Instituto Universitário de Lisboa

Departamento de Ciências e Tecnologias da Informação

Electricity price forecasting utilizing machine learning in MIBEL

João Filipe Ferreira Januário

Orientador:

Professor Doutor Luís Miguel Martins Nunes

ISCTE-IUL

Co-orientador:

Doutor Nuno Pinho da Silva,

R&D NESTER

Janeiro, 2019

Abstract

Short term electricity price forecasts have become increasingly important in the last few decades due to the rise of more competitive electricity markets throughout the globe. Accurate forecasts are now essential for market players to maximize their profits and hedge against risk, hence various forecasting methodologies have been applied to electricity price forecasting in the last few decades. This dissertation explores the main methodologies and how accurately can three popular machine learning models, SVR LSTM and XGBoost, predict prices in the Iberian market of electricity. Additionally, a study on input variables and their relationship with the final price is made.

Keywords

Machine Learning, Electricity, Clearing Market, Prediction, Input Variables

Contents

List of Figures

List of Tables

Abbreviations

AvE – Average error

AvPE – Average Percentage Error

ANN – Artifical Neural Network

AR – AutoRegressive

ARIMA – AutoRegressive Integrated Moving Average

ARMA - Autoregressive moving average

BP – Backpropagation

CRISP-DM - Cross-Industry Standard Process for Data Mining

EPF – Electricity Price Forecasting

ERBFN – Enhanced Radial Basis Function Network

FWV - Fourier Wave Filtering

GARCH - Generalized autoregressive conditional heteroskedastic

LM – Levenberg Marquardt Algorithm

LSTM – Long Short-Term Memory

LS-SVM – Least-Squared Support Vector Machines

LTPF – Long Term Price Forecasting

MAPE – Mean Absolute Percentage Error

MIBEL – Iberian Electricity Market

MLP- Multi-Layer perceptron

MTPF – Medium Term Price Forecasting

OMIE – Iberian Electricity Market Operator

RBF – Radial Basis Function Network

RES – Renewable Energy Sources

RNN – Recurrent Neural Network

SLP – Single-Layer Perceptron

SOM – Self Organizing Map

SR – Scaling Range

STPF – Short Term Price Forecasting

SVM – Support Vector Machine

SVR – Support Vector Regression

TF – Transfer Function

WMAPE – Weekly Mean Absolute Percentage Error

WT- Wavelet Transformation

1- Introduction

Forecasting is a term used to define the ability to predict a future event or variable based on the past study of variables that affect the objective of the forecast. In a competitive liberalized power market, it's vital for all participants to be able to forecast the electricity price with high accuracy to formulate strategies that allow them to bid and sell energy with the lowest risk possible.

Since the early 1990s, various countries began a process of liberalization of energy trade leading to more competitive energy markets, one example being the Iberian market of electricity (MIBEL). Formed in 2007 as a collaboration between the governments of Portugal and Spain, this is an interconnected market with the goal of distributing electricity in the two countries while benefitting the final consumer by promoting a competitive and free market.

The MIBEL includes three levels of trading one of them being long term and the remaining two being short term, the long term option is in the form of bilateral contracts and can span from a few months to a few years, while day-ahead electricity market and intraday electricity market sessions, also known as spot trading, are short term (Pastor, Pinho, & Esteves, 2018) (Szkuta, Sanabria, & Dillon, 1999).

The short-term options consist of a market where participants bid in a shared pool and a market operator clears the price, this makes the price volatile, but it gives all consumers and providers an equal opportunity to maximize the value of their actions in MIBEL.

The long-term option is used when companies prefer to avoid the volatility of daily market prices and want to hedge against the risk of participating in the short-term market. The main objective of the electricity market is to decrease the cost of electricity by promoting competition and transparency.

As the market is interconnected energy prices are usually the same in both Portugal and Spain but there might be occasions where they differ due to the physical capability of the network (Pastor et al., 2018).

Short-term trading options are highly dependent on the demand and available offers of energy which in turn makes the short-term electricity market extremely volatile. This causes the ability to predict prices before-hand to be very valuable in order to maximize the value that the market participants get from the day-ahead electricity market. Providers want to maximize their profit and consumers want to get the highest amount of energy possible at the cheapest price.

Day-ahead trading is especially important due to storing energy for later use not being a trivial matter, as the manufacturing of large-scale batteries is generally considered not economically viable. At the moment the most frequent way of storing energy in MIBEL is through water reservoirs by utilizing excess energy to pump water that is then stored to generate electric power at a later time.

Long term predictions are also valuable, but they are mostly used for strategic decisions such as: 1-) modifying the amount of energy that is traded between regions. 2-) expanding energy generation. 3-) re-planning the distribution network (Szkuta et al., 1999).

The electricity market has been studied extensively over the years, and certain aspects such as load forecasting have had great breakthroughs (Abdel-Aal, 2006) (Dang-Ha, Bianchi, & Olsson, 2017). Electricity price forecasting presents some characteristics that create difficulty when building prediction models, namely: 1-) it is highly affected by the calendar, especially by weekends and holidays where the prices vary when compared to normal weekdays, creating outliers throughout the year. 2-) It is highly affected by the sources of energy in the network. 3-) It affected by seasons because in drier seasons like the summer there is less hydropower energy being generated as opposed to the more humid seasons. 4-) The pricing can vary a lot in the same day, as early morning hours usually present much cheaper prices than afternoon hours. 5-) It is affected by the player's bidding strategies that, naturally, are not known ahead by the market.

This dissertation will be focused on predicting short term electricity prices in the MIBEL market, using machine learning algorithms which have been demonstrated to have good results when utilized to predict time-series problems (Aggarwal, Saini, & Kumar, 2009) (Jones et al., 1989) (Ruta & Gabrys, 2007) (Ahmed, Atiya, El Gayar, & El-Shishiny, 2010).

Machine learning has been applied to time-series prediction since the '80s (Jones et al., 1989), and since then these methods have increasingly grown in popularity. Being applied not only to electricity price forecasting but various different real-world scenarios, including but not limited to financial market predictions, environmental state predictions and reliability forecasting (Sankar & Sapankevych, 2009).

In the latest decades machine learning has proven to have equal or better results than classical statistical models in several different problems due to: 1-) The increase in theoretical understanding of the models. 2-) The number of varied models that have been developed. 3-) The increasingly ease of access to computing power and open-source machine learning libraries (Ahmed et al., 2010).

Several machine learning models have been applied to predict electricity clearing prices in other markets (Weron, 2014), however to the best of our knowledge it is not an extensively explored methodology in MIBEL.

Algorithms that have shown promising results in previous works, namely SVR, LSTM, and XGBoost, were selected in order to compare how each one performs in the MIBEL. Additionally, a baseline was established utilizing AR (1) and ARIMA. Before starting to describe the practical work, an analysis is made to understand which approaches have been proven useful when predicting time series, and more specifically, electricity price time series.

Objectives

This dissertation will be aimed at studying to what extent is it possible to forecast short term electricity prices utilizing machine learning models. The maximum time frame considered will be up to one week and as a comparative baseline other models that have shown good results will be used. It will also be interesting to compare the difference in the quality of the results in the various models as the time horizon is extended, for example, to compare the day-ahead forecast to the 1-week ahead forecast.

The price of electricity is affected by several variables and the importance of each one to the final price is not yet fully understood, variables like the season, the time of the day, power generation, and others are suspected to be highly important to the final price. As such a study will be made as to how the variables identified as potentially important affect the final market price. For this dissertation, the Iberian electricity market will be studied exclusively and all historical data is aggregated by technology and provided by *R&D Nester*.

In short. the following questions can be identified to be resolved in this work:

- What are the most important variables that affect electricity pricing?
- Is it possible to forecast short term electricity prices in the MIBEL with “good” results using artificial intelligence models?
- How do the results compare to other types of models?
- How does the time horizon affect the forecast quality?

Document Structure

Chapter 2 starts as by describing the methodology utilized in this work followed by giving a brief introduction to the MIBEL market. The remainder of the chapter presents the related work done in electricity price forecasting with a special focus on machine learning.

The theoretical background can be found in Chapter 3 where the main concepts behind the models relevant to the literature and to this dissertation are explained.

Chapter 4 consists mainly of an extensive statistical analysis of all the relevant data to be used in the models.

In Chapter 5 all the experiment definition and results can be found.

In chapter 6 the conclusions can be found as well as a suggestion for further work.

An appendix can be found after chapter 6 that contains all the relevant information about the training of each model.

2- Literature review and business understanding

Development Methodology

In this dissertation, the *Cross-Industry Standard Process for Data Mining* (CRISP-DM) (Wirth, 2000) will be used as a development methodology. This framework splits the development process into six separate phases, *Business Understanding, Data Understanding, Data Preparation, Modeling Evaluation, and Deployment.*

This is a flexible framework in which the developer can go back and forth between phases as needed to better adjust the final model.

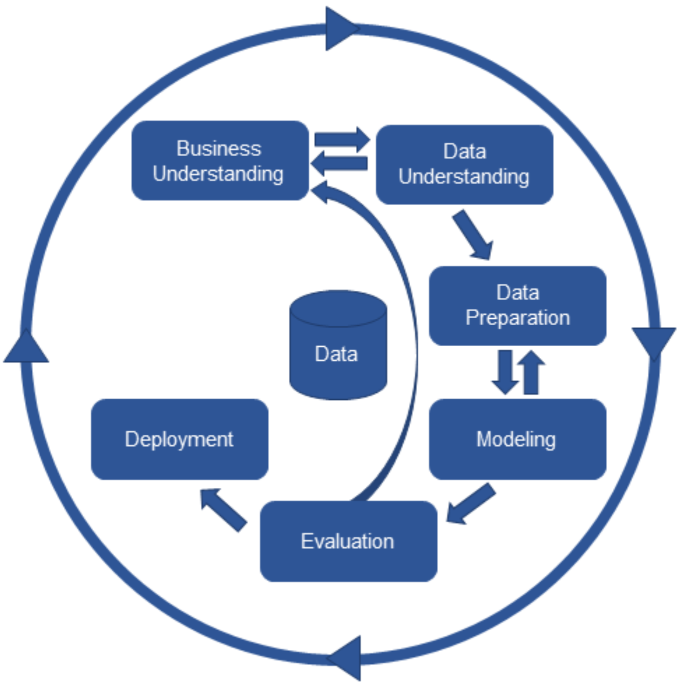

Figure 1 - CRISP-DM Framework (Sharma & Bradford, 2017)

The outer circle shown in *Figure 1* represents the idea that *data mining* can always be bettered, even after deployment the model can always be adjusted with new data to further improve its forecasting ability. Inside the circle, the different phases are represented, each phase can be defined as such:

Business Understanding*:* Initial phase of the project where the developer needs to understand the context of the problem that needs to be solved in order to identify potential issues.

Data Understanding: Data exploration phase, in this phase the training data needs to be studied, in order to identify potential issues (null values, outliers), and to identify subsets that might have unexpected patterns of relationship between each other.

Data Preparation: Treating the data by solving the problems identified in the previous phase. In addition to this, some data treatment techniques may be experimented in this phase, like normalizing the data or removing certain variables.

Modeling: In this phase, the model is created or adjusted, the model parameters are calibrated in order to achieve the best results.

Evaluation: After the model from the previous phase is finished training and predicting a data set, the results need to be carefully analyzed, if the results aren't good enough then one of the previous phases needs to be adjusted, otherwise the developer moves on to the deployment phase.

Deployment: The deployment phase presents the finished product, usually in the form of a report.

The business understanding phase can be mostly seen in chapters 1, 2 and 3, where the motivation, objective, context, and scope of the problem are described, along with a description of how the MIBEL works as well as the theoretical background of the machine learning algorithms that will be applied.

The data understanding phase relates to chapter 4, where the statistical analysis of every relevant variable found in the business understanding phase can be seen.

The following phases mostly relate to the practical work presented in chapter 5, where information about the data preparation, model parameters, and result evaluation is found.

Business Understanding

The day-ahead market in MIBEL follows a marginal price model. This model aims to keep electricity price as low as possible while completely satisfying the daily market demand.

The agents bid supply and demand offers for the 24 hours of the next day. The selling offers are sorted in ascended order while the buying offers are sorted in descending order. Sessions are closed at 11:00 (GMT +0) in the previous day (D-1). The MIBEL spot market operator OMIE announces the clearing prices for each hour of the next day, taking into account the curve intersection between the supply curve (generated from aggregated supply bids) and the demand curve (generated from aggregated demand bids). Additionally, some adjustments might be needed to make sure the physical network is able to handle the load without fail (Pastor et al., 2018).

Non-renewable energy providers need to pay coal or gas in order to generate electricity. Renewable energy providers on the other hand, generate electricity from sources in nature like the wind or the sun and as a result they do not have to pay for their primary energy sources. This results in renewable energy providers being able to provide cheaper offers than non-renewable providers.

As the market follows a marginal price model this means that the last bid to match the demand will define the final price for each hour. That price is then the price that all transactions occur independently of the initial bidding price. This means that all providers will satisfy demand at the exact same price as long as their offer is lower or equal to the intersection between the supply and demand curves. As renewable energy providers sell at cheaper prices than non-renewable energy providers, they enter the market first and can sometimes completely fulfill the market demand.

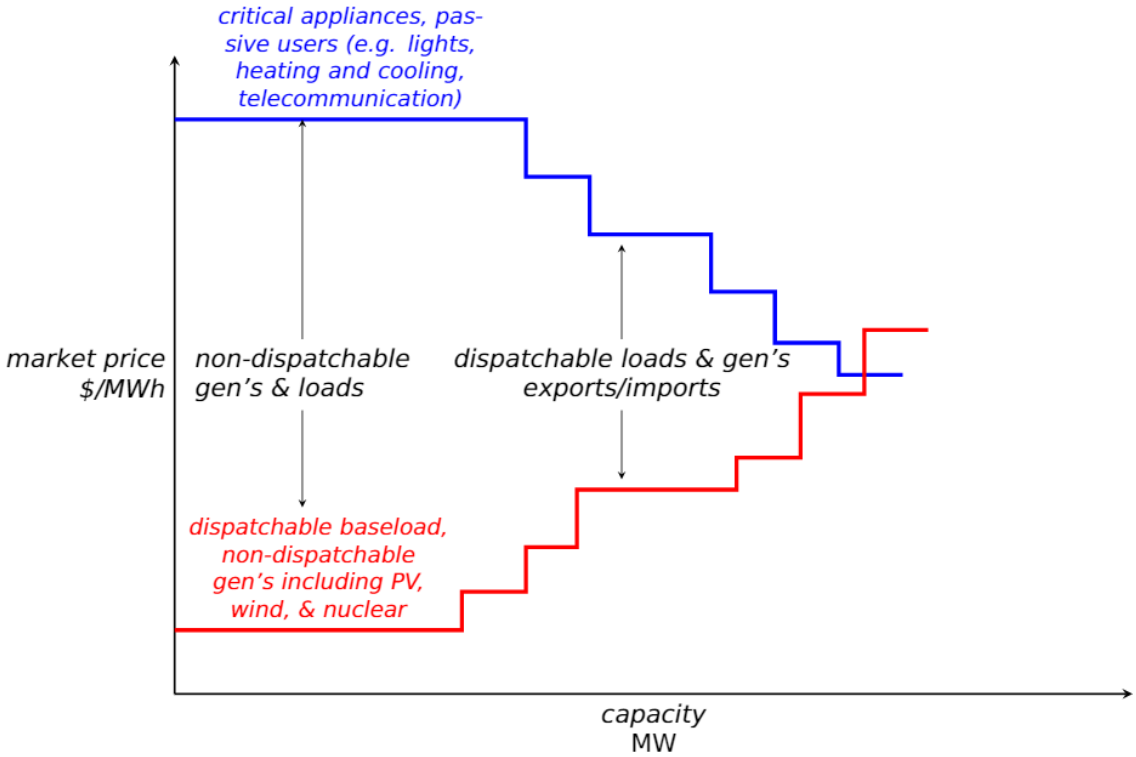

Figure 2 - Example of spot market clearing price (neighbourpower.com/blog/solar-deregulated-power-market/)

As can be seen in *Figure 2* non-dispatchable load generation such as wind and solar have the lowest supplier prices and critical appliances such as light, heating or telecommunication have the highest demand prices. The intersection between the two curves is where the marginal price for each hour is defined. This process occurs for every hour.

Literature Review

The purpose of an organized electricity market is to match the supply and demand of electricity to determine the market-clearing price (Weron, 2014). In the MIBEL participants can conduct their business in the day-ahead electricity market or in the form of bilateral contracts between companies.

In the day-ahead electricity market, the participants bid supply and demand for the 24 hours of the next day to a common pool, the sessions for the day-ahead market are always closed in the previous day at 11:00 (GMT +0). The spot market operator OMIE announces the clearing price for each hour of the next day with the selling offers being sorted in ascending order while the buying offer being sorted in descending order.

In addition to the day-ahead electricity market, bilateral contracts are utilized by companies to hedge against the day-ahead electricity market risk and to make sure they are able to satisfy their needs in the long term, in bilateral contracts the buyer and the seller negotiate directly and agree on the distribution of a fixed amount of energy at a fixed price during a certain amount of time.

Publications on electricity price forecasting started roughly around the year 2000, prior to this date there is almost no literature on this topic, steadily increasing for the next few years and having doubled in 2005 and then tripled in 2006 when comparing to 2002 numbers. For the next few years there was a steady increase in publications about this topic, with a slight drop-off in numbers in 2010, and then a huge increase in 2012 and the following years (Weron, 2014).

Articles utilizing machine learning techniques for electricity price forecasting have increased tremendously in the past decade, when compared to the prior decade where most articles focused on developing statistical models. This is due to the increasing ease of access to computing power and open software machine learning libraries, which has caused the research community to have a renewed interest in this topic as it is yet an unsolved problem that can potentially be solved or at least vastly more understood utilizing machine learning algorithms.

It is not obvious that machine learning algorithms are able to outperform other methods of prediction, mainly due to each study utilizing different datasets, different software implementations and different evaluation models (Weron, 2014), as such this section will include studies on other methodologies besides machine learning.

Electricity price forecasting

Accurate forecasts don't guarantee profits and there is always risk in trading in the MIBEL market due to the high volatility of prices (Aggarwal et al., 2009). Since the price is directly related to the amount of energy in the network traded between suppliers and consumers there is also an interest in the prediction from a viewpoint of the network managers in order to better plan how the network will operate.

Electricity price is a time-series, which means that the series is defined as a consistent sequence of electricity prices over a period of time. Electricity price forecasting is the capability to predict the price of energy at one or more points in time. This prediction is complex seeing as there are multiple variables that affect the final price, and some of them like load forecasting are also complex to predict.

Predictions can be classified into short-term price forecasting (STPF), medium-term price forecasting (MTPF) or long-term price forecasting (LTPF) (Singh, Husain, & Mohanty, 2016). Short-term predictions allow companies to formulate strategies to optimize their participation in the spot market, while medium-term and long-term strategies allow companies to adjust their overall strategy. This may include the overall level of production in the case of suppliers, bilateral contracts with other companies, planning investments, among others (Aggarwal et al., 2009).

The horizon of STPF is between an hour to a week, while MTPF and LTPF range from a few weeks to a few months and a few months to several years respectfully. The purpose of each horizon will be detailed further in the next subsections.

Short-term price forecasting

Short-term forecasts are essential for the players that participate in the spot market. Producers need to forecast energy prices in order to formulate their strategy for participation in the market but also to better optimize the scheduling of their electrical resources to maximize profits.

Investment decisions in renewable generation in the current European regulatory framework with reduced subsidies may leverage short-term price forecasting to simulate markets and compute realistic cash flows in market simulations. For example, short-term price simulation can be used as input in the R&D Nester's renewable portfolio simulator (Pastor et al., 2018).

Due to energy generally not being a storable resource, consumers need to be able to take advantage of the cheapest moments to get the maximum amount of usable energy. This means that consumers need to have an active participation on the market to satisfy their daily needs while wanting to minimize their risk as much as possible (Catalão, Mariano, Mendes, & Ferreira, 2007).

Medium-term and Long-term price forecasting

Medium-term and long-term price forecasting are very important to the overall electricity market. Most notably, these help strategize when to form bilateral contracts with other companies in order to maximize their profits while satisfying their needs. MTPF and LTPF are also important for other activities such as, generation expansion planning, maintenance scheduling and overall investing (Torbaghan, Motamedi, Zareipour, & Tuan, 2012). For network managers, these predictions help them to better plan how the network will need to change over time and to monitor distribution safety.

Doing medium and long-term predictions is an incredibly complex task, seeing as short-term predictions are not yet fully understood. The time horizon is usually much longer and since electricity price is volatile predicting them accurately over a long period of time is much more difficult than in STPF. In addition to this, since the market liberalization happened somewhat recently along with the investing in renewable energies, the existing historical data with quality is limited which makes having good results with MTPF and LTPF compared to STPF very difficult (Torbaghan et al., 2012).

Input Variables

The best choice of input variables is still an open area of research, and there have been as many as 40 different input variables utilized by different researchers throughout the years.

The most widely used variable is the historical data of electricity prices, being utilized in practically every work related to electricity price forecasting (Aggarwal et al., 2009). It is apparent that prices exhibit seasonality on the daily, weekly and possibly at the yearly levels. The last one not being relevant for short-term price forecasting but the first two have to be taken into account. This is because peak hours during the day have considerably higher prices than hours in the middle of the night and there is also a sizable difference in the price curve between weekdays and weekends or holidays (Weron, 2014).

Another variable that affects pricing in the spot market is system load. This is the level of demand and consumption to be expected in the system as it is the basic level of supply and demand in the spot market.

Weather variables, such as temperature, wind speed, precipitation, and solar radiation are also suspected to be important to the final market price. These variables affect the quantity of energy that can be generated from renewable energy sources. This causes the price to have a clear pattern of pricing between seasons as can be seen in *Figure 3*. Dryer seasons like the summer have considerably less precipitation than wetter seasons like in the winter, which greatly affects the generation of hydraulic energy. This effect is significant enough that some researchers argue that the prediction should be made by training separate models for each season, instead of one model for various seasons to achieve better results (Niu, Liu, & Wu, 2010).

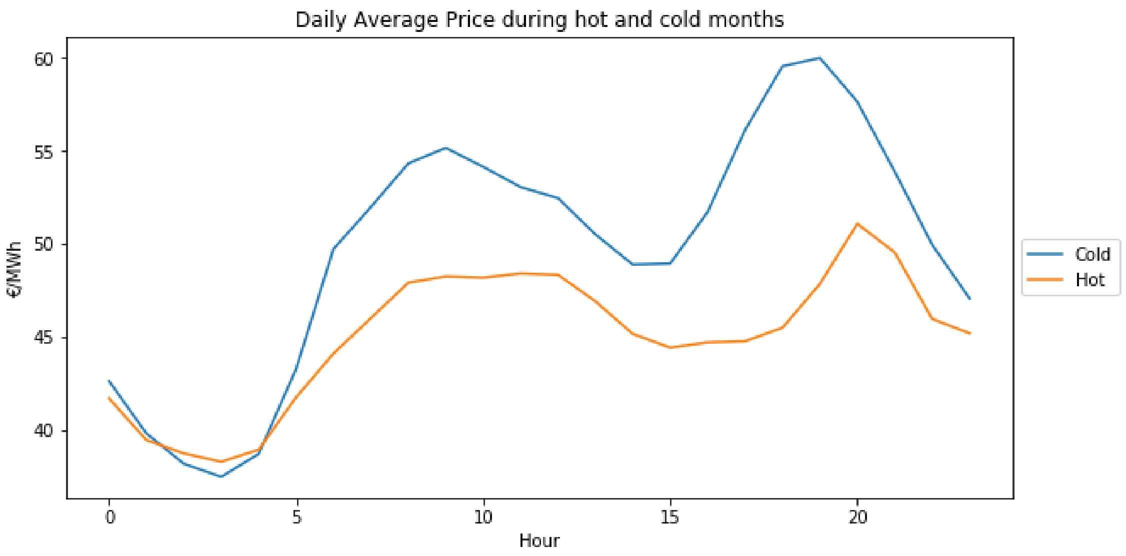

Figure 3 - Average price of electricity for a day during cold and warm months in MIBEL

Fuel costs, especially oil, natural gas, and coal to a lesser extent, are also suspected to have an impact on spot market pricing. Scheduled maintenances, outages, or other types of failures in power grid components might also be responsible for shifts in pricing (Weron, 2014).

In addition to previously stated variables, there are other events that cause shifts in pricing. For example, an important football game might draw a lot of additional power to certain areas over a period of time. This creates outliers, meaning sudden price spikes or drops in the data.

Electricity price forecasting methodologies

Electricity price forecasting has had an increasing interest since the early 2000s and many researchers have tried to contribute to solve this problem. As such there are various methodologies that have been applied to electricity price forecasting problems with varying degrees of success. The methodologies which have been proven to have the best results are: 1-) multi-agent models. 2-) statistical analysis methods. 3-) machine learning algorithms. (Aggarwal et al., 2009). In this section, each of these methodologies will be briefly described.

Multi-agent models

Multi-agent simulations consist of models which simulate how a system operates by generating heterogeneous agents that behave like real-life agents in a set environment. In this case, it can be done by simulating companies with different bidding strategies and making them interact with each other and analyze how the price shifts over time by companies trying to meet the supply and demand in a simulated market (Ventosa, Baíllo, Ramos, & Rivier, 2005).

These agent-based models are extremely flexible seeing as you can study how a different parameter in the bidding strategy can affect the overall result quite easily. On the other hand, developing these simulations require a lot of knowledge of how companies operate in the market. This is because there are several different components that need to be defined including but not limited to the number of companies in the simulation, their bidding strategies and how they interact with each other.

Relying on these types simulation to have precise quantitative results is very risky, as it relies too much on the developed simulation being perfectly modeled, as such these types of simulation generally focus on qualitative problems rather than quantitative ones.

Statistical analysis

Statistical methods try to forecast electricity price by using a mathematical combination of historical electricity pricing data, but certain methods can also include other relevant variables like weather forecasting or production and consumption figures.

There are several stochastic models utilized in electricity price prediction, the most commonly found models being: 1-) Auto Regression. 2-) Moving Average. 3-) ARMA. 4-) ARIMA. 5-) GARCH. These models all take into account the history of pricing, but there are other multivariate models like TF and ARMA with exogenous variables that take into account other variables that might affect electricity price (Aggarwal et al., 2009).

These types of methods have performed poorly when dealing with outliers. Several researchers have tried to improve upon these models to better deal with outliers, seeing as electricity historical data is extremely volatile and can have outliers for no apparent reason. Some authors recommend filtering out outliers by replacing them with a more usual value, either by taking an average of the week or by taking the average price of nearby neighbors, before applying the stochastic model to the data in order to achieve better results(Weron & Misiorek, 2008)(Janczura, Trück, Weron, & Wolff, 2013).

There exist several ways to detect outliers. Some simple methods include taking weekly price average and variance and then considering prices that are too distant from those values as outliers. More sophisticated methods include recursive filters (Weron & Misiorek, 2008) or wavelet filtering (Stevenson, 2001). It is not recommended that a fixed price threshold is considered when detecting outliers seeing as electricity prices shifts somewhat significantly over the various seasons and that needs to be taken into account (Fanone, Gamba, & Prokopczuk, 2013).

Machine Learning

Machine learning has been an increasingly popular field since the ’80s, increasing in popularity in the last decade due to the ease of access to computational power and open source algorithms. These types of computational algorithms have been shown to have an immense ability to handle complexity and non-linearity amongst various different types of problems (Obermeyer & Emanuel, 2016).

There are many different types of machine learning algorithms that have been developed for general use. In the case of prediction of a non-discrete variable, commonly known as regression, artificial neural networks (ANN) and support vector machine (SVM) are the main classes of machine learning techniques (Weron, 2014) and have been extremely effective when applied to some time-series forecasting problems in the past(Ahmed et al., 2010).

Despite generally being versatile these algorithms have some weaknesses. If the dataset is not correctly balanced or they are poorly configured they may overfit, that is lose the ability to generalize due to memorizing in too much detail the training dataset instead of capturing relationships between the data (Razak et al., 2015). Moreover, there are currently so many algorithms, most with many different configuration variables, that it can become very time consuming to find a good solution for each problem.

