to improve their efficiency in the energy use and decrease their CO2 emissions. [12]

Remarking this last observation, it is possible to understand the importance of the energy markets, especially the electric one, owing to the possibilities that a flexible and a better electricity market can provide. A market that allows the integration of the renewable energies, the consumer’s actions (for instance demand response), and a strong interconnexion capacity, namely a develop market can afford the fact of decrease the CO2 emissions and the energy use without neglecting the economic growth.

This kind of possibilities have been made relevant by the EC [13] and it is part of a new electricity market concept. This new strategy begins with giving more decision power to the consumers to participate actively in the market with the purpose of decrease their electricity bill [14]. Also, intelligent energy systems which are part of the smart grids will have an increasing importance in the future empowering the consumers for deciding what it is the best moment to consume electricity or not.

About the technology in the electricity market, the European Union wants to increase the participation of all the technologies available in equal conditions. It is important to mention that it is necessary invest and support the renewables energies but based on the market mechanics. Finally, the importance of the market coupling is set as a mainstay for the next years with the purpose to approach the technical actions in the reality to the financial actions in the market.
1.4.2. The Swedish and Spanish Case

Attending to the comparison between the Swedish and Spanish electricity market there can be found a lot of similarities between both. On the one hand, the processes that involve the operation of the markets are the same: generation, selling, transport and distribution [15] and the main organizations make the same tasks. On the other hand, the market working way is a little bit different, for instance, in Sweden there are a lot of grid companies which manage the grid by each local area, while in Spain there is only one [16]. Also, the regulation band in Spain is done by adjustments in the supply with other generators and in Sweden it is also done using the bidding strategy [17].

Regarding to the technology in Sweden most of the electricity is produced by nuclear and hydropower power plants (84%), [18] while in Spain [19] the mix is more varied and any technology has more than the 25% but with a significant wind power installed.

Taking a closer look in wind power generation, another of the hindrance that it has is the size and the technology available in the current local grids. This is a common thing in both markets, in fact, in Spain must reinforced them to surpass the 10000MW of power capacity installed [20]

About the policies created in Sweden for increase the wind energy generation in 1991 a bonus of 25% was introduced, which was increased until 35% in 1993. The second investment support period was from 1998 to 2002 with a subsidy of 15% trying to increase 0.5 TWh annually. In the Spanish case, the first objectives were fixed in the Policy Plan for the Promotion of Renewable Energy' (PFER) 1991-2000, with the goal of 175MW. Well, the capacity installed in 2000 was 2500MW which was an underestimation possibilities world record.

This fact can be explained due to the special tariffs than the green energies had during this years, and because of the grind connection investment support too. This connection investment was up to 30% since 1980, it was not until the creation of the special regime (which guarantees the right to grid connection, a standard five years’ purchase contract and a defined (revisable) price per delivered kWh). A Royal Decree from 1998 (2818/1998) introduced different tariffs and environmental premiums for different green technologies. Win energy generation tariff was between 6,2 and 6,6 eurocents/kWh from 1999 to 2004. [21]

Nevertheless, the tariff system lead to Spain to a huge deficit and also to a lack of wind development companies in comparison with other countries as Germany [22], the 58% of the total installed capacity of wind power was done for already existing companies.

As regards to the non-renewable energy resources there is a huge difference between Sweden and Spain. Spain as well as other countries (France, Germany and UK) decided to support their coal industry, which is a disadvantage for the renewable energies in competitive terms [23]. Also, another blockage generated by the non-renewable energies is the use of nuclear energy.

Besides the nuclear power is about the 40% of the total electricity generation, Sweden had it a nuclear phase-out policy that wanted to strike out the nuclear power generation by 2010. This aim was not achieved but was one of the energy political driven forces that
lead to increase their renewable generation. However, Spain approved in 2011 a law that allow the nuclear power plants to work more than 40 years if the CSN (Security Nuclear Council) gives its endorsement. [24]

It should be highlighted that besides the differences between both markets, the flow dispatch follows the same methodology. The ED (Economic Dispatch) is the main problem in the system operation and has the aim of satisfying all the system constrains with the minimum cost. Basically, ED is a huge optimization problem that considers multiple number of variables and different kind of them (linear, nonlinear characteristics etc.). [25]

1.4.3. Forecasting Methods

With the purpose of managing the electricity markets in the best way possible, different mathematical methods have been used for predicting the energy demand and the price of the electricity. Predicting these variables, future strategies about the construction of new power plants can be done and new investments about interconnexion capacity or which technology is the best for building it can be chosen. Moreover, the electricity price forecast allows the energy producers to adjust their power plants to the optimal benefit point at every moment, as well as in the future would determinates the possibilities of the consumers to achieve a less expensive invoice. Definitely, forecasting is essential for make possible the power flow dispatch [26]

Electricity prices and energy load (demand) are highly volatile by nature, mainly because the energy supply and the energy demand must be balanced in real time. Due to the difficulty of storing the electricity, the market must be fast adapting to the real needs [27]. In order to forecast the electricity demand and prices the Time Series Analysis is used. Normally, it is possible to distinguish two kinds of models: the linear and nonlinear models. But independently of the chosen method the main objective to make an accurate model is to properly select the major drives that affect to the demand and the price. This can be: number of plants, fuel prices, temperature, hydropower reserve, wind speed, plant outages, emission cost, transmissions etc.

One of the most popular linear models are the ARIMA models that have already been used for forecasting a lot of commodities as oil price, gas etc. with enough good results. Currently, there are so many examples of the using of ARIMA models for forecasting, for instance the simpler version of ARIMA models (the AR models) which have been used by Nordpool in Norwegian systems for forecasting the prices of the week. Also in Spain and California markets, ARIMA models have been used for predicting the next day prices. [28] [29]

Moreover, there is the possibility of mixing different forecasting methods to improve the accuracy or just in order to use the best part of these methods depending on the situation. Different kind of simulations have been done mixing the ANN and ARIMA models with the purpose of integrate the impact of the renewables energies in the predictions. [30]

Besides the forecasting possibilities, it is important to talk about TSA, one of the most important issues before trying to do EPF and ELF is knowing if the time series that is
going to be analysed is stationary or not. If the TS is not stationary there are many transformations that can be applied in order to make it stationary. Normally this is done applying some of the Box-Cox transformations [31] [32]

Sometimes, ARIMA models are not enough to analyse time series data, that is because some TS have strongly seasonal components or strong cyclical components. For that kind situations, there are another models more appropriated, the GARCH models and SARIMAX models. In fact, it has been observed that seasonality and local circumstances are more important than other variables than the temperature [33]

As ANN models, GARCH models can be combined with ARIMA models to decrease the MAPE and the RMSE of the forecasting models [34]. The main advantage of the GARCH models in front of the ARIMA, is that they are more proper to capture the dynamic part of the TS. In addition, GARCH models requires four phases: data preparation, model identification, parameter estimation and diagnostic checking. [35]

1.4.4. Literature review conclusion

Energy is a topic currently non-solve, besides all the advances developed with the creation of the EM and with the project of making a joined market, there is still being gaps in the energy development. Not every country is concerned about the energy efficiency or about emissions produced by the energy generation and the global energy use is growing up year by year.

Regarding to the Swedish and Spanish EMs, it should be noted that these markets are from high-income countries and that is the reason because they are making several efforts in order to increase TREP. However, the literature review about the different energy policies in both countries is large, and there are so many different conclusions depending of the paper reviewed. This is one of the main driving force of this report, trying to explain why having Sweden a higher energy use rate per person and a more TNREG, the prices are lower than in Spain.

Finally, it has been showed that there many statistical forecasting techniques that can be chosen. Every model mentioned has its strengths and weaknesses but in the end, there are only models, they are not perfect and because of that, combined methods has been developed in order to get more accurate forecasts. Despite this trend, the future of forecasting is uncertainly but probably forecasting techniques would be used more frequently in the future.

1.5 Approach

The comparison has been made using the following structure: data collection, data analysis and data comparison.

As regards to the forecasting model:

- Results of the simulation
  - Hourly based model
  - Monthly based model
2 Theory

The following section is going to be subdivided in two main parts. The first one is about how the electricity market works and its main organizations. The second one is about time series analysis and the model that has been selected to forecast the spot electricity prices.

2.1 Electricity Market Operation

The system operation includes the necessary activities to guarantee the safety and the continuity supply of the electricity, as well as the correct coordination between the electricity generators and the transport grid. Ensuring that the electricity produced by the generators will be carried until the distribution networks in the quality requirements ordered by the current legal normative. The system operation is the technical operation of the EM and it is the duty of the TSO.

However, the electricity market operation needs not only technical management, but also the financial management. Financial management includes the possibility of offer trading, clearing and settlement services in day ahead and intraday markets. Moreover, the financial operation must integrate the OTC and the ETD contracts and inform the system operation about them to settle the electricity balance. The financial operation is a must of the MO.

Figure 2.1: Electricity System Time Line: Market and Physical.
2.1.1. Transport System Operator

2.1.1.1. Red Eléctrica de España

Besides the obligations stated above, the TSO is the manager of make the load forecasting and handle the generators and the power lines in real time, achieving that the programming supply in the power plants coincides with the consumer’s electricity demand. Otherwise, the TSO send the necessary orders to the generator to adjust their production, increasing or decreasing the electricity generation, so that sufficient generation margins are maintained to face possible losses of generation or changes in expected demand.

In Spain, the TSO is Red Eléctrica Española (REE) while in Sweden is SK (Svenska Krafnät). The Spanish TSO manages the imbalance of the EM with the AS markets, allowing the electricity generation planning freely established by balance responsible parties in the day-ahead market and bilateral contracts, and subsequently in the intraday markets to be adjusted to the quality, reliability and security requirements of the electricity system. ASM are used as the available tool for resolving technical constraints of the system, the allocation of ancillary services and the management of deviations. [36]

Moreover, REE is the responsible for annual power demand evolution predictions for long term. As has been mentioned several times in this report, forecasts are essential to plan the electricity grid development and the electricity generation planning.

