Simulation of the Zero and Non-zero bids in the 2030 Day Ahead MIBEL Market

Mahmoud Moustafa Elsabbahi Saber

Thesis to obtain the Master of Science Degree in

Energy Engineering and Management

Supervisors : Prof. Rui Manuel Gameiro de Castro
Dr. Fernando Jorge Lopes

Examination Committee

Chairperson: Prof. Susana Isabel Carvalho Relvas
Supervisor: Prof. Rui Manuel Gameiro de Castro
Member of the Committee: Prof. João Hermínio Ninitas Lagarto

December 2021

Declaration

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Acknowledgments

This Master's Thesis is dedicated to my family who have always supported me and never stopped believing in me. I am profoundly grateful to my parents for all their unconditional love, support and guidance over the years.

I would also like to thank Professor Rui Castro for his continuous guidance, advice and help during the development of this work.

Not to forget, a special thanks to my Portuguese family who made my journey much easier over the past two years.

Last but not least I would like to thank all my friends who helped and supported me whenever I needed them.

Abstract

Nowadays we live in a world that is rapidly heading towards a greener and more sustainable life. Actions have been taken by many countries to support decarbonization, improving efficiency and more use of renewable energy over fossil fuel. From its side the European commission established a very ambitious plan for the near and far future. Targets have been set for 2020 and been accomplished, yet more to be reached by 2030 and all the way to 2050. These targets are backed by national plans to be achieved by individual countries within the EU. In this paper Portugal and Spain's electricity market (MIBEL) is investigated to study the effect of the massive penetration of renewable energy on the future electricity market. Focusing on the supply curve, the main objective of this paper is to forecast the quantities of electricity bid in the 2030 day-ahead market, to study the Iberian market behaviour to the increasing renewable penetration, and its effect on the quantities that are bid at the day ahead market. The quantities are categorized into zero and non-zero price segments. As a final result of this study quantities bid by generation players were forecasted by an artificial neural network as a first step of predicting the supply curve. Results show that the quantities of electricity bid at zero price will increase significantly, which supports the speculation that renewable electricity will have a bigger share in the MIBEL energy mix by 2030.

Keywords: Long-Term Forecast, MIBEL, Renewable Energy, Supply and Demand Curves, Artificial Neural Networks, Day-Ahead Market .

Resumo

Nos dias de hoje o mundo caminha cada vez mais rápido para uma vida mais verde e sustentável. Vários países têm tomado medidas de apoio à descarbonização, melhorando a eficiência e o uso de energia renovável em vez de combustíveis fósseis. Por sua vez, a Comissão Europeia estabeleceu um plano bastante ambicioso não só para o futuro próximo como para um distante. Foram estabelecidas e cumpridas metas para 2020, e ainda outras metas a serem alcançadas até 2030 e 2050. Estes objectivos são apoiados por planos nacionais a serem alcançados por cada país dentro da UE. Neste artigo investiga-se o mercado de eletricidade de Portugal e Espanha (MIBEL) com foco no efeito da penetração maciça das energias renováveis no futuro mercado da eletricidade. Focando na curva de oferta, o principal objetivo deste trabalho é prever as quantidades de eletricidade licitadas no mercado do dia 2030 para estudar o comportamento do mercado ibérico face ao aumento da penetração das energias renováveis e o seu efeito nas quantidades licitadas no dia mercado à frente. As quantidades são categorizadas em segmentos de preço zero e diferente de zero. Como resultado final deste estudo, as quantidades licitadas por jogadores de geração foram previstas por uma rede neural artificial como uma primeira etapa de previsão da curva de oferta. Os resultados mostram que as quantidades de eletricidade licitadas a preço zero irão aumentar significativamente, o que corrobora a especulação de que a eletricidade renovável terá uma maior participação no cabaz energético do MIBEL até 2030.

Palavras-chave: Previsão a Longo-Prazo, MIBEL, Energias Renováveis, Preços “Spot” de Eletricidade, Redes Neuronais Artificiais, Mercado Diário.

Table of Contents

List of Tables

List of Figures

Glossary

ABSM Agent-Based Simulation Models

AF Activation Function

ANN Artificial Neural Networks

APREN Associação Portuguesa de Energias Renováveis

ARCH Auto-Regressive Conditional Heteroskedastic

ARIMA Auto Regressive Integrated Moving Average

ARMA Auto-Regressive Moving Average

ARX Auto-Regressive (Exogenous)

CI Computational intelligence

CNN Convolutional Neural Networks

CV(RMSE) Coefficient of Variation of Root-Mean Squared Error

DAEM Day Ahead Electricity Market

ENTSOE European Networks of Transmission System Operators for Electricity

GRU Gated Recurrent Unit

LSTM Long Short-term Memory

MAE Mean Absolute Error

MAPE Mean Absolute Percentage Error

MAPE Mean Absolute Percentage Error

MAPE Mean Absolute Percentage Error

MLP Multi-Layer Perceptron

MRS Markov Switching Model

MSE Mean Squared Error

OMIE Operador del Mercado Ibérico

PNIEC Plan Nacional Integrado de Energía y Clima

PR Parameter Rich

PS Parsimonious Structural

RE Renewable Energy

REE “Red Eléctrica de España”

REN “Redes Energéticas Nacionais”

RMSE Root Mean Squared Error

RNC Roteiro para a Neutralidade Carbónica

RNN Recurrent Neural Networks

SARIMA Seasonal Auto-regressive Integrated Moving Average

SETAR Self-Exciting Threshold Auto-Regressive

SFE Supply Function Equilibrium

SPCM Strategic Production-Cost Model

SVM Support Vector Machine

TAR Threshold Autoregressive

TARX Threshold Autoregressive (Exogenous)

TSO Transmission System Operator

Chapter 1

1 Introduction

1.1 Framework and Motivation

Concerns over climate change and global warming are increasing day after day. Our planet has been suffering from the human activity for decades, leading to changes in its climate patterns. These changes are unfortunately beyond the natural climate variability and can lead to disastrous irreversible consequences. The main cause of this threat is greenhouse gases being emitted into our atmosphere. These emissions are mainly driven by combustion of fossil fuels, agricultural practices, waste treatment and industrial processes. All of these drivers directly impact our planet's climate by increasing the global temperature and rising sea levels leading to extreme weather conditions that in return effects the ecosystem, the economy, the human society and health.

Therefore, to tackle these threats the majority of countries signed the Paris agreement in 2015 aiming to reduce the greenhouse gases in effort to limit the global temperature rise to 2 degrees Celsius above the preindustrial levels. From its side also the European commission (EC) set some very ambitious targets for 2020, 2030 and 2050 [1] [2] [3]. The 2020 climate and energy package targets have been met and now the road to the 2030 package has started aiming to achieve the following targets:

- At least 55% cuts in greenhouse gas emissions (from 1990 levels).
- At least 32% share of the EU's final energy produced by Renewable Energy Source.
- At least 32.5% improvement in energy efficiency with respect to the projections of the expected energy use in 2030.

The EC also took some initiatives including the European Green Deal that stands for three main objectives:

- No net emissions of greenhouse gases by 2050.
- Economic growth decoupled from resource use.
- No person and no place left behind.

Figure 1 represents the global greenhouse emissions by sector, and as shown the majority of the emissions are being produced from the energy sector and this is why there is a huge global drive towards using more sustainable energy sources and increasing renewable electricity in every country's energy mix.