In addition to the previously mentioned algorithms, there are also other types of techniques that can be useful when developing a machine learning model. Clustering algorithms, such as K-means, can be extremely useful to split the dataset into data groups that have more relationships between each other. This can then be utilized to find unexpected patterns of relationship between the data and better treat outliers and as a result improve the final forecasting.

As shown in Table 1, from a small number of studies exclusively utilizing machine learning models, it is possible to distinguish a multitude of different datasets, time periods studied, different prediction periods and several different methods of evaluation. This seems to support Weron's (2014) claims that there is no model that obviously outperforms all the others in every situation in an EPF context.

Table 1 - Examples of Machine Learning Models Applied to EPF

Paper	**Model**	**Training Data**	**Predicted Period**	**Preprocessing technique**	**Results**
(Catalão et al., 2007)	MLP trained with LM	2002 Spanish Market, 2000 Californian Market	1 week	-	AvPE 3%-9%
(Szkuta et al., 1999)	MLP trained with BP	Victorian Electricity Market, October 1996-May 1997	1 week	-	Daily AvE 2.18-11.09
(Razak et al., 2015)	MLP trained with LM(1); LS-SVM(2)	Ontario Power Market, 2003-2006	1 week	Data normalization	WMAPE(%) 11.48-22.56(1); 10.11-18.12(2)
(Singh et al., 2016)	MLP trained by LM(1); Custom Model with 4 NN(2)	New South Wales electricity market	1 day	WT(2)	WMAPE(%)- 5.30-9.32(1); 4.78-7.18(2)
(Yamin, Shahidehpour, & Li, 2004)	MLP trained with BP	Californian Market - 1999	1 week	Outliers removed	WMAPE(%)- 11-13

(Wu, Zhou, Yu, Zhu, & Yang, 2004)	MLP trained with BP	South Chinese Market, 2003	10 days	Noise Filtration using FWV	AvPE 8%
(Lin, Gow, & Tsai, 2010)	ERBFN	Pennsylvana-New Jersey-Maryland Market,2002	1 week	-	MAPE(%) 5.5622
(Sansom, Downs, & Saha, 2003)	SVM	NSW State Electricity Market, 1998	7 days	-	MAPE(%) 25.8 (over a 9 week period)
(Niu et al., 2010)	SVM	Pennsylvana-New Jersey-Maryland Market,2002	1 week	Clustering and SR [0,1]	Error Distribution of [-10%,10%]
(Yao, Song, Zhang, & Cheng, 2000)	RBF	UK Electricity Market, 1997	1 week	WT and different models for each weekday	AE(%) – 3.46-7.44

AvE – Average error; **AvPE** – Average Percentage Error; **ERBFN** – Enhanced Radial Basis Function Network; **FWV** - Fourier Wave Filtering; **LM** – Levenberg Marquardt Algorithm; **LS-SVM** – Least-Squared Support Vector Machines; **MAPE** – Mean Absolute Percentage Error; **MLP-** Multi-Layer perceptron; **RBF** – Radial Basis Function Network; **SVM** – Support Vector Machine; **SR** – Scaling Range; **WMAPE** – Weekly Mean Absolute Percentage Error; **WT-** Wavelet Transformation;

It was also noted that some studies are limited to specific predicted time-periods to avoid special weeks prone to outliers such as weeks containing holidays in the middle of the week or where a big event is occurring. This is also pointed out as a problem in (Aggarwal et al., 2009). One such example is (Razak et al., 2015) where LS-SVM is shown to have a slightly better forecast accuracy than an MLP trained with LM, yet the models were only applied to two different weeks of the year and only one month apart from each other.

It is not obvious that the MLP model wouldn’t outperform the LS-SVM in other weeks with slightly different patterns.

In order to do a comprehensive test in the MIBEL market, this dissertation will compare the most promising models found in literature. Furthermore, a comparison to other statistical models will also be made in order to see how the results differ while utilizing the exact same data. At this point in time, there is no study found to the best of our ability that compares different machine learning models and classical models in the context of the Portuguese market.

3- Theoretical Background

Machine Learning Algorithms

As previously stated, and shown in table 1, the two most popular methods for electricity price forecasting are ANN and SVM. Every ANN can be classified into a more specific type based on its architecture and training method. This section will go over the theoretical background of each algorithm used in this dissertation, namely LSTM, which is a type of ANN, SVR which are support vector machines applied to regression problems, and XGBoost which is an algorithm based on decision trees.

Artificial neural networks

ANN have several advantages over classic mathematical models most notably they have: 1) High tolerance to noise in the data. 2) The ability to understand the relationship between variables that aren't yet fully understood. 3) Form learning patterns that allow these algorithms to make good predictions without memorizing the data (Han, 2006).

These algorithms are also efficient when dealing with vast amounts of data due to the possibility of parallelizing the processing that is needed to train the algorithm. This might not always be possible when dealing with classical mathematical models (Weron, 2014).

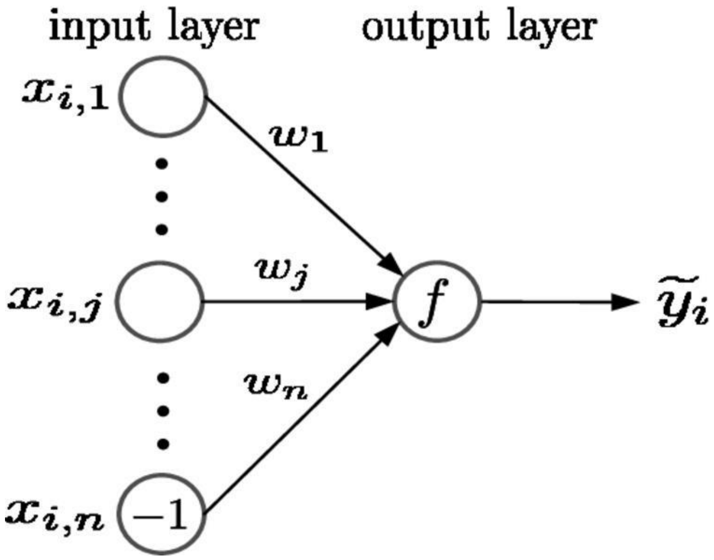

Figure 4- Example of an SLP (Wang, Zhang, Tao, & Wang, 2018)

The simplest feed-forward neural network is called a *single-layer perceptron*, this network contains no hidden layers and is therefore equivalent to linear regression.

As can be seen in *Figure 4* initial inputs, *xi,* is given to the network and comes from the data, the input layer has as many nodes as input variables.

The weight, *w*, represents how much a connection between two neurons weighs, they are initiated with a random seed and are then trained using a learning algorithm.

The activation function, *f,* calculates the output of each node. There exist many different activation functions found in literature like threshold, sigmoid, radial basis, and linear functions as an example. Further detail can be found in *Figure 5*.

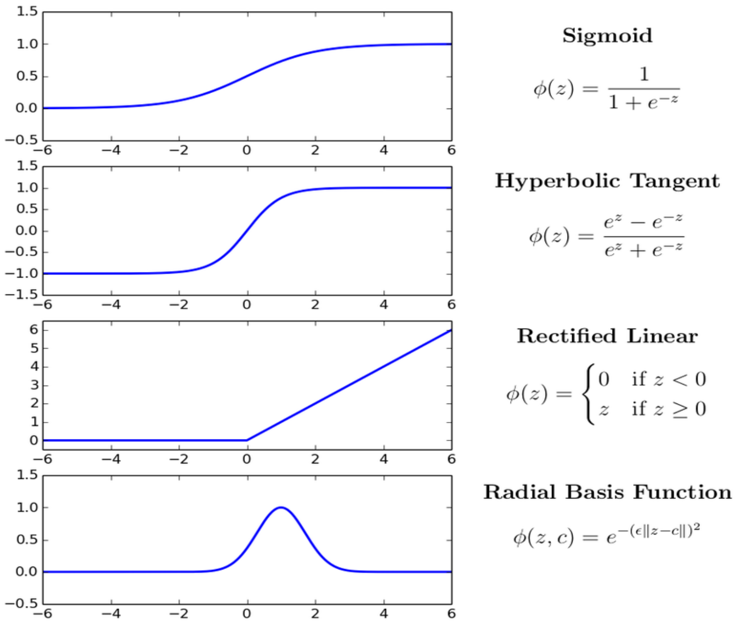

Figure 5 - Activation Functions (Hughes & Correll, 2016)

Additionally, each node also has a bias, commonly referred to in the literature as *Θ*, which is a constant term added to the value of the node. Finally, *Y* is the final output of the network.

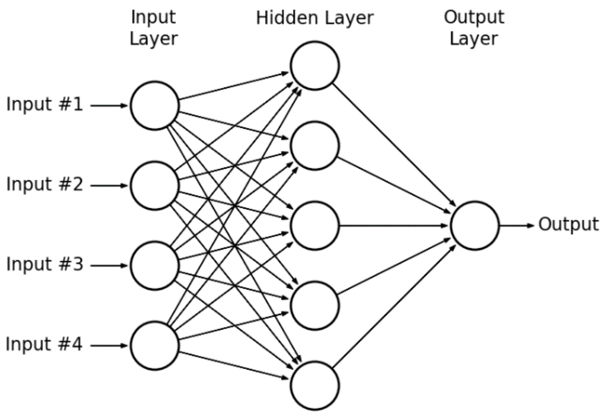

Figure 6 - Example of an MLP (Mohamed, Negm, Zahran, & Saavedra, 2015)

Figure 6 represents what is called a multi-layer perceptron. The units in each layer are connected to the next layer, but not each other, until they reach the output layer, which may have one or more nodes depending on the type of problem the ANN is solving. This is the most commonly known and used type of ANN (Osório, Gonçalves, Lujano-Rojas, & Catalão, 2016).

Finding the optimal model of the network to solve a specific problem is often through trial and error, the models can differ in the number of layers, the number of nodes in each layer, activation functions, and weight training models.

A Recurrent Neural Network (RNN) is a more complex type of ANN. These types of networks excel at problems with sequential data like time-series (Graves, Mohamed, & Hinton, 2013) due to utilizing the previous hidden state for the next prediction as is exemplified by *Figure 7*.

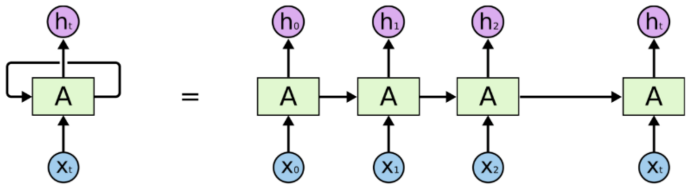

Figure 7 - Example of a Recurrent Neural Network (icode9.com/content-4-141370.html)

This architecture allows the network to utilize all previous information to predict the next data point as the current hidden state will be a function of all previous hidden states, which intuitively seems to provide better results in time-series problems than Feed-Forward neural networks due to these not taking the previous hidden states into account.

Long Short-Term Memory Networks

Long Short-Term memory networks, LSTM in short, are a type of recurrent network introduced in (Hochreiter & Schmidhuber, 1997), which try to solve the vanishing and exploding gradient problems present in classic RNN architecture (Hochreiter, 1998).

These networks contain special units named memory blocks, these blocks contain memory cells with self-connections that act as mini-layers storing the temporal state of the network, in addition to multiplicative units called gates that control the flow of information (Sak, Senior, & Beaufays, 2014).

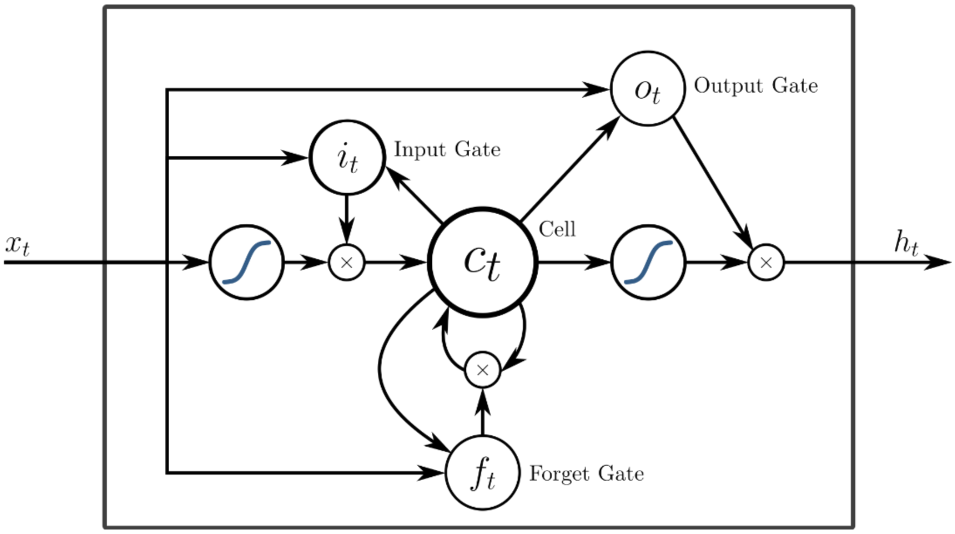

Figure 8 - Example of an LSTM unit
(wagenaartje.github.io/neataptic/docs/builtins/lstm/)

As can be seen in *Figure 8* the more commonly used LSTM unit consists of three gates, the input gate, the output gate and the forget gate.

The input and output gates control the flow of information through the network, as the names suggest the input gate control the input into the memory cell and the output gate controls the output into the rest of the network.

The forget gate is a later addition to this architecture (Cummins, Gers, & Schmidhuber, 1999) and it allows to adaptively reset the memory of the network as the context from previous information is needed.

ANN Training

Before being able to forecast efficiently, networks need to be trained. Feed-forward networks are usually trained in a supervised manner, meaning there is a training set available which contains the inputs and the expected outputs properly labeled.

When learning it is expected that the network constructs an input-output mapping(Catalão et al., 2007), adjusting the weights and biases in each iteration with the goal of minimizing the error between the produced output and the desired output. Learning consists then of a minimization process where the error is minimized until an acceptable criterion for convergence is reached.

It is crucial that the network is not trained to achieve an error of 0 during training. This means that the network has been over trained and while it succeeds tremendously with the training data-set it will have a poor ability to generalize when forecasting new data that it has never seen before (Jain, Mao, & Mohiuddin, 1996).

There are several ways to train a network, but by far the most popular learning algorithm in an EPF context is backpropagation (Aggarwal et al., 2009). In the backpropagation training algorithm, the input is passed through every layer until a final output is calculated. The output is then compared to the expected value and an error is calculated, which is then propagated back through the network adjusting the weights and biases of each node as necessary to better minimize the error. This is then repeated to every data entry of the training data set and if the network is well configured and the data set has the correct amount of information the network will learn how to predict other inputs based on the training data.

The standard backpropagation learning algorithm is a steepest descent algorithm that tries to minimize the sum of square root errors (Catalão et al., 2007). The mean squared root error that is back propagated to the input layer is commonly defined as:

$$MSE = \frac{1}{2}\sum_{i=0}^{i} \|Yi - Di\|^2$$

Yi -Real Value, *Di* -Network Output

(Jain et al., 1996)(Szkuta et al., 1999)

Support Vector Machines

Support vectors are a widely used tool that has been applied with success in several pattern recognition and classification problems but also in non-linear regression problems such as electricity price forecasting (Weron, 2014).

This algorithm can be traced back to (Cortes & Vapnik, 1995) statistical learning theory and has been since then widely popularized in research literature.

Support Vectors transforms the data into a high-dimensional space and then tries to find simple linear functions that form boundaries between the data allowing for a decision to be made.

In Figure 9 an illustration of a Support Vector Machine used for classification can be found. Initially an input space is given that has an extremely complex decision boundary. Those inputs are then transformed utilizing a Kernel, which is the function that maps lower-dimensional data into higher dimensions. The data is then separated by a hyperplane and a classification can be made, for instance points above the hyperplane belong to one class and points below it to another.

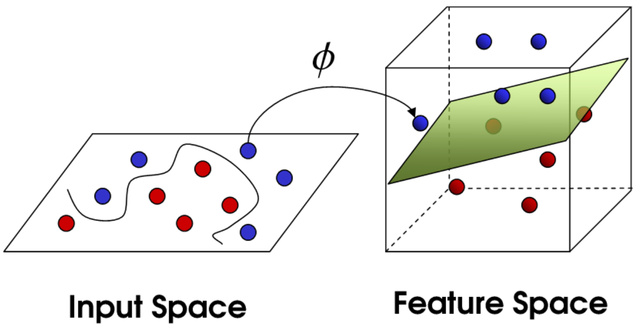

Figure 9 - Example of an SVM classifier (dataanalyticspost.com/Lexique/svm/)

For Support Vector Regression, the principle is the same as for classification as it maintains all the characteristics that are associated with the Support Vector algorithm, with only a few minor differences. The major difference between the two is obviously the output. In classification the output is the class that the features belong to, in regression the output is a real number with infinite possibilities. Also, in the case of regression, a margin of tolerance (epsilon) is defined as can be seen in *Figure 10*.

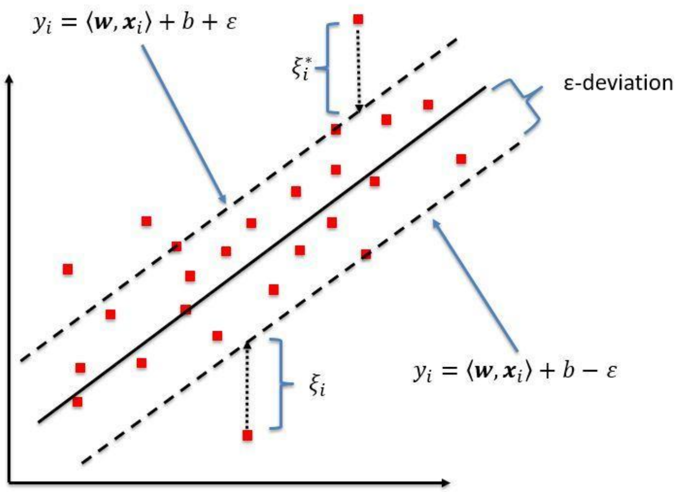

Figure 10 - Schematic of a one-dimension SVR model (Kleynhans, Montanaro, Gerace, & Kanan, 2017)

Utilizing Support Vectors is a two-step process. First, a sub-section of the data is utilized to train the algorithm and after it being trained it tries to predict the rest of the data, repeating this step as adjustments to the algorithm is necessary. In an EPF context, while not nearly as popular as ANN, SVMs have shown some promising results as can be seen in *Table 1*.

In one of the first researches utilizing SVM in EPF context, a direct comparison between MLP and SVM was made. SVM was shown to have the same forecasting accuracy as MLP while requiring less time to be trained (Sansom et al., 2003). Another recent research also showed that an SVM outperformed a simple ANN in terms of accuracy and efficiency (Razak et al., 2015).

SVM however are typically utilized in a hybrid system to achieve the best results which add complexity to the model. One of the most popular hybrid SVM models in EPF is utilizing SOM classifiers to cluster hourly electricity price and then applying an SVM for each cluster (Niu et al., 2010). Another hybrid model combining ARMAX models and least-squares SVM shows an improvement over simpler SVM models (Chaâbane, 2014).

Despite these promising results, there's not a lot of research into SVM when compared to ANN. While some specific scenarios were shown to have excellent results utilizing an SVM or a hybrid model where SVM was involved, it is not possible to conclude that generally, SVM models outperform ANN models in an EPF context because there is not enough research on the topic.

XGBoost

XGBoost is a scalable implementation of gradient boosted decision trees(Friedman, 1999), available in the form of an open-source library for multiple popular programming languages and frameworks. This system has been widely used in the past few years winning multiple machine learning competitions in websites such as Kaggle (Chen & Guestrin, 2016).

A decision tree consists of multiple decision nodes, decisions, and leaves, which represent the final prediction. The topmost node corresponds to the best-found predictor of the target variable, a simple visual representation can be found in *Figure 11*.

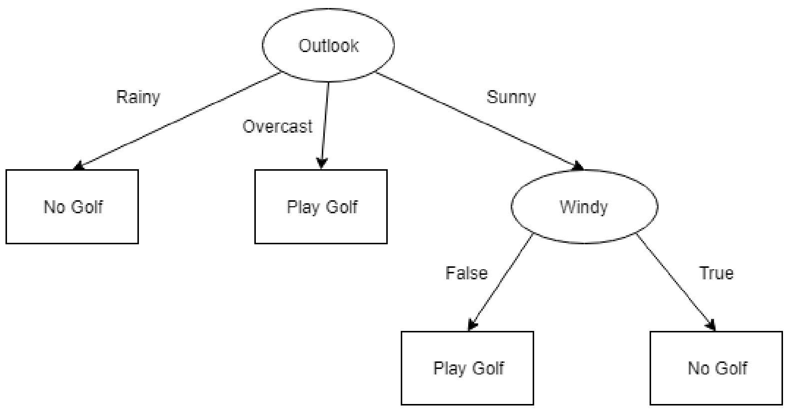

Figure 11 - Simple decision tree example

In order to improve the results of classical decision trees, ensemble methods are often used, which combine several different decision trees to improve the predictive performance of the model.

Boosting is one of the most commonly found ensemble techniques (Chen & Guestrin, 2016) and it consists of fitting consecutive decision trees and the final prediction for a given example is the sum of predictions from each tree, this can be described mathematically as:

$$Y_i = \sum_{k=1}^{K} f_k(x_i) \ , f_k \in \mathrm{F}$$

K- Number of trees

$f_k(x_i)$ – Prediction value for given example in a given tree

F- Space of all decision trees

In order for the algorithm to adjust its set of functions in each iteration, the following objective function is minimized:

$$L(\phi) = \sum_i l(Yi, Ri) + \sum_k \Omega(f_k)$$

l- Function that measures the difference between the prediction and the real value

Ω – Penalizes the complexity of the model

$$\Omega(f) = \Upsilon T + \frac{1}{2}\lambda\|w\|^2$$

T – Number of leaves

w – Leaf weights

In Gradient boosting a gradient descent algorithm is utilized to optimize the loss functions of each tree. In regression this loss function can be based on the residuals between the predicted value and the real value. This means that on each tree iteration, the objective is to improve on the residual error from the last tree, taking account information from every prior tree.

Gradient boosted decisions trees can potentially suffer from the same problems as other algorithms that utilize gradient descent, such as overfitting or local minima and those problems need to be taken into account when utilizing this algorithm.

Grid Search

As can be understood from reading the previous sub-chapters, every algorithm has many different parameters than can be tweaked in order to increase the accuracy of the predictions.

As it would be impractical to test every possible parameter in each algorithm, one possible solution is to utilize a grid search. By defining jumps in each parameter every possible combination of those parameters is tested as an individual model and then compared to the others. This can potentially be extremely time-consuming if small jumps in each parameter are defined and so it is important to understand what each parameter means and how a certain increase or decrease will possibly affect the model.

To give a practical example, in SVM one of the most important hyper-parameters is C, which, usually, takes values from anywhere between 0.1 to 100 in literature. If every value between 0.1 and 100 was to be tested in increments of 0.1 it would take impossibly long to test every combination of C. In this case it would be much more practical to take jumps by increments of 10 or even slightly higher which would compute much faster but still maintain around the same level of accuracy.

Grid Search can then be defined, in short, as a method to perform hyper-parameter optimization for a given model (Syarif, Prugel-Bennett, & Wills, 2016)

.

4- Statistical analysis

From the literature review, it is clear that electricity prices tend to present commonly found patterns on a monthly, weekly and daily levels. In this section these trends will be explored in the data utilized for this dissertation. The data utilized spans from 2015 to 2017 in Portugal and was aggregated by technology.

It is expected to find common patterns found in literature such as prices variations in colder/hotter months, afternoon/morning hours, workdays/weekends, as well as possibly identifying other patterns. This section directly relates to the data understanding phase of the CRISP-DM methodology, explained previously in chapter 3.

Additionally, it’s extremely important to understand the significance of other variables like consumption levels, renewable energy generation and energy generated from fossil fuel and their relationship with the final market price.

Calendar

In literature, it is commonly accepted that the hour is a major factor to take into consideration when trying to predict electricity prices, as obviously the needs and demands vary greatly throughout the day especially when comparing middle of the night hours when few people are awake to peak activity hours.

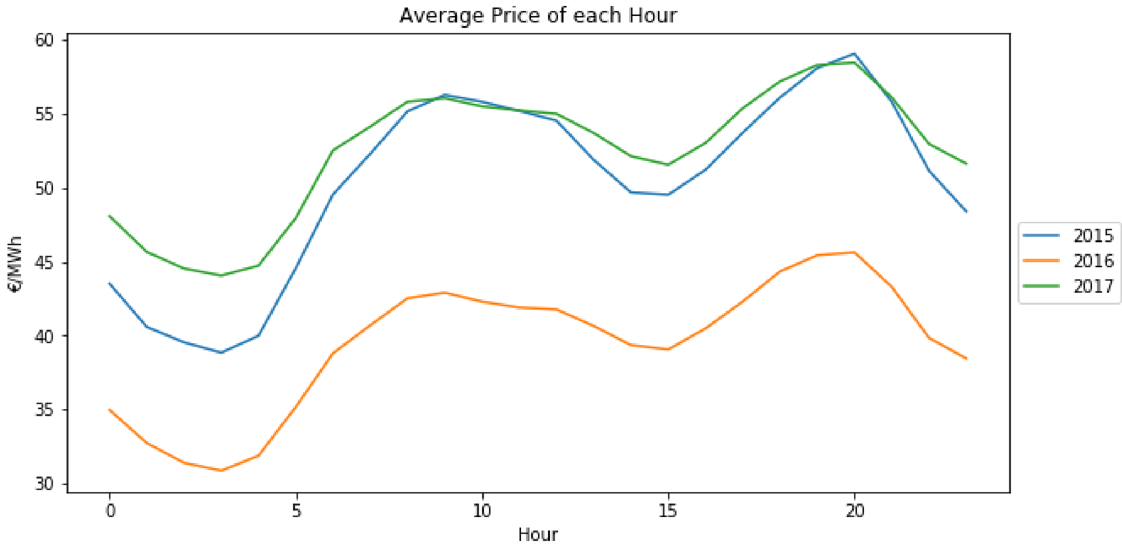

Figure 12 - Average Price of each hour in 2015,2016 and 2017

As can be seen in *Figure 12***Figure *12***, electricity price presents a very clear relationship with the different hours throughout the day. Generally, prices tend to drop during morning hours, dipping to its lowest at about 3, and during the beginning of the afternoon at about 12. On the other hand, they tend to start increasing at about 5 and 16, hitting their daily high at about 20. This pattern is clearly present in every year as the shape of the curve is about the same in all the years studied despite overall prices being higher or lower depending on the year.

Electricity pricing has also shown key differences in different days of the week mainly due to holidays, workdays and weekends and the different needs correspondent to each one.