Also, all the market producers and consumers that are connected to the electrical grid has to pay a fee to obtain access to the grid. These fees are composed of the access term (which depends of: the amount of power that the consumer wants, the voltage that the consumer requires and the tariff period) and the operation term (which correspond to the payment for the system operation tasks).

To get more in deep in the ancillary services or adjustment markets, there are formed by the following ones:

I. Resolution of technical constraints:

Services with the purpose of resolve this constrains limiting and modifying, if it is required, the supply schedules of the power stations or pumped storage generation that allow technical constraints identified to be resolved with the lowest cost for the system, and the subsequent rebalancing of generation and demand to offset schedule modifications incorporated to resolve the identified technical constraints. This constrains can be caused by the unfulfillment of the security conditions, the lack of reserve power, the absence of reserve of secondary and / or tertiary regulation, the capacity shortage for the voltage control and the insufficient power capacity to restore the normal functioning. This service is regulated by the P.O. 3.2 (Operation Procedure Proposal 3.2)

II. Adjustment Services

Services needed to secure the energy supply in the safety, quality and reliability required conditions (network frequency stability and voltage value in the electrical transport nodes). Are as follows:
**Additional Upward Reserve Power.** Ancillary service, of an optional nature, managed and remunerated by means of market mechanisms and whose aim is providing the electricity system with the required level of upward power reserve, taking into consideration the power reserve available in the forecasted schedule of the day-ahead horizon.

**Secondary Control Band.** Ancillary, of an optional nature, whose purpose is to maintain the generation-demand balance, correcting deviations with respect to the planning power exchange of the ‘Spain’ Control Block, and the system frequency deviations. Its action horizon is between 20 seconds and 15 minutes. This service is remunerated by means of market mechanisms via two concepts: availability (control band) and usage (energy). [37]

**Tertiary Control.** Ancillary service, of an optional nature, that, if subscribed to, is accompanied by the obligation to bid, and is managed and remunerated by means of market mechanisms. Its purpose is to resolve the deviations between generation and demand and the restoration of the secondary control band reserve used. The tertiary control band reserve is defined as the maximum variation of power generation that a generation unit can perform within a maximum of 15 minutes, and which can be maintained for at least 2 hours.

![Figure 2.2: ENTSO-E’s recommendations for frequency control.](image)

**III. Divergence Management**

The divergence management is an optional nature service managed and remunerated by means of market mechanisms. Divergence management has the aim of solving the deviations between generation-demand identified after each intraday market session closure until the effectivity horizon of the next session.
2.1.1.2. Svenska Kraftnät

Svenska Kraftnät is the TSO for electricity in Sweden. As a TSO, it has the overall responsibility to keep the balance between electricity supply and demand.

Likewise, that REE, the balance management is based on planning how much energy is going to be used in the electricity system. Both sides of the market (suppliers and the TSO) make forecasts of the amount of energy that will be consumed by each hour. After, the responsibility of delivering the correct amount of electricity falls to the suppliers. In operation, the balance is done by automatic and manual up- and down-regulation of generation or demand.

Due to the Swedish weather, and the quantity of hydropower that the country has, the hydro reserve forecasting and the temperature forecasting are essential in Sweden. Weather effects the amount of electricity produced and consumed and the difference between the coldest and the warmer day could be a 40%.

In the same way that REE, SK is the responsible for receiving the payments of the grid access fees. The connection tariff consists of [38]:

- Capacity fee. This fee depends of the geographical latitude.
- Hourly usage fee. Calculated using the product of a coefficient of loss for the connection node, the input or output of energy at this point and the current price of the energy loss.
- Initial connection fee. Only required when the investment is significant and it is only going to benefit a few market players.

Besides the duties of an TSO, SK is the responsible of creating the rules and the agreements for the electricity trading in all de country and between the rest of the Nordic countries.

Regarding the methods that SK has in order to restore the supply-demand balance there mainly three:

I. Technical management.

When there is instability in the physical grid because unplanned events like a switching of a power station or a transmission lines, or because an unplanned low temperature the TSO must be prepared for managing this load variability. In order to do that, Fingrid (Finland TSO) and SK have acquired a total amount of 2600MW peak load reserves. 600MW are placed in Finland and 2000MW in Sweden. The sharing of both reserves is in accordance with the current legislation in both countries. [39] [40] [41]

Also, in the case that the power reserve will be not enough, and there is a risk of a power shortage or a blackout. The TSO can cut down the electricity transmission to the consumers. However, that situation never happened yet.
II. Counter – Trade.

When the transmission flow needs to be reduced between to Swedish areas, an increase of the supply can be ordered in the deficit area meanwhile in the surplus zone the supply is decreased. Using this service, SK manage the imbalance between each region in Sweden. [42]
III. Physical balance maintained by trade and balance regulation.

Before delivering the electricity, the balance providers can trade the electricity in order to plan their physical balances. Balance services assume control over balance management at the delivery hour. SK accepts bids [volume (MW) and price (SEK-MWh)] from generator that can deal with a quickly (less than 10 mins) increase or decrease in their production level. Also, the consumers can submit bids for increasing or decreasing their energy use level. Always the balance service actives only the most favourable bids each hour.

2.1.2. Market Operator

In theory, there is not constrains that in a liberalized market every player could trade the energy using any kind of bilateral agreements at every time. However, almost every electricity market has created an organized short-term market (normally called spot market) managed by an entity called the market operator. In fact, this is the main task of the MO, in Sweden the electricity volume traded using the spot market in 2011 was approx. the 70% while in Spain was the 87,17% the same year.

The Swedish MO is Nordpool Spot AS and Spanish MO is OMIE. Nasdaq OMX and OMIP are respectively the platforms which manage the future exchanges in each country and which are permanently coordinated with the MO.

2.1.2.1. OMIE and OMIP

As stated above, the management of the spot market is carried out by OMIE (Iberian Spanish Market Operator). The negotiation of the spot market is built on a daily pool, with strategic bidding for all the hours of the following day with the same price in Spain and Portugal (except Market Splitting).

After the daily pool, there are several intraday market sessions where players can negotiate the electricity for the different hours of the day covered by the day-ahead market. This negotiation is done by a pool too.
The day-ahead market as a part of the EM, has the purpose of perform the electricity exchange for the next day. The market agents send the sales and purchase bids, which are included in the bidding strategy for each 24 periods of the next day.

All the Generators above 50MW that not have a bilateral agreement must present offers to the day ahead market. Generators below 50MW and special system (CHP and Renewables except normal hydropower) can take part optionally.

Normally the intraday market is used as balance market for the day-ahead markets. Allowing market players to have more accuracy in their generation or electricity use plans, this market is currently structured into the six sessions showed in the graph:
And the structure of all the time steps in the market can be observed with detail in the graph:

**Figure 2.8: Time Line of System and Market Operation (OMIE).**

On the other hand, the Iberian Electricity Market (MIBEL) is formed by other entity called OMIP (Iberian Portuguese Market Operator) which is the ETD market for Iberian and non-Iberian products. In the future, it is expected that both organizations (OMIE and OMIP) will fusion in only one the OMI (Iberian Market Operator). Also, there exists a Clearing House and Central Counterparty in all the operations done by OMIP, this organization is named OMIClear and has the duty of clear trades on the OTC markets. Moreover, OMIClear can clear trades of another markets that have, as underlying assets, energy based products or similar. The structure is basically the introduced below:

**Figure 2.9: OMIP Main Participants and Organizations.**

The main products traded in OMIP are also the following ones:
Although the main volume of electricity is traded in the spot market, it is possible to combine all the products available in the EM, achieving a decrease of the electricity cost. An example is showed in the figure [43]:

Figure 2.10: OMIP Future Products.

Figure 2.11: Mixed Electricity Supply: Spot Market (OMIE) and Future Exchanges (OMIP).
2.1.2.2. Nordpool Spot AS and Nasdaq OMX

The Nordic electricity market, Nordpool Spot AS is the manager of the electricity trading for physical supply, that is, the spot market. Market players submit the sales and purchase bids to Nordpool and the system price (spot price) and the area prices are calculated for each hour a day ahead of the supply period.

One of the main differences between Sweden and Spain is that Sweden has four bidding areas, from Lulea (SE1 area) to Malmö (SE4 area). The price of each bidding area is calculated by the amount of electricity delivered, the demand of electricity and the transmission capacity between all the areas. Only when all the transmission capacity is used for electricity trade between the two bidding areas, they will have different prices in each area.

As well as OMIE, Nordpool manages the spot market of other countries, in fact Nordpool is the manager of practically all the Nordic countries (Sweden, Norway, Estonia, Lithuania, Denmark and Finland).

Nordpool is mainly formed by two different markets: Elspot and Elbas, the first one is the day ahead-market in which market players can send their bids no later than 12:00 PM the day before to the day delivery. The second one is the intraday market that provides continuous power trading 24 hours a day, covering individual hours, up to one hour prior to delivery.

The gap between the day's Elspot price-fixing round and the actual delivery hour of the concluded contracts can be quite long (36 hours at the most). As consumption and sales situations change, a market player may find a need for trading during these 36 hours.

However, like the Spanish EM, Sweden has its own place for doing futures contracts. Nasdaq OMX includes futures, options, DS futures and Electricity Price Area Differential contracts. An approach to the electricity trade in the Nordic region can be observed in the graph:

![Diagram of Nordic Electricity Market](Figure 2.12: Main organizations in Swedish Electricity Market)
2.1.3. **Strategic Bidding and Market Clearing Price (EUPHEMIA Algorithm)**

As has been explained in the section about the respective market operators of Sweden and Spain, the market players send their purchase and sell bids to the system operator. The Strategic Bidding is the process where the buyers and the sellers try to maximize their profit. In Auction theory, this process depends directly of the auction types that are used in the market.

The EM is based in a multi-unit auction, each unit is a MW sold and bought by a market player. There are mainly two kinds of multi action: the uniform-price auction (used in Nordpool Spot and OMIE) and the pay-as-bid auction (using in APX market). However, game theory demonstrates that the final price is quite similar [44].