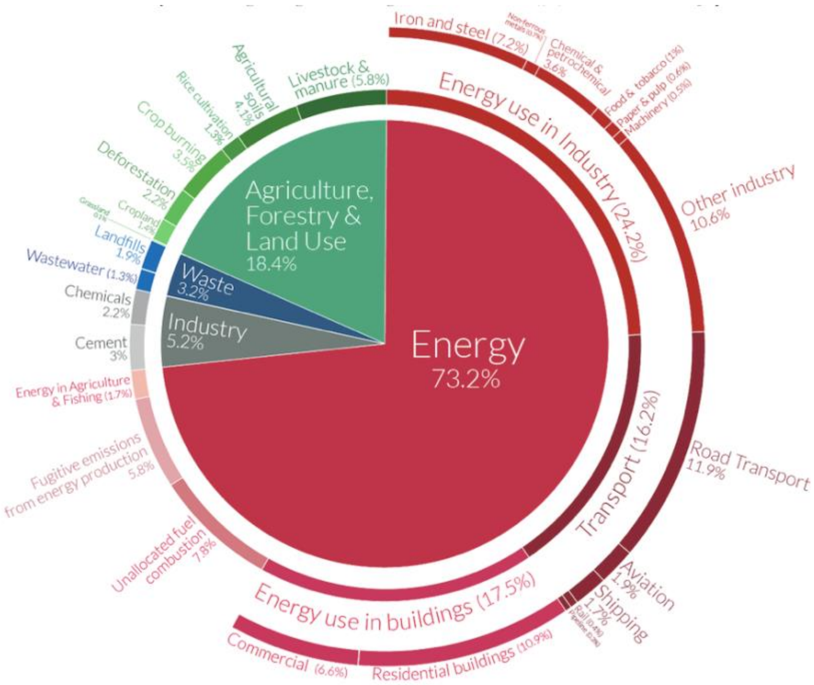


Figure 1 - Global greenhouse emissions by sector 2016 [62] .

Considering the energy sector contribution to the greenhouse emissions it becomes obvious that energy systems have to be changed and developed towards more sustainable energy sources and less fossil fuel. In another way energy systems should be decarbonized. These changes will include an energy mix that is able to supply the increasing future electricity demand with renewable electricity, and definitely this massive renewable penetration will have direct impacts on several market features like:

1. Electricity price
2. Market players biding behavior
3. Investments

This is why it is seen to be very important to study and assess how the future electricity markets will behave to such changes. There is lots of previous studies and research that focused mainly on price time series forecasting and neglect the core principles that controls the price which is the supply and demand curves also known as purchase and sales curves. This mechanism of the market is what really determines the price of electricity and the bidding behavior of market players. Therefore, in this research a new methodology is developed as a first step for the aim of modelling the future electricity prices for the day-ahead MIBEL electricity market in a different way, by forecasting the supply and demand curves which are also referred to as bid and ask curves. Once we are able to model the two curves, the electricity price can be obtained by simply getting the intersection point between both curves. The information provided from the two curves will also be very useful for further research regarding bidding behavior and structure, not to forget the prediction of extreme price events.

1.2 Objectives

The central purpose of this thesis is to study the day-ahead MIBEL Electricity Market and develop a model that can forecast the future supply curve by 2030. Due to some market data limitations the objective is simplified, and the work is focused on modeling the supply curve of the day-ahead market by forecasting the quantities of electricity bid at different price segments, zero and non-zero prices.

To achieve such purpose the following objectives are set:

- Study of the Iberian day-ahead electricity market (MIBEL).
- Collection, processing and analysis past hourly market data from different sources.
- Construction, training and validation of an artificial neural network using the gathered data and the identified variables
- Forecast the zero and non-zero bids of the supply curve of the day-ahead market.
- Analyse the results and check if it complies with the objectives 2030.

1.3 Thesis Structure

The present thesis dissertation is divided into eight chapters, which are:

1. Introduction.

2. Iberian Electricity Market – An overview on the MIBEL electricity market to provide a better understanding of its history, characteristics and future projections.

3. Literature Review – A review of the published studies regarding the problem under investigation.

4. Methodology – An overview on the methodology used throughout this work.

5. Theoretical Framework and Implementation – Presents the theoretical foundations of artificial neural networks and practical implementation of the proposed forecasting models.

6. Models' Validation – Analysis over the models' forecasting accuracy to measure the model's reliability.

7. Results and Discussion – Provides an overview over the forecasted projections of the zero and non-zero quantities of the 2030 supply curves. At the same time, a careful analysis over the simulated results is conducted.

8. Conclusions – Main conclusions and contributes of the developed study, future work proposals and improvements.

Chapter 2

2 Iberian Electricity market - MIBEL

In 1998 Portugal and Spain took the first step in building the Iberian electricity market which had the aim of integrating both countries electricity systems. The integrated market came to operation in 2007 bringing benefits not only to the consumers from both countries and allowing all participants to have a free access and equal rights [5] but also on the European scale, since it is a step towards building up the internal European electricity market.

The MIBEL market is operated by Operador del Mercado Ibérico (OMIE) [6], which is owned by the Spanish society OMEL and the Portuguese society OMIP, both owning equal shares. For each country there is a Transmission System Operator (TSO). In Portugal the TSO is Redes Energéticas Nacionais (REN) [7], and in Spain Red Eléctrica de España (REE) [8].

The organisational structure of the Electricity market reflects a vertical chain of activities that can be categorised into three Main points:

- Energy Production (wholesale - MIBEL).
- Energy Transportation (Transmission and Distribution).
- Retail Market.

The electricity transport activities which are represented in transmission and distribution are based on the use of an existent network that allows the transportation of electricity from production facilities to the end user. These networks are characterized as natural monopolies. Regulations are subjected over the use of these networks allowing third parties to have access through payment of a regulated tariff. On the other hand, the electricity production (wholesale market) and the retail market are open to competition, which can be justified by the introduction of more efficient techniques to manage resources that are involved in these activities.

The MIBEL is currently composed of:

1. Derivatives Market (OMIP): It is a trading platform for which future buying and production commitments are established. Physical (energy) or financial (money) settlements are allowed in the derivatives market.
2. Spot Market (OMEL): Is a market that deals with the daily trades resembled in the day-ahead market and intraday adjustment market, where the buying and selling are set for the day following the trade.
3. Ancillary services Market: Operates in real time aiming to balance the market by setting an equilibrium between electricity production and consumption.
4. Bilateral trading Market: Where electricity selling and buying trades are arranged for diverse time horizons.

2.1 Day-ahead electricity market (DAEM)

Since our case under study is focused on the day-ahead market (Spot market) and specifically the supply curve, the reader should have a better understanding of the day-ahead market in terms of its characteristics and working mechanism.

The spot market is by far the most important market in the MIBEL structure, since it is where most of the electricity is traded. Given its importance, this work will be focused only on the MIBEL spot market, and any reference made to a market feature will be regarding this specific market.

In the spot market electricity is generally traded in a daily auction for each hour of the following day. The closing hour of the spot market is 10:00 am on the day before the supply, and the clearing prices are announced at 11:00 am. The clearing price is the point where the supply curve meets the demand curve. These curves are composed of the supplier's bids and the consumers offers. Each bid and offer include the quantity of electricity that a market agent is willing to sell or buy in [MWh], and the price in [€/MWh]. After all market agents submit their bids and offers, the supply bids are arranged in ascending order and demand offers are arranged in descending order. The two curves are then plotted against each other, and the intersection point is determined. This point is what so called clearing price or spot price. The corresponding quantity of electricity refers to the amount of electricity traded at a specific hour of the following day. It is important to note that all the suppliers that submit a bid at a price higher than the clearing price will not be able to participate, as well as all consumers that offer to buy electricity at a price lower that the clearing price will not be accepted in the market [5]. It is also important to note that if the quantity of the traded electricity exceeds the interconnection capacity between Portugal and Spain the market splits, and different prices of electricity take place in each country as a solution to this congestion.

As explained above, each market agent has to make a/an bid/offer to the market operator. These bids/offers can be simple or complex depending on their content. Simple bids/offers are those which express an amount of energy and an equivalent price. On the other hand, complex bids/offers are those which not only express an amount of energy and an equivalent price but also include complex conditions to be taken into consideration during the matching process and they are the one taken into consideration to calculate the electricity price [5].

These complex conditions can be at least one of the following:

- Condition of indivisibility: states that bid block is indivisible meaning that if the bid is matched it will be accepted in the market as a whole and not a fraction.
- Minimum income condition: this condition states that the seller should earn a minimum income so that the bid block can be considered in the matching process.
- Scheduled stop condition: in the event that bids are not matched due to the application of the previous minimum income condition, they can be treated as simple bids.
- Production capacity variation or load gradient condition: for each production unit a maximum and a minimum variation in energy is set, limited by the hourly maximum production capacity of the unit.

Figure 2 represents the supply and demand curves for hour 15 of day 15/01/2021 [9]. As shown the sale bids are in light green and matching sale bids are in dark green color. The intersection of the matching sales offers with the matching purchase offers indicates the clearing price.