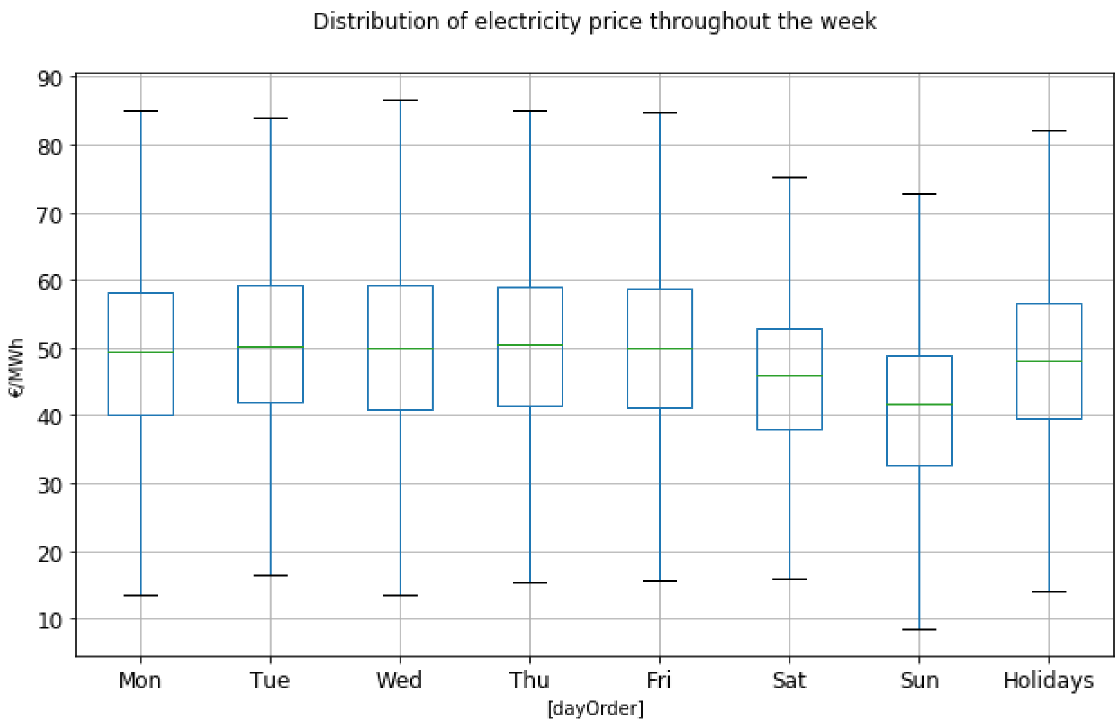

Figure 13 - Boxplot of Electricity Prices from Monday to Sunday and Holidays

Generally, as can be seen in *Figure 13*, prices tend to be higher during weekdays than when compared to weekends, with Sunday being the day that presents the lowest prices on average. Furthermore, Saturday seems to be the day where prices are more stable as it shows the least amount of standard deviation compared to all other days. During the week all 5 workdays show very similar patterns.

During holidays average prices tend to be slightly lower than prices during weekdays which suggests that during public holidays prices tend to behave more closely to Saturdays than weekdays.

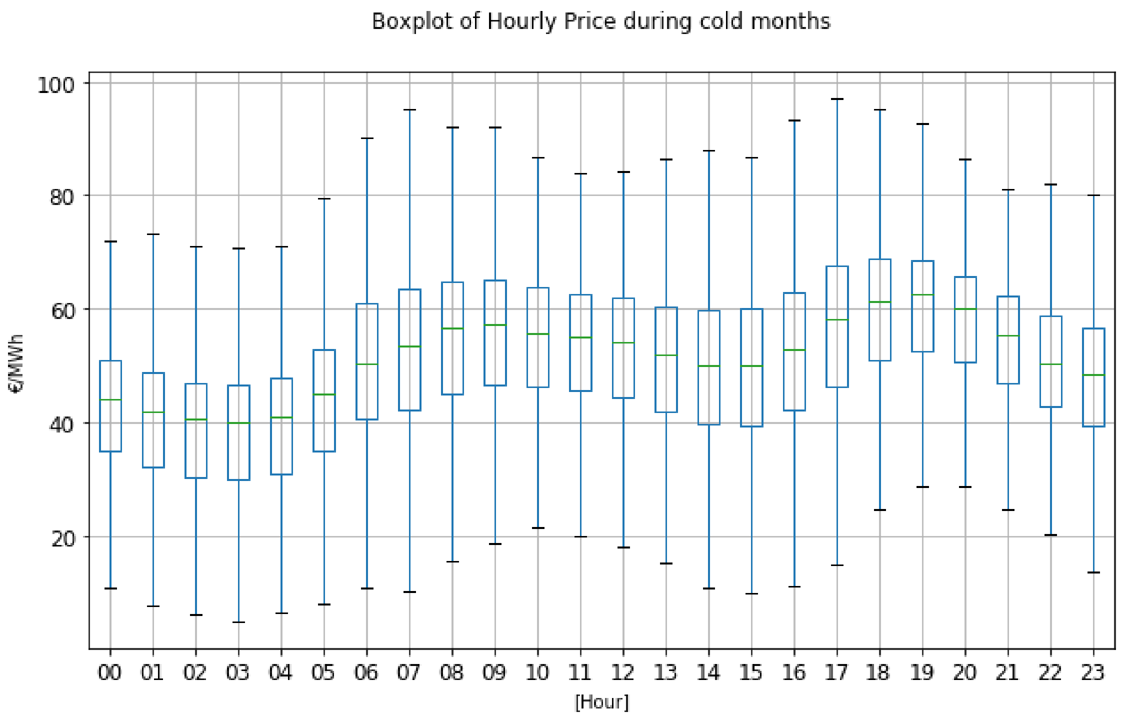

Figure 14 - Price distribution during cold months (October to March)

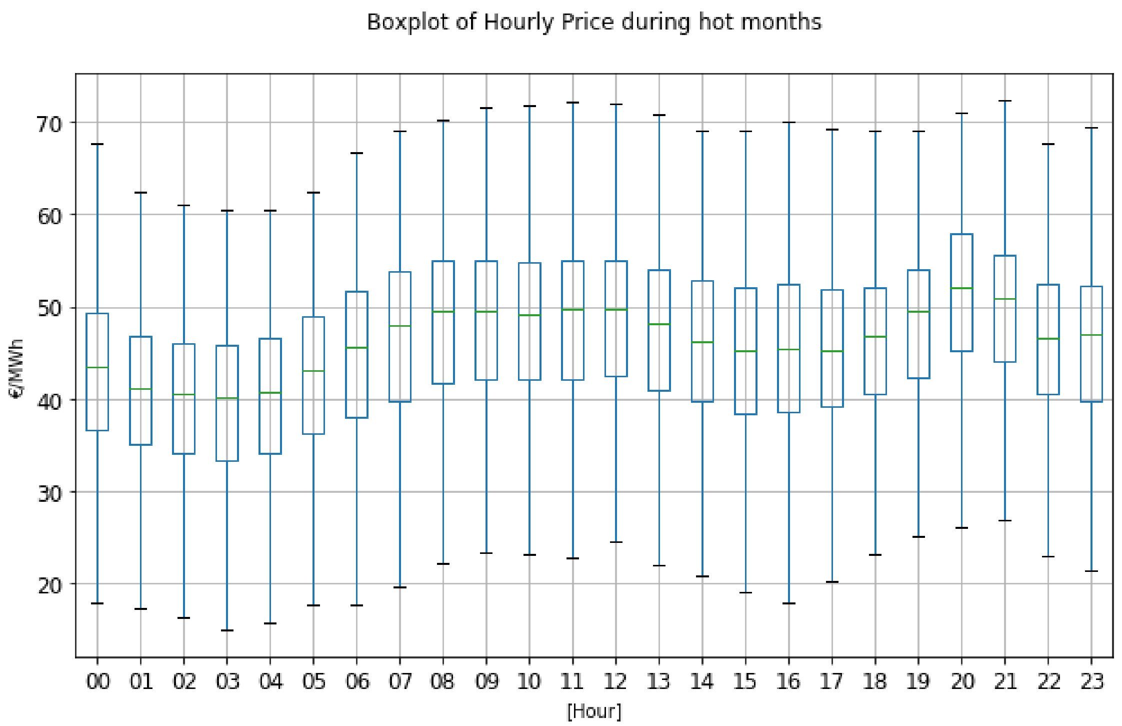

Figure 15 - Price distribution during hot months (April to September)

Daily prices also present some key differences in the different seasons of the year, as can be seen in *Figure 14* and *Figure 15*. During early morning hours it doesn't seem to be significant what season of the year it is, as prices are extremely similar in both cold and hot months. However, prices in cold months quickly rise when compared to their hotter counterpart, resulting in higher average prices throughout the rest of the day.

The overall series shape doesn't present many differences, except that hot months seem to be relatively more stable throughout the day, not presenting a lot of variation between the lowest and highest points. This further confirms that hotter months will tend to be easier to predict than colder months. In addition to this, one other relevant difference between the two series is that the highest peak price happens a few hours later (20:00) in hotter months than colder months (18-19), while the lowest point happens at the same time (at about 3).

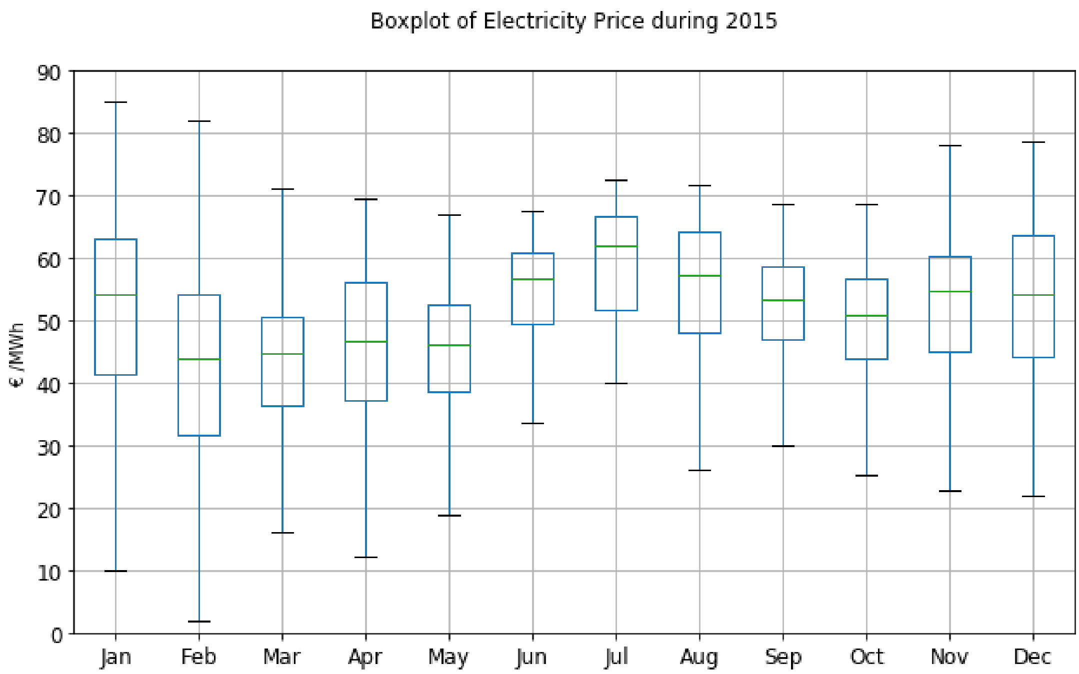

Figure 16 - Monthly distribution of Electricity Price during 2015

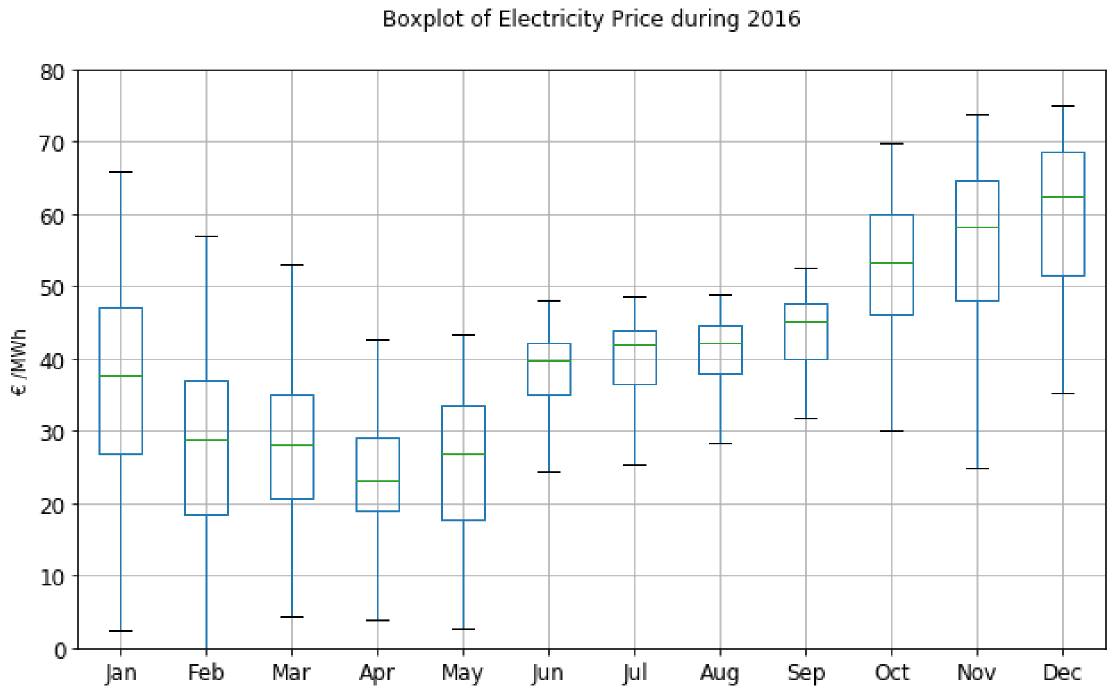

Figure 17 - Monthly distribution of Electricity Price during 2016

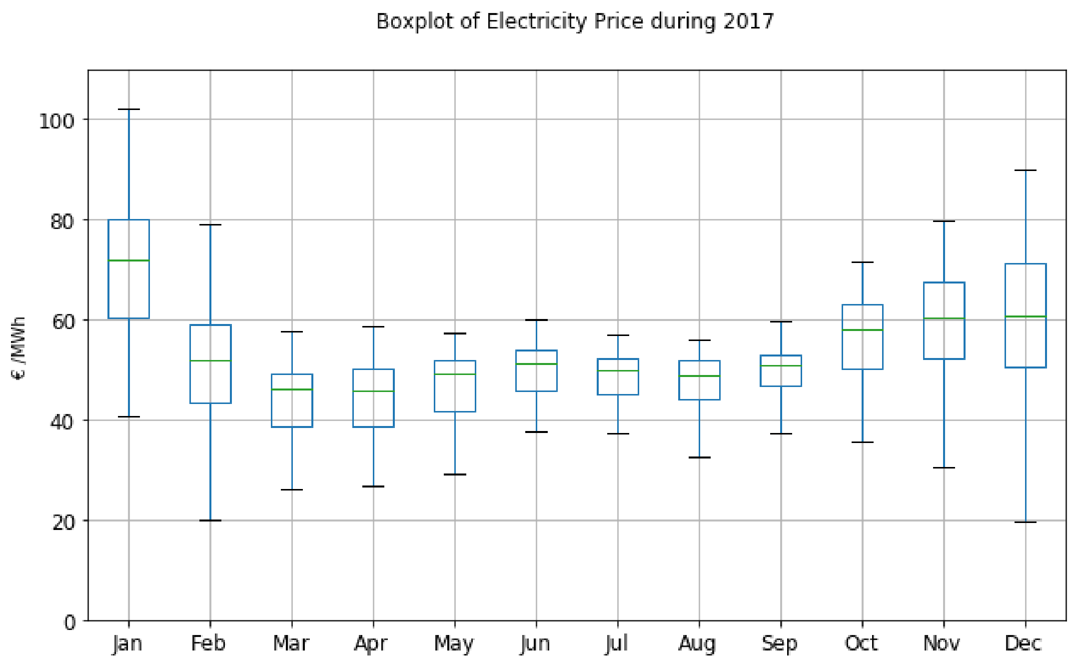

Figure 18 - Monthly distribution of Electricity Price during 2017

As can be seen in *Figure 16*, *Figure 17* and *Figure 18* prices during colder months (from October to March), tend to have a much higher standard deviation than prices during hotter months (from April to September). This effect is especially noticeable in the year of 2016, where colder months have increased volatility when compared to 2015 and 2017.

Table 2 - Mean, Standard Deviation, Minimum, and Maximum of hotter months

Year	Mean	Standard Deviation	Minimum	Maximum
2015	52.11	10.90	10.00	72.48
2016	35.30	10.43	2.79	58.00
2017	47.78	6.09	26.60	60.15

Table 3 - Mean, Standard Deviation, Minimum, and Maximum of colder months

Year	Mean	Standard Deviation	Minimum	Maximum
2015	48.71	13.25	2.00	85.05
2016	43.57	17.35	0.00	75.00
2017	57.19	13.94	8.00	101.99

From the analysis of *Table 2* and *Table 3* it is possible to verify in greater detail that as generally found in literature, colder months have a highest average price than hotter months despite not always having the highest average price, as can be seen in 2015 where hotter months had a higher average price than colder months.

Additionally, it seems correct to assume that hotter months present a much more stable price series when compared to colder months, as every standard deviation was smaller in hotter months when compared to their colder counterpart.

Furthermore, in every year studied, the highest and lowest peak of the price series are always present in the coldest months, this naturally results in a higher difficulty in predicting the price series during cold months than hotter months due to the increased volatility shown in the data. This means that seasonality through the year is an extremely important factor that has a direct and easily identifiable relationship with how and when the price will shift throughout the months and it needs to be taken into account when predicting electricity price series.

Consumption

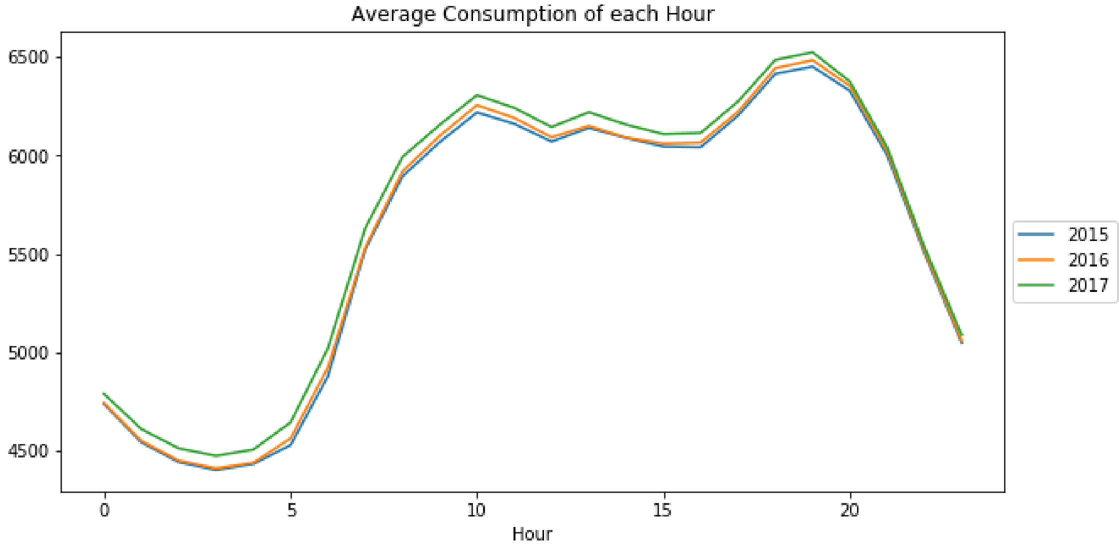

Figure 19 - Average Consumption of each Hour

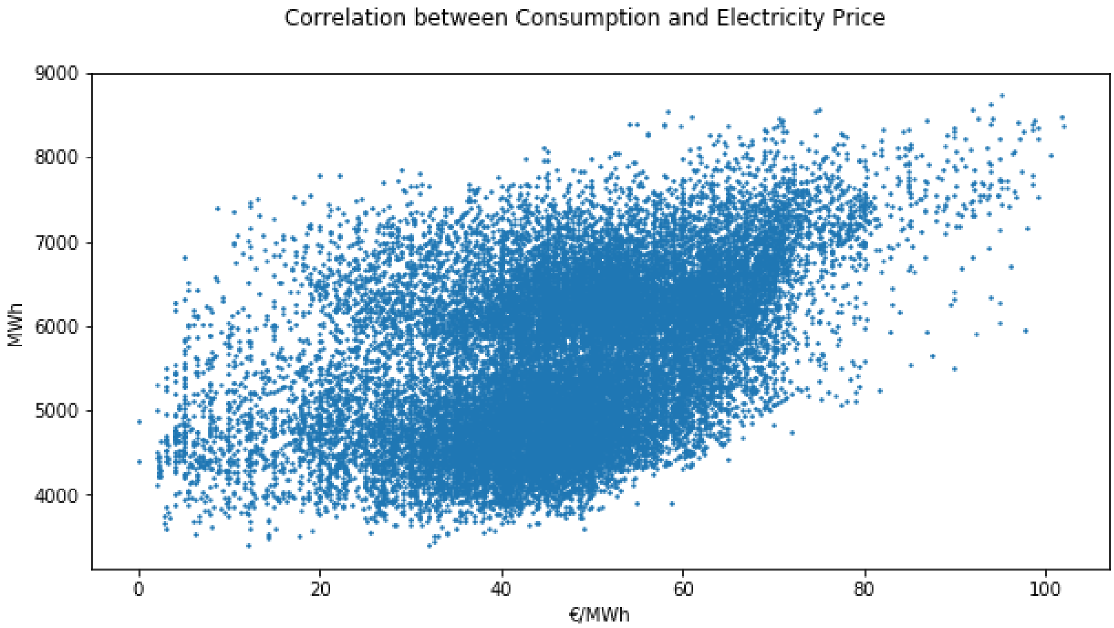

Figure 20- Correlation between Consumption and Electricity prices

Analyzing *Figure 19* it's clear that the overall shape of the series is about the same as the price series in *Figure 12*, with the key difference that consumption levels don't seem to vary at all between each year while prices, especially in 2016, varied a lot. This is probably a good indicator that consumption levels don't have a strong correlation with the final market price, at least not directly. By analyzing *Figure 20* where it can be seen that a low price does not necessarily mean low consumption levels and vice-versa the previous assumption is further confirmed. The correlation between these two variables is 0.47 which is to be expected given the previous explanation.

In *Figure 21* the hourly distribution can be seen in further detail via a *boxplot*, and it becomes even clearer that consumption levels present a very strong correlation with the calendar variables, dipping to low levels in morning hours and generally rising to high levels during more active hours.

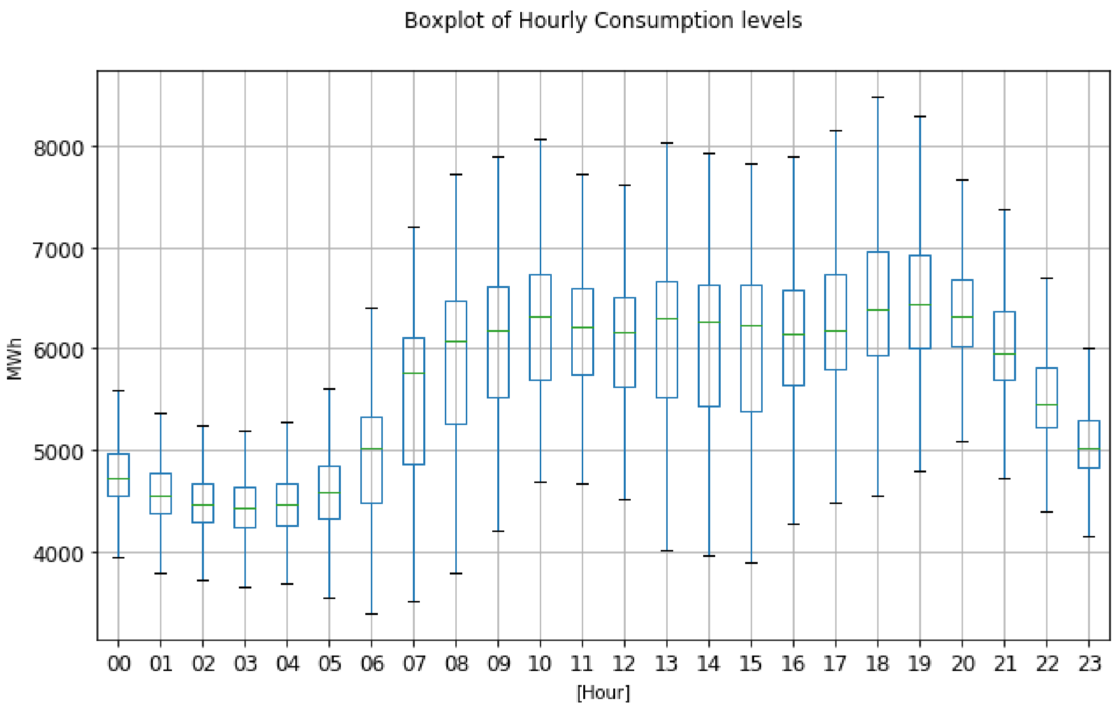

Figure 21 - Distribution of Hourly Consumption levels

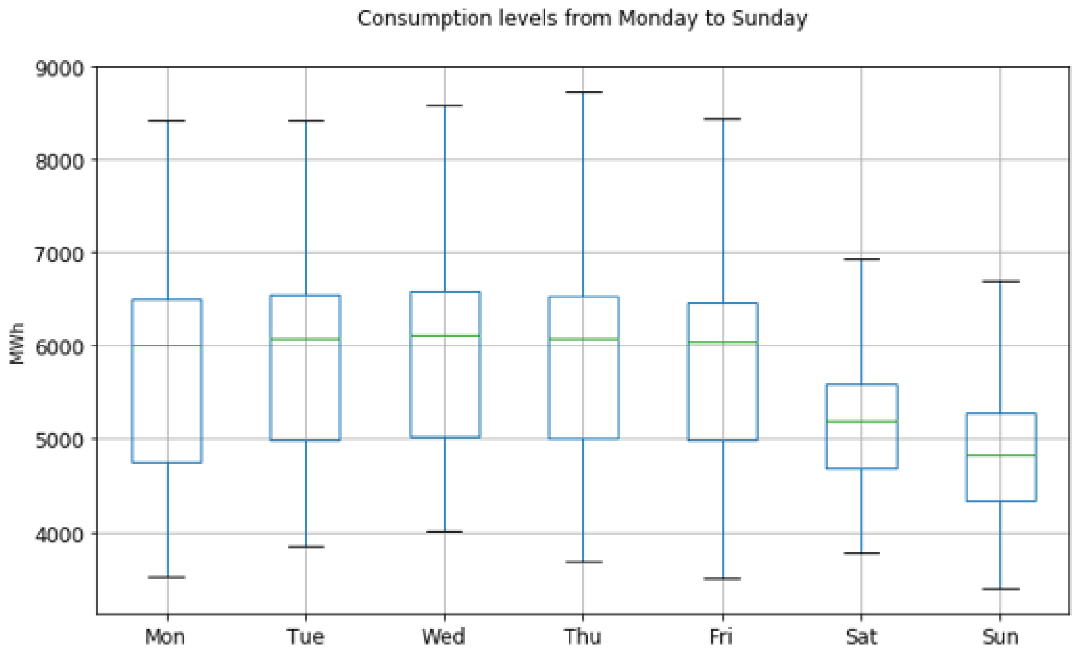

Figure 22 -Distribution of Consumption from Monday to Sunday

Analyzing *Figure 22* the correlation between the period of time and consumption levels becomes even more evident as it follows the exact same pattern as electricity market prices in *Figure 12*, being stable through the week with not much difference between each work-day and then dropping on the weekends, Sunday being the day where it is at its lowest.

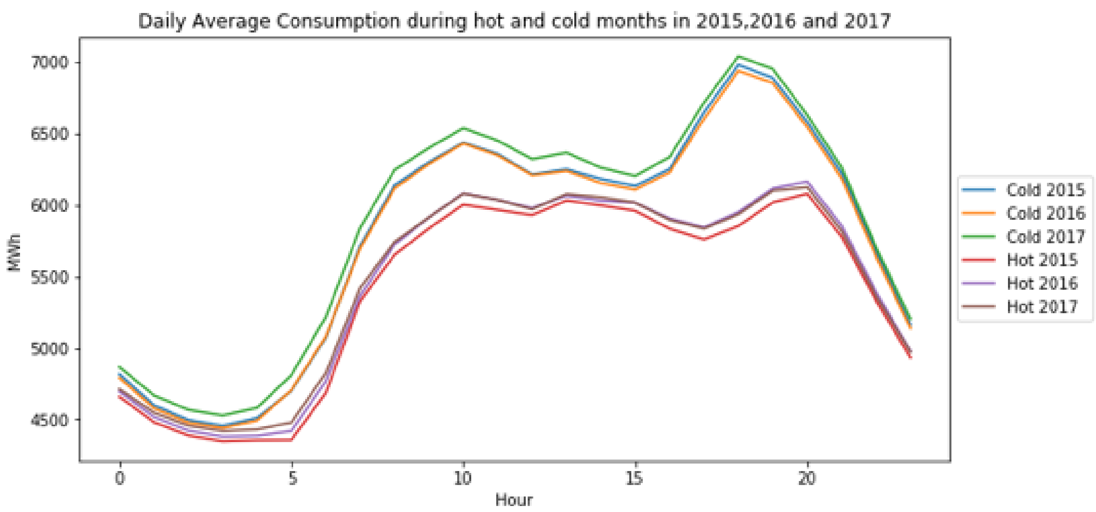

Figure 23 - Daily Average Consumption during hot and cold months

Figure 23 shows the average level of consumption during each year for hot and cold months, and it is clear that despite the year all consumption levels are extremely similar as long as it's the same season. This is contrary to the pattern shown by electricity prices, shown in *Table 2* and *Table 3*, where the average price in 2015 during hot months was higher than the average price during cold months, and while there were significant differences between the average price of each year and each season, those differences are not reflected in the consumption levels.