![Figure 2.13: Uniform-Action MCP vs Pay-As-Bid Auction MCP.](image)

The electricity demand is covered with the cheaper generators and later the more expensive options are being incorporated until the cross between demand and supply. Therefore, the final price of the electricity is the price of the more expensive MWh that has been sell it each hour.

EUPHEMIA is the algorithm used for calculating the market clearing price in both EMs. As has been mentioned above, the incorporation of generators is stopped when the demand and supply curves will cross, this is called the MCP, Market Clearing Price.
The main objective of EUPHEMIA is to maximize the social welfare, it returns the market clearing prices, the matched volumes, and the net position of each bidding area as well as the flow through the interconnectors.

\[
Max \sum_{n} \left\{ \int_{0}^{d^a} D^a(x)dx - \int_{0}^{s^a} S^a(y)dy \right\} \tag{2.1}
\]

Where:

- \(a\) represents area
- \(d^a\) is the demand in area \(a\)
- \(D^a\) is the demand function in area \(a\)
- \(s^a, S^a\), are the analogue for the supply in area \(a\) and the supply function.
- \(N\) is the number of areas

The full description of the welfare maximization problem exceeds the length and the objective of this report and can be read it in [45] and in [46]
2.1.4. Market Splitting

Besides all the tools that the EM must deal with the network congestion, sometimes all the transmission capacity is completely used. When this take place, the market is divided in different geographical areas, running EUPHEMIA separately.

In Spain, the market is divided into Portugal and Spain and in Sweden, the situation is a little bit complex because the Nordpool integrate more countries than OMIE. The Nordpool use to be divided in 6 or 7 regions, normally two or three areas within Norway, Sweden, Finland and Western and Eastern Denmark. However, nowadays Norway can be divided in 5 areas and the incorporation of Estonia, Lithuania and Latvia has been reached the areas to 14 regions. When the market splitting happens not only prices, volumes and power exchange bids information has to be included, but also the information about the physical electricity generation planning.

Figure 2.15: Bidding Areas in Noordpool Spot.
2.2 Time Series Analysis

A time series is defined as sequence of observations, measured in a determined moment and chronological ordered. The time series analysis is the different techniques used to study the causal relationship between two variables that change over time and that there are influenced by themselves. Normally, this analysis is orientated to described the data and the futures values of this series (forecasting techniques).

2.2.1. Stationarity

Mainly, the TS can be classified in two main different groups, stationarity time series or non-stationarity time series. Although both groups can be analysed, due to the nature of the electricity prices and the electricity loads, and the results obtained in the Literature Review chapter, only the description of stationarity TS is going to be carried out.

A TS is stationarity when is stable along the time, that is, when the average and the variance are constant (or practically) during the time. Also, there are two different of stationarity processes, the weak or wide-sense stationarity and the strictly stationarity processes.

Formerly:

**Stationarity process**: a process $X_t$, defined on $-\infty < t < \infty$, is stationary if $X_{t+c}$ has the same distribution as $X_t$ for all $-\infty < c < \infty$; equivalently this means that all of its finite dimensional distributions depend only on time differences.

\[
E(X_t) = \mu < \infty; \forall t \quad \text{Ec. [2.2]}
\]
\[
\text{Cov}(X_t, X_{t-j}) = \gamma_j < \infty; \forall t, \forall j \quad \text{Ec. [2.3]}
\]

**Weak Stationary process**: a process $X_t$ is weakly stationarity, or covariance-stationary if:

\[
E(X_t)^2 = \mu < \infty; \forall t \quad \text{Ec. [2.4]}
\]
\[
E(X_t) = \mu; \forall t \quad \text{Ec. [2.5]}
\]
\[
\text{Cov}(X_{t_1}, X_{t_2}) = \left(X_{t_1+j}, X_{t_2+j}\right) \forall t_1, t_2, j \quad \text{Ec. [2.6]}
\]

(depend only of the difference between the two times)

2.2.2. Classical Method (Decomposition)

A TS is fundamentally formed by four components that are not directly observable, generally it is only possible to get estimations. These mainly components are:

**Trend**: which represents the principal behaviour of the TS. In the easy way, the trend can be understood as the average evolution during an extended period. The trend can be deterministic (a static function e.g. a line or polynomial function) or can be stochastic.
Deterministic trend:

\[ X_t = a + dt + u_t \]

Ec. [2.7]

Where \( a; d \) are parameters

\( t \) is the temporal index

\( u_t \) is a stationarity process with \( \mathbb{E}(u_t) = 0 \)

\( \mathbb{V}(u_t) = \sigma^2 < \infty \)

Being \( \{ \mathbb{E}(X_t) = a + dt \text{ or a polynomial function} \}

\( \mathbb{V}(X_t) = \mathbb{V}(u_t) = \sigma^2 \)

The non-stationarity source is the average and the fluctuation is stationarity around a deterministic trend.

Stochastic trend:

- Random walk

\[ X_t = X_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{WN}(0, \sigma^2) \]

Ec. [2.7]

Where \( \mathcal{WN} \) is a white noise process

\( \mathcal{WN} \) is the unweighted sum of the WN elements

Being \( \{ \mathbb{E}(X_t) = 0 \}

\( \mathbb{V}(X_t) = \sigma^2 \)

The non-stationarity source is the variance.

- Random with drift

\[ X_t = m + X_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{WN}(0, \sigma^2) \]

Ec. [2.8]

Being \( \{ \mathbb{E}(X_t) = tm \}

\( \mathbb{V}(X_t) = t\sigma^2 \)

The non-stationarity source is the variance and the average.

Cycle: is the fluctuation around the trend in a huge lapse of time. The factors that cause the fluctuation are not clear. One example of cyclical component can be founded in the evolution of the temperature during the years.

Seasonal: is a periodic movement within a short period that can be delimited.

Random (residual): is the part formed by the data spread without any pattern, normally random data is practically unpredictable. Every data that are not part of the trend, cycle and seasonality components is classified has a random data.
The identification of this components originated a method called classical decomposition, usually divided into the additive decomposition and the multiplicative decomposition. Basically, the time series is parted in a seasonal component, a trend-cycle component and a residual component that includes every data that can not be in the other two parts.

**Additive decomposition:**

\[ X_t = S_t + T_t + E_t \]

Ec. [2.9]

**Graph 2.1:** Additive Decomposition applied to Nordpool Spot Prices, frequency =672 (monthly)

**Multiplicative decomposition:**

\[ X_t = S_t \cdot T_t \cdot E_t \]

Ec. [2.10]

Being \( X_t \) the data at period \( t \), where \( S_t \) is the seasonal component at period \( t \), \( T_t \) is the trend-cycle component at period \( t \) and \( E_t \) is residual or random component.

**Graph 2.2:** Multiplicative Decomposition applied to Nordpool Spot Prices, frequency =168 (weekly)
Another way to use the multiplicative decomposition, is to first transform the observed data using box-cox transformations and later, when the TS seems to be stable over the time use the additive model. In fact, when a log transformation is used:

\[
X(\lambda) = \begin{cases} 
\frac{X^{\lambda-1}}{\lambda}, & \text{if } \lambda \neq 0; \\
\log X, & \text{if } \lambda = 0;
\end{cases}
\]

Ec. [2.11]

For negative y-values

\[
X(\lambda) = \begin{cases} 
\frac{(X+\lambda_2)^{\lambda_1-1}}{\lambda_1}, & \text{if } \lambda_1 \neq 0; \\
\log(X + \lambda_2), & \text{if } \lambda_1 = 0;
\end{cases}
\]

Ec. [2.12]

The multiplicative decomposition can be:

\[X_t=S_t \cdot T_t \cdot E_t \text{ equivalent to } X_t=\log S_t+\log T_t+\log E_t.\]

Ec. [2.13]

2.2.3. Statistical Models: Box-Jenkins Method

Box-Jenkins method was created by George Box and Gwilym Jenkins, the full proposition was published in 1970, in their textbook: Time Series Analysis: Forecasting and Control.

The procedure is based in the assumption that a time series can be approximated using an ARMA model if is stationarity or and ARIMA model if is not stationarity. The method is based in the three following steps:

- Model Identification and Model selection: the first step includes the stationarity test of the studied variables, the seasonality detection, the necessary box-cox transformations and finally the model order selection depending of the PACF and ACF plots.
- Estimation: in this step, the parameters that forms the model need to be the best ones, obtaining with the model to closest values to our original observed data. In this report, AIC (Akaike's Information Criterion) and MLE (Maximum likelihood estimation) has been used to determinate the best model.
- Model Checking: Evaluate the model proposed with different techniques like calculating the RMSE or MAPE. A more serious checking can be done using cross-validation methods or quantitative analysis of residuals.
2.2.3.1. Model Identification:

The first step to detect the stationarity of the TS analysed, and if differencing or box-cox transformations are required, is using a unit root test. One of the most popular is the Augmented Dicker-Fuller (ADF) test:

\[ X'_t = \phi X_t - 1 + \beta_1 X'_{t-1} + \beta_2 X'_{t-2} + \ldots + \beta_k X'_{t-k}, \]  

Ec. [2.14]

where \( X'_t \) denotes the first-differenced series, \( X'_t = X_t - X_{t-1} \) and \( k \) is the number of lags to include in the regression. If the original series needs differencing, then the coefficient \( \hat{\phi} \). If \( X'_t \) is already stationary, then \( \hat{\phi} < 0 \)

There is also another test like the Kwiatkowski-Phillips-Schmidt-Shin (KPSS), but for simplicity coding reasons the author of this report choose the ADF test.

If the result stationarity hypothesis is rejected, different techniques should be applied to transform the data in a stationarity one. It can be founded more information about the critical values of the ADF test in the table.

**Table 2.1: Values for DF t-distribution.**

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Without trend</th>
<th>With trend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1%</td>
<td>5%</td>
</tr>
<tr>
<td>( T = 250 )</td>
<td>-3.46</td>
<td>-2.88</td>
</tr>
<tr>
<td>( T = 500 )</td>
<td>-3.44</td>
<td>-2.87</td>
</tr>
<tr>
<td>( T = \infty )</td>
<td>-3.43</td>
<td>-2.86</td>
</tr>
</tbody>
</table>
Having done the ADF-test there are two principal techniques in order to make the TS stationarity, the first one has been mentioned several times in this report and is the box-cox transformation. The second one is taking differences of the TS.