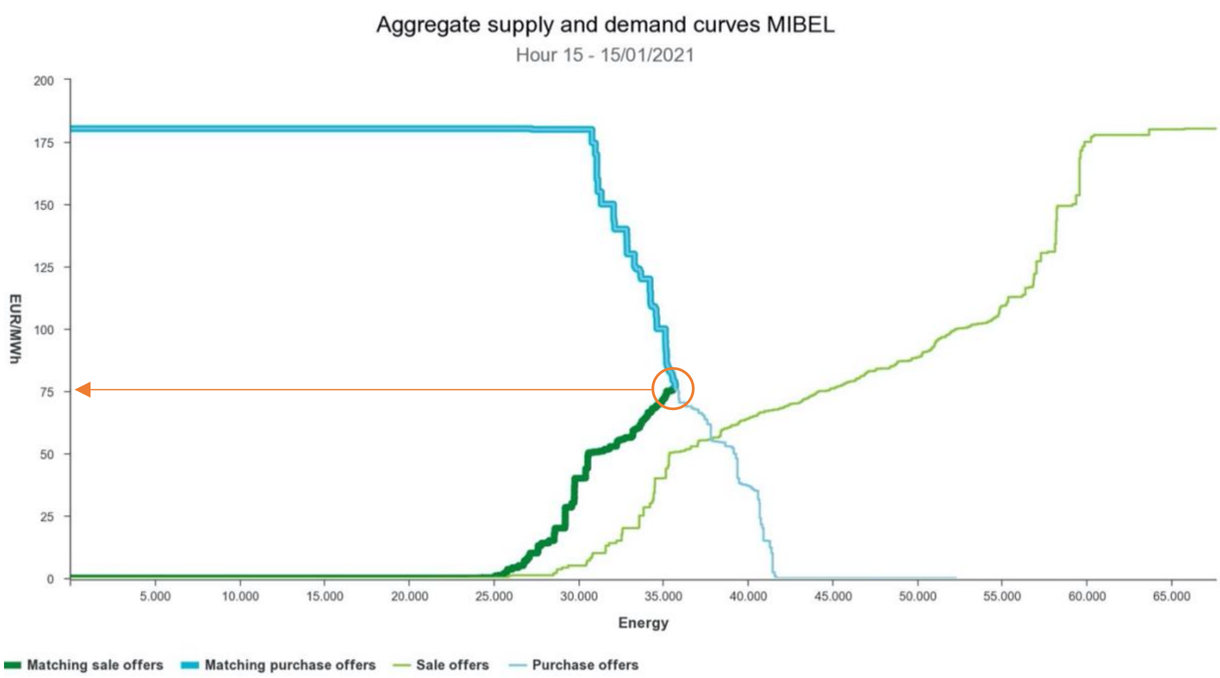


Figure 2 - Supply and demand curves for hour 15 of day 15/01/202, [9]

2.2 Iberian generation portfolio

In this section the reader should gain more closure on the Iberian electricity mix in-terms of each country's national capabilities separately.

2.2.1 Portugal

Starting with Portugal the total installed capacity till August 2021 is 21.2 GW with an increase of 0.8 GW from the previous year 2020. Table 1 illustrates the installed capacity distribution in MW for Portugal with reference to August 2021.

Table 1 Portuguese installed capacity distribution up to August 2021 [7]

INSTALLED CAPACITY

	Aug 21 [MW]
Hydro	8 917
Wind	5 246
Solar	1 113
Biomass	703
Natural Gas	4 578
Coal	576
Other Thermal	28
TOTAL	**21 160**
Pumping	3 148
Consumption	

In figure 3 bellow a comparison between the annual production of electricity in Portugal between year 2019 and 2020. It can be seen that in 2020 the renewable production supplied 60% of the total national consumption, which is compared to a 51% in the previous year. It is also important to note that these percentages in 2020 are the highest ever recorded values for renewable sources annual generation in Portugal.

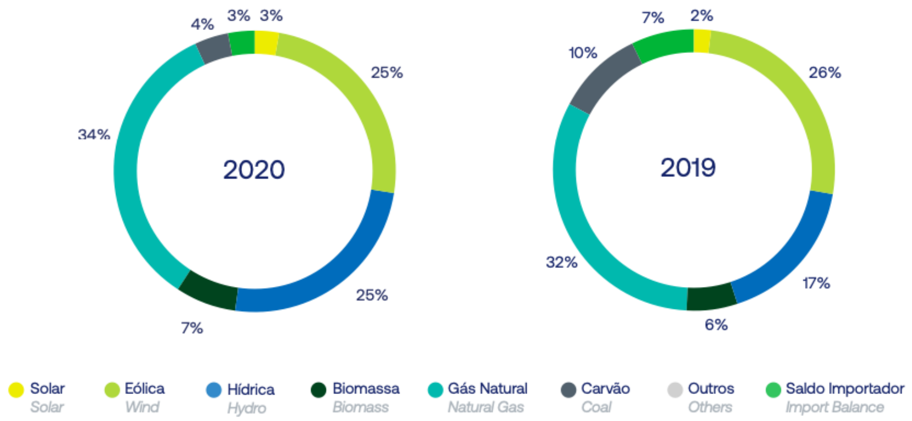


Figure 3 Portuguese energy generation structure comparison between 2019 and 2020 [10]

Since the annual data for year 2021 are not published yet, an estimation of the generated electricity for the period starting January till August 2021 is previewed, Portugal generated around 31,595 GWh of electricity, from which 68.8 % were generated from renewable sources.[11] The figure bellow illustrates the contribution of each technology to the electricity production.

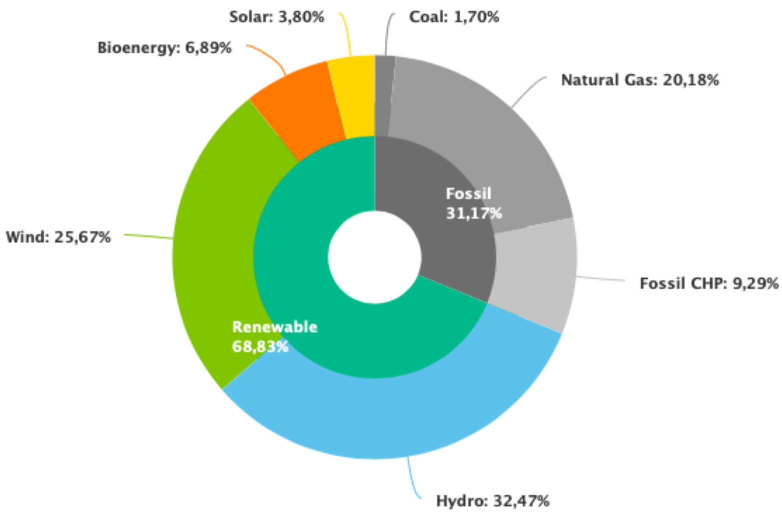


Figure 4 Portuguese total Energy generation from January till August 2021 [11]

2.2.2 Spain

When it comes to Spain, the installed power capacity up till August 2021 is equivalent to 112.2 GW. Which is 6.6 GW higher than the previous year 2020. Figure 5 illustrates the installed capacity distribution for Spain.