It is clear then that both consumption level and electricity prices are strongly connected to the calendar variables such as the current hour of the day or the weekday but knowing the expected consumption level at a certain hour is not enough to be able to accurately predict the price, as lower consumption levels can still potentially hit high price points and vice-versa. It seems then extremely important to understand the source of the energy that was generated at each point in time and not just the overall energy expended.

Renewable Energy Sources

Hydropower

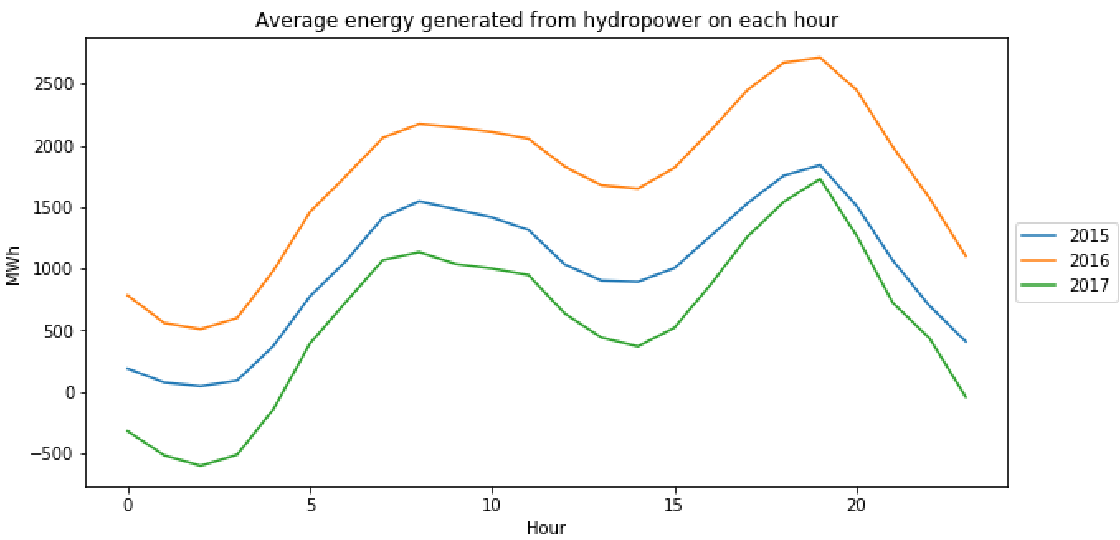

Figure 24 - Average energy generated from hydropower sources on each hour

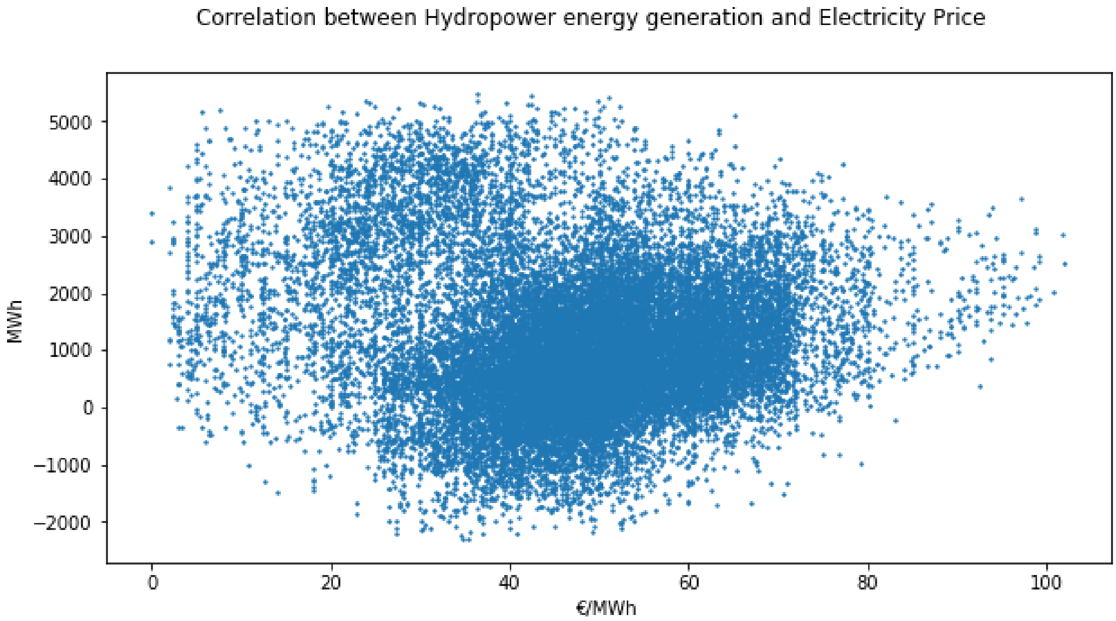

Figure 25 - Correlation between Energy generated by hydropower and electricity price

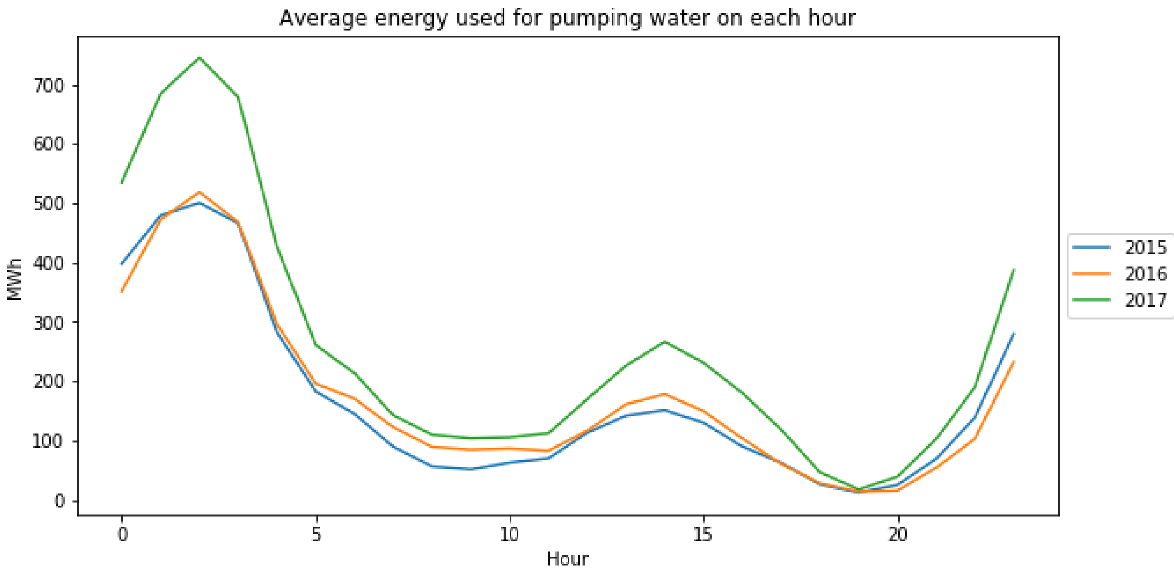

Figure 26 - Average energy used for pumping water on each hour

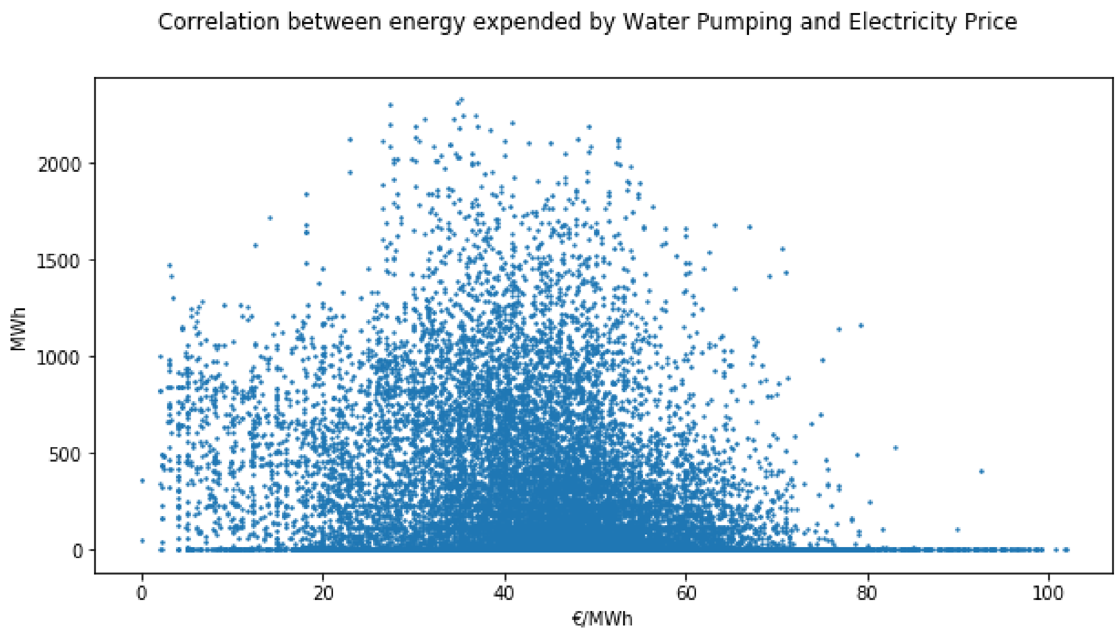

Figure 27 - Correlation between Water Pumping and Electricity Price

Figure 24 represents the added values of hydropower energy generated by hydropower plants with no reservoirs, reservoirs and mini-hydropower plants, which can generate up to 10 MW/h.

Mini hydropower plants have the lowest contribution to the overall generation, as both reservoirs and hydropower plants generate about 10x more energy each, but it's still significant enough to include in the data analysis.

Negative energy occurs when plants utilize energy to store water for generating energy at a later time. This usually occurs at night due to prices being cheaper, as can be seen in *Figure 26.* This means that a high level of consumption for water pumping is usually related to cheaper prices. *Figure 27* confirms this assumption, as it shows that starting at about the 50€/MWh mark there is much less consumption being utilized for water pumping when compared to lower price points. The lower end of the price range is lot more distributed, meaning that higher levels of water pumping are probably a good indicator to identify when prices are going to be low, as it is likely that water is pumped during periods where price is low.

Water pumping is not however a good predictor for the exact value of price, as for example a consumption level of 0 can be related to any price between 0€/MWh to 100€/MWh.

The curves in *Figure 24* tend to have the same pattern as the consumption curves in *Figure 19* as it would be expected that more energy is generated on peak hours as opposed to less active hours.

More interestingly, 2016 had a significantly higher hydropower energy generation than both 2015 and 2017 which can partially explain why that year had a significantly lower overall price, as energy from renewable sources is cheaper than fossil fuels. It is still important to note however that electricity prices in 2015 and 2017 are very similar. While 2017 did have the highest average price overall, the difference is not as significant as *Figure 24* would imply, as there is even a point where the average price in 2015 surpasses 2017, which never happens for hydropower energy generation. This means that while hydropower energy generation does seem to have a big impact on the price, it is not the only variable that needs to be considered.

Finally, looking at *Figure 25* it's possible to see that when energy generated from hydraulic sources is at its highest, the price is always in the lower range and at the very low price range, between 0€/MWh and 20€/MWh negative energy almost never occurs which seems to further indicates that high energy generated from hydraulic sources is a good contributor to bringing the final market price down.

Wind Energy

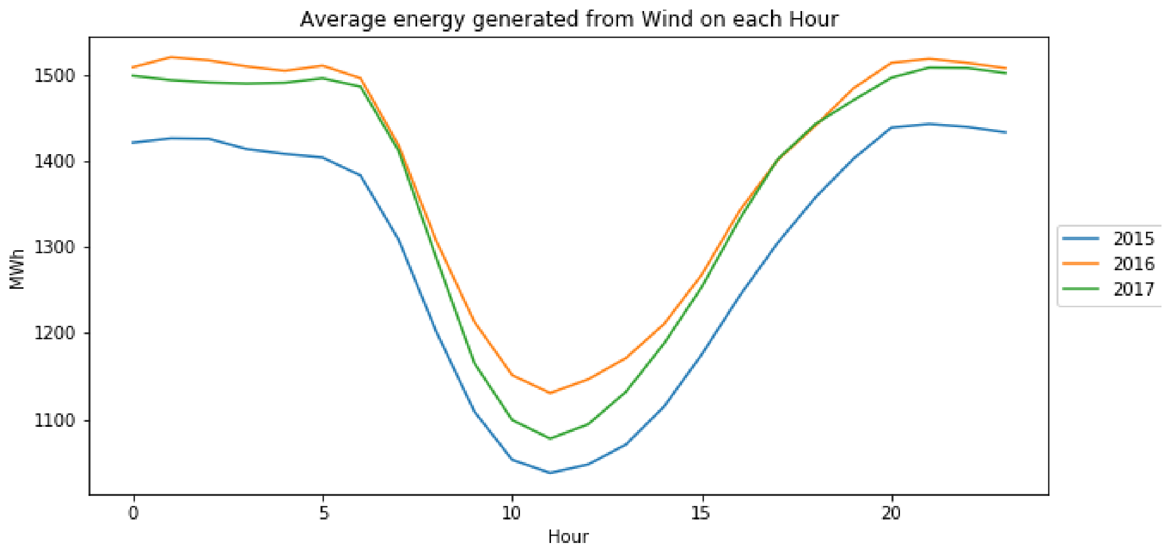

Figure 28 - Average Energy Generated from Wind on each hour

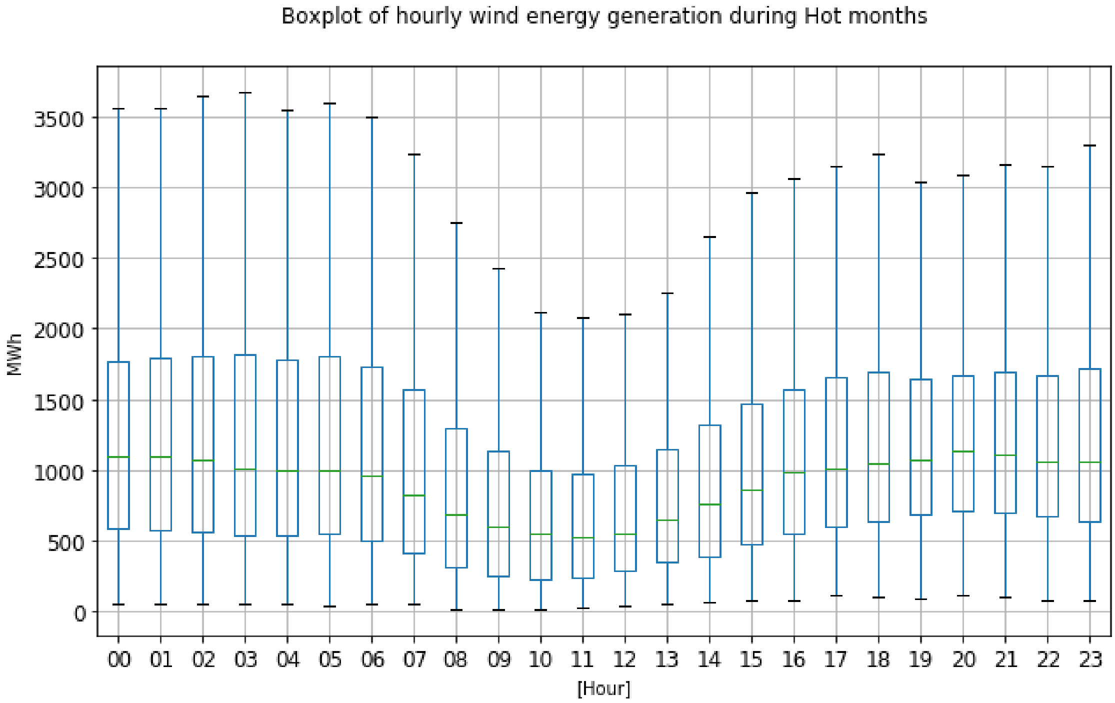

Figure 29 - Distribution of hourly Wind energy generation during Hot months

.

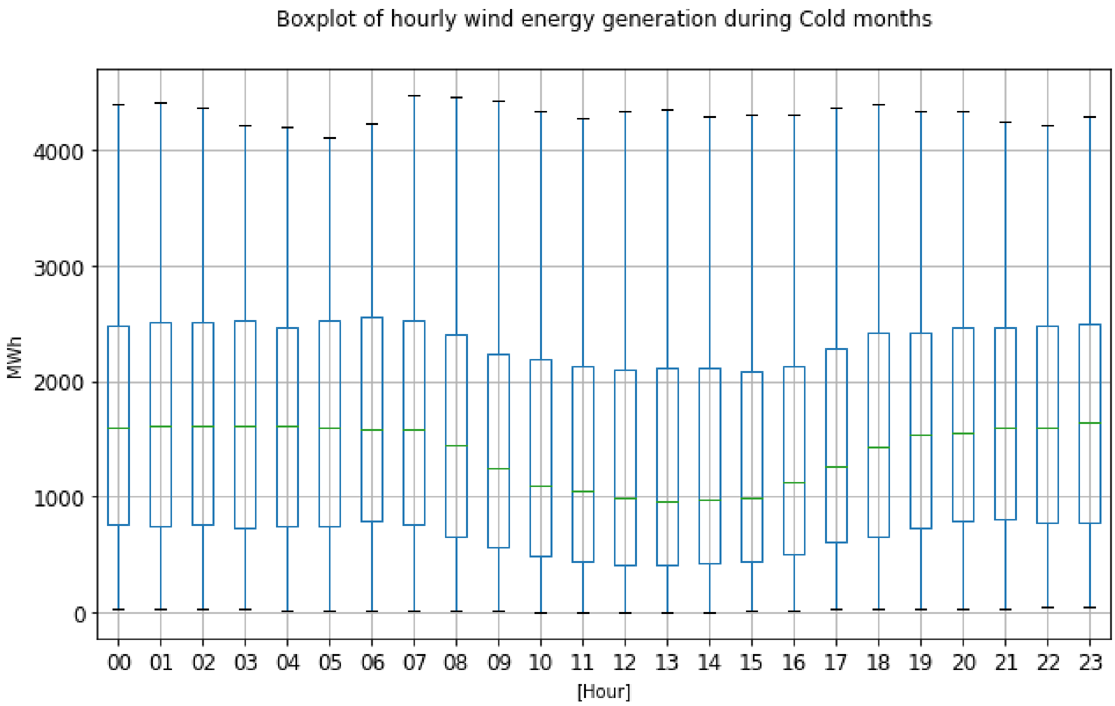

Figure 30 - Distribution of hourly Wind energy generation during cold months

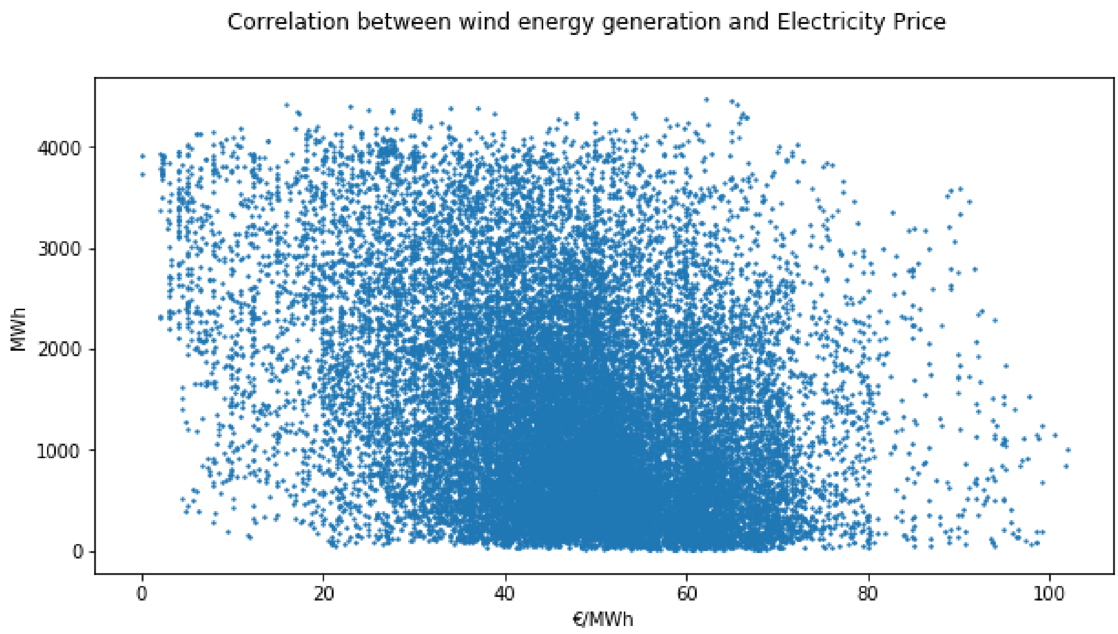

Figure 31 - Correlation between Wind Energy Generation and Electricity Price

Figure 28 shows the energy generated from wind turbines, similarly to *Figure 24*, 2016 was the year with the highest amount of energy generated, although the difference is not as significant as energy generated from hydraulic sources to other years, it is an added reason for significantly cheaper overall prices in 2016.

Additionally, the energy generated in 2017 was slightly higher than in 2015, which is the opposite of energy generated from hydraulic sources. This is also a good indicator of the importance of these two variables, as both price curves ended up being similar in *Figure 12*. However, the highest difference of the two curves at any point is only about 100MWh, which is a lot lower than the highest difference in *Figure 24* which is about 500MWh. This means that these two variables are not enough to explain the final market price, as even adding wind energy to hydropower energy, 2015 still presents a consistently higher overall energy generation than 2017.

It is also important to note that the shape of the curve is completely different to the consumption levels presented in *Figure 19*. The lowest point of energy generation in *Figure 28* occurs in the middle of the day, while the highest points occur in the middle of the night, this is the opposite behavior of the other patterns analyzed so far. One possible explanation for this is that Eolic turbines are usually installed on mountain ridges, in these areas winds tend to be stronger at night than during the day. This effect is especially noticeable during the summer, and as can be seen by comparing *Figure 29* to *Figure 30*. During hot months Wind energy generation during the day has less standard deviation than during the night making the energy generation consistently lower, while during cold months the difference between night and day is not as noticeable.

It is then concluded that when Wind energy generation is high, it's extremely unlikely for prices to be high as well as can be seen in *Figure 31*, as most high peaks of energy generation will occur during the night, where price is usually much cheaper, this being especially true during the summer.

Biomass

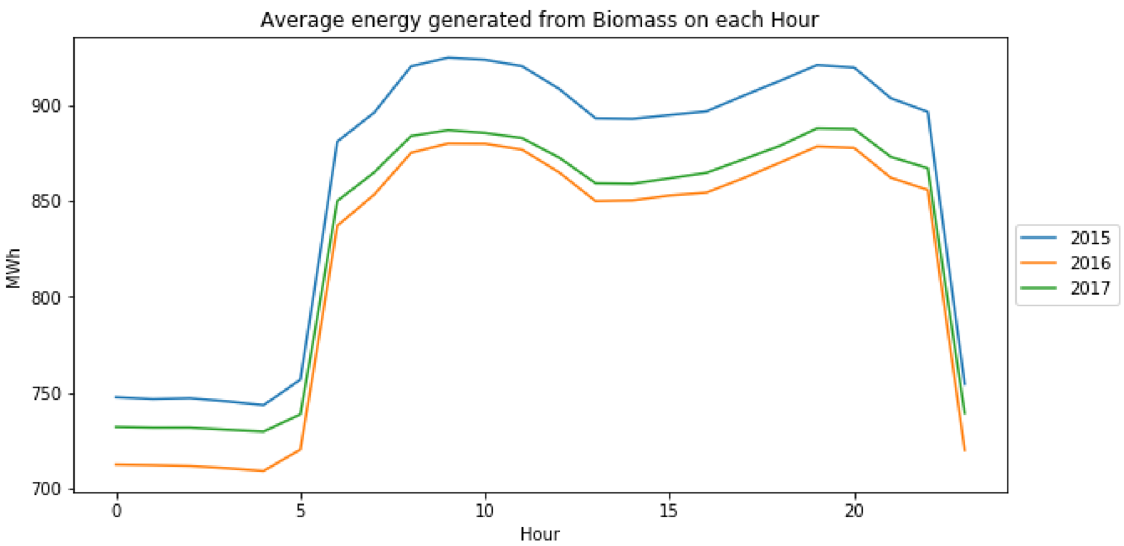

Figure 32 - Average energy generated from Biomass on each hour

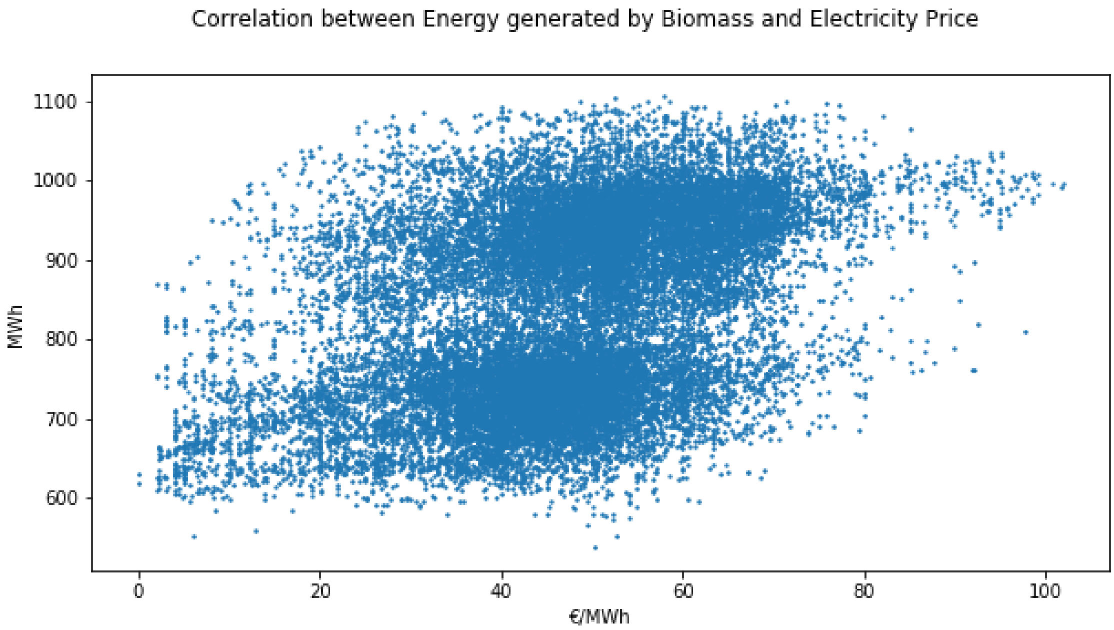

Figure 33 - Correlation between Biomass Energy Generation and Electricity Price

Figure 32 shows the average energy generated hourly by biomass throughout the three different years being analyzed. As opposed to the previous variables analyzed, all 3 years are very similar, as the point of highest difference is only about 50MWh higher in 2015 when compared to the following years. Additionally, the overall energy contribution from biomass is shown to be lower than the energy generated from both water and wind power, as the highest average point, is only about 900MWh.

In relation to the final market price, it's possible to see in *Figure 33* that it is very rare for prices to be extremely low when energy generated from biomass is high. This can be explained by the lowest price points occurring generally during early morning hours, where energy generated from biomass is at the lowest as those are generally the lowest points of expected consumption.