Normally, the log box-cox transformation helps the TS to be stable in variance:

**Graph 2.3: NordPool Spot Average Monthly Price.**

![Graph 2.3](image)

**Graph 2.4: NordPool Spot Average Monthly Log-Price.**

![Graph 2.4](image)

However sometimes the TS has a strong trend component, so in that case taking differences can help stabilize the mean of a time series by removing changes in the level of a time series, and so eliminating trend and seasonality.
Which looks clearly more stationary than the original series, in fact showing the results of a ADF test:

### Table 2.2. ADF-Test for First Difference TS and the original TS

<table>
<thead>
<tr>
<th>Results of Dickey-Fuller Test:</th>
<th>ts_log_diff</th>
<th>original ts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Statistic</td>
<td>-4.740148</td>
<td>-2.400734</td>
</tr>
<tr>
<td>p-value</td>
<td>0.000071</td>
<td>0.141529</td>
</tr>
<tr>
<td>#Lags Used</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Number of Observation Used</td>
<td>55</td>
<td>59</td>
</tr>
<tr>
<td>Critical Value</td>
<td>-3.555273</td>
<td>-3.546395</td>
</tr>
<tr>
<td>Critical Value</td>
<td>-2.915731</td>
<td>-2.911939</td>
</tr>
<tr>
<td>Critical Value</td>
<td>-2.59567</td>
<td>-2.593652</td>
</tr>
</tbody>
</table>

Differencing is a good technique to make the TS stationary, however it should be pointed out that it is really important to avoid the over-differencing, which means differencing more than is required, resulting in the addition of complexity and extra correlation.

Differenced TS can be written as:

\[
X'_t = X_t - X_{t-1} \quad \text{Ec. [2.15]}
\]

Sometimes the first differenced TS is not stationary at all and it might be required another difference to obtain a stationary series.

\[
X''_t = X'_t - X'_{t-1} = (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) = X_t - 2X_{t-1} + X_{t-2} \quad \text{Ec. [2.16]}
\]
There is also possible to take seasonal differences like that:

$$X'_t = X_t - X_{t-m}$$  \hspace{1cm} \text{Ec. [2.17]}

*where m is the number of seasons*

When stationarity and seasonality have been addressed, the next step is to clarify the order of the model. This order is done by the p, q and d parameters which are:

- p (Number of Auto Regressive terms): lags of dependent variable.
- q (Number of Moving Average terms): terms are lagged forecast errors in prediction equation.
- d (Number of differences): the data values have been replaced with the difference between their values and the previous. Basically, they are the number of no seasonal differences.

For doing that, the ACF and the PACF functions and plots could be used after the stationarized by difference.

The ACF is defined as:

$$\rho_j = corr(X_t,X_{t-j}) = \frac{Cov(X_t,X_{t-j})}{\sqrt{V(X_t)} \sqrt{V(X_{t-j})}}$$  \hspace{1cm} \text{Ec. [2.18]}

The PACF is defined as:

$$\pi_j = corr(X_t,X_{t-j}/X_{t-1}X_{t-j+1}) = \frac{Cov(X_t - \bar{X}_t, X_{t-j} - \hat{X}_{t-j})}{\sqrt{V(X_t - \bar{X}_t)} \sqrt{V(X_{t-j} - \hat{X}_{t-j})}}$$  \hspace{1cm} \text{Ec. [2.19]}

The ACF plot shows the autocorrelation which measures the relationship between $X_t$ and $X_{t-j}$ for different values of j. If $X_t$ and $X_{t-1}$ are correlated, then $X_{t-1}$ and $X_{t-2}$ must also be correlated. Both values $X_t$ and $X_{t-2}$ are connected, so the information of $X_{t-2}$ can be used for forecast $X_t$.

To avoid this issue, partial autocorrelations can be used. These measures the relationship between $X_t$ and $X_{t-1}$ overcoming the effects of other time lags. For different values of the autocorrelation, they can be calculated as:

$$\alpha_k = k\text{th partial autocorrelation coefficient}$$  \hspace{1cm} \text{Ec. [2.20]}

= the estimate of $\phi_k$ in the autoregression model
Where the confidence intervals are for an ARMA model:

\[ X_t \pm 1.96 \sqrt{\hat{h}_t} \sigma \]  

Ec. [2.21]

There are some guidelines to identify the models, unluckily any real process follows a perfect AR or MA model so simulation and model estimation is required to create the more best model.

2.2.3.2. Model Estimation:

The main approaches to fitting Box–Jenkins models are nonlinear least squares and maximum likelihood estimation. However, Akaike’s Information Criterion (AIC) can be used for selecting predictors for the regression.

AIC is defined as:

\[
AIC = -2 \log(L) + 2(p + q + k + 1)
\]

Ec. [2.22]

where \( L \) is the Likelihood Function

\[
k = 1 \text{ if } c \neq 0 \text{ and } k = 0 \text{ if } c = 0
\]

For ARIMA models:

\[
AICc = AIC + \frac{2(p + q + k + 1)(p + q + k + 2)}{T - p - q - k - 2}
\]

Ec. [2.23]

A more negative value of the AIC is an indicator of a better model.

As regards to the MLE, in ARIMA models is very similar to the least squares estimates, which is obtained minimizing the following function:

\[
\sum_{t=1}^{T} e_t^2 \text{ being } e \text{ the "random" error}
\]

Ec. [2.23]
Formerly the $L$ function is defined as:

$$\mathcal{L}(\theta|x_1, x_2, ..., x_n) = \prod_{i=0}^{n} f(x_i|\theta)$$  \text{Ec. [2.24]}

And the MLE:

$$\hat{\theta}_{MLE} = \arg\max_{\theta \in \Theta} \mathcal{L}(\theta|x_1, x_2, ..., x_n)$$  \text{Ec. [2.25]}

being $\mathcal{L} = \log \mathcal{L}_N(\theta)$  \text{Ec. [2.26]}

The principle of MLE is taking as an estimation of the studied parameter, the value that shows the maximum probability of the observed data.

The likelihood is based in giving a certain “credibility” to an estimator or an estimation. In probabilistic terms, it might be possible to speak of the likelihood as the probability that a given sample occurs, or a specifically sample happens if the estimation is true (which means that the estimator is correct).

Note however, that the likelihood function is NOT a probability function: for instance, in general, it does not integrate to 1 (with respect to $\theta$).

2.2.3.3. Model Checking:

In order to estimate how accurate is the forecasted values of the model created, it has been used the MAPE and the RMSE values. They can be defined as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{X}_i)^2}$$  \text{Ec. [2.27]}

$$\text{MAPE} = \frac{100}{n}\sum_{t=1}^{n} \left| \frac{X_i - \hat{X}_i}{X_i} \right|$$  \text{Ec. [2.28]}

Where:

- $n$ is the total number of data.
- $X_i$ is the value for the TS in the instant $i$.
- $\hat{X}_i$ is the value forecasted in the instant $i$. 
2.2.4. AR Model

An autoregressive model is defined autoregressive if the endogenous variable of a period $t$ is explained by the observations of itself corresponding to previous periods, adding, as in the structural models, an error term.

The autoregressive models are abbreviated as AR models, after which the order of the model is indicated as: AR (1), AR (2) etc. The model orders describe the number of observations retarded of the TS analysed that take part in the following equation:

$$X_t = \phi_0 + \phi_1X_{t-1} + \phi_2X_{t-2} + \cdots + \phi_pX_{t-p} + \varepsilon_t$$  \hspace{1cm} \text{Ec. [2.30]}

where

- $\phi_p$ the coefficients of the model
- $\varepsilon_t$ is the error term

Normally the error term is called white noise when this term accomplishes that:

$$E(X_t) = 0; \forall t$$  \hspace{1cm} \text{Ec. [2.31]}

$$Var(X_t) = \sigma^2 = cte$$  \hspace{1cm} \text{Ec. [2.32]}

$$Cov(X_t, X_{t-j}) = 0; \forall t, \forall j$$  \hspace{1cm} \text{Ec. [2.33]}

Economically the equation can be written as:

$$\phi_p(L)X_t = \phi_0 + \varepsilon_t$$  \hspace{1cm} \text{Ec. [2.34]}

Where $\phi_p(L)$ is the Lag polynomial operator, that is:

$$\phi_p(L) = 1 - \phi_1L - \phi_2L^2 - \phi_pL^p$$  \hspace{1cm} \text{Ec. [2.35]}

Where in turn $L$ is the lag operator, that operates on an element of a time series to produce the previous element. It is usual to see in other literature that the lag operator is called sometimes the backward shift operator $B$.

2.2.5. MA Model

A moving average model is one that explains the value of a variable in a period $t$ as a function of an independent term and a succession of errors corresponding to prior periods, conveniently weighted. These models are usually denoted by MA, followed, as in the case of AR models, with the order in the parenthesis. The general equation is:

$$X_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_q\varepsilon_{t-q}$$  \hspace{1cm} \text{Ec. [2.35]}

where

- $\theta_p$ the coefficients of the model
- $\varepsilon_t$ is the error term
- $\mu$ is the independent term
In the same way than AR models, the equation can be written shorted as:

\[ X_t = \theta_q(L)\varepsilon_t + \mu \]  
Ec. [2.36]

2.2.6. ARMA Models

If a TS has both characteristics of AR and MA processes at the same time, it can be described using and ARMA model. ARMA models have regression terms (p) and moving average terms (q).

\[ X_t = c + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \]  
Ec. [2.37]

2.2.7. ARIMA Models

ARIMA models are a generalization of the ARMA models, usually ARIMA models are used when the data seems to be not stationary or when non-seasonal differencing has been applied several times, transforming the original TS in a stationary one.