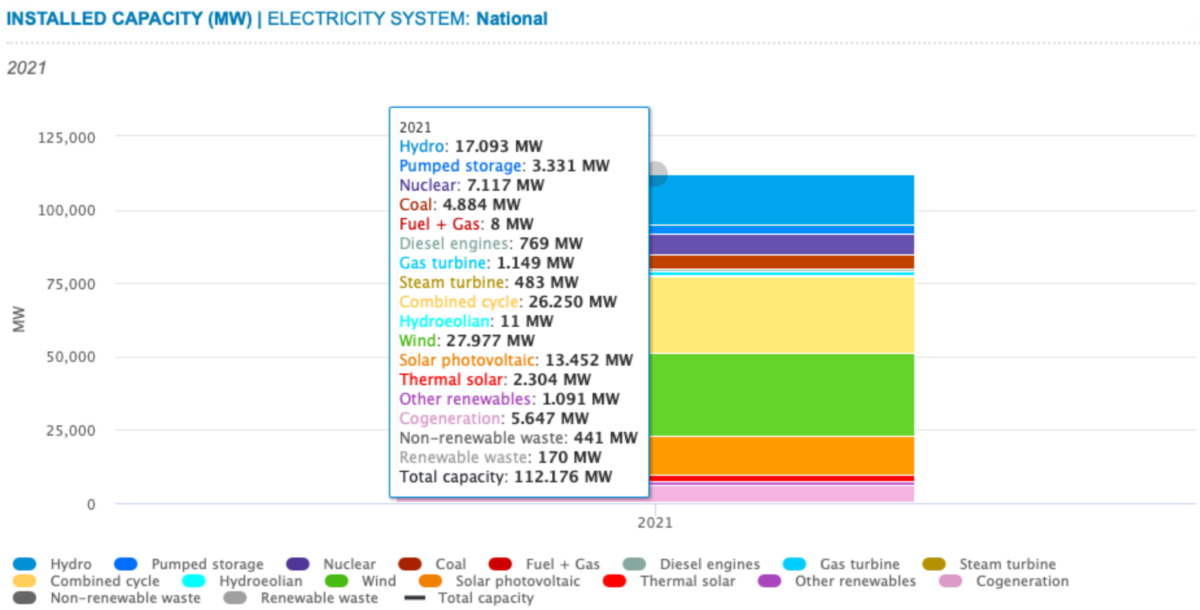
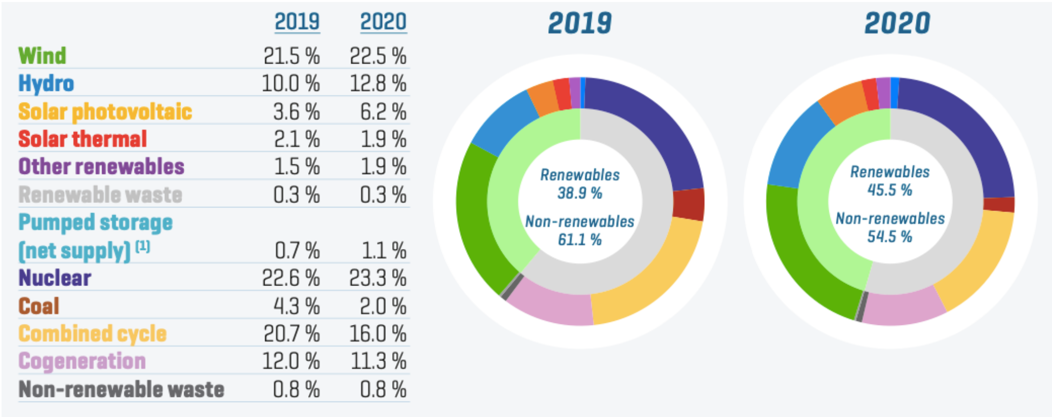


Figure 5 Spanish installed capacity distribution from January till August 2021 [8]

In the figure bellow another comparison between the annual production of electricity in Spain between year 2019 and 2020. It can be seen that in 2020 the renewable generation supplied 45.5% of the total national consumption, which is compared to a 38.9% from the previous year.



	2019	2020
Wind	21.5 %	22.5 %
Hydro	10.0 %	12.8 %
Solar photovoltaic	3.6 %	6.2 %
Solar thermal	2.1 %	1.9 %
Other renewables	1.5 %	1.9 %
Renewable waste	0.3 %	0.3 %
Pumped storage (net supply) [1]	0.7 %	1.1 %
Nuclear	22.6 %	23.3 %
Coal	4.3 %	2.0 %
Combined cycle	20.7 %	16.0 %
Cogeneration	12.0 %	11.3 %
Non-renewable waste	0.8 %	0.8 %

Figure 6 Spanish Energy Generation structure comparison between 2019 and 2020 [12]

Again, an estimation of the generated electricity for the period starting January till August 2021 is previewed for Spain in figure 7. The country's power system generated around 183,533 GWh of electricity, from which 48.5% were generated from renewable sources. Showing an increase in the shares of renewables of 4.5% compared the same time of the previous year 2020.

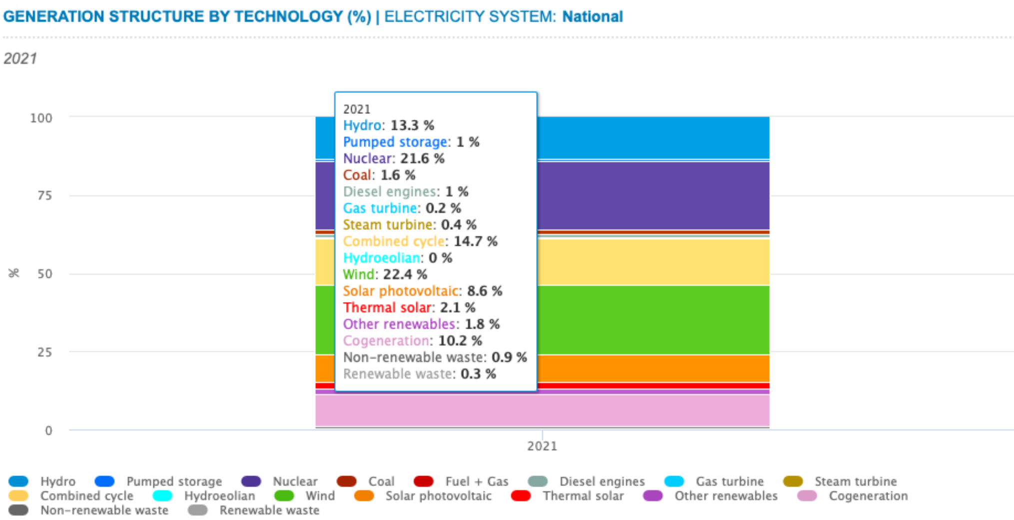


Figure 7 Spanish Energy Generation structure for 2021 [8]

Chapter 3

3 Literature Review

Modelling of electricity markets is an effective means of testing and evaluating the market design prior to its deployment, and therefore limiting problems before they occur. Also, it is an important tool as it can provide answers to a variety of concerns related to complex market scenarios and what if situations. Not only that, but it is very essential to have accurate forecasts since utility companies rely on these forecasts to operate. Moreover, long term planning of future investments or political initiatives and programs, like the European targets for implementing more renewable energy in the power system, all require reliable techniques and models to simulate the markets and forecast the electricity price to be able to take solid decisions and achieve the desired targets.

In the following sections of the literature review, the main categories of modelling techniques are reviewed. Shedding the light on the main approaches and methods used in electricity price forecasting and markets modelling. Next, different forecasting horizons are explored. Studies covering short, medium and long term forecasts are discussed. Last but not least, several studies that follow a similar path as the one executed in this thesis are reviewed to provide an intuition regarding forecasting of supply and demand curves.

3.1 Overview of modelling methods

Due to the fact that electricity markets are becoming more and more liberalized, and data is more disclosed, modelling of the electricity markets in general and electricity prices in particular is being more complex than before. Therefore, a wide variety of forecasting models were developed during the past decade to tackle this issue and provide insights on the future markets and their behaviour. Multiple methods and ideas have been developed and tested for price forecasting with a varying degree of success. For example, in [13] Weron categorized electricity price models into 5 different categories that can be reviewed as follows:

3.1.1 Multi-Agent Models:

Multi-agent models are models that simulate the operation of a system of heterogenous agents taking into consideration their interactions between each other and then forecast the price by matching the supply and demand. Unlike other models multi-agent focuses more on the optimization and equilibrium between the supply and demand to predict prices. Which is the reason why the multi-agent technology is one of the leading modelling approaches to simulate the electricity markets due to its ability to simulate complex and dynamic situations. Each agent has the ability to represent an entity taking in consideration the entity´s behaviour, goals and other agent's actions. This makes multi-agent tools more complex than standard methods [14].