On the other end of the graphic, however, it can be seen that there are multiple points where both the price and energy generated is very high. This opposes what was previously stated throughout this chapter, as an increase in generation from renewable energies sources should bring the price down. While it is possible to see some slight correlation with less frequency of prices over 70 €/MWh and high energy generation levels it is not enough to state that a high energy generation from biomass will result in a lower price, especially as prices above 90€/MWh, occur multiple times with biomass energy generation being almost at its maximum. This is a good indicator that biomass probably has a small, but not irrelevant, effect on the final market price.

Solar Energy

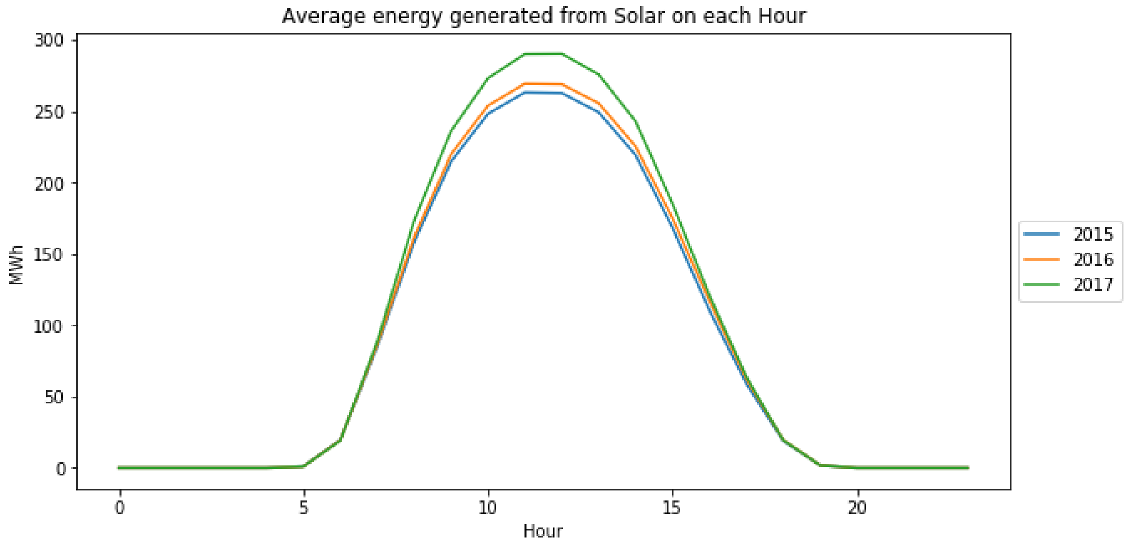

Figure 34 - Average energy generated from Solar Panels on each hour in 2015,2016 and 2017

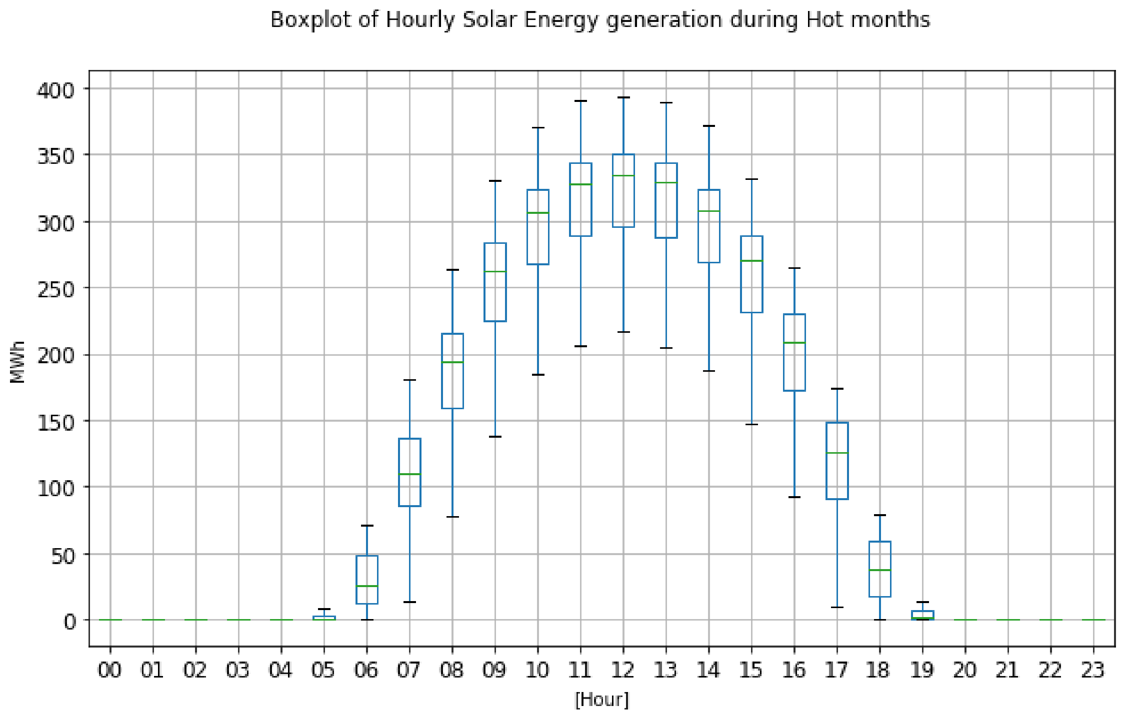

Figure 35 - Distribution of hourly Solar energy generation during Hot months

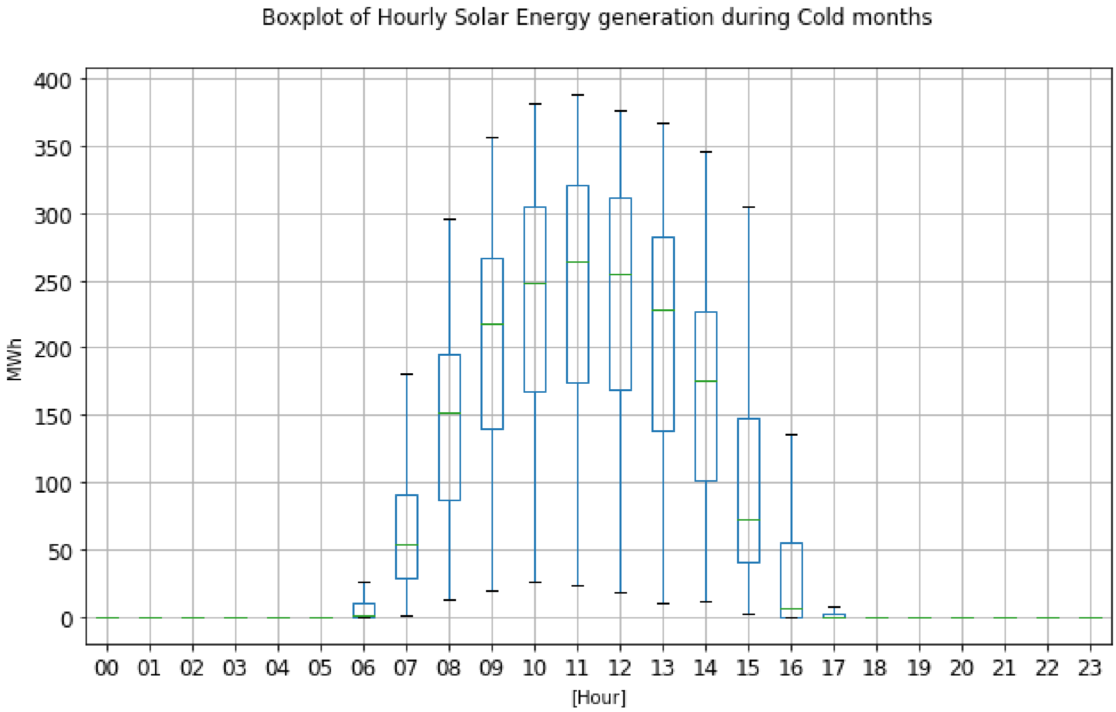

Figure 36 - Distribution of hourly Solar energy generation during Cold months

Solar energy presents the overall lowest contribution from renewable energy sources as can be seen in *Figure 34*, as the highest point in the graph is only about 280MW/h, which is much lower than the other variables analyzed so far.

It is possible to see a slight increase every year, as 2015 was the year with the lowest average energy generation, followed by 2016 and 2017 being the year with the highest. This is possibly due to an increasing number of solar panels generating energy, but the difference does not seem significant enough to affect the overall final price, as it is only a minor increase on a very low overall energy generation.

Figure 35 and *Figure 36* show the seasonal difference in solar energy generation. As expected, hotter months tend to have a more significant contribution to the overall energy generation than colder months. This is not only due to the energy generated from solar panels being consistently higher in hotter months but also to the number of hours where energy is generated being higher as well, as in cold months energy generation from solar sources stops at around 17:00 while in hotter months it only stops at 19:00 due to the increased hours of sun exposure.

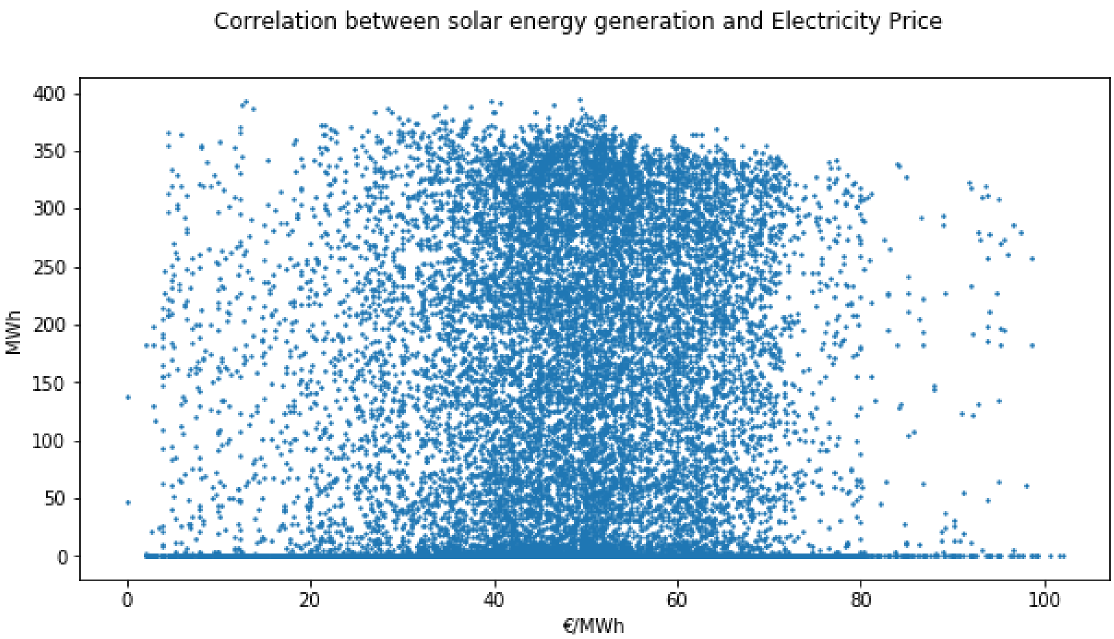

Figure 37 - Correlation between Solar energy generation and electricity price

Due to the overall low contribution of solar energy, it is unlikely that it has a significant impact on the final market price, and as can be seen in *Figure 37*, there doesn't seem to be any significant correlation present between the two variables, as both high and low prices occur with high and low solar energy generation.

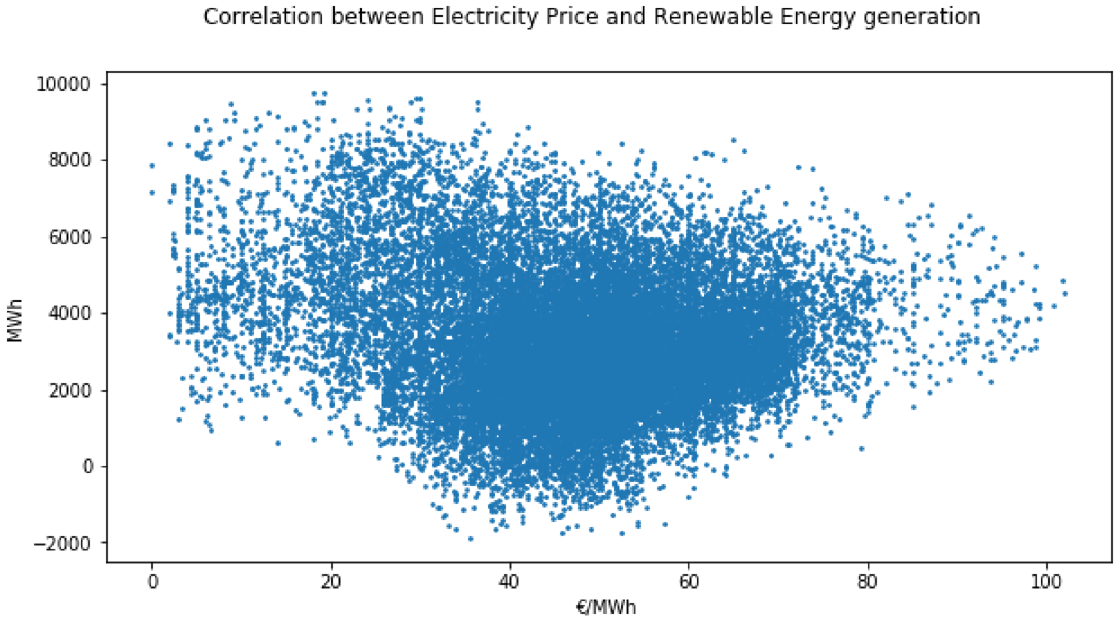

Figure 38 - Correlation between Electricity Price and Renewable energy generation

In *Figure 38* it’s possible to see the added values of all renewable energy sources analyzed in this section and its relation to the final market price. It is verifiable that very high energy generation from renewable energy sources always results in very low prices. This is due to the fact that marginal cost of electricity from RES has lower marginal cost than that of non-RES electricity generation and MIBEL follows a marginal price model.

On the other end of the graph, it is possible to see that there exist some points in time where both energy and prices are high. Comparing this data to *Figure 20* , it can be seen that in the higher price range there exist several points in time where total consumption exceeds 7000 MWh, which never happens in that price range for renewable energy generation. This means that when consumptions levels are unusually demanding, energy from renewable energy sources is not enough to fulfill the total requirements of the market. As non-renewable energy needs to be utilized to fulfill the market demands the price goes up, this will be further explored in the next sub-section.

Finally, around the middle price range in *Figure 38* it’s possible to see multiple points where energy generation is at a negative point, which can be directly related to *Figure 26* and *Figure 27*, as energy is sometimes used to pump water for future use, usually during night hours when consumption is low.

Non-Renewable Energy Sources

In MIBEL, non-renewable energy comes in the form of coal and natural gas.

Coal

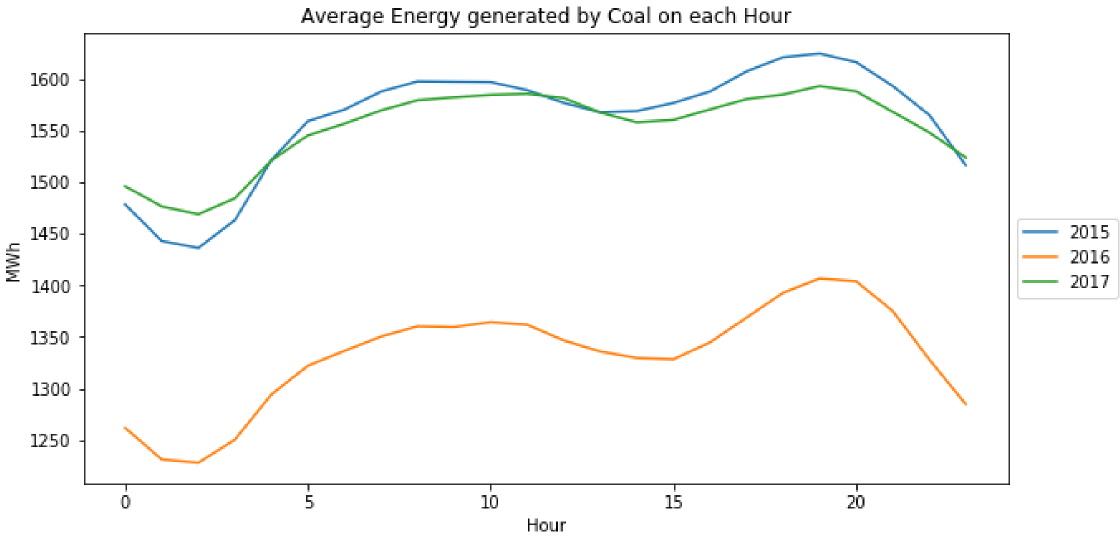

Figure 39 - Average energy generated by Coal on each Hour during 2015, 2016 and 2017

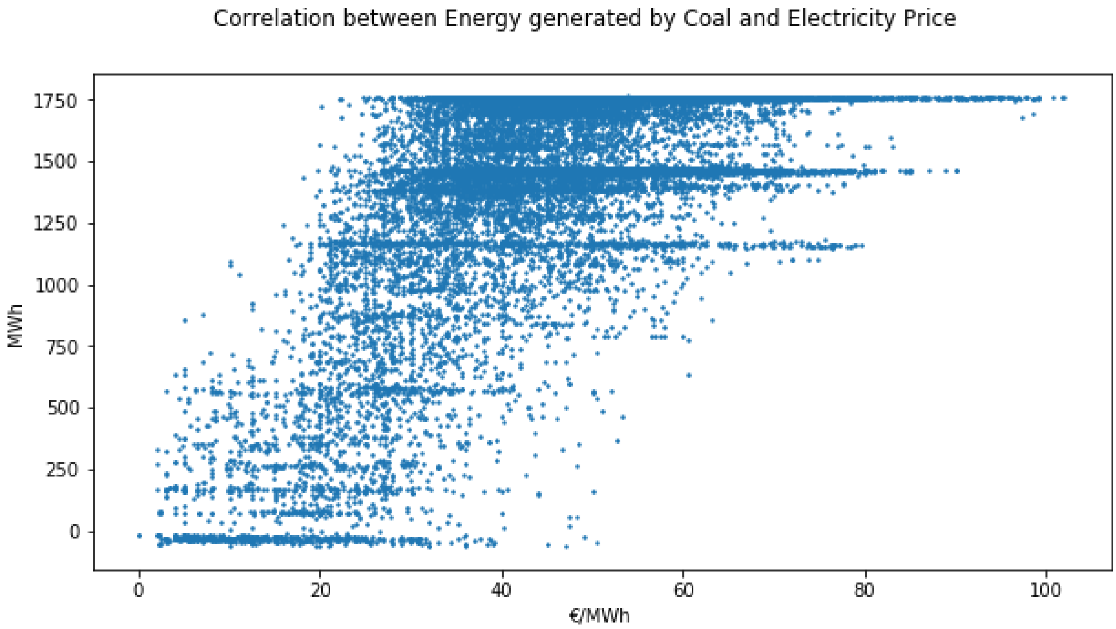

Figure 40 -Correlation between energy generated by Coal and electricity price

Coal is an expensive form of non-renewable energy, and so it is expected that the more coal is utilized to generate energy the higher the final market price will tend to be.

Analyzing *Figure 39*, it's possible to see that the energy generated by coal has very similar curve shapes to the final market price shown in *Figure 12*, as both coal energy and final market price were significantly lower in 2016 when compared to the other two years, showing potential for a strong correlation of these two variables. Additionally, energy generated by coal clearly has a strong relationship with renewable energies.

The more energy generated by renewable sources, the less energy is needed from non-renewable sources to fulfill demands of consumption. Coal was much less used in 2016 possibly due to this year having a greater contribution from renewable energies. hydropower energy was much higher in that year as can be seen in *Figure 24*, which reduced the need of coal which caused the average price to be lower.

In *Figure 40* it is possible to confirm the strong correlation between the usage of coal to generate energy and the final market price, as low prices occur exclusively when energy fueled by coal is very low, and high prices only occur when energy generated by coal is high, making energy generated by coal on a given hour a good predictor to the final market price. The calculated correlation between these two variables is 0.63 which is a further indicator that coal generation is a good variable to take into account for the predictions models.

Natural Gas

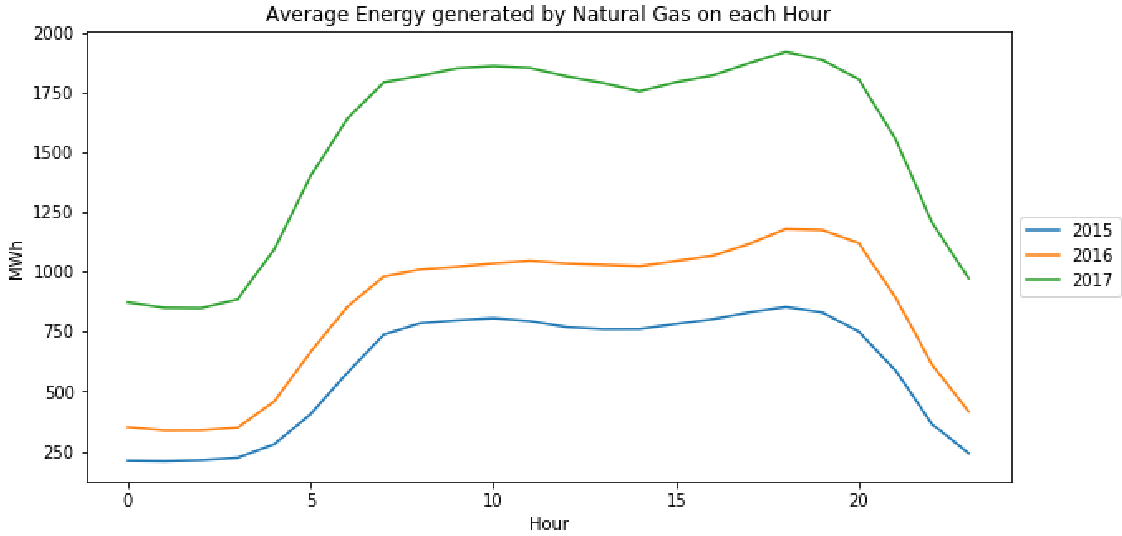

Figure 41 - Average energy generated by natural gas on each hour

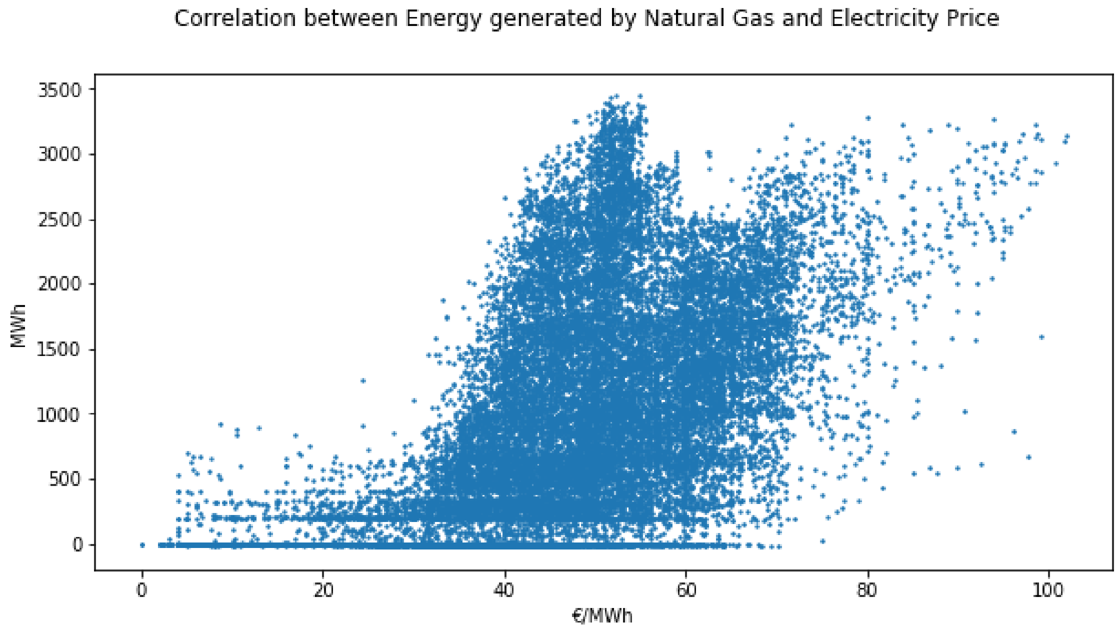

Figure 42 - Correlation between energy generated by natural gas and electricity price

Natural gas is the other relevant source of non-renewable energy in the MIBEL, as can be seen in *Figure 41* this type of energy has been the source of considerable investment, as the average energy generated from natural gas increases every year, especially comparing 2016 to 2017.

Again in *Figure 42*, it's possible to see the correlation between low energy generated from non-renewable sources and low prices, as these only occur when the energy generated from natural gas is below 1000MWh. On the other end, it's possible to see that high prices generally occur when the energy generated from natural gas is high, which in turn means that the energy from renewable sources will be lower than usual which naturally causes prices to go up.

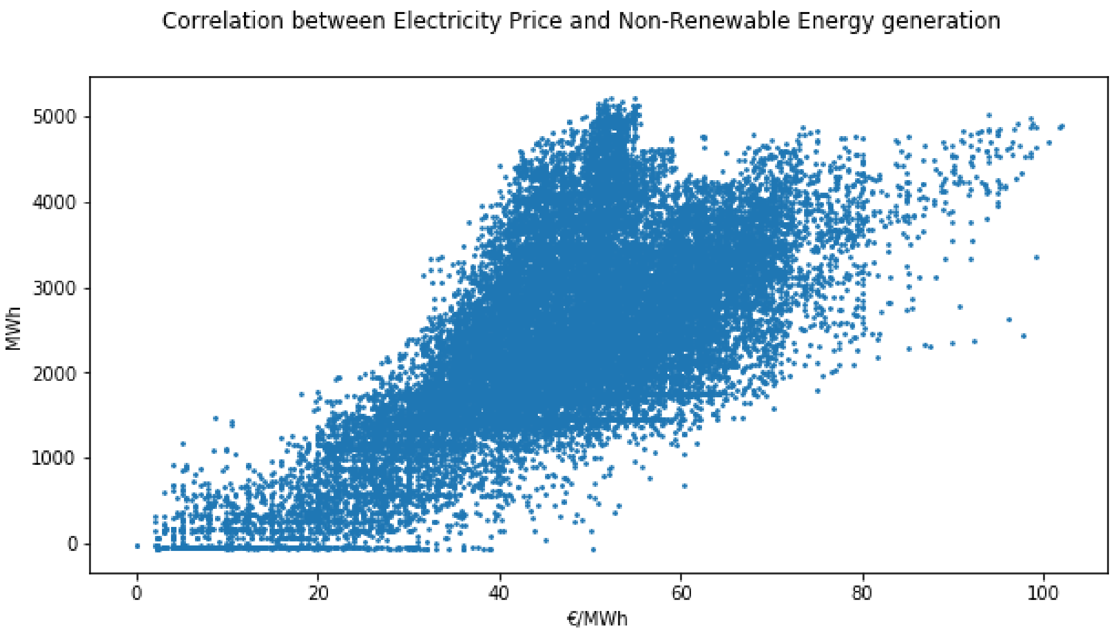

Figure 43 - Correlation between Electricity Price and Non-Renewable Energy generation

Figure 43 shows the added values of coal and natural gas and its relationship with the final price. As it is expected due to the marginal price model of the day-ahead market in MIBEL, the graphic shows a positive correlation between energy generation and electricity price. This is because non-renewable energy enters the market with higher prices than renewable energy, so the more non-renewable energy it is utilized to fulfill the market requirements the higher the price usually is.

As can be seen and as expected from the previous analysis of *Figure 38* in the low-price range there is very little requirement from non-renewable energy sources to fulfill market needs which naturally brings the price down.

On the high price range, it's possible to verify the assumption that high prices always require high energy generation from non-renewable energy sources. This also helps to explain why in *Figure 38* there are some points where energy from renewable sources is high in the high price range, as in *Figure 20* in that price range there are some points with an exceptionally high consumption levels which need both renewable and non-renewable energy sources to be fulfilled which brings the price up.

In the middle price range, it's possible to see both high and low levels of energy generated which also happens in *Figure 38* which probably indicates that non-renewable energy is highly dependent on total consumption and the total amount of energy that can be satisfied by renewable sources, as these usually enter the market with lower prices and as such have priority over non-renewable sources.