Being an ARMA model as Ec. [2.37] and being \( \Delta \) the difference operator (integration) defined by:

\[ \Delta^d X_t = X_t - X_{t-1} = (1 - L)^d X_t \]  
Ec. [2.38]

The shorter equation of ARIMA models can be written as:

\[ \left( 1 - \sum_{i=1}^{p} \phi_i L^i \right) \Delta^d X_t = \left( 1 + \sum_{i=1}^{q} \theta_i L^i \right) \varepsilon_t \]  
Ec. [2.39]
3 Method

As the theory statement, the method chapter is going to be subdivided in two parts: one of them is about the comparison between the two markets, and the other one about the forecasting methods.

3.1 Comparison between the two markets

The comparison between the Swedish EM and the Spanish EM has been developed has a case study. EXCEL has been used to represent the results in several graphs and charts. The comparison can be considered as a kind of meta-analysis because as has been mentioned in the limitations chapter the data has been recollected from multiples sources.

First of all, a general overview has been carried, the average prices of the kWh in each sector (industry and households) and its respectively components have been described. Also, the evolution of them during the year 2014 and 2016 has been graphed. The exchanged rate selected between € and SEK was 1EUR=9.634SEK.

After that, a comparison between the Renewable Mix and the non-renewable mix has been developed. Also, the evolution of the electricity mix and of the energy use has graphed. Then the different energy policies about renewable energy has been described.

Finally, the forecasting models are presented with its respectively calculation of the MAPE and the RMSE.

3.2 Forecasting method (TSAFTTools.py and FNRVTTool.m)

Two main codes have been built in python and MATLAB respectively. The python one, has been coded using the following libraries:

- Pandas (for importing the data).
- Numpy (for managing and operating the arrays).
- Statsmodels (for using the AR, MA and ARIMA models).
- Matplotlib (for representing the corresponding graphs).

The code [8.1] called TSAFTTools.py is an script that perform all the techniques described in the chapter 2. First, TSAFTTools.py import the data from a CSV file and create an array with two columns: the date and the SEK (or €) – MWh.

Secondly, it implements a stationarity test for each variable that contains the TS analysed, being the original one or another that has been transformed. Also, different techniques are applied to the TS (moving average techniques, exponential weighted average techniques, box-cox transformations and differencing) in order to make them stationary.

Thirdly the code decomposes the TS which seems more stationary to observe if the trend or the seasonality still having considerable effects. Also, the ACF and PACF plots are showed helping the code user to estimate the p and q parameters. The AIC and the \( \hat{\ell} \) value are also displayed in the command prompt.
Fourthly an AR, MA and ARIMA model are fitted to the TS that is more stationary, which means, the TS with the smallest p-value in the ADF test. The RSS is showed in the top of each AR and MA graphs and when the ARIMA model is fitted, the RMSE is calculated. All these numerical steps are accompanied by its corresponding graphs. However, the selection of the model order is the duty of the code user, interpreting the statistical results by himself.

Then, FNRVTool.m (Fixing NaN and Repetitive Values Tool) has been used in order to correct the output data. Due to a NaN (not a number) problem the output array of the forecasting values is 24 times bigger than the input array, the FNTool.m takes averages each 24 values (because they are repeated), having finally the output data corrected.
4 Results and Discussion

4.1 Electricity use per capita

The Graph 4.1 shows the expected higher electricity use per capita of Sweden over Spain. That fact can be explained mainly because two things, the first one is that the Swedish GDP per capita is higher than the Spanish one. Normally this is a clearly indicator of the country development and both indicators (GDP and MWh per capita) are directly elated with several progress indicators like: HDI (Human Development Index), life expectancy, public healthcare etc.

The second one, which is the most obvious, is the lower temperatures that take place in Sweden. Normally the temperature is one of the strongest modifiers to the electricity load. Besides the higher temperatures in Spain During the summer, the Swedish weather is considered more extreme than the Spanish one.

4.2 Average Price of the kWh and Price Components by sectors

The Graph 4.2 shows the firs interesting aspect of that comparison, despite the higher Swedish electricity use per capita, the average prices of the Swedish kWh are lower than the Spanish ones in both sectors (industrial and households). In fact, the differences in the Industry sector could be almost the 50% as it is showed in the Graph 4.3.

Graphs 4.4. and 4.5 display another odd observation, besides the Spanish public opinion believes and the Spanish news say [47] [48] the higher cost is the supply one, also the global percentage without taxes is higher in Spain than in Sweden. It should be pointed out another characteristic thing, in the Industrial Sector practically there is not too much external taxes apart from the Network Cost. In addition, in the Swedish case the Network cost is the same for both consumers in %. These facts appear in the Graph 4.6 and Graph 4.7.

4.3 Electricity Generation and Power Capacity Evolution

With the Graph 4.7 and 4.8 it begins to see part of the main problems of the SsES. As it can be observed during 1990 until 2010 the Spanish electricity generation was almost multiplied per two*! However, the Swedish was quasi constant, in fact the average % variation was 2.5%. Also, another problem was that this generation increase was covered using principally Natural gas and Oil, and only after the financial crisis (2008) the electricity use goes down.

As regards to the Graph 4.10, it can be seen that the Swedish capacity evolution follows more or less the same trend than the generation (being quasi constant). However, since 2008 the installation of wind power and district heating the Swedish capacity raises until 10-15%.

Graph 4.11 shows one of the huge disasters of the Spanish Electricity system: the overcapacity. In 13 years the installed capacity was doubled which was totally unnecessary. In fact, according to REE the maximum power demand was 45.450MW in
2007 which is clearly lower than the total amount of capacity [49]. Moreover, the Generation/Capacity ratio has a negative trend, which means that this capacity is being underused.

As it can be seen above in Graph 4.13 and Graph 4.14 almost two thirds part of the Swedish electricity capacity installed is from a renewable source, and the 87% is CO2 free, having only the 13% of the installed capacity CO2 emissions.

Watching the Graph 4.14 the problem appears again, not only the SsES has overcapacity but also the 46% of the total capacity installed is from fossil fuels. This is because the CHP “boom” caused by two main factors: the first one was including these power plants in the special system (see chapter 4.5). The second one was explained before, and was because the supply increase satisfied by Gas showed in the Graph 4.8.

Graphs 4.16, 4.17, 4.18, 4.19 display the Electricity Generation mix in both countries and the supply % from Renewable and Non-Renewable resources. It should be pointed out that the Spanish electricity generation in 2014 was more mixed than the Swedish one.

4.4 Main Renewable Generation Technologies:

4.4.1. Wind Power

The plotted Graphs 4.20 and Graph 4.21 describe the evolution of the wind power generation in Sweden and Spain respectively. It should be highlighted that despite the generous subsidies given to Spanish wind power (read chapter 4.5), the evolution of the Swedish generation seems to be exponential instead of the Spanish one. In the same way, the Swedish P/C ratio (Graph 4.22) appears exponential at the beginning and even the fluctuations still growing up.

About the Spanish P/C this seems to be a line, however it is important to consider that it has been hard to find reliable data before 1997. Also, looking at the Spanish generation (Graph .24) it is easy to see that the trend is practically linear and a standstill has been appeared since 2012 (when the bonus for the renewable power plants was suspended, see chapter 4.5).

Other interesting indicator displayed in Graph 4.25 is the C/NT, showing that the new wind turbines are more powerful than the first ones having the last 15 years a positive trend.

* It has been impossible to calculate some parameters about the generation evolution because two facts: the first one was that only the data from 2006 to 2017 had a .xls, .csv or .txt format. The second one was because all the data from the past years was scanned (in .pdf) and there was enough time for computerize all of it.
4.4.2. Biomass for Electricity

According to REE [50] the total biomass for electricity capacity installed in Spain was practically inexistent (less than 1%), and the contribution to the electricity supply was less than the (0.5%). For that reason, the author of this report has considered that is not significant to represent the evolution of this resource. Also, over the total installed capacity of fuels (excluding nuclear), biomass is less than 2%.

Attending to the Swedish case the evolution of biomass use for electricity generation was huge and started approx. in 1990 (Graph 4.26). In fact, since 1990 the electricity supply from this resource has been increased in 700%! (Graph 4.28) Moreover the % of fossil fuels (excluding nuclear) in 2014 electricity generation was less than 3%” instead of the Spanish one that was higher than 26% (CHP + Oil and rest of fuels, Graph 4.18). The Swedish fossil fuels in % over the total electricity generation can be observed in Graph 4.27.

4.4.3. Hydropower

Regarding hydropower sector, it should be noted that is the most stable in Swedish case. In fact, the evolution of the generation did not change (the absolute average change in the last 25 years was the 8% over the total average hydropower generation and less than 4% over the total electricity average generation). As regards to the installed capacity the variation was less than the 1.5%. These evolutions are displayed in Graph 4.29 and 4.31

Looking at the Spanish case it is noticed a different trend, the total hydropower generation had an absolute average change of 20%” this can be explained mainly because of the introduction of special system hydropower generation. In addition, the increase of the special hydropower installed capacity was an 12%” over the total since 1990. Graph 4.30 and 4.32 shows this development.

4.4.4. Solar Power

As well as the biomass in the Spanish case, the Swedish solar power has not been represented due to understandable reasons. The first one is that the installed capacity is less than 0,5% and also the contribution to the electricity generation has not been found.

Attending to the Spanish case it is observe another of the problems that raise the deficit during the last 10 years. The solar electricity generation in 2006 was practically zero but was increasing until 2012 (Graph 4.33 and graph 4.34) when the subsidies stopped. This is a big issue because Spain put too much economical effort installing solar power when its price was high.
4.5 Subsidies Summary:

4.5.1. Swedish Case: Green Certificates

The electricity certificate is a market-based support system for renewable electricity supply. The system was introduced in May 2003 and its purpose is increasing the TNREG and achieving the most cost-efficient renewable energy system. It is expected that the certificates system will be extended until 2030.

Each MWh produced by renewable energies can be granted by SK using an electricity certificate. These certificates can be sold in the EM to the consumers, mainly using as an intermediate the electricity suppliers who must fulfil the compulsory quota of electricity certifications. Basically, is like a “regulation” of part of the energy demand. This quota is calculated proportionally according to the total energy demand. Only the Energy Intensive Industry is exempt for buying this kind of energy.