Wholesale electricity price forecasting used to be a straightforward process. This process is composed of matching the required demand to the supply, which is calculated from the summation of the existing and planned generation and stacked in order of the unit's operating cost. These models named cost-based models, had the ability to forecast hour by hour electricity prices. Being compatible only for regulated and stable structured markets that faced little price uncertainty and not much competition. This can be justified as they ignored the strategic market practices. To overcome these issues the equilibrium approach added strategic bidding considerations to the cost-models, enabling the prediction of market prices without the need of historical prices. but only market concentration and known supply costs. From here the adaptative agent-based models started to take over due to their ability to address other features of the markets that the equilibrium models used to neglect. In [15] three main modelling trends were identified. These trends are: simulation, equilibrium and optimization models. Based on this classification, optimization models focused on optimizing bidding strategies and maximizing profits for one of the competing firms in the market. Since it is not in the scope of electricity price forecasting it will not be reviewed in detail. For a more comprehensive discussion about that topic check (Ventosa, Baíllo, Ramos & Rivier, 2005) [15]. On the other hand, equilibrium models like (Nash-Cournot, supply function equilibrium) involved the overall electricity market behaviour and market competition amongst all players. When the problem of concern is too complex to be solved with an equilibrium model framework, simulation models are used as an alternative.

3.1.1.1 Equilibrium Models

In this framework electricity is considered a homogenous good and the decisions of the suppliers is used to determine the market equilibrium. A drawback to this framework is that it tends to predict higher prices of electricity than the ones observed. The concept of conjectural variations was introduced to address this problem. Stating the fact that competitors produce more electricity as a reaction of the high prices. See (Day, Hobbs & Pang, 2002) [16].

The second method of equilibrium models is the supply function equilibrium (SFE). SFE models the electricity price as the equilibrium of supply and demand curves in the wholesale market. To calculate the SFE, a set of differential equations are solved unlike the Nash-Cournot framework where algebraic equations are used. The use of both of these equations raised some concerns about their numerical traceability. Also, in order to speed the computations, the demand is aggregated into blocks which in turn disregards the extreme point from the analysis, creating limitations when forecasting electricity prices for the purpose of risk management [13].

Another uncommon equilibrium approach is the strategic production-cost model (SPCM). It was proposed in (Batlle & Barquin, 2005) [17] as a modified version of the traditional production-cost model. The SPCM included the bidding strategies of the agents in the market in the modelling process. For each agent the profit is maximised, taking into account the predicted behaviour of other market competitors and the agent's costs. In comparison with the Nash-Cournot framework and the SFE, the main improvement is the computational speed which makes it a good choice for real time analysis.

3.1.1.2 Agent-based Simulation Models

As explained above the equilibrium methods usually rely on a system of either algebraic or differential equations which are often very hard to solve and eventually resorting to heuristics to solve the problem [15] [16]. Also, these modelling approaches suffer from some limitations regarding how market competition between players can be represented. From here agent-based simulation models (ABSM) found a way to have an advantage over the other mentioned equilibrium models, as ABSM don't suffer from these limitations at the same time they are not much harder to compute.

In paper [18], one of the primary applications of ABSM was discussed to model the strategic behaviour studied in electricity markets. The study tested several market designs that came handy with the changes that took place in England and whales electricity markets, and it concluded that the hourly bidding using the pay as bid system resulted in the highest prices on the other hand the daily bidding accompanied with uniform pricing resulted in the lowest prices.

A multi-agent simulation tool was developed and introduced to the scientific community in 2003 called MASCEM (Multi-Agent Simulator for Competitive Electricity Markets) [19]. This simulation tool was designed to study the behaviour and evolution of the electricity markets. As mentioned before MASCEM is a multi-agent tool meaning that market entities are represented as agents, such as consumers, generators, trades and market operators. Each of these agents set their own decision rules and objectives. The simulation tool also works as a decision support tool as it provides several negotiation options to the user which can be simulated indicating the best negotiations. MASCEM allows simulation of the day ahead market (pool) including 24 negotiations per day and the agents participating are: sellers, buyers, trades, market operator and network operator. Also, bilateral contracts, balancing market and ancillary services are simulated.

Nowadays Agent-based models are often used as an element of a bigger more complex hybrid electricity price forecasting system, instead of being the main forecasting model. For example, in [20], a monitoring system consisting of a multi agent simulator and a price forecasting module that delivers the inputs to the simulator, these two components work together to form a hybrid system that is able to predict future market scenarios, clearing and production information.

Even though agent-based models are very flexible analysing tools, these degrees of freedom also have a drawback since lots of assumptions are made and justified, which increase modelling risk.

3.1.2 Fundamental Model

Fundamental models mainly focus on the impacts of important economical and physical factors on the electricity price that are present in production and trading processes of electricity. Fundamental drivers such as (demand, production, installed capacities, system parameters and weather conditions) are assumed and the inputs are modelled and predicted usually by the means of statistical, computational intelligence or reduced form methods [13]. Fundamental approaches are often considered in the literature as a component of a hybrid model with neural networks, regression or timeseries. These hybrid models usually use fundamental factors like fuel prices, weather conditions, demand, Co_2 prices. etc. as inputs [21] [22]. Generally, there are two subcategories of fundamental models which are:

- Parameter rich models (PR)
- Parsimonious structural models (PS)

Both models are explained in more details in the section below. For fundamental models, both PR and PS face some challenges within the practical implementation. The first challenge is data availability. Economic and physical data for a specific market can sometimes be hard to obtain. Another challenge is the assumptions made to relate these economic and physical factors to the marketplace. Results obtained of such models are sensitive to these assumptions and a biased assumption can therefore result in inaccurate forecast. Moreover, the more complex the model is the more effort required to model these parameters.

3.1.2.1 Parameter Rich Models

This first sub-category of fundamental models are PR models. These models are usually developed exclusively to entities and their details are not revealed to the public. The majority of the results published are related to power markets that are hydro-dominant. In [23] a study on the old Norwegian power systems was developed by using a supply and demand model. Parameters like weather conditions and hydro inflow were used to explain the formation of the clearing price.

3.1.2.2 Parsimonious Structural Models

Parsimonious models are the second sub-category of fundamental models. They are known to be simple models with great explanatory predictive power. An example of PS model can be found in [24] beginning with an empirical analysis of the supply and demand curves, the study introduced a non-linear Ornstein–Uhlenbeck model for predicting spot prices. One advantage of that model is that it can exhibit price spiks, even though it is a pure diffusion jumpless model.

(Coulon & Howison, 2009) [25] used a stochastic process to develop a fundamental model for spot price electricity. The drivers used were fuel prices, generation capacity availability and demand. also, a parametric form was used for the bid function which maps these price drivers to the electricity price. Using the observed bid data, they found connections between the movements of bids and the corresponding fuel prices.

3.1.3 Reduced-form Models

The main purpose of Reduced-form models is to replicate the main characteristics of electricity prices, such as marginal distributions, future points, price dynamics and correlations between commodity prices, usually on a daily time scale. They are also widely used for risk management systems and pricing derivatives. However, if a price mechanism is used that is not sufficient for capturing the main features of electricity prices, the model's conclusions are likely to be flawed. Also, if the model is too complex, the computational burden will prevent its online use in trading departments [13]. Reduced-form models can be divided into two approaches:

3.1.3.1 Jump-Diffusion Models

Jump diffusion models are finite frequency exponential Lévy models with a finite frequency of jumps. They serve as prototypes for a wide range of complicated models. They've been used to simulate option pricing in finance for a long time. The jump diffusion models are made up of two parts: a jump and a diffusion. The Brownian motion determines the diffusion term, while the Poisson process determines the jump term. The jump portion allows to model the underlying asset's price jumps that occur suddenly and unexpectedly.

Surprisingly, reduced-form models have been observed to perform reasonably well when forecasting volatility or price spikes.

One of the biggest flaws of jump-diffusion models is that they can't show repeated spikes at the same frequency as market data. Additionally, spike clustering can be seen on both a daily and hourly time frame [13].

3.1.3.2 Markov Regime-Switching Models

The (Hamilton, 1989) [26] Markov switching model, often known as the regime switching model (MRS), is one of the most commonly used nonlinear time series models. Multiple structures (equations) are used in this model to characterize the time series behaviour in various regimes. This model can capture more complex dynamic patterns by allowing switching between these components. The switching mechanism is controlled by an unobservable state variable that follows a first-order Markov chain, which is a novel aspect of the Markov switching model.

In a fairly natural approach, unlike jump-diffusion models MRS models allow for consecutive spikes. After a spike, returning prices to the ‘normal' regime is also simple, as the regime-switching process allows for temporal changes in the model dynamics [13].