Weather

From the analysis of the previous sub-chapter, it's possible to conclude that the most contributing renewable energy sources are Wind and Water energy. Both these sources of energy are highly dependent on the weather, as Wind energy is only generated if there exist strong winds in the area where Wind turbines are installed, and hydropower energy is potentially higher as more water exists from raining.

Temperature or solar radiation could also be important factors to consider but as could be seen in the previous sub-chapter, energy from solar sources contributes much less than most other sources of renewable energy and so these variables are not expected to have a strong effect on the predictive capacity of the models due to them affecting a variable that does not contribute much to the overall energy generation.

In this subchapter, the relationship between real weather data and the energy generated from renewable energy at those points in time will be explored in order to evaluate how well the available weather data can be ultimately related to the final market price.

Wind

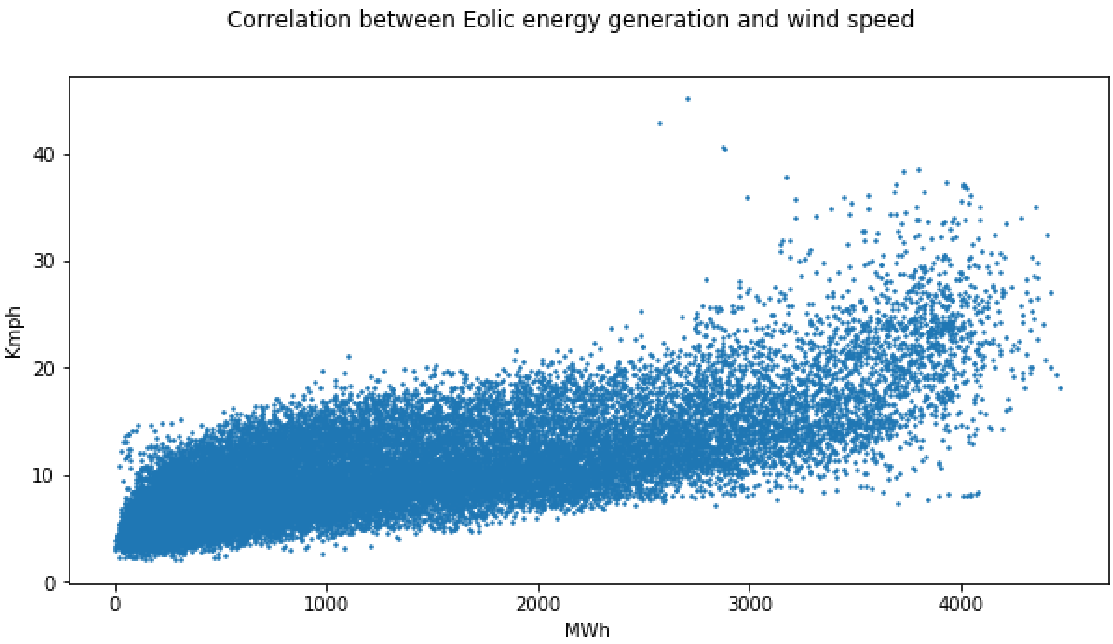

Figure 44 - Correlation between Wind energy generation and wind speed

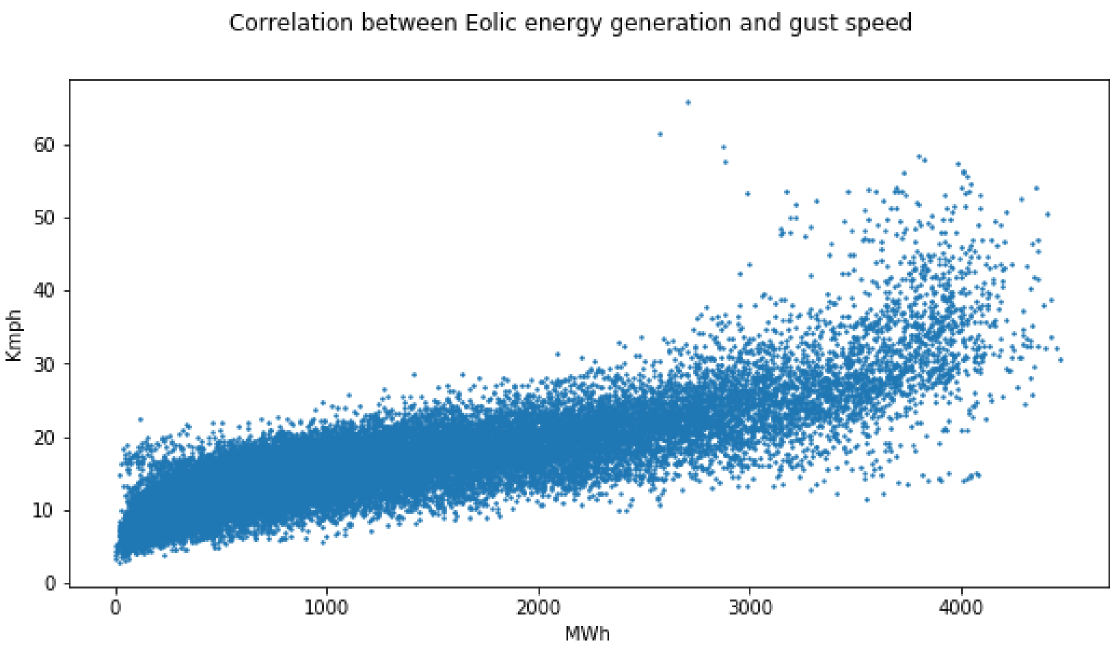

Figure 45 - Correlation between Wind energy generation and gust speed

Data from wind was generated considering the contribution of each district of Portugal to the overall Wind energy contribution and then averaging each district weather variables with that value in consideration. As such the highest contributing districts like Viseu will have a higher impact to the final value as opposed to less contributing districts like Santarem.

Data from wind comes in two different forms, wind speed, and gust speed. A gust of wind is a short-term burst that reaches much higher velocity than average winds, these occur much more commonly around mountains or hills where Eolic turbines are usually installed. Due to that reason, it might be important to distinguish between average wind speeds and gusts as these are more likely to occur around Eolic turbines.

In *Figure 44* and *Figure 45* it's possible to see the relation between these two variables and the wind energy generated. It is possible to see a clear correlation between high wind speeds and high wind energy generation, as low energy generation mostly occurs during times where wind speed is low and high energy generation when wind speeds are high. It is also important to note the difference between wind and gust speed. *Figure 45* shows that gusts present a stronger correlation to wind energy generation when compared to *Figure 44* as was initially expected due to gusts occurring more frequently in areas where Eolic turbines are usually installed.

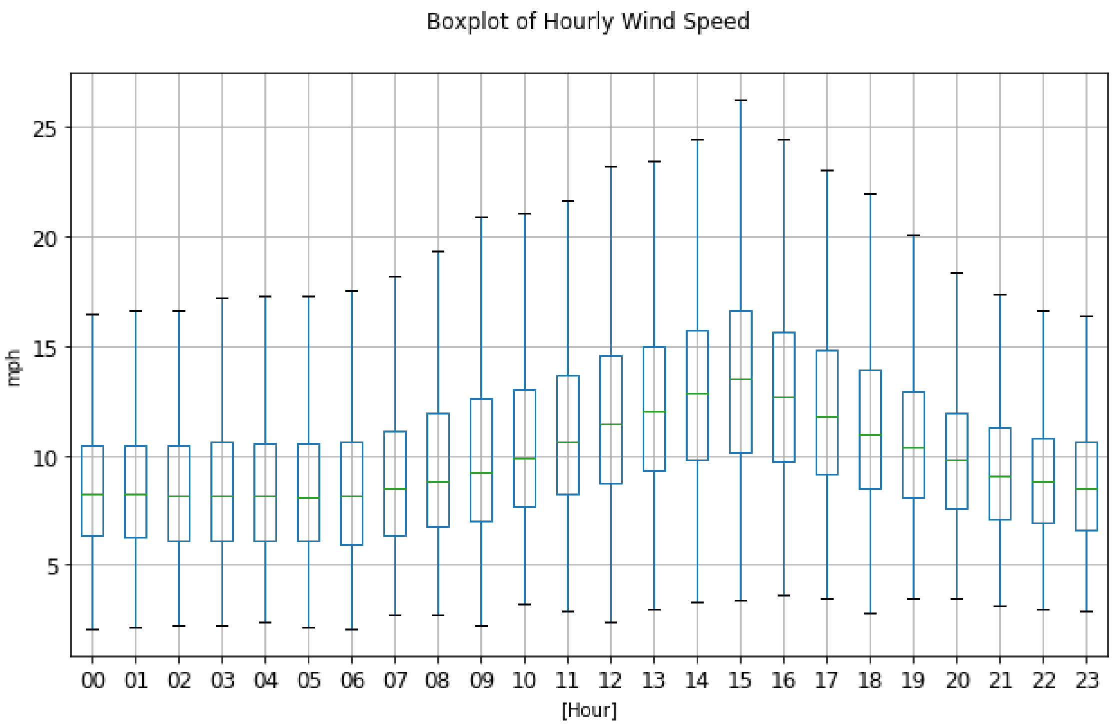

Figure 46 - Hourly distribution of wind speed

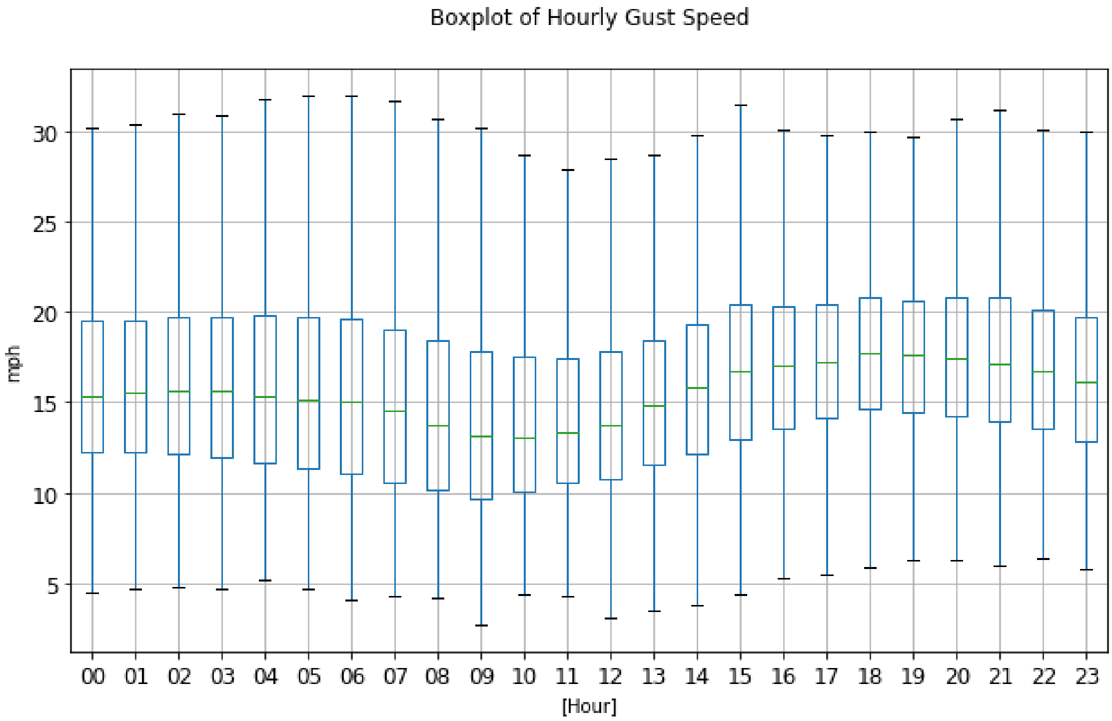

Figure 47 - Hourly distribution of Gust Speed

Looking at *Figure 46* and *Figure 47* and comparing to *Figure 28* it's possible to see that gust speed seems to have a much more similar distribution to wind energy generation than wind speed. Gusts speeds are generally higher during the night than during the day which is exactly the same as wind energy generation. On the other hand, wind speeds do not seem to follow this pattern as they are generally higher during afternoon hours. This is further indication that gust speed seems to have a much stronger correlation to wind energy generation when compared to wind speed. Calculating these two values, it is seen that wind speed has a correlation with price of -0.28, while gust speed has a correlation of -0.43. This is further indication that gust speed is much more related to the final price than wind speed.

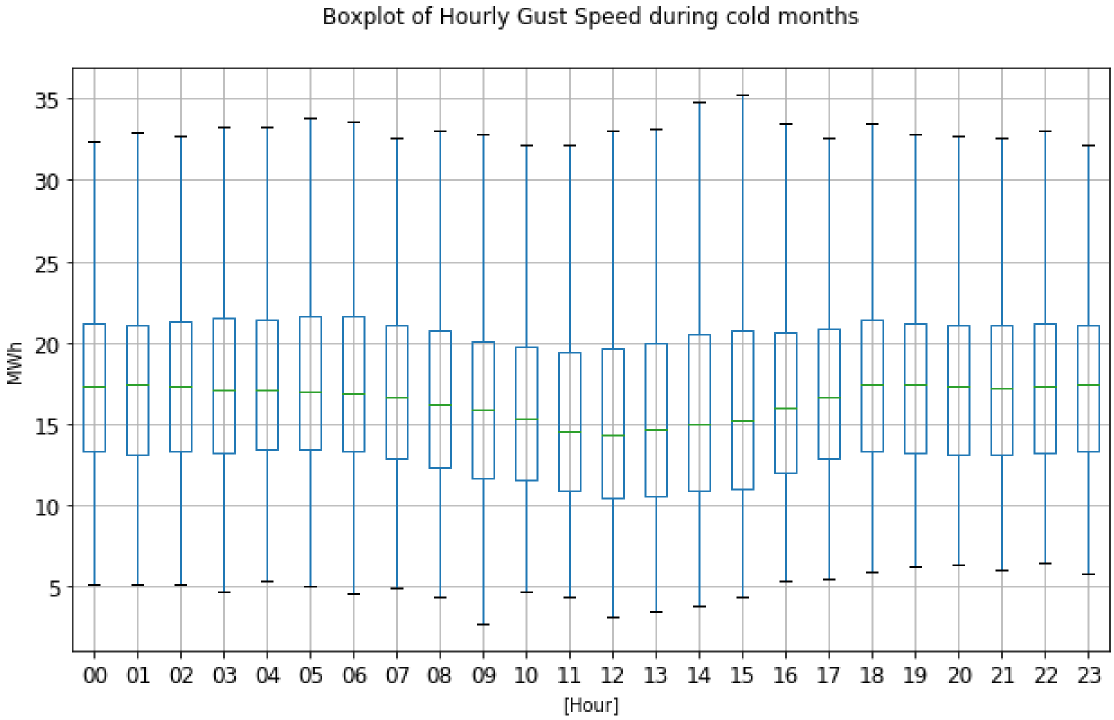

Figure 48 - Distribution of Hourly gust speed during cold months

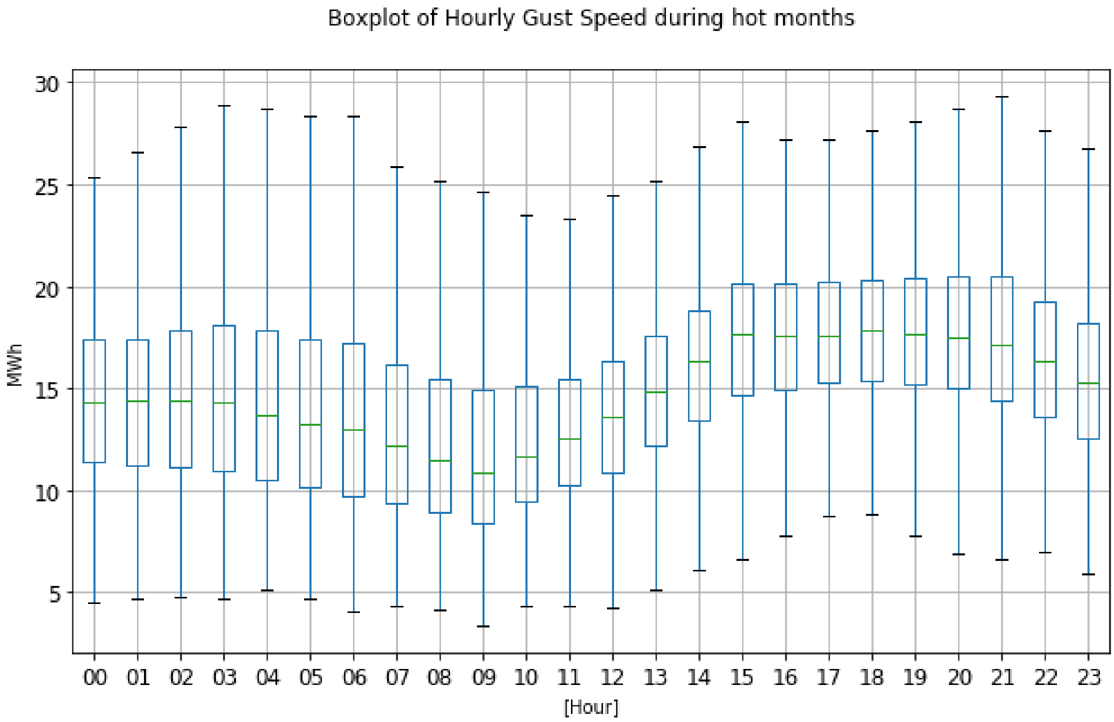

Figure 49 - Distribution of Hourly gust speed during hot months

Comparing *Figure 48* and *Figure 49* to *Figure 29* and *Figure 30* it is possible to see in further detail that gust speeds and wind energy generation present very similar behaviors. During hot months gust speeds during the beginning of the day are significantly lower than during the night which is exactly the same behavior as presented in wind energy generation. The same is true for colder months, as the difference in gust speeds and wind energy generation is much more stable throughout the entire day.

Temperature

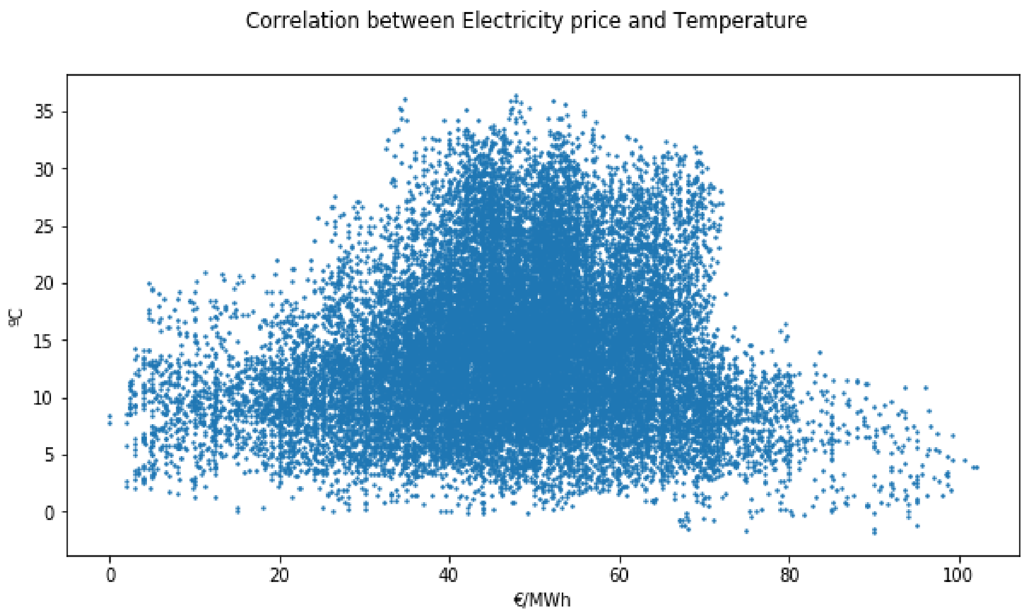

Figure 50 - Correlation between Electricity price and Temperature

In *Figure 50* it is possible to see the correlation between electricity price and temperature. As was previously seen, electricity price tends to follow seasonal patterns with colder months usually having more volatile prices and a higher standard deviation. This can be seen in the data presented as high prices occur exclusively when temperatures are low and low prices occur mostly on low temperatures as well.

As a predictor for the exact price however, temperature does not seem to be a good metric, as prices when temperatures are low can range from anywhere 0€/MWH to 100€/MWh and even in higher temperatures prices can range anywhere from 5€/MWH to 70€/MWh. This makes it very hard to directly relate a given temperature in a given day to the price that is to be expected. Even in the highest temperatures the price range is still very large. This is possibly due to the same temperature occurring in both cold and hot months at different times of the day. For example, 20ºC might be the temperature during peak hours in the winter and early morning hours in the summer.

As such, it is clear that temperature has some relation with the seasonality of the prices, but as an exact price predictor there seem to be much more promising metrics analyzed so far.

Humidity and precipitation

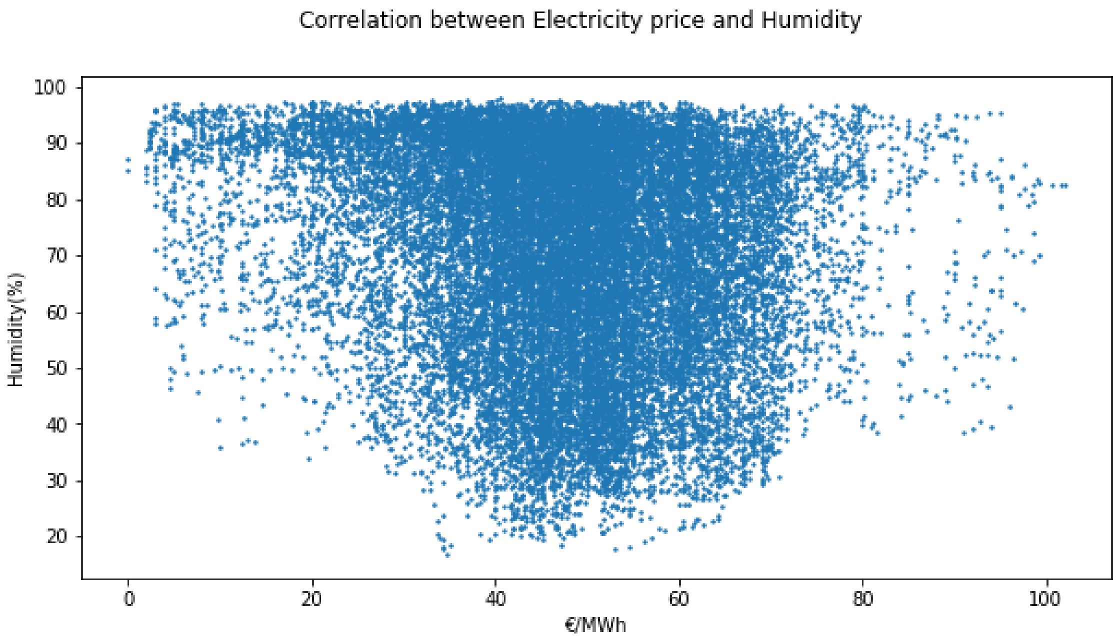

Figure 51 - Correlation between Electricity price and Humidity

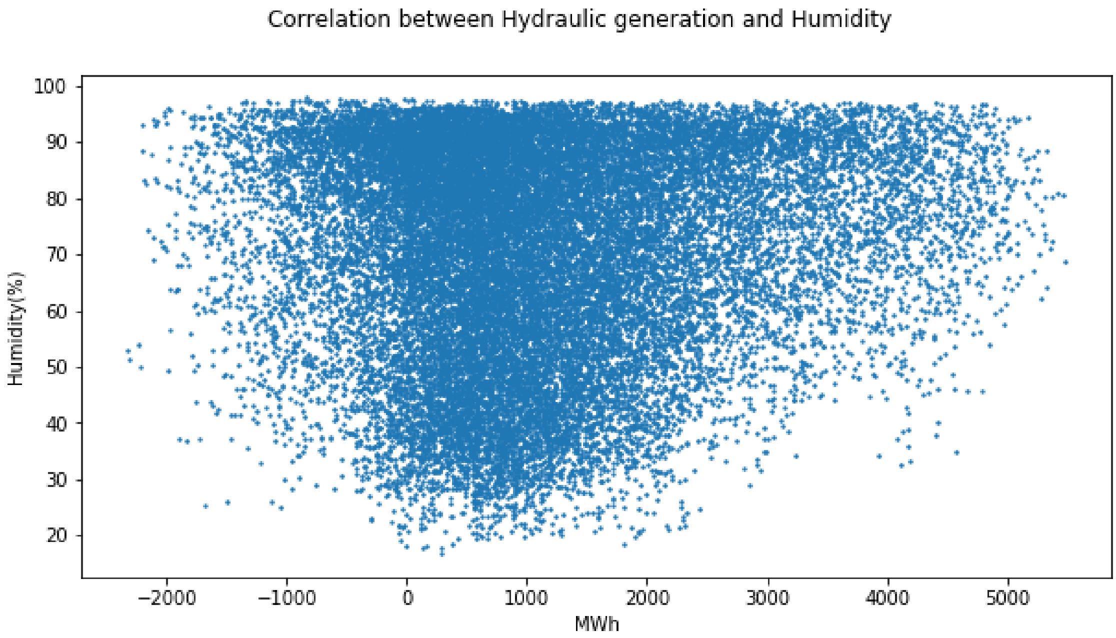

Figure 52 - Correlation between Hydraulic Generation and Humidity

Figure 51 shows the correlation between price and humidity. High humidity values, over 90%, represent times when it was raining. It is possible to see that for high humidity values it is more frequent for prices to be lower rather than higher as more precipitation allows for an higher generation of hydraulic power, this can be further seen in *Figure 52*. However, for humidity values lower than 90% there does not seem to be a strong relationship to the price, as both high and low prices appear almost equally frequent.

As the vast majority of data is humidity values below 90%, it is hard to correlate humidity to the final market price, as there is only a slight change of pattern in the very topmost humidity values that are vastly outnumbered by lower values.

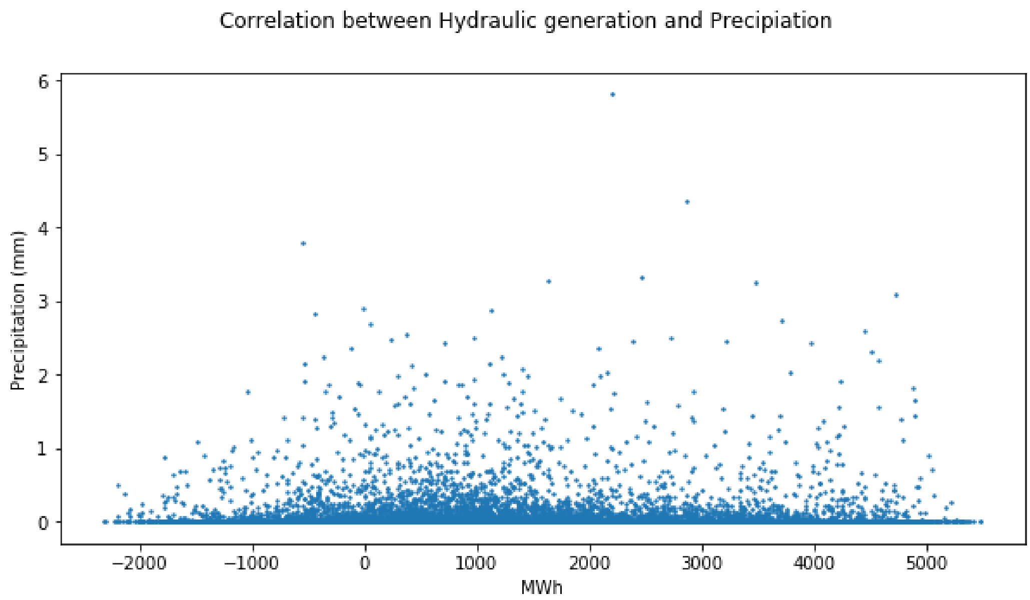

Figure 53 - Correlation between Hydropower Energy Generation and Precipitation

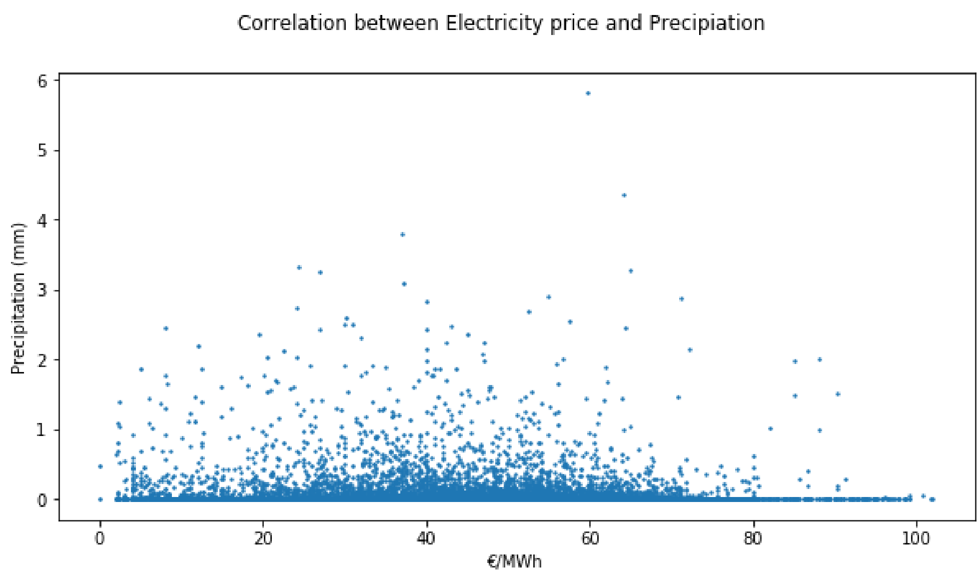

Figure 54 - Correlation between Electricity price and Precipitation

As can be seen in *Figure 53* and *Figure 54*, precipitation presents one major issue has an input variable which is having an extremely unbalanced dataset. This is because in the vast majority of hours it does not rain which causes the majority of entries in the precipitation dataset, about 93% of the data, to be below 0.1mm. Just taking into account values where average precipitation is exactly 0mm, that accounts for about 82% of the dataset.