Currently the renewable energies included in this measure are:

- Wind Power.
- Solar Power.
- Wave Power.
- Geothermal.
- Hydropower.
- CHP (only to peats).

It is important to remark that the hydropower green certificates are a little bit special and not every hydropower plant is included. Concretely only the following ones:

I. Hydropower stations that due to legislation or major renovations no longer are economically feasible to operate.
II. Small scale hydropower (1.5MW per unit before 2013)
III. Reutilization of hydropower abandoned stations.

The graph 4.35 represents the electricity generation under the green certifications and its evolution by technology.

4.5.2. Spanish Case: special and ordinary system, fixed tariff and premium system

Nowadays the Spanish electricity generation technologies are divided in two by the law. The first group is the ordinary system which includes nuclear power plants, fuel power plants, coal power plants, gas power plants and large-scale hydropower. The second one is the special system in which includes all the renewable energy power plants and CHP plants under 50MW.

The ordinary power plants are traded in the EM without any additional bonus, instead of that the special system has two options to sell its electricity [49]:

[49]
• Premium system: regulated rate, the 100% of the existing installations in February 2012, solar PV and the 28% of wind power are under this system.
• Market + Subsidy: 72% of the wind power installations choose this option.

The summary of this bonuses is in table 4.1, however it can be observed that the average regulated price of the Solar electricity kWh is, at least, five times more expensive than the wind power one. In 2014 the RD 2014 modify the way of calculating the regulated rate, the premium system will be called now Specific Retributive System and has two main terms:

I. Investment retribution
II. Operation retribution

4.6 ARIMA Results

4.6.1. Nordpool Spot Prices

Due to the extension of this work the author of this report will only analyse the model accuracy and some parameters than find remarkable.

Following the Box-Jenkins method, the model considered for Nordpool spot prices in 2015-2016 was an ARIMA (1,1,0). The summary of the model and all its statistical data can be found in the appendix 7.4. As it can be seen in the Graph 4.36 the model proposed follows the original data, having more or less the same seasonality and trend. However, the MAPE is really high, concretely is around 20% which is inaccurate if the literature about this topic is taken in account. (Table 4.2)

Nevertheless, this is probably because two fundamental issues: despite of the powerful tool that the ARIMA models are, when seasonality is too strong and the spikes are numerous the models do not fit properly. In fact, the spikes can be enough strong for destroying the stationarity during a period. For instance, when the highest spikes prices were eliminated (only three values) the MAPE decrease in 1% which is a huge decrease taking in account that 3 values are only the (0,017%) of the sample data.

The second one is because no external factors have been considered, that is, non-exogenous variables as temperature, labour factor, holidays, hydropower reservoir and interconnexion capacity.

Moving forward to the next year, the model proposed an ARIMA (1,1,0) has been resulted in a more accurate one with a MAPE of 7.74% (which is pretty good) and with an RMSE of 37.033 which is really low for the amount of data. (Table 4.3)

It can be hard to believe that using the same kind of model the accuracy of the second one be fairly better than the model from 2015-2016, but this is completely understandable. The fact of having less sample data, also does that it considers less spikes values which as has been mentioned above, have a remarkable effect in the model accuracy.
4.6.2. OMIE Spot Prices

Following the Box-Jenkins method, the model considered for OMIE spot prices in 2017 have been an ARIMA (1,1,1) which obtained the lowest AIC criterion. The summary of the model and all its statistical data can be found in the appendix 7.4. As it can be seen in the Graph 4.36 the model proposed is not really good as well as the model from the Nordic spot prices in 2015-2016. Despite the model follows more or less the trend the spikes and the volatility leads the model to fail with a MAPE of 20.24%. (Table 4.4) and Graph.
5 Conclusions

5.1 Study Results

The study carried out sheds some light about the reality of both EMs, the main problems of the Spanish electricity systems are showed and are the following ones:

I. The overcapacity caused by the too much generous subsidies to the special system. Including the CHP power plants in the special system was a mistake, these power plants have a high operation and maintenance.

II. Solar power capacity and Solar electricity generation are at a standstill. As the table 4.1 shows the subsidies to this energy were completely over dimensioned, and now without any bonus and the regulatory frame in Spain they are not attractive for new investments. However, watching the P/C ration it is possible to intuit than despite of the intermittent generation of solar power, the amount of total power is used.

III. The financial crisis in 2008 leads (finally) Spain to decrease his electricity generation, however if efficiency measures would be more important to the Government probably the electricity use would follow the same trend than the Swedish one.

IV. Spain Electricity system has a huge dependence from fossil fuels and this is mainly because the CHP power installed and because the Government did not promote the biomass enough, the Swedish case is an example of success in the biomass use for electricity generation. Also, implementation of biomass would lead Spain to decrease its CO\textsubscript{2} emissions.

V. Nordpool spot is more stable in prices than OMIE, having a common market and bigger interconnexion capacity allows the Swedish Electricity market to share economical efforts in frequency and power reserve regulation.

VI. Despite the successful use of biomass, Sweden has a huge dependence from Nuclear power, it is true that this energy is CO\textsubscript{2} free in terms of generation but it has additional environmental risk perfectly well known.

VII. Currently, Swedish solar power is under development but it is really difficult to estimate the real potential of this energy in this country due to the lack of sun intensity.

VIII. Green certificates have given a freer option to the electricity producers and they have been revealed as a more cost-efficient option than the Spanish subsidies. In fact, the highest capacity evolution has been in wind power, which had a lower €/MWh than the solar power in the past years. Moreover, green certificates move the risk to the companies and not to the Government, achieving a less deficit.

IX. Higher forecast accuracy would have been a determinant factor in the Spanish System to not install its wasteful backup power. In addition, higher forecast accuracy would have avoided the installation of some CHP power plants that would cause less CO\textsubscript{2} emissions during the last years.
X. Spikes problem produce strong effects in the forecast accuracy. However, this issue can be solved using three techniques:
   a. Splitting the data in two groups: the first would be those data that not rise the average $\pm 3$ times the value of the standard deviation $\sigma$, the second one would be the spikes data. Then a model will be developed for each data and finally the results would be aggregated.
   b. Using SARIMAX models with exogenous data as the temperature, labour factor etc. and with corresponding seasonal value.
   c. Using a modern approach as Artificial Neural Networks, which has not been considered due to its complexity.

5.2 Outlook & Perspective

Electricity price forecasting and Electricity Load Forecasting will be determinant in the future of renewable energies. Also forecasting techniques can be applied to predict the CO2 impact. However, not only advances in this method are required, but also the TRNEP must be sustained by the correct government policies.

The future electricity mix has to allow to the consumers to use these tools to control their energy use their future bill, Demand Response could be a good policy which depends directly of the forecast accuracy.

The Net Balance proposed by different solar associations can be another measure in order to decrease the system cost and also to increase the participation of renewable energies.

In spite of the advantages of the green certificates a future research is necessary in order to know if is the quota would be set too “high” the effect would be similar than a too generous fixed tariff. Moreover, if a EU subsidy will be implemented in the future the form should be as a green certificate or as a premium system? If the EU electricity common market is built can be the auction mechanism (as the German one) or the Renewables Obligation from UK (similar to the fixed tariff) be a better solution?

Future research is needed to consider all these variables, and to identify the most effective methods for increasing the TNREG without achieving a huge deficit. Similarly, more research is needed to determinate the most effective type of interventions that the consumer will use in the future.
6 References


[7] Price Coupling of Regions – PCR Presentation to Presentation to 5th SG meeting South-West Electricity REM, Lisbon.


7 Appendix

7.1 Comparison Graphs

Graph 4.1: Swedish vs Spanish use per capita.

Graph 4.2: Spanish vs Swedish Sectors Average Electricity Prices.
Graph 4.3: Difference between Spanish and Swedish Electricity Average Prices by sector.

Graph 4.4: 2014 Households Swedish Price Decomposition (%).

Graph 4.5: 2014 Households Spanish Price Decomposition (%).

Graph 4.6: Swedish Industrial Price Decomposition (%).

Graph 4.6: Industrial Spanish Price Decomposition (%).
Graph 4.7: Swedish Electricity Generation Evolution (Last 25 years).

Graph 4.8: Spanish Electricity Generation Evolution (Last 40 years, External Source: IEA).

Graph 4.9: Swedish Electricity Capacity Evolution (Last 20 years).
Graph 4.10: Spanish Generation, Capacity and P/C Ratio Evolution (Last 13 years, External resource: REE)

Graph 4.11: 2014 Swedish Power Capacity % by Technologies.

Graph 4.12: 2014 Swedish Electricity Power Share (%).

Graph 4.14: 2014 Spanish Electricity Power Share (%).

Graph 4.15: 2014 Swedish Electricity Generation (%) by Technologies.
Graph 4.19: Swedish Wind Power Generation Evolution.

Graph 4.20: Spanish Wind Power Generation Evolution.

Graph 4.21: Swedish Wind Generation/Wind Capacity Ratio.
Graph 4.22: Swedish Nº Turbines and installed Wind Power Capacity Evolution

Graph 4.23: Spanish Wind Power Capacity Evolution.

Graph 4.24: Swedish Wind Power Capacity-Number of Turbines Ratio.
Graph 4.25: Swedish Fuel Use for Electricity Generation.

Graph 4.26: % of Electricity Generation of Non-Biomass Fuels over Total Generation

Graph 4.27: Swedish Biomass Use for Electricity Generation Evolution.
Graph 4.28: Swedish Hydropower Generation Evolution.

Graph 4.29: Spanish Hydropower Generation Evolution.

Graph 4.30: Swedish Hydropower Capacity Evolution.
Graph 4.31: Spanish Hydropower Capacity Evolution.

Graph 4.32: Spanish Solar Capacity Evolution.
Graph 4.33: Spanish Solar Generation Evolution.