Reduced-form models aren't expected to precisely forecast hourly prices, but they should be able to recover the key characteristics of power spot prices on a daily time scale. These models are often used for derivatives pricing and risk analysis because they provide a simplified but reasonably realistic depiction of price dynamics.

3.1.4 Statistical Models

Statistical methods forecast the current price by combining historical prices with historical or present values of external variables, such as consumption, generation, fuel prices and weather condition. Additive and multiplicative models are the two most common types. They differ in whether the expected price is the product (multiplicative) of a number of factors or the sum (additive) of a number of components. The multiplicative is significantly more well-known. However, the two are related: a multiplicative model for prices can be converted into an additive model for log-prices [13].

The inability to model non-linear processes is a common criticism of this sort of approach. Their accuracy is determined not just by the algorithm, but also by the quality of the data used, which is critical for including major electricity price determinants. These approaches have a poor performance during spike times. Despite receiving some negative reviews in financial markets, they have proven to be a reliable method for forecasting electricity prices in power markets.

Statistical models are classified as technical analytical tools by certain writers. Technical analysts don't try to calculate an asset's underlying or fundamental value; instead, they look at price charts for patterns and indicators that can predict how well it will perform in the future. While the effectiveness and utility of technical analysis in financial markets is sometimes questioned, the methodologies have a better chance in power markets due to the seasonality of electricity price processes during regular, non-spiky periods.

Statistical models can be sub-categorized into:

3.1.4.1 Similar-day and Exponential Smoothing Methods

The similar-day method is a prominent benchmark model in electricity price forecasting. It works by scanning historical data for days that have similar features to the expected day and using those historical values as price estimates for the future. Day of the week, day of the year, holiday type, weather, and consumption data are all examples of similar features that can be used. The forecast could be a linear combination or a regression approach that includes numerous similar days instead of a single similar-day price [27].

Exponential smoothing is a relatively simple benchmark that is very popular in load forecasting but less common in electricity price forecasting (see, for example, Taylor, 2010) [28]. It's a practical forecasting approach in which the prediction is derived using an exponentially weighted average of previous observations. The idea here is that forecasts are not computed from consecutive previous observations alone, but an independent smoothed trend and seasonal component can be added.

3.1.4.2 Regression Models

One of the most commonly used statistical techniques is regression. Regression is used to learn more about the correlations between a dependent or criterion variable and several independent or predictor variables. The sum-of-squares of the variances between observed and predicted values is reduced in multiple regression, which is based on least squares. Multiple regression, in its most basic form, assumes that the connection between variables is linear. Despite the abundance of forecasting approaches, regression models remain one of the most popular electricity price forecasting methods [13].

3.1.4.3 Auto-Regressive Time Series Models

The Auto-Regressive Moving Average (ARMA) model is a typical time series model that accounts for the random nature and time correlations of the phenomenon under investigation. The present value of the price X_t in the ARMA(p,q) model is expressed linearly in terms of its p previous values (autoregressive component) and the noise's q previous values (moving average component) [13].

3.1.4.4 ARX-type Time Series Models

The Auto-Regressive time series models discussed in the previous section, relate the signal under examination to its own history without explicitly using information from other time series. However, the current and historical values of numerous exogenous elements, most notably generation capacity, demand profiles, and ambient meteorological conditions, have an impact on electricity pricing. Time series models ARX containing exogenous (X), or input variables can be utilized to capture the relationship between prices and these fundamental factors [13].

It can be difficult to tell the difference between regression and ARX-type models. But it can be said that the models are classed as regression models if the number of fundamental regressors is big. However, if the autoregressive structure is complex, they should be classified as ARX-type models.

3.1.4.5 Threshold Autoregressive Models

Threshold models (TAR) are not only utilized in time series analysis, but they are also applied in other areas of statistics. The general idea is that when the values of a variable exceed a specific threshold, a process may act differently. That is, when values exceed a threshold, a different model may be used than when values are below the threshold. (Tong & Lim, 1980) [29] proposed the first TAR model in 1980. The regime is assumed to be defined by the value of an observable variable v_t in relation to a threshold value T. In (Weron & Misiorek, 2006) [30], the predictive power is evaluated for a TAR and TARX (with the system-wide demand as the exogenous variable) models. The study is conducted on the California market. The price for hour 24 on the previous day is used as the threshold variable v_t in the TAR(X) models, and the threshold level is estimated for each hour in a multi-step optimization approach with ten equally spaced starting points spanning the whole parameter space. The out-of-sample predicting performances were significantly below acceptable levels, and the models even failed to outperform the naive approach.

3.1.4.6 Heteroskedasticity and GARCH-type models

Previously explained Linear AR(X)-type models are assumed to be homoscedastic, with a constant variance and covariance function. From an empirical standpoint, financial time series including electricity spot prices exhibit different sorts of nonlinear dynamics, the most important of which is the time-series' significant dependency on its own past. Some of these series' non-linearities are due to a non-constant conditional variance, and they are characterized by the clustering of big shocks, or heteroskedasticity [1]. (Engle's, 1982) [31] Auto-Regressive Conditional Heteroskedastic (ARCH) model was the first formal model to successfully handle the problem of heteroskedasticity. The conditional variance of the time series is represented by an autoregressive process in this model, which is a weighted sum of squared prior observations. In real applications, the calibrated model's order turns out to be quite large. However, if we let the conditional variance depend not only on the previous values of the time series, but also on a moving average of past conditional variances, the resulting model allows for a more concise representation of the data.

3.1.5 Computational Intelligence.

Computational intelligence (CI) is a very diverse set of nature-inspired computational techniques that have been developed to solve issues that traditional methods (e.g., statistical) cannot efficiently handle. CI combines components of learning, evolution, and fuzziness to produce approaches capable of adapting to complicated dynamic systems, and hence may be considered "intelligent" in this sense. CI is an interdisciplinary field that includes Artificial Neural Networks (ANN), Fuzzy Systems, Evolutionary Algorithms, and hybrid paradigms. These approaches are versatile and can handle complex and non-linear issues, with multiple researchers claiming that they perform well in forecasting electricity prices. The ability to handle complexity and non-linearity is a major strength of computational intelligence techniques. In general, CI methods outperform the statistical techniques outlined in Section 3.1.4 at simulating the aspects of power prices. At the same time, their adaptability is also considered sometimes a flaw. The ability to adapt to non-linear, spiky behaviour does not always imply improved point forecasts.

According to (Hobbs et al,1998) [32], when compared to standard forecasting techniques, ANN are frequently the most accurate forecast tool, especially when dealing with nonstationary, nonlinear, discontinuous, and complex problems. In the same research, a survey of 18 electric utilities and 5 gas utilities is conducted to assess forecasting accuracy in energy planning using ANN. Economic gains were found from all utilities that utilize ANN on a daily basis, such as time savings because ANN are simple and fast to operate, and improvements in forecasting accuracy, allowing for a better and more precise management process and avoiding unnecessary money lost.

3.1.5.1 Feed-forward Neural Networks

A single-layer perceptron is the most basic feed-forward neural network, it has no hidden layers and is comparable to a linear regression. A linear combination of the inputs generates the forecasts. The used weights (which correspond to the regression coefficients) are chosen using a learning algorithm, which minimizes some cost function, such as the mean squared error. The non-linear multi-layer perceptron (MLP) can be created by adding an intermediate layer with hidden nodes. This is the most common type of feed-forward network, which contains neurons grouped into layers with unidirectional connections between them. that is, the outputs of one layer's nodes are inputs to the next layer's nodes.

(Kohzadi et al.,1996) [33] and (Zou et al.,2007) [34] compared FFNN to conventional time series models such as Auto Regressive Integrated Moving Average (ARIMA) when forecasting wheat and live cattle prices. According to the study, ANN outperforms statistical time series models, with one justification being the non-linearity and high volatility seen in data, which the linear ARIMA model cannot capture.