Looking at the relationship of precipitation with hydraulic generation in *Figure 53* it can be seen that the more negative hydropower energy generation is the less common it is to be raining.

Analyzing *Figure 54* there seems to be a higher tendency for lower prices to happen when it is raining, as opposed to prices above 70€/MWh where the instances where it rained were very few.

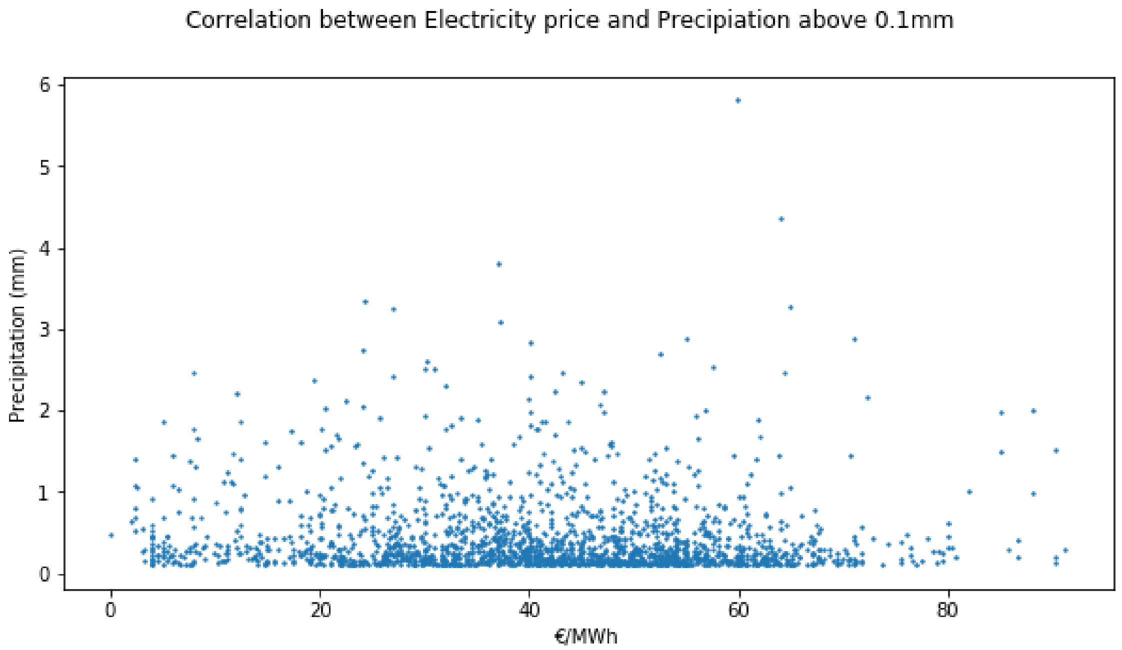

Figure 55 - Correlation between Electricity price and Precipitation above 0.1mm

Overall it seems like precipitation is a significantly worse variable to use as an input feature when compared to some previous variables like gust speed. It can be clearly seen in *Figure 55* that prices are more likely to not be on the higher end when precipitation is above 0.1mm. However, it is hard to utilize that information to accurately predict a specific price as a precipitation of 1mm for example can be directly related to any price between 7€/MWh and 80€/MWh. Furthermore, the extremely unbalanced dataset presents a serious challenge when utilizing precipitation has an input feature.

5- Defining the models and results

Establishing a baseline

Before modeling and running experiments with machine learning algorithms it is important to establish a comparative baseline with other types of models already established in literature to better evaluate the performance and results.

Two different statistical analysis models were used to establish a baseline, a simpler model in AR (1) and a more complex model in ARIMA. In order to evaluate the models, all of them were tested to predict 50 randomly chosen days as the day-ahead from August 2015 to December 2017 additionally a week-ahead point was also predicted to compare the loss of accuracy between day-head to week-ahead predictions.

It is important to note that due to how the MIBEL market works, we cannot use every point of data before the predicted target hour as an input to the models. This is due to the closing hour of the market: 11:00 of day D. meaning that if we want to predict D+1 at most the data available is going to be until the previous hour of the market closing for D+1 which corresponds to day D at 10:00, assuming a prediction of D+1 at the last possible moment. This means that the models will be trying to predict the next 12-36 hours after the last hour of possible offers on day D.

The metrics used for result evaluation are the mean absolute percentage error (MAPE%) and root-mean-square error (RMSE)

$$MAPE = \frac{100\%}{n} \sum_{p=0}^{n} \left| \frac{Av - Fv}{Av} \right|$$

n – Number of total predicted points

Av - Actual Value

Fv – Forecasted Value

$$RMSE = \sqrt{\frac{\sum_{p=0}^{n}(Fv - Av)^2}{n}}$$

n – Number of total predicted points

Av - Actual Value

Fv – Forecasted Value

AR(1)

AR(p) is an autoregressive model that utilizes past observations, denominated as the lag variables, in order to predict future values. To do this, it measures the correlation between the output value and the provided inputs and if a strong correlation exists, either positive or negative, it is a good indicator that good predictions can be expected. The higher the correlation, the more likely that past variables will be able to predict the future. In this case, the input and the output are the same, electricity price. As such it is important to study the autocorrelation of this variable, that is, the correlation between the variable and itself in previous time steps. AR(1) is the first-order process, which means that the current value is based on the immediately preceding value.

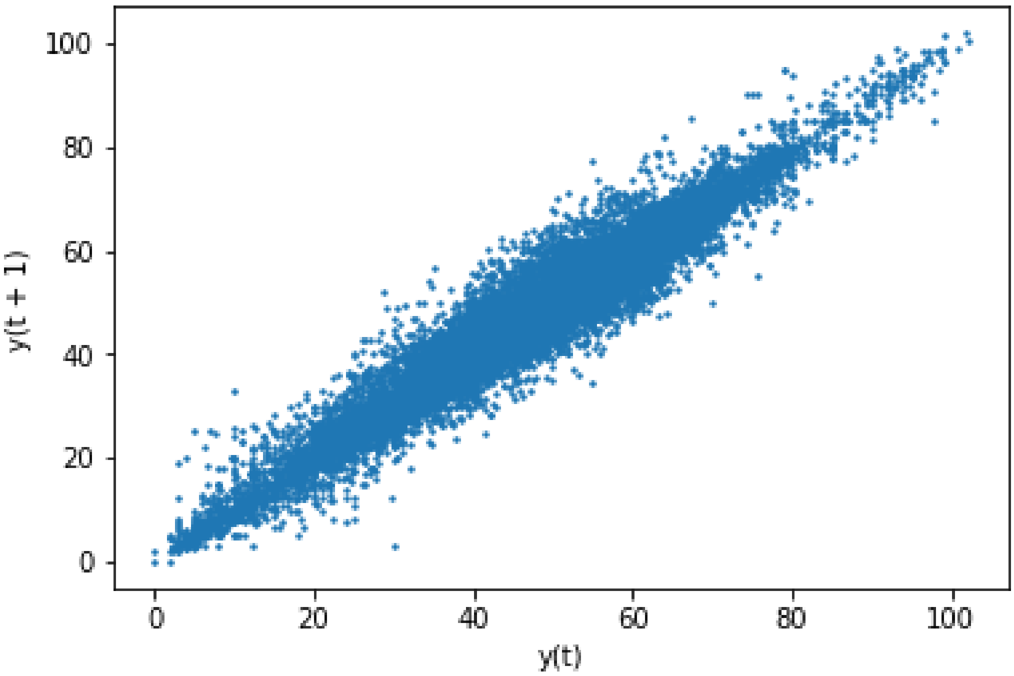

Figure 56 - Autocorrelation of electricity price with the next hour price

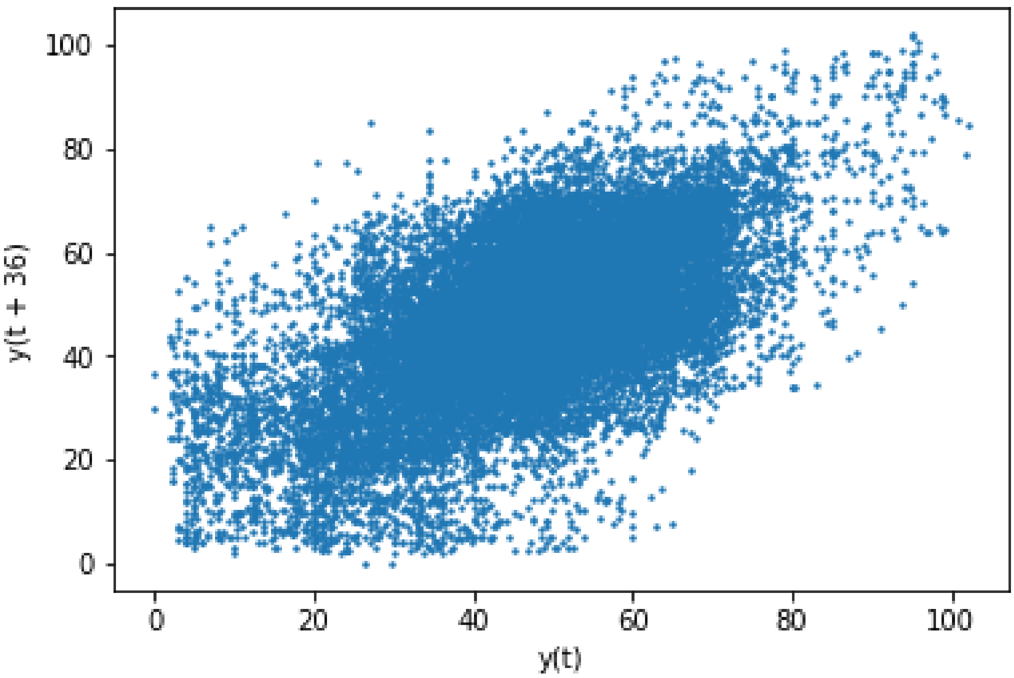

Figure 57 - Autocorrelation of electricity price after 36 hours

As can be seen in *Figure 56*, electricity price presents a very strong autocorrelation within a 1-hour timeframe. This means that prices usually do not have extremely big variations from one hour to the next, which can contribute to making a more accurate prediction. Naturally, as the time frame is increased the correlation gets weaker. This can be seen in *Figure 57* where the autocorrelation gets significantly weaker within a 36-hour timeframe when compared to *Figure 56*, it's still worth to note however that some correlation is still clearly present, especially as prices get higher.

The electricity price at time *t* defined as *Xt* has the following model in AR(1) (Hua, Li, & Li-zi, 2008):

$$x_t = c + \phi \cdot x_{t-1} + u_t$$

In order to simplify the implementation of the model, a python library named *statsmodels* was used whereby inputting data from the previous week to D (day before the targeted day for prediction), the above equation parameters were automatically measured for each hour of D+1 and a prediction for 50 random days was made, achieving the following results:

Table 4 - AR(1) Results

Time Period	MAPE(%)	RMSE
Day-Ahead Prediction	26.53% ± 8.16	11.15 ± 4.12
Week-Ahead Prediction	41.68% ± 9.23	16.25 ± 4.81

AR (1) is by far the simplest algorithm used in this work with almost instantaneous runtime and for such a simple algorithm the final results are not too bad. The week-ahead prediction proved to be significantly worse than the day-ahead prediction which was to be expected for a linear regression model.

While the RMSE values are not very high, it is important to keep in mind that electricity prices do not have a very high standard- deviation as can be seen in **Table *2*** and **Table *3***, and so these results are not usable in a real scenario, as it is essentially guessing the price within the standard deviation. Nonetheless, this is a good indicator that more complex algorithms might have some usability in a real-world scenario as it is expected that results improve as more input features other than the previous price get added into the models.

ARIMA

ARIMA is an acronym that stands for AutoRegressive Integrated Moving Average. This model is a class of statistical models commonly used to analyze time series.

The model utilized for the comparative baseline was created using the models of (Contreras, Espínola, Nogales, & Conejo, 2003) as an example. The creation of these models consists of four steps.

Step 1 is optional and it consists of transforming the data to achieve a more stable mean and variance, in order to do this the library *StandardScaler* was utilized, as ARIMA models have been well documented to have bad performance when exposed to outliers and high variance in the data and it would make the baseline unreliable if this step was skipped over.

Step 2 and 3 consist of choosing and validating the parameters of the model. In order to do this the python library *pmdarima* was utilized to test every possible combination of p, d, and q between values of 0 and 5 against the available data. **p** corresponds to the order of the autoregressive model, it is commonly referred to as the lag order as it consists of the number of lag observations to include in the model. **d** is the integrated part of the model and corresponds to the degree of differencing. Finally, **q** corresponds to the order of the moving-average model and sets the error of the model as a linear combination of the error values of past observations. Any one of these parameters can take the value of 0 which means that an ARIMA model can be configured to perform the functions of an ARMA (AutoRegressive Moving Average), Ar (AutoRegressive) or MA(Moving Average) models.

Step 4 consists of utilizing the trained model to predict the target variables, again using the *statsmodels* library, the better fitting model which was p=2, d=1 and q=3 as this was the model that outputted the lowest AIC value (Akaike Information Criterion) which measures how well a model fits to the data while taking into account the overall complexity of the model. This model was then applied to the same 50 days chosen randomly in the previous experiment, achieving the following results:

Table 5 - ARIMA results

Time Period	MAPE(%)	RMSE
Day-Ahead Prediction	15.21% ± 5.06	6.91 ± 2.96
Week-Ahead Prediction	22.92% ± 6.94	11.52 ± 3.98

ARIMA models are significantly more complex than AR (1) and as such it is expected to show significant improvement. As it is possible to verify comparing **Table 5** to **Table 4**, analyzing the day-ahead prediction the MAPE(%) was reduced by more than 10% and the RMSE by about 4 which for a time-series where the values do not vary a lot is a good improvement as it means that predictions are much more consistently closer to the real values. The week-ahead prediction still has a lot of room for improvement as the RSME is still very similar to the values of the electricity price standard deviation.

Defining the models

As can be seen in *Table 1* there have been several different machine learning algorithms applied to electricity price forecasting, with different training methods, applied to different time periods and electricity markets and with different result metrics. This makes comparing results from different works and finding the best performing one a difficult task. As such, this work applied the most promising algorithms found from the literature review in the same conditions to try to find out the best performing algorithm in the MIBEL market.

There are several possible ways to train machine learning algorithms, the most common ones being either feeding random samples from the entirety of the available data or feeding sequential data up to a certain point in time.

As this work is based on predicting data from a time-series, context from the sequential past data is crucial, as such the data selected for training the algorithms is sequential. For each predicted hour of the day-ahead a different model was trained with data up to the last 36 hours of the final predicted point. The data goes back to a maximum of 5 months, depending on the month that's being predicted, as some months may vary a lot between each other as was previously analyzed at the beginning of chapter 4.

To keep the models realistic, only data until 10:00 of D as the market closes at 11:00 is available and the day ahead prediction begins at 23:00 of D and ends at 22:00 of D+1. An example of the model used to predict the last hour of the day-ahead can be seen in *Figure 58*, where t represents the closing hour of the market, 11: 00, and X represents the number of hours that the training data goes back to, up to a maximum of 3720 which represent 5 months.

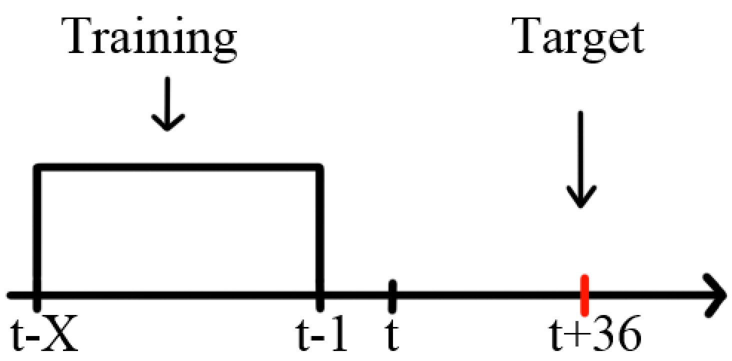

Figure 58 - Example of data used to predict the last hour of the day-ahead

In order to find the best performing parameters for each algorithm, the *GridSearch*CV library was applied to every algorithm used in this work, which tests a number of different specified values in the parameters and scores them against each other. Due to the added complexity of some parameters, and as a result an added increase in the run time, each set of parameters was applied to 60 random days during hyper-parameter testing. Detailed information about the tests can be found in the appendix.

During this work several input variables were tested, it is important to note that a lot of the variables analyzed in chapter 4 such as total consumption levels or energy generated from renewable sources are not known at the time of the prediction and so it wouldn't be realistic to utilize them as inputs for the models. As such, in order to keep the models realistic, consumption and wind energy values are based on real predictions. Weather data was extracted from a public API. Wind energy prediction and gust speed, while they ultimately present the same variable, are both utilized to get a sense on how much accuracy is lost by utilizing weather data as an input feature instead of existing predictions. The description of each input variable can be seen in *Table 6*.

Table 6 -Input Variables

Symbol	**Description**
P(t-x)	Electricity price x hours ago (t represents the target prediction time, x varies between 13 and 37 depending on the model)
DoW	Day of the week (specified numerically)
H	Hour of the day (specified numerically, i.e. Midnight- 0, 1 AM – 1, 3 PM- 15)
C	Expected Consumption
WEP	Wind Energy Prediction
GS	Gust Speed

Long Short-Term Memory

Artificial neural networks are by far the most commonly found algorithm in electricity price forecasting literature. For this work an Artificial Recurrent Network (RNN) named

Long Short-Term Memory (LSTM) was used. This is a type of recurrent neural network that can better utilize long-term information, in comparison to classic RNN, as is needed to improve the forecast accuracy. It has shown excellent results in time-series forecasting problems. More detailed information can be found in the theoretical background chapter.

For the practical implementation, the *lstm* module from the *keras* library in *python* was utilized, a hyperbolic tangent(*tanh*) as the activation function inside the LSTM module. On the outer NN layer, backpropagation was used as a training method and with an added bias on each node. From the Grid Search the best layer composition (in terms of runtime and overall score quality), was found to be a hidden layer composition with 100 neurons. For adjusting the weights, the- Adam optimization algorithm (Kingma & Ba, 2014) was found to be the best performing.

Table 7 - Results for day-ahead predictions with LSTM

Input variables	**MAPE(%)**	**RMSE (€/MWh)**
P(t-x) + H	15.44 ± *5.31*	7.01 ± *3.02*
P(t-x) + H + DoW	15.12 ± *5.12*	6.89 ± *2.93*
P(t-x) + H + DoW + C	14.31 ± *5.02*	6.22 ± *2.84*
P(t-x) + C	15.03 ± *5.09*	6.84 ± *2.96*
P(t-x) + H + DoW + C + GS	13.86 ± *4.36*	5.97 ± *2.19*
P(t-x) + H + DoW + C + WEP	13.23 ± *4.13*	5.62 ± *2.05*

Table 8 - Results for week-ahead predictions with LSTM

Input variables	**MAPE(%)**	**RMSE (€/MWh)**
P(t-x) + H	23.36 ± *7.11*	11.95 ± *4.11*
P(t-x) + H + DoW	22.75 ± *6.95*	11.56 ± *3.96*
P(t-x) + H + DoW + C	22.11 ± *6.78*	10.76 ± *3.75*

P(t-x) + C	22.84 ± *6.99*	11.63 ± *4.04*
P(t-x) + H + DoW + C + GS	21.96 ± *6.11*	10.56 ± *3.41*
P(t-x) + H + DoW + C + WEP	21.03 ± *6.08*	10.21 ± *3.32*

XGBoost

XGBoost has gained immense popularity in the last few years, being initially released in 2014, and has since been used to win many machine learning competitions and is generally found to be one of the best performing machine learning algorithms. For this work the python XGBoost library was used, the best performing parameters were found to be 1000 estimators, a learning rate of 0.01 and a max depth of 4 on each tree.

Table 9 - Results for day-ahead predictions with XGBoost

Input variables	**MAPE(%)**	**RMSE (€/MWh)**
P(t-x) + H	14.78 ± *4.99*	6.51 ± *2.82*
P(t-x) + H + DoW	13.94 ± *4.78*	6.04 ± *2.75*
P(t-x) + H + DoW + C	12.98 ± *4.14*	5.55 ± *2.03*
P(t-x) + C	13.87 ± *4.69*	5.98 ± *2.67*
P(t-x) + H + DoW + C + GS	12.79 ± *4.02*	5.40 ± *1.91*
P(t-x) + H + DoW + C + WEP	11.91 ± *4.00*	5.19 ± *1.89*

Table 10 - Results for week-ahead predictions with XGBoost

Input variables	**MAPE(%)**	**RMSE (€/MWh)**
P(t-x) + H	22.32 ± *6.56*	11.05 ± *3.91*

P(t-x) + H + DoW	22.16 ± *6.51*	10.86 ± *3.80*
P(t-x) + H + DoW + C	21.89 ± *6.23*	10.42 ± *3.33*
P(t-x) + C	22.21 ± *6.59*	10.91 ± *3.93*
P(t-x) + H + DoW + C + GS	21.29 ± *6.10*	10.03 ± *3.29*
P(t-x) + H + DoW + C + WEP	20.74 ± *6.04*	9.81 ± *3.23*

Support Vector Regression

SVR was implemented using the SVR module from the sklearn python library. The best performing parameters were found to be a C of 70 and an epsilon of 0.1. The kernel utilized was RBF as it showed to be the best performing in early testing. This was to be expected as this is a commonly used kernel in previous EPF works (Weron, 2014) and it has shown to be the most appropriate for forecasting financial variables (Papadimitriou, Gogas, & Stathakis, 2014).

Table 11 - Results for day-ahead prediction with SVR

Input Variables	**MAPE(%)**	**RMSE (€/MWh)**
P(t-x) + H	16.26 ± *4.98*	7.98 ± *3.05*
P(t-x) + H + DoW	15.76 ± *4.90*	7.56 ± *2.99*
P(t-x) + H + DoW + C	15.42 ± *4.11*	7.00 ± *2.91*
P(t-x) + C	15.71 ± *4.82*	7.51 ± *2.97*
P(t-x) + H + DoW + C + GS	15.32 ± *3.98*	6.94 ± *2.76*
P(t-x) + H + DoW + C + WEP	15.03 ± *3.64*	6.61 ± *2.68*

Table 12 - Results for week-ahead prediction with SVR

Input Variables	**MAPE(%)**	**RMSE (€/MWh)**
P(t-x) + H	23.74 ± *7.28*	12.31 ± *4.83*
P(t-x) + H + DoW	23.03 ± *7.05*	11.76 ± *4.53*

P(t-x) + H + DoW + C	22.36 ± *6.58*	11.09 ± *4.11*
P(t-x) + C	22.86 ± *6.94*	11.64 ± *4.51*
P(t-x) + H + DoW + C + GS	22.14 ± *6.35*	10.85 ± *3.94*
P(t-x) + H + DoW + C + WEP	21.58 ± *6.31*	10.56 ± *3.92*

Analyzing the results, it can be seen that SVR showed the worst performance of the three algorithms tested, while XGBoost showed to be the best performing. This is true for both day-ahead and week-ahead predictions. It is important to note that for the week-ahead predictions the results were closer between the three algorithms than the day-ahead predictions.

Comparing to the baseline algorithms, ARIMA initially outperformed SVR and LSTM while AR (1) was immediately outperformed by every algorithm. SVR and LSTM eventually outperformed ARIMA as more variables were added to the models. This is an advantage over the ARIMA model.

In the next chapter a more detailed analysis can be found of what these results mean to the initial proposed questions, as well as a brief summary of the entire work.

6- Conclusion

Summary

Electricity price forecasting is a complex problem that is extremely important in today's electricity market both for providers and consumers. This problem has been studied by several researchers in various electricity markets, using various techniques, from classic mathematical models to complex machine learning algorithms, the last ones being especially prominent in more recent years. This works focuses on predicting electricity price in the Iberian electricity market utilizing machine learning algorithms.

The work began by exploring the related work in the field, several bodies of works that predict electricity price can be found in the literature with various different approaches, including classical mathematical, multi-agent and machine learning models.

Due to each different electricity market having its set of working rules and the fact that there doesn't exist a universally accepted scoring metric, comparing the results of different works proved to be a difficult task. Literature spans to dozens of different electricity markets, predicting different time-frames and often being evaluated with different metrics, an example of this can be found in *Table 1*.

In regard to input variables, there are dozens of different input variables utilized in the different works, some based on practical data, others purely theoretical and as such unrealistic in a real scenario. This adds another layer of complexity to comparing different works, as a promising input variable in one market doesn't necessarily directly translate to other markets, as they are not only bound to different rules, but some markets are more dependent on certain types of energy compared to others.

From the literature review, due to the difficulties that were found and explained so far in this chapter, it was concluded that it would not be possible to take only one high performing algorithm, as different algorithms proved to be effective in different contexts.

In the initial phase of testing SVR and MLP were the chosen algorithms. After further testing and researching, LSTM was concluded to generally outperform MLP in the context of this problem. XGBoost also showed promising results and had not yet been applied to electricity price forecasting to the best of our knowledge.

As such, three different algorithms were selected for the final experiments, support vector machines (SVR), LSTM, and XGBoost, additionally to define a baseline, a simple mathematical algorithm, AR(1) and a more complex algorithm that showed good results in some works, ARIMA, were utilized.

The next step consists of analyzing the data available and seeing which variables could be relevant for the last phase of the work, five groups of variables were analyzed consisting of time, consumption, renewable energy sources, non-renewable energy sources and weather variables, this can be seen in chapter 4.

A strong pattern with time variables was found early on, where electricity price clearly shows hourly, daily and monthly shifts very consistently, as prices are usually cheaper on weekends than during work-days, cheaper in late-night hours than in the afternoon and more stable during hotter months compared to colder months.

Consumption levels also presented a very consistent pattern when analyzed taking the calendar into account as it consistently shifts at the same times of day during each season. However a high or low level of consumption does not necessarily directly reflect on the price, as it is clearly shown in *Figure 12* and *Figure 19*. In these figures it can be clearly seen that consumptions levels stayed about the same during 2015, 2016 and 2017 while prices, varied significantly, especially in 2016. This indicated that consumption levels by themselves were not enough to accurately predict the final price and it would be important to understand the sources of the energy and their effect on the price.