Graph 4.34: Electricity Generation evolution by type of power in the Swedish electricity green certificate

Table 4.1 Summary of article 2 of RD 661/2017

<table>
<thead>
<tr>
<th>Group</th>
<th>Technology</th>
<th>Subgroup</th>
<th>Power</th>
<th>Time</th>
<th>Regulated Price c€/kWh</th>
<th>Reference bonus c€/kWh</th>
<th>Upper Limit c€/kWh</th>
<th>Lower Limit c€/kWh</th>
</tr>
</thead>
<tbody>
<tr>
<td>b.1</td>
<td>Solar PV</td>
<td>b.1.1</td>
<td>P &lt; 100 kW</td>
<td>First 30 years</td>
<td>48,8743</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 kW &lt; P &lt; 10 MW</td>
<td>First 30 years</td>
<td>46,3348</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10 &lt; 50 MW</td>
<td>First 30 years</td>
<td>25,4997</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thermal Solar</td>
<td>b.1.2</td>
<td>First 25 years</td>
<td></td>
<td>29,8957</td>
<td>28,1894</td>
<td>38,1751</td>
<td>28,1936</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Since 2012</td>
<td></td>
<td>23,9164</td>
<td>22,5515</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.2</td>
<td>Wind</td>
<td>b.2.1</td>
<td>First 20 years</td>
<td></td>
<td>8,127</td>
<td>2,0142</td>
<td>9,4273</td>
<td>7,9103</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Since 2012</td>
<td></td>
<td>6,7921</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>b.2.2</td>
<td></td>
<td></td>
<td>9,3557</td>
<td>18,2009</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

y = 1682.8x - 3E+06
R² = 0.939
7.2 Python Code: TSAFTools.py (Time Series Analysis and Forecasting Tools)

```python
# -*- coding: utf-8 -*-

Created on Mon May 8 18:53:09 2017
Proof of Concept number 8-2
@author: Edgar Bahilo Rodriguez
@University: Högskolan i Gävle and UPV (Polytechnic University of Valencia)
Partially based in codes available in:
- Introduction to Time Series Forecasting With Python by Dr. Jason Brownlee
- Forecasting Spot Electricity Market Prices Using Time Series Models by Dawit Hallu Mazengia (Chalmers University of Technology)
[MATLAB CODE]
- Lecture 17: Time Series Analysis by Pavlos Protopapas (Harvard University)

Changelog:
  Loading data: done
  Data-check: done
  Stationarity test: done
  Transformation:
    - Log: done
    - Root square: pending
    - Cube root: pending
  Smoothing methods:
    - Exponential: done
    - Moving Average: done
  Differencing: done
  Decomposing: done
  ACF and PACF plots: done
  AR model: done
  MA model: done
  ARIMA model: done
  Fitting model: done
  RMSE: done

CAUTION!!!!: The code does not select the best model, you have to try the different orders (p,d,q) by your own and also you have to select the corresponding transformed Time Series to fit the model. Read First: Introduction to Time Series Analysis and Forecasting 1st Edition by Douglas C. Montgomery, Cheryl L. Jennings, Murat Kulahci, or whatever documentation related with time series analysis.

```
from statsmodels.tsa.arima_model import ARIMA

#Good Looking Graphs
plt.style.use('bmh')
rcParams['figure.figsize'] = 18,8

#Reading and checking the data (important to change the date parse if is required)
#USE YOUR OWN DATA!
dateparse = lambda dates: pd.datetime.strptime(dates, '%d-%m-%Y')
data = pd.read_csv('../Data/Hourly/elspot-prices_2015-2016_hourly_sek_peakcorrected.csv', index_col='Day', date_parser=dateparse)
print(data.head())
print('Data Types:')
print(data.dtypes)
print(data.index)
ts = data['SEK-Mwh']
plt.plot(ts)
plt.xlabel('Time')
plt.ylabel('SEK-MWh')
plt.show()

#FIRST STATIONARITY TEST: NO DATA TREATMENT

def test_stationarity(timeseries):
    #Determining rolling statistics
    rolmean = pd.rolling_mean(timeseries, window=24)
    roldiv = pd.rolling_std(timeseries, window=24)

    #Plot rolling statistics:
    orig = plt.plot(timeseries, color='red', label='Original')
    mean = plt.plot(rolmean, color='green', label='Rolling Mean')
    std = plt.plot(roldiv, color='blue', label='Rolling Std')
    plt.legend(loc='best')
    plt.title('Rolling Mean & Standard Deviation')
    plt.show(block=False)

    #Perform Dickey-Fuller test:
    print('Results of Dickey-Fuller Test:')
    dftest = adfuller(timeseries, autolag='AIC')
    dftestout = pd.Series(dftest[0:4], index=['Test Statistic', 'p-value', '#Lags Used', 'Number of Observations Used'])
    for key, value in dftest[4].items():
        dftestout['Critical Value (%s)' % key] = value
    print(dftestout)

    #If is necessary, fixing NAN values problem
    ts2 = ts[~np.isnan(ts)]
test_stationarity(ts2)

#ESTIMATING AND ELIMINATING TREND
# Log Transformation

```python
ts_log = np.log(ts2)
plt.plot(ts_log, color='blue')
plt.title('Log of sample data')
plt.xlabel('Time')
plt.ylabel('Log (SEK - MWh)')
plt.show(block=False)
```

# Smoothing Methods

## Moving AVERAGE

```python
moving_avg = pd.rolling_mean(ts_log, 24)
plt.plot(ts_log)
plt.plot(moving_avg, color='purple', label='SEK-MWh moving average')
plt.legend(loc='best')
plt.show(block=False)
ts_log_moving_avg_diff = ts_log - moving_avg
ts_log_moving_avg_diff2 = ts_log_moving_avg_diff[np.logical_not(np.isnan(ts_log_moving_avg_diff))]
```

```python
print(ts_log_moving_avg_diff2.head(24))
ts_log_moving_avg_diff2.dropna(inplace=True)
test_stationarity(ts_log_moving_avg_diff2)
```

## Exponentially weighted moving average

```python
expwighted_avg = pd.ewma(ts_log, halflife=24)
plt.plot(ts_log)
plt.plot(expwighted_avg, color='black', label='SEK-MWh expwighted')
plt.title('expwighted_avg vs ts_log')
plt.legend(loc='best')
plt.show(block=False)
ts_log_ewma_diff = ts_log - expwighted_avg
test_stationarity(ts_log_ewma_diff)
```

## Differencing until order 2

```python
ts_log_diff = ts_log - ts_log.shift()
plt.plot(ts_log_diff, label='SEK-MWh ts_log_diff')
plt.legend(loc='best')
plt.show(block=False)
ts_log_diff.dropna(inplace=True)
test_stationarity(ts_log_diff)
ts_log_diff2 = ts_log_diff - ts_log_diff.shift()
ts_log_diff2.dropna(inplace=True)
plt.plot(ts_log_diff2, label='SEK-MWh ts_log_diff2')
plt.legend(loc='best')
plt.show(block=False)
test_stationarity(ts_log_diff2)
```

# CAUTION with the freq: 24 daily, 168 weekly, 672 monthly, 8064 yearly.

```python
decomposition = seasonal_decompose(ts_log_diff2, model='additive', freq=672)
trend = decomposition.trend
seasonal = decomposition.seasonal
residual = decomposition.resid
plt.subplot(411)
```
```python
plt.plot(ts_log_diff2, label='Original')
plt.legend(loc='best')
plt.subplot(412)
plt.plot(trend, label='Trend')
plt.legend(loc='best')
plt.subplot(413)
plt.plot(seasonal, label='Seasonality')
plt.legend(loc='best')
plt.subplot(414)
plt.plot(residual, label='Residuals')
plt.legend(loc='best')
plt.tight_layout()
plt.show(block=False)

# Testing Descomposition

results_AR = ARIMA(ts_log_diff2, order=(1, 1, 0)).fit(disp=-1)
plt.plot(results_AR.fittedvalues, linestyle=':', color='red', label='SEK-MWh Forecasted')
plt.title('Forecasted AR residuals vs ts_log_diff2 price, RSS: %.4f' % sum((results_AR.fittedvalues-ts_log_diff2)**2))
plt.legend(loc='best')
plt.show(block=False)

# MA MODEL

model = ARIMA(ts_log, order=(0, 1, 1)).fit(disp=-1)
plt.plot(ts_log_diff2)
```

62
plt.plot(results_MA.fittedvalues, linestyle='-', color='red', label='SEK-MWh Forecasted')
plt.title('Forecasted MA residuals vs ts_log_diff2 price, RSS: %.4f' % sum((results_MA.fittedvalues-ts_log_diff2)**2))
plt.legend(loc='best')
plt.show(block=False)

#ARIMA MODEL
model = ARIMA(ts_log, order=(1, 1, 0))
results_ARIMA = model.fit(disp=-1)
print(results_ARIMA.summary())
plt.plot(ts_log_diff2)
plt.plot(results_ARIMA.fittedvalues, linestyle='-', color='red', label='SEK-MWh Forecasted')
plt.title('Forecasted ARIMA residuals vs ts_log_diff price, RSS: %.4f' % sum((results_ARIMA.fittedvalues-ts_log_diff2)**2))
plt.legend(loc='best')
plt.show(block=False)

predictions_ARIMA_diff = pd.Series(results_ARIMA.fittedvalues, copy=True)
predictions_ARIMA_diff.head()
predictions_ARIMA_diff_cumsum = predictions_ARIMA_diff.cumsum()
predictions_ARIMA_diff_cumsum.head()
predictions_ARIMA_log = pd.Series(ts_log.ix[0], index=ts_log.index)
#Acumulative problem fixed with excel (the output vector has 24 times the size of the input vector)
#Fixed with discretization in excel

predictions_ARIMA_log = predictions_ARIMA_log.add(predictions_ARIMA_diff_cumsum, fill_value=0)
predictions_ARIMA_log.head()
predictions_ARIMA = np.exp(predictions_ARIMA_log)
plt.plot(ts2, color='blue', label='SEK-MWh Original')
plt.plot(predictions_ARIMA, linestyle='-', color='orange', label='SEK-MWh Forecasted')
plt.title('Original Electricity Prices vs Forecasted Prices. RMSE Value: %.4f' % np.sqrt(sum((predictions_ARIMA-ts2)**2)/len(ts)))
plt.legend(loc='best')
plt.xlabel('Time')
plt.ylabel('SEK-MWh')
plt.ylim([0,800])
plt.show()
7.3 FNRVTool.m (Fixing NaN and Repetitive Values Tool)

clear all;
clc;
load('./Data_Output_Corrected/Data3.csv'); #choose your own data

m = length(Data3);
v = zeros (1,24);
c = 1;
media = [];

for i = 1:m
    if c > 24
        c = 1;
    end

    v(c) = Data3(i);

    if c == 23
        media = [media, mean(v)];
        v = zeros (1,24);
    end

    c = c+1;
end
7.4 ARIMA (1,1,0) Nordpool Hourly Prices 2015-2016.