3.1.5.2 Recurrent Neural Networks

Feed-forward networks are classed as static since they produce only one set of output values from a given input, rather than a sequence of values. They are also memoryless, meaning that their response to an input is unaffected by the prior state of the network. On the other hand, Recurrent networks, are dynamic systems. The neuron outputs are computed when a new input pattern is provided. The inputs to each neuron are adjusted as a result of the output feedback, causing the network to enter a new state. Recurrent neural networks (RNN), like feedforward and convolutional neural networks, learn from training input. They are differentiated by their "memory", which allows them to impact the current input and output by using information from previous inputs. While typical deep neural networks presume that inputs and outputs are independent of one another, the output of RNN is dependent on the preceding items in the sequence. While future occurrences might also be useful in determining the outcome of a given sequence [13].

3.1.5.3 Fuzzy Neural Networks

Fuzzy neural networks are an example of a hybrid approach that combines a neural network's learning capacity with fuzzy logic's noise-handling capability [35] [36]. In its most basic form, a fuzzy neural network is a three-layer feedforward network with a fuzzy input layer (fuzzification), a hidden layer containing the fuzzy rules, and a final fuzzy output layer (defuzzification). Fuzzy sets are confined within (fuzzy) connections between layers, though a five-layer network with sets included in the second and fourth levels can occasionally be found. (Hong & Hsiao, 2002) [37] demonstrated one of the first applications of fuzzy logic to electricity price forecasting, classifying historical data into three clusters (peak, medium, and off-peak) and then employing a recurrent network for forecasting.

3.1.5.4 Support Vector Machines

The support vector machine (SVM) is a classification and regression tool based on Vapnik's [38] statistical learning theory from 1995. SVM makes a non-linear mapping of the data into a high-dimensional space, then employs basic linear functions to generate linear decision boundaries in the new space, in contrast to ANNs, which attempt to build complex functions of the input space. SVM has the advantage of providing a single solution that is defined by the global minimum of the optimized function, as opposed to several solutions associated with local minima, like ANNs do. In addition, they rely less on heuristics and have a more flexible structure [39]. SVM has been widely used in pattern classification and non-linear regression applications. SVM classifiers can be used to forecast future trends after they have been trained. The term prediction has a distinct meaning in the context of SVM, as [40] point out. 'Prediction' refers to a two-step supervised classification process: A SVM is trained as a classifier using a subset of the data, and then used to predict the remaining data in the data set.

The applications of SVM in energy price forecasting are typically those of elements in hybrid systems. However, (Sansom, Downs, & Saha, 2002) [41] compare an MLP and an SVM with the same inputs and conclude that the SVM delivers more consistent forecasts and requires less time for optimal training.

3.2 Forecasting horizons

When it comes to energy price forecasting, there are time-range classifications in terms of how far into the future estimates are made. The literature on this topic is not clear, being ambiguous and with varied interpretations among the authors. There are three classifications of time horizons: short-term, medium-term, and long-term.

Weron [13] describes short-term as period ranging from a few minutes to a few days, with the majority of it being tied to day-ahead market forecasts. The medium-term prediction spans a few days to a few months and is frequently used for risk management and derivatives pricing. The long-term forecast spans a few months to several years and is frequently associated with strategy, planning, and investment.

Ziel et al. [42] propose an alternative definition, indicating that short-term refers to time spans up to one month, medium-term to one year, and long-term to time periods more than one year. The forecast horizon must be defined in order to run a good model. Techniques and algorithms differ across all three-time frames.

During the literature research, it was discovered that the number of published algorithms and techniques for short-term forecasts vastly outnumbers the number of published algorithms and approaches for long-term forecasts. This is consistent with the literature assessment conducted in [42], which states that the proportion of long-term forecast approaches and published articles is irrelevant when compared to short-term forecast techniques. Only 8% of the 710 papers examined were related to mid and long-term forecasting, with the number of research directly dedicated to long-term forecasting being substantially smaller than that of medium-term forecasting.

3.2.1 Short Term Forecasting

The number of techniques and algorithms available for performing short-term projections is tremendous. Some of them will be presented in this subsection, ranging from simple models like FFNN to hybrid models that integrate two or more techniques.

To estimate short-term electricity prices, Catalo et al. [43] presented a FFNN. This is the most well-known and basic neural network. This algorithm modifies the weights using the backpropagation algorithm, comparing the anticipated output with the real output (supervised learning), in order to minimize the error, given only historical market prices as input. Prior 42 days of historical pricing are necessary to anticipate the price at day "D." The model was tested in the Spanish and Californian markets, with results compared to the ARIMA and NAVE models. The results reveal that FFNN outperforms the other two models in every scenario.

For the Turkish power market, Ugurlu et al. [44] present an excellent review comparing alternative neural network architectures and statistical models. Traditional methods such as Markov, Nave, Self-Exciting Threshold Auto-Regressive (SETAR), and Seasonal Auto-regressive Integrated Moving Average (SARIMA) are compared to computational intelligence methods such as Convolutional Neural Networks (CNN), Long Short-term Memory (LSTM), and Gated Recurrent Unit (GRU) methods in this study. When CI techniques are compared to traditional methods, the results reveal that they clearly outperform them. Another interesting finding is that error is lower in the autumn and winter months than in the spring and summer, illustrating the impact of temperature and seasonality.

3.2.2 Medium Term Forecasting

In the power market, medium-term price forecasting is important. Accurate medium-term energy price forecasting is crucial in many applications, including maintenance scheduling, generation expansion planning, and bilateral contracting. Because of the wide forecasting horizon, the reliance of medium-term power prices on different variables, and the scarcity of explanatory data, medium-term price forecasting is a difficult undertaking.

Medium-term load forecasting models based on climate parameters are provided in [45] and [46]. They forecast load demand in Spain and Greece, respectively, by combining climate elements such as humidity and adjusted temperature data in the form of the number of days in which the temperature is above or below a marginal value for each month and thereby forecasting load consumption.

[47] divides the paper's research into statistics and fundamental models. The dependent parameters for the demand and supply sides are then computed directly using both forecasted climatic data and past supply and demand data, using favourable aspects of both models. Although this model is correct, not all of the data they utilized is publicly available, particularly in the new deregulated privatized market, on both the supply and demand sides. Furthermore, their model is based on climate factors that are prone to being misleading due to the unpredictability of weather variations, particularly in the recent three years, which increases the chance of mistake.

3.2.3 Long term forecasting

When compared to short-term forecasting, long-term forecasting has gotten less attention in the literature, with few studies on the subject. One reason is the unpredictability surrounding long-term pricing drivers such as future fuel costs, policies, political interference, technical changes, energy mix, grid operations/developments, and so on. Long-term electricity price behaviour is strongly dependent on electrical system investments, as well as the evolution of numerous elements such as demand, subsidies, fuel costs, carbon prices, support schemes, green taxes, energy mix and grid investments.

Mohammadi [48] studied the association between long-term electricity pricing and the three most popular fuel sources (coal, natural gas, and oil). This assessment was made on the electricity market in the United States, utilizing data from 1960 to 2007. The study's findings demonstrate that power prices are closely connected with coal prices, with coal being the primary price driver among the three fuel sources discussed. Natural gas, while showing some link, is not as strong as coal, and it is ultimately found that oil has no long-term impact on electricity pricing.

Kotur et al. [49] propose an FFNN for anticipating long-term electricity prices. ANNs utilizing historical prices as input to forecast future prices is not an effective technique in long-term ranges, according to this study; more information is required. The author employs physical attributes and real-world inputs from the electrical system, including as generation from convectional and non-conventional technologies, imports and exports, demand, seasonal and daily time indicators. The model is applied to the British power market, where the forecast is made for 15000 samples on an hourly basis, yielding a forecast of nearly two years.

Rahul [50] presented a revolutionary method for long-term load forecasting at the hourly level. The model is based on a Recurrent Neural Network composed of (LSTM-RNN) cells. The suggested model is found to be extremely accurate, with a Mean Absolute Percentage Error (MAPE) of 6.54 within a confidence interval of 2.25 percent, making it suitable for offline training to anticipate electrical load over a five-year period.

Hossein [51] investigated the problem of long-term load forecasting for the New England Network case study using a variety of commonly used machine learning methods such as FFNN, SVM, RNN, generalized regression neural network, k-nearest neighbors, and Gaussian Process Regression. The outcomes of these strategies are compared using the mean absolute percentage inaccuracy (MAPE).