The next logical step was then to analyze renewable and non-renewable energy sources, wherein short it was found that in general, the more non-renewable energy sources are needed to fulfill the market needs, the higher the market price is expected to be.

In renewable energy sources, the two highest contributors were found to be hydraulic and wind energy. As for non-renewable energies, coal was found to be the highest contributor to high prices.

The last group of variables analyzed were weather variables, in an attempt to correlate observable weather data to the final market price. As values of consumption of each source of energy are not known at the time of the prediction.

Two variables were analyzed related to wind energy, gust speed and wind speed. Gust speed was found to have very similar patterns to wind energy generation, which is a good indicator that it could be used to predict the final price. This is contrary to what was initially expected, as eolic turbines are built with a preventive system that causes them to shut down during very strong wind speeds and so it was expected that strong gusts would shut down the turbines which should make gust speed have a weak or even a negative correlation to energy generation. After further researching, it was concluded that wind turbines generally only cut-off their power generation at wind speeds above 90 km/h as can be seen in *Figure 59*. As can be seen in *Figure 45* the highest registered average gust speed is only about 68 km/h, which means that it's extremely rare for gust speeds to reach a point where they trigger the shutdown mechanisms of wind turbines. In this figure it is also possible to see the cut-in speed in action, as energy generation only begins on gust speeds slightly above 5km/h.

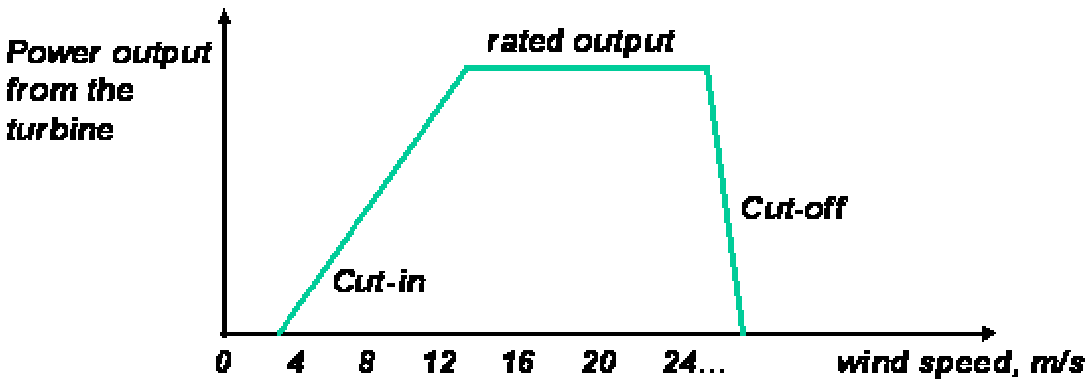

Figure 59 - Wind speed cut in and cut out speeds
(energy.kth.se/compedu/webcompedu/webhelp/S9_Renewable_Energy/B2_Wind_Energy/C1_Introduction_to_Wind_Power/ID33_files/Power_curve.htm**)**

After the analysis, the algorithms mentioned previously were applied to the data, where XGBoost slightly outperformed LSTM and both outperformed SVR, both in the day-ahead and week-ahead predictions.

Comparing to the baseline all algorithms outperformed AR (1) independently of the number of input variables. ARIMA was only initially outperformed by XGBoost, as LSTM and SVR only started outperforming ARIMA as more input variables were added to the models.

From analyzing how the input variables affected the results, it's interesting to note that having the consumption or the day of the week and hour provides very similar results, which further confirms the strong relationship that exists between consumption and time variables mentioned earlier. Furthermore, the best results are consistently given by having consumption, time variables and gust speed, which indicates that the real gust data that was gathered does have some effect on the final price.

Objective Discussion

As stated in the introduction, this work had four main questions which it tried to answer. In this section each of those questions will be discussed more deeply.

Starting with the most important variables, from this work it is clear that in the MIBEL, renewable energies have a large impact on the final pricing. This is clearly shown in chapter 4, where it can be seen that hydraulic and eolic generation especially present very clear patterns in relation to the final market price. It is important to note that not all renewable energy sources have an impact on price, as for example solar energy generation has a very small contribution to the network compared to other sources.

Calendar variables like weekdays or time of the day are also extremely important as they show very strong correlation with the expected consumption levels, which greatly affects pricing.

The last major variable of note is previous hour pricing, as can be seen in *Figure 56* electricity price has a very strong autocorrelation with the previous hour as it is very rare for the price to go up or down by a lot from one hour to the next. This means that if the price of *t-1* is known, predicting *t* becomes much easier. This strong autocorrelation goes down quickly however as can be seen *Figure 57* where by increasing the difference from 1 hour to 36 hours, the price is shown to vary much more.

The previous point is related to the next question, how does the time prediction horizon affect the model results. In chapter 5 it is clearly seen that the results of the next day-prediction are significantly better than the results of the week-ahead prediction. This is mainly due to the autocorrelation of pricing being much weaker after a week when compared to the timeframes necessary for day-ahead prediction, as after a week the price could have completely varied, making the knowledge of last week price not very valuable.

Furthermore, a simplification was made in this work where the weather variables utilized were not the forecasted values but the actual registered values, so in a real-world scenario the week-ahead predictions would possibly be even worse as the error in forecasting week-ahead weather is higher than day-ahead.

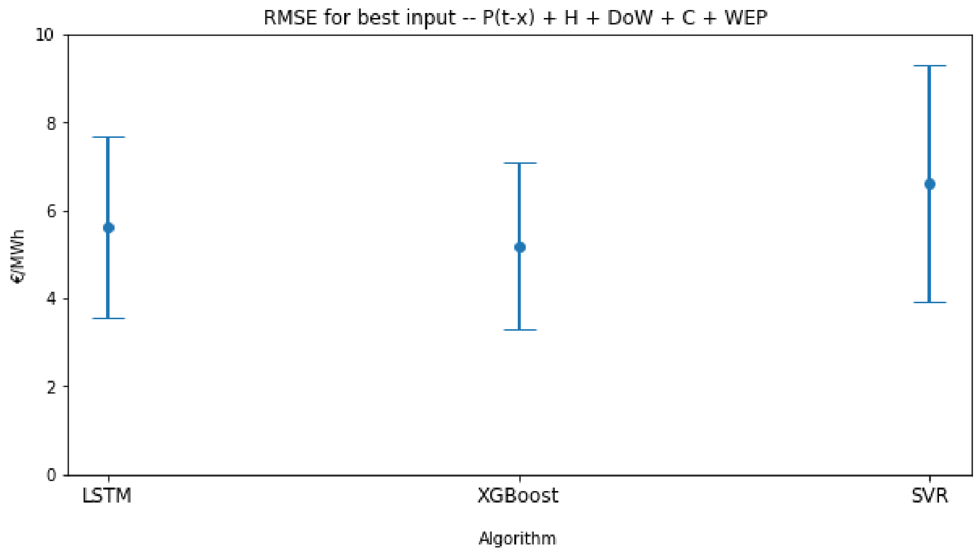

Figure 60 – Best input RMSE values for day-ahead LSTM, XGBoost and SVR

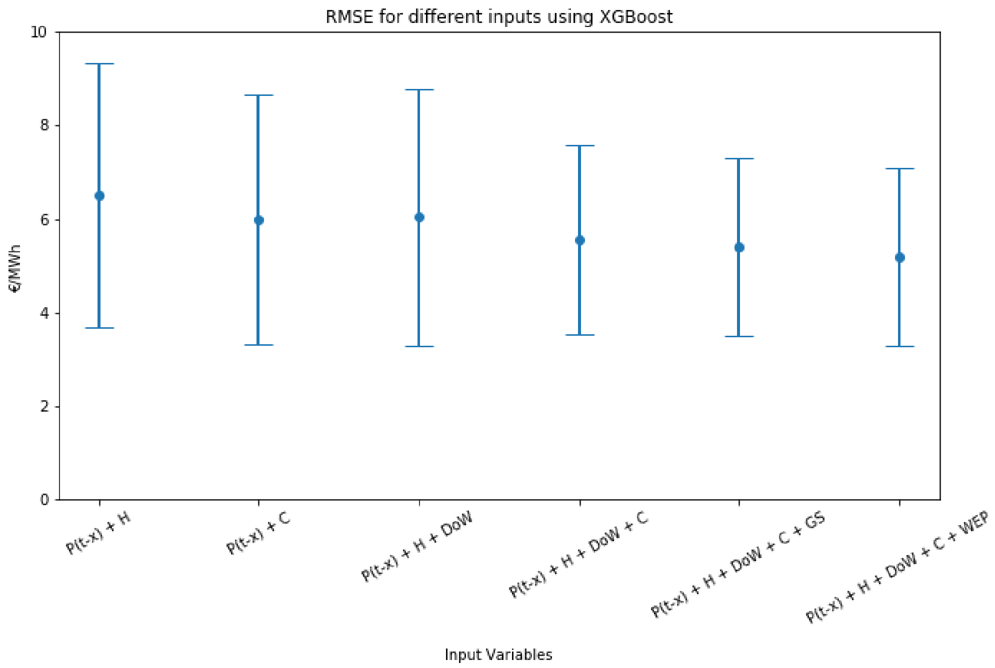

Figure 61 - Day-ahead RMSE for different inputs using XGBoost

Analyzing the day-ahead results shown in *Figure 60*, it is clear that there is still some room for improvement as XGBoost, the best performing model, presented an RMSE of 5.23. The other two models utilized slightly underperformed when compared to XGBoost but their results are also promising.

Analyzing the error for different inputs shown in *Figure 61* it can be seen that knowing some information about energy generation, in this case wind energy generation in the form of gust speed or wind energy prediction , improves the predictive capability of the models and makes them more stable, as both models that used wind related variables presented a lower average RMSE and standard deviation than the rest.

It is also interesting to note that knowing either the consumption levels or the hour and the day of the week provided similar results yet knowing the three variables improved the results comparatively. Because price follows a clear trend with calendar variables, knowing only the consumption but not what time or day it is, may result in error when predicting the price. On the other hand, knowing what time it is but not knowing the consumption also results in some error in the predictions, because while price usually follows trends between night and day it is still important to know the expected consumption levels as for example two Mondays at the same hour can still have very different consumption levels which may affect the price.

Comparing the day-ahead predictions to the week-ahead predictions it is clear that week-ahead predictions are not close to being good enough to be utilized in a real-world scenario. This is an indicator that for medium-term predictions to be viable there needs to be another type of data preparation and model construction.

Comparing these results to the baseline models, AR (1) which is a very simple statistical model was clearly outperformed by every other model. However, ARIMA was only initially slightly outperformed by XGBoost. This indicates that statistical models are a still viable methodology for predicting electricity prices in the MIBEL, even when taking advanced machine learning models into account.

In short, the answers to the initial questions can be summarized as such:

- Day-ahead predictions in the MIBEL are a realistic goal in the short-term with the type of models utilized in this work. Medium-term predicitions with goods results do not seem to be realistic utilizing the type of models and data preparation as shown in chapter 5.
- Advanced machine learning models like LSTM do not seem to be obviously better than complex statistical models like ARIMA as it was only initially outperformed by XGBoost. SVR and LSTM models needed more features added to them in order to outperform ARIMA.
- Time-horizon grealy affects the prediction results, as just in a week difference the results got significantly worse.
- The most important variables in the MIBEL were found to be calendar variables, which greatly affect consumption, and renewable energy generation, mainly hydraulic and eolic sources. Non-renewable energy sources also greatly affect price but they go hand-in-hand in the previously two mentioned variables. For very-short term predictions, previous hour prices were also found to be very important.

Future Work

From the work in this dissertation it is clear that renewable energy sources greatly affect the final market price. As such, an important piece of information missing from the models utilized, that can potentially vastly improve the results, is hydropower energy generation. From the analysis done in chapter 4 it is clear that utilizing just the humidity or precipitation values is not enough to predict hydropower energy generation, as it does not rain in the majority of hours in Portugal. This results in an extremely unbalanced dataset. Furthermore, some of the hydraulic generation comes from stored water, so that needs to be taken into account as well.

Other sources of renewable energy, while not as contributing as eolic or hydraulic, could also be taken into account like biomass and solar energy. Developing accurate models that can predict these values, especially hydropower energy generation, in an hourly timeframe, should be a priority for future researches.

From a model standpoint, XGBoost showed the best results from the models tested. As such it would be interesting to see a work focused exclusively on improving this model, with a deeper hyper-parameter tuning and further data optimizations. Splitting the model into more specific models that only deal with daytime or nighttime data, workdays or weekends, or other patterns found in this work could also potentially greatly improving the predictive capabilities of the models. This would obviously come with an increased computational cost and model complexity as a trade-off for potentially better results.

As stated in the previous sub-chapter, accurate short-term predictions seem to be a realistic goal utilizing the type of models and data preparation that were used in this dissertation. However, for medium-term and long-term predictions it seems that a lot of work still needs to go into developing models that work with real-world data.

Models that try to predict medium or long-term values seem to not be able to utilize previous prices as a training feature, as the prices from one week to the next can shift drastically. Furthermore, utilizing weather data to predict renewable energy generation in the long term is not feasible, as there is a lot of error associated with weather predictions longer than a week’s time. If short-term predictions are further improved these could potentially be utilized in a more complex model for medium-term predictions.

Appendix

GridSearchCV

In this section all the scorings for each parameter combination can be found for the three algorithms tested. Like previously stated, the tested data is equivalent to 60 randomly picked days. The validation was made using 5-fold. The scoring metric for parameter tuning utilized was the Coefficient of determination (R^2) which measures the proportion of the variance for a dependent variable that's explained by the independent variables in the model (Aggarwal et al., 2009).

SVR

		Epsilon		
		0.01	**0.1**	**0.5**
C	**0.1**	(1)	(2)	(3)
	1	(4)	(5)	(6)
	30	(7)	(8)	(9)
	70	(10)	(11)	(12)

	100	(13)	(14)	(15)

Table 13 - Parameters tested for SVR model

Model	**Score**
(1)	0.22 ± 0.02
(2)	0.23± 0.03
(3)	0.22± 0.02
(4)	0.22± 0.03
(5)	0.46± 0.08
(6)	0.46± 0.09
(7)	0.35± 0.09
(8)	0.48± 0.1
(9)	0.45± 0.09
(10)	0.49± 0.08
(11)	0.55± 0.09
(12)	0.50± 0.11
(13)	0.47± 0.1
(14)	0.53± 0.09
(15)	0.48± 0.11

Table 14 - Scoring for each SVR model tested

LSTM

Optimizer			
hidden_nodes		**relu**	**adam**
	10	(1)	(6)
	30	(2)	(7)
	50	(3)	(8)
	100	(4)	(9)
	200	(5)	(10)

Table 15 - Parameters tested for LSTM model

Model	Score
(1)	0.19 ± 0.03
(2)	0.23± 0.03
(3)	0.31± 0.04
(4)	0.37± 0.03
(5)	0.38± 0.04
(6)	0.28± 0.04
(7)	0.33± 0.05
(8)	0.39± 0.1
(9)	0.50± 0.07
(10)	0.50± 0.08

Table 16 - Scoring for each LSTM Model

XGBoost

n_estimators	learning_rate	max_depth	Score
10	**0.001**	**2**	0.12 ± 0.03
100	**0.001**	**2**	0.17 ± 0.05
500	**0.001**	**2**	0.19 ± 0.05
1000	**0.001**	**2**	0.21 ± 0.05
2000	**0.001**	**2**	0.21 ± 0.06
10	**0.01**	**2**	0.14 ± 0.03
100	**0.01**	**2**	0.22 ± 0.04
500	**0.01**	**2**	0.26 ± 0.06
1000	**0.01**	**2**	0.31 ± 0.06
2000	**0.01**	**2**	0.30 ± 0.07
10	**0.1**	**2**	0.14 ± 0.04
100	**0.1**	**2**	0.21 ± 0.04

500	**0.1**	**2**	0.27 ± 0.07
1000	**0.1**	**2**	0.31 ± 0.07
2000	**0.1**	**2**	0.31 ± 0.07
10	**0.001**	**4**	0.20 ± 0.04
100	**0.001**	**4**	0.22 ± 0.08
500	**0.001**	**4**	0.27 ± 0.06
1000	**0.001**	**4**	0.30 ± 0.05
2000	**0.001**	**4**	0.29 ± 0.07
10	**0.01**	**4**	0.24 ± 0.05
100	**0.01**	**4**	0.41 ± 0.07
500	**0.01**	**4**	0.48 ± 0.08
1000	**0.01**	**4**	0.59 ± 0.08
2000	**0.01**	**4**	0.58 ± 0.09
10	**0.1**	**4**	0.22 ± 0.03
100	**0.1**	**4**	0.39 ± 0.05
500	**0.1**	**4**	0.43 ± 0.04
1000	**0.1**	**4**	0.51 ± 0.05
2000	**0.1**	**4**	0.52 ± 0.05
10	**0.001**	**6**	0.19 ± 0.04
100	**0.001**	**6**	0.22 ± 0.08
500	**0.001**	**6**	0.25 ± 0.04
1000	**0.001**	**6**	0.28 ± 0.05
2000	**0.001**	**6**	0.29 ± 0.06
10	**0.01**	**6**	0.22 ± 0.03
100	**0.01**	**6**	0.38 ± 0.06
500	**0.01**	**6**	0.44 ± 0.06
1000	**0.01**	**6**	0.55 ± 0.07
2000	**0.01**	**6**	0.54 ± 0.06
10	**0.1**	**6**	0.21 ± 0.03
100	**0.1**	**6**	0.36 ± 0.05
500	**0.1**	**6**	0.42 ± 0.04
1000	**0.1**	**6**	0.51 ± 0.06
2000	**0.1**	**6**	0.52 ± 0.06

Table 17 - Scoring for each XGBoost model tested

References

Abdel-Aal, R. E. (2006). Modeling and forecasting electric daily peak loads using abductive networks. *International Journal of Electrical Power and Energy Systems*, *28*(2), 133–141. https://doi.org/10.1016/j.ijepes.2005.11.006

Aggarwal, S. K., Saini, L. M., & Kumar, A. (2009). Electricity price forecasting in deregulated markets: A review and evaluation. *International Journal of Electrical Power and Energy Systems*, *31*(1), 13–22. https://doi.org/10.1016/j.ijepes.2008.09.003

Ahmed, N. K., Atiya, A. F., El Gayar, N., & El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econometric Reviews*, *29*(5), 594–621. https://doi.org/10.1080/07474938.2010.481556

Catalão, J. P. S., Mariano, S. J. P. S., Mendes, V. M. F., & Ferreira, L. A. F. M. (2007). Short-term electricity prices forecasting in a competitive market: A neural network approach. *Electric Power Systems Research*, *77*(10), 1297–1304. https://doi.org/10.1016/j.epsr.2006.09.022

Chaâbane, N. (2014). A novel auto-regressive fractionally integrated moving average-least-squares support vector machine model for electricity spot prices prediction. *Journal of Applied Statistics*, *41*(3), 635–651. https://doi.org/10.1080/02664763.2013.847068

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.

Contreras, J., Espínola, R., Nogales, F. J., & Conejo, A. J. (2003). ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems*, *18*(3), 1014–1020. https://doi.org/10.1109/TPWRS.2002.804943

Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, *297*(20), 273–297. https://doi.org/10.1111/j.1747-0285.2009.00840.x

Cummins, F., Gers, F. A., & Schmidhuber, J. (1999). Learning to Forget: Continual Prediction with LSTM. *ICAN'99 Int. Conf. on Artificial Neural Networks*, *2*, 850–855. https://doi.org/10.1197/jamia.M2577

Dang-Ha, T. H., Bianchi, F. M., & Olsson, R. (2017). Local short term electricity load forecasting: Automatic approaches. *Proceedings of the International Joint Conference on Neural Networks*, *2017-May*, 4267–4274. https://doi.org/10.1109/IJCNN.2017.7966396

Fanone, E., Gamba, A., & Prokopczuk, M. (2013). The case of negative day-ahead electricity prices. *Energy Economics*, *35*(May 2008), 22–34. https://doi.org/10.1016/j.eneco.2011.12.006

Friedman, J. H. (1999). Greedy Funtion Approximation: A Gradient Boosting Machine.

IMS 1999 Reitz Lecture.

Graves, A., Mohamed, A., & Hinton, G. (2013). SPEECH RECOGNITION WITH DEEP RECURRENT NEURAL NETWORKS Alex Graves. *IEEE International Conference*, (3), 6645–6649.

Hochreiter, S. (1998). The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions. *Internation Journal of Uncertainty, Fuziness and Knowledge-Based System*, *6*(2), 107–116.

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation 9(8):1735-1780*, *9*(8), 1–32.

Hua, Z., Li, X., & Li-zi, Z. (2008). *Electricity price forecasting based on GARCH model in deregulated market*. 1–410. https://doi.org/10.1109/ipec.2005.206943

Hughes, D., & Correll, N. (2016). *Distributed Machine Learning in Materials that Couple Sensing, Actuation, Computation and Communication*. 1–80. Retrieved from http://arxiv.org/abs/1606.03508

J. Han, M. K. (2006). *Data Mining : Concepts and Techniques*. 703. https://doi.org/10.1093/nar/gku1019

Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, *29*(3), 31–44. https://doi.org/10.1109/2.485891

Janczura, J., Trück, S., Weron, R., & Wolff, R. C. (2013). Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling. *Energy Economics*, *38*, 96–110. https://doi.org/10.1016/j.eneco.2013.03.013

Jones, R. D., Lee, Y. C., Barnes, C. W., Flake, G. W., Lee, K., Lewis, P. S., & Qian, S. (1989). Function approximation and time series prediction with neural networks. *1990 IJCNN International Joint Conference on Neural Networks*, 649–665 vol.1. https://doi.org/10.1109/IJCNN.1990.137644

Kingma, D. P., & Ba, J. (2014). *Adam: A Method for Stochastic Optimization*. 1–15. Retrieved from http://arxiv.org/abs/1412.6980

Kleynhans, T., Montanaro, M., Gerace, A., & Kanan, C. (2017). Predicting top-of-atmosphere thermal radiance using MERRA-2 atmospheric data with deep learning. *Remote Sensing*, *9*(11), 1–16. https://doi.org/10.3390/rs9111133

Lin, W. M., Gow, H. J., & Tsai, M. T. (2010). An enhanced radial basis function network for short-term electricity price forecasting. *Applied Energy*, *87*(10), 3226–3234. https://doi.org/10.1016/j.apenergy.2010.04.006

Mohamed, H., Negm, A., Zahran, M., & Saavedra, O. C. (2015). Assessment of Artificial Neural Network for bathymetry estimation using High Resolution Satellite imagery in Shallow Lakes : Case Study El Burullus Lake . *International Water Technology Conference*, (March), 434–444.

Niu, D., Liu, D., & Wu, D. D. (2010). A soft computing system for day-ahead electricity price forecasting. *Applied Soft Computing Journal*, *10*(3), 868–875. https://doi.org/10.1016/j.asoc.2009.10.004

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the Future- Big Data, Machine Learning and Clinical Medicine. *N Engl J Med*, *375*. https://doi.org/10.1016/j.physbeh.2017.03.040

Osório, G. J., Gonçalves, J. N. D. L., Lujano-Rojas, J. M., & Catalão, J. P. S. (2016). Enhanced Forecasting Approach for Electricity Market Prices and Wind Power Data Series in the Short-Term. *ICIASF 2005 RecordInternational Congress OnInstrumentation in AerospaceSimulation Facilities*, 166–175. https://doi.org/10.3390/en9090693

Papadimitriou, T., Gogas, P., & Stathakis, E. (2014). Forecasting energy markets using support vector machines. *Energy Economics*, *44*, 135–142. https://doi.org/10.1016/j.eneco.2014.03.017

Pastor, R., Pinho, N., & Esteves, J. (2018). Market-based bidding strategy for variable renewable generation in the MIBEL. *15th International Conference on the European Energy Market (EEM), Lodz*.

Razak, I. A. B. W. A., Abidin, I. B. Z., Siah, Y. K., Rahman, T. K. B. A., Lada, M. Y., Ramani, A. N. B., … Ahmad, A. B. (2015). Support vector machine for day ahead electricity price forecasting. *AIP Conference Proceedings*, *1660*. https://doi.org/10.1063/1.4915865

Ruta, D., & Gabrys, B. (2007). Neural network ensembles for time series forecasting.

Proceedings of International Joint Conference on Neural Networks,Orlando,Florida,USA,August 12-17,2007, 1235. https://doi.org/10.1145/1569901.1570067

Sak, H., Senior, A., & Beaufays, F. (2014). Long Short-Term Memory Recurrent Neural Networks Architectures for Large Scale Acoustic Modeling. *INTERSPEECH*. https://doi.org/10.1016/j.edumed.2017.01.001

Sankar, R., & Sapankevych, N. I. (2009). Time Series Prediction Using Support Vector Machines: A Survey. *IEEE Computational Intelligence Magazine*, *4*(2), 24–38. https://doi.org/10.1109/MCI.2009.932254

Sansom, D., Downs, T., & Saha, T. (2003). Evaluation of support vector machine based forecasting tool in electricity price forecasting for Australian national electricity market participants. *Journal of Electrical and …*, *22*(3), 227–233. Retrieved from http://espace.library.uq.edu.au/view/UQ:9850

Sharma, C., & Bradford, M. (2017). Empirical Evaluation of various Deep Neural Network Architectures for Time Series Forecasting. *MSc Research Project*.

Singh, N., Husain, S., & Mohanty, S. R. (2016). An improved WNN for day-ahead electricity price forecasting. *4th Students Conference on Engineering and Systems, SCES 2015*, 1–6. https://doi.org/10.1109/SCES.2015.7506450

Stevenson, M. (2001). *Filtering and Forecasting Spot Electricity Prices In The Increasingly Deregulated Aust ralian Electricity Market*. 3.

Syarif, I., Prugel-Bennett, A., & Wills, G. (2016). SVM parameter optimization using grid search and genetic algorithm to improve classification performance. *Telkomnika (Telecommunication Computing Electronics and Control)*, *14*(4), 1502–1509. https://doi.org/10.12928/TELKOMNIKA.v14i4.3956

Szkuta, B. R., Sanabria, L. A., & Dillon, T. S. (1999). Electricity price short-term forecasting using artificial neural networks. *IEEE Transactions on Power Systems*, *14*(3), 851–857. https://doi.org/10.1109/59.780895

Torbaghan, S. S., Motamedi, A., Zareipour, H., & Tuan, L. A. (2012). Medium-term electricity price forecasting. *2012 North American Power Symposium, NAPS 2012*. https://doi.org/10.1109/NAPS.2012.6336424

Ventosa, M., Baíllo, Á., Ramos, A., & Rivier, M. (2005). Electricity market modeling trends. *Energy Policy*, *33*(7), 897–913. https://doi.org/10.1016/j.enpol.2003.10.013

Wang, J., Zhang, X., Tao, X., & Wang, J. (2018). EPSLP: Efficient and privacy-preserving single-layer perceptron learning in cloud computing. *Journal of High Speed Networks*, *24*(3), 259–279. https://doi.org/10.3233/JHS-180594

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, *30*(4), 1030–1081. https://doi.org/10.1016/j.ijforecast.2014.08.008

Weron, R., & Misiorek, A. (2008). Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting*, *24*(4), 744–763. https://doi.org/10.1016/j.ijforecast.2008.08.004

Wirth, R. (2000). 13 CRISP-DM : Towards a Standard Process Model for Data Mining. *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, (24959), 29–39. https://doi.org/10.1.1.198.5133

Wu, W., Zhou, J.-Z., Yu, J., Zhu, C.-J., & Yang, J.-J. (2004). Prediction of spot market prices of electricity using chaotic time series. *Machine Learning*, (August), 26–29.

Yamin, H. Y., Shahidehpour, S. M., & Li, Z. (2004). Adaptive short-term electricity price forecasting using artificial neural networks in the restructured power markets. *International Journal of Electrical Power and Energy System*, *26*(8), 571–581. https://doi.org/10.1016/j.ijepes.2004.04.005

Yao, S. J., Song, Y. H., Zhang, L. Z., & Cheng, X. Y. (2000). Prediction of system marginal prices by wavelet transform and neural networks. *Electric Machines and Power Systems*, *28*(10), 983–993. https://doi.org/10.1080/07313560050129855