Table 4.2: MAPE, RMSE, AVERAGES of Nordpool spot prices 2015-2016

<table>
<thead>
<tr>
<th>Hours</th>
<th>Original Data (SEK/MWh)</th>
<th>Forecasted (Corrected) (SEK/MWh)</th>
<th>O-F</th>
<th>ABS(O-F) *100</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>244.91</td>
<td>255.98</td>
<td>-11.07</td>
<td>4.52</td>
</tr>
<tr>
<td>2</td>
<td>243.76</td>
<td>255.96</td>
<td>-12.2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>241.47</td>
<td>255.94</td>
<td>-14.47</td>
<td>5.99</td>
</tr>
<tr>
<td>4</td>
<td>245.1</td>
<td>255.89</td>
<td>-10.79</td>
<td>4.4</td>
</tr>
<tr>
<td>5</td>
<td>247.49</td>
<td>255.96</td>
<td>-8.47</td>
<td>3.42</td>
</tr>
<tr>
<td>6</td>
<td>254.18</td>
<td>256.01</td>
<td>-1.83</td>
<td>0.72</td>
</tr>
<tr>
<td>7</td>
<td>254.76</td>
<td>256.15</td>
<td>-1.39</td>
<td>0.54</td>
</tr>
<tr>
<td>8</td>
<td>255.33</td>
<td>256.16</td>
<td>-0.83</td>
<td>0.32</td>
</tr>
<tr>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>17520</td>
<td>284.13</td>
<td>303.16</td>
<td>-19.03</td>
<td>6.7</td>
</tr>
</tbody>
</table>

MAPE (%) | RMSE | AVERAGE OF ORIGINAL | SIGMA ORIGINAL | AVERAGE OF FORECASTED
---------|------|---------------------|----------------|-------------------
21.42524238 | 49.01050863 | 224.746721 | 79.54839774 | 233.9886694 |

Graph 4.35: Log of Nordpool Spot Prices 2015-2016 (Box Cox transformation).
Graph 4.36: First non-seasonal difference of Nordpool Spot Prices 2015-2016

Graph 4.37: Stationary test of the first differenced series graphical results: average and variance variation

Graph 4.38: ACF and PACF plots of the first differenced series from Nordpool Spot Prices 2015-2016
ARIMA Model Results

Model: ARIMA(1, 1, 0) Log Likelihood 23544.919
Method: css-mle S.D. of innovations 0.063
Date: Tue, 30 May 2017 AIC -47083.838
Time: 13:06:08 BIC -47060.525
Sample: 01-01-2015 HQIC -47076.162
- 12-31-2016

| Coef  | Std. Err | Z    | P>|z| | [95.0% Conf. Int.] |
|-------|----------|------|-----|------------------|
| const | 2.725e-06| 0.001| -0.003 | 0.998 | -0.002 | 0.002 |
| ar.1  | 0.5136   | 0.006| 79.236 | 0.000 | 0.501 | 0.526 |

Graph 4.39: ARIMA Model Summary: Coefficients, AIC, Likelihood and Confidence intervals.

Graph 4.41: Nordpool spot Forecasting prices vs Original Prices
7.5 ARIMA (1,1,0) Nordpool Hourly Prices 2017

Table 4.3: MAPE, RMSE, AVERAGES of Nordpool spot prices 2017

<table>
<thead>
<tr>
<th>Hours</th>
<th>Original Data (SEK/MWh)</th>
<th>Forecasted (Corrected) (SEK/MWh)</th>
<th>O-F</th>
<th>ABS(O-F) *100</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>244.91</td>
<td>255.98</td>
<td>-11.07</td>
<td>4.52</td>
</tr>
<tr>
<td>2</td>
<td>243.76</td>
<td>255.96</td>
<td>-12.20</td>
<td>5.00</td>
</tr>
<tr>
<td>3</td>
<td>241.47</td>
<td>255.94</td>
<td>-14.47</td>
<td>5.99</td>
</tr>
<tr>
<td>4</td>
<td>245.10</td>
<td>255.89</td>
<td>-10.79</td>
<td>4.40</td>
</tr>
<tr>
<td>5</td>
<td>247.49</td>
<td>255.96</td>
<td>-8.47</td>
<td>3.42</td>
</tr>
<tr>
<td>6</td>
<td>254.18</td>
<td>256.01</td>
<td>-1.83</td>
<td>0.72</td>
</tr>
<tr>
<td>7</td>
<td>254.76</td>
<td>256.15</td>
<td>-1.39</td>
<td>0.54</td>
</tr>
<tr>
<td>8</td>
<td>255.33</td>
<td>256.16</td>
<td>-0.83</td>
<td>0.32</td>
</tr>
<tr>
<td>2764</td>
<td></td>
<td>303.16</td>
<td>-19.03</td>
<td>6.70</td>
</tr>
</tbody>
</table>

MAPE (%) | RMSE | AVERAGE OF ORIGINAL | SIGMA ORIGINAL | AVERAGE OF FORECASTED
---------|------|---------------------|----------------|-----------------|
7.745571969 | 37.03301712 | 291.2428407 | 38.21954218 | 270.61 |

Graph 4.40: Log of Nordpool Spot Prices 2017 (Box Cox transformation).
Graph 4.41: First non-seasonal difference of Nordpool Spot Prices 2017

Graph 4.42: Stationary test of the first differenced series graphical results: average and variance variation

Graph 4.43: Stationary test of the first differenced series graphical results: average and variance variation
### ARIMA Model Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Model:</td>
<td>ARIMA(1, 1, 0)</td>
<td>Log Likelihood</td>
<td>5527.980</td>
</tr>
<tr>
<td>Method:</td>
<td>css-elle</td>
<td>S.D. of innovations</td>
<td>0.023</td>
</tr>
<tr>
<td>Date:</td>
<td>Tue, 30 May 2017</td>
<td>AIC</td>
<td>-11049.959</td>
</tr>
<tr>
<td>Time:</td>
<td>12:21:59</td>
<td>BIC</td>
<td>-11032.166</td>
</tr>
<tr>
<td>Sample:</td>
<td>01-01-2017</td>
<td>HQIC</td>
<td>-11043.534</td>
</tr>
<tr>
<td></td>
<td>- 04-26-2017</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| coef  | std err | z     | P>|z| | [95.0% Conf. Int.] |
|-------|---------|-------|-------|------------------|
| const | 2.357e-05 | 0.001 | 0.020 | 0.984 | -0.002 | 0.002 |
| ar.1  | 0.4772 | 0.017 | 28.558 | 0.000 | 0.444 | 0.510 |

### Graph 4.44: ARIMA Model Summary: Coefficients, AIC, Likelihood and Confidence intervals.

![Graph 4.44 ARIMA Model Summary](image)

### Graph 4.45: Nordpool Spot Forecasting prices vs Original Prices

![Graph 4.45 Nordpool Spot Forecasting prices vs Original Prices](image)
7.6 ARIMA (1,1,1) OMIE Hourly Prices 2017 additional results.

Table 4.4: MAPE, RMSE, AVERAGES of OMIE spot prices 2017

<table>
<thead>
<tr>
<th>Hours</th>
<th>Original Data €/MWh</th>
<th>Forecasted (Corrected)</th>
<th>C-F</th>
<th>ABS(C-F)*100</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>62.41</td>
<td>56.4725343</td>
<td>-5.937465704</td>
<td>9.51</td>
</tr>
<tr>
<td>2</td>
<td>47.73</td>
<td>59.07262766</td>
<td>11.34262766</td>
<td>23.76</td>
</tr>
<tr>
<td>3</td>
<td>45.42</td>
<td>58.8584834</td>
<td>13.4384834</td>
<td>29.59</td>
</tr>
<tr>
<td>4</td>
<td>43.63</td>
<td>58.82068497</td>
<td>15.19068497</td>
<td>34.82</td>
</tr>
<tr>
<td>5</td>
<td>43.37</td>
<td>58.79045147</td>
<td>15.42045147</td>
<td>35.56</td>
</tr>
<tr>
<td>6</td>
<td>43.45</td>
<td>58.78582357</td>
<td>15.33582357</td>
<td>35.30</td>
</tr>
<tr>
<td>7</td>
<td>46</td>
<td>58.78699287</td>
<td>12.78699287</td>
<td>27.80</td>
</tr>
<tr>
<td>8</td>
<td>55.38</td>
<td>58.82944566</td>
<td>3.449445663</td>
<td>6.23</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>2764</td>
<td>284.13</td>
<td>303.16</td>
<td>-19.03</td>
<td>6.70</td>
</tr>
<tr>
<td>MAPE</td>
<td>RMSE Corrected</td>
<td>Average</td>
<td>Sigma</td>
<td>Average</td>
</tr>
<tr>
<td>20.242%</td>
<td>270.08</td>
<td>49.92769677</td>
<td>15.01336822</td>
<td>51.45603168</td>
</tr>
</tbody>
</table>

Graph 4.46: Log of OMIE Spot Prices 2017 (Box Cox transformation).
Graph 4.47: First non-seasonal difference of OMIE Spot Prices 2017

Graph 4.48: Stationary test of the first differenced series graphical results: average and variance variation

Graph 4.49: Stationary test of the first differenced series graphical results: average and variance variation
Graph 4.50: ARIMA Model Summary: Coefficients, AIC, Likelihood and Confidence intervals.

Graph 4.51: OMIE spot Forecasting prices vs Original Prices