LianLian presented an efficient approach for day-ahead power price forecasting (EPF) based on a recurrent neural network model with (LSTM). For sequential data, the applied approach is capable of learning features and long-term dependencies of previous information on current forecasts. The proposed strategy has been successfully implemented to the Australian market in the Victoria region as well as the Singapore market [52].

3.3 Forecasting of Supply and Demand Curves

Almost all of the previously mentioned models and techniques for forecasting of electricity price mainly focus on the price time series and neglect the root mechanism of determining the price, which is the supply and demand curves, which represent the quantities of electricity traded in an exchange. These two curves do not only contain all the information needed to determine the price but also additional information on other prices for other market volumes. Which is important when it comes to determining extreme price movements. Modelling and forecasting the electricity prices by using real auction data is considered to have lots of potential yet not fully explored. In this section relevant studies will be reviewed discussing different forecasting approaches of future electricity markets, focusing on supply and demand curves approach to forecast the electricity prices and predict extreme price movements.

In [53] one of the first research that developed a model that uses real auction data of an electricity market along with an Ornstein-Uhlenbeck process to produce a model that can account for price spikes.

In [54] a promising study modelled the supply and demand curves to obtain the clearing price. In this study it was assumed that the demand curve followed a linear function, but the supply curve followed a non-linear function to build a price quantity model. Factors as gas prices, gas supply and temperatures were used to approximate the market curves.

A paper researched and developed a forecasting method to predict the bidding curve of generation players in the Iberian electricity market, MIBEL [55]. This forecasting model is constructed as a two-step artificial neural network (ANN) prediction model. The first step model works on the prediction of the amount of energy to be bid at zero price for a certain hour. The second step model involves the prediction of rest of the bidding curve. The accuracy of the trained ANN model is assessed by the determination coefficient R, Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE).

Study [56] proposes another approach for predicting the electricity price by forecasting supply and demand curves. This approach includes modelling and predicting of hourly supply and demand curves and the location of the intersection point to obtain the equilibrium volume and market price. This methodology is developed by using functional data analysis methods like parametric and non-parametric functional auto regressive models (FAR). This study is performed on the Italian electricity market (IPEX). In order to build the hourly demand and supply curves all individual bids and offers are considered for every auction. The model starts by accumulating raw bids and offers to obtain the empirical demand and supply curve which is then converted to smooth functions using a basis function.

A functional auto regressive model is used for the prediction based on past observed curves. To evaluate the prediction sample Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are calculated.

In [57] a study conducted on the Italian electricity market presents a methodology to simulate the future electricity market by using hourly generation bids data sets. This method is capable of providing a deeper insight on the bidding behaviour of the generation participants, which gives an advantage over the historical time series forecasting of electricity price. Based on the future forecasted demand and supply the clearing price can be determined by the intersection of both curves.

Another study [58] researched the German and the Austrian day ahead electricity market, aimed to develop a model to forecast electricity prices by using the supply and demand curve approach instead of directly forecasting the electricity price time series. The model can be referred to as the X-model which combines the perceptions of market structure and econometric analysis. A stochastic model forecasts the bid volume of each price class. Finally, the supply and demand curves are calculated and by getting the intersection point the clearing price can be obtained.

Chapter 4

4 Methodology

This chapter provides an overview over the methodology used in this research and guides the reader through the different phases to give a better understanding of the developed model.

As mentioned before in chapter 1, this paper focuses on the wholesale day-ahead MIBEL market. Real hourly market data coming from hundreds of generators has been collected and processed. This hourly data is then fed into an artificial neural network for the initial aim of creating a model that can predict the quantities and prices of the electricity bid into the market, in other words the supply curve. Due to the in-availability of some required market data, the objective was simplified to model the supply curve but instead of categorizing bids in terms of their technology type, bids are categorized based on two price segments: zero and non-zero price. Therefore, the model's outputs are the total quantities of electricity bid at zero and non-zero price segments. Which is a first step towards the main goal of modelling the supply curve. The proposed methodology can be described by the following steps:

1. Data collection.
2. Data treatment and preparation.
3. Choosing the future scenario and projection to use.
4. Artificial neural network model (ANN).
5. Model validation and final results.

Figure 8 represents a schematic view of the proposed methodology.

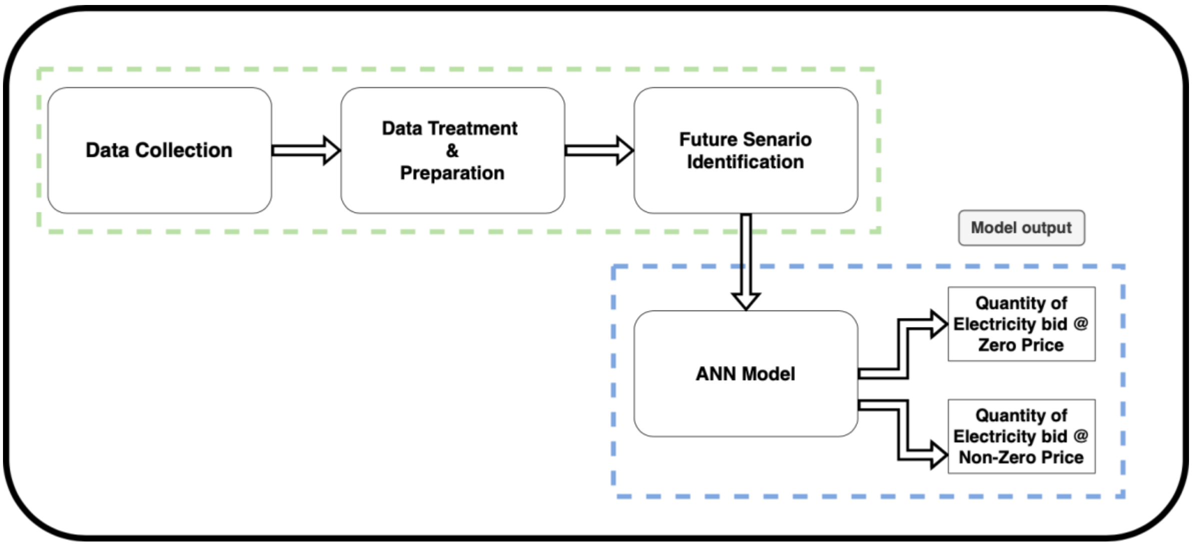


Figure 8 Schematic view of the proposed methodology

4.1 Data collection

As it was mentioned before in chapter 2, the MIBEL market is operated by Portugal and Spain. Therefore, to evaluate and model the hourly supply curve of the day-ahead market it is mandatory to gather information with reference to both countries. Historical hourly bids and historical hourly electricity production per technology were collected from OMIE [6], which is the Spanish market operator. OMIE only provided two years and half of hourly data. From 2019 till June 2021. In this period over 900 files containing around 81,000,000 data samples are collected and processed. Samples of the historical data files are shown in the tables below.

In table 2 a sample of a file containing the hourly generated electricity in (MWh) categorized by technology type for the day 14/09/2019 is previewed.

Table 2 Sample of a file containing hourly generated electricity by technology for day 14/09/2019

OMIE - Mercado de electricidad									
Energía horaria por tecnologías (MWh)									
Fecha	Hora	CARBÓN	NUCLEAR	HIDRÁULICA	CICLO COMBINADO	EÓLICA	SOLAR TÉRMICA	SOLAR FOTOVOLTAICA	COGENERACIÓN/RESIDUOS/MINI HIDRA
14/09/2019	1	832.0	6 977.4	823.1	5 679.9	10 255.6	60.1	7.9	5 155.9
14/09/2019	2	726.0	6 883.4	555.3	4 314.0	9 816.5	30.1	7.9	4 902.4
14/09/2019	3	655.0	6 722.4	546.5	4 012.0	9 358.4	45.1	7.9	4 935.9
14/09/2019	4	475.0	6 701.4	468.5	3 929.7	8 943.4	45.1	7.9	4 942.8
14/09/2019	5	475.0	6 721.4	468.8	3 930.2	8 635.9	45.1	7.9	4 966.6
14/09/2019	6	475.0	6 742.4	470.6	3 849.6	8 415.4	45.1	7.9	5 004.6
14/09/2019	7	445.0	6 763.4	594.1	4 036.0	8 267.6	45.1	14.7	4 979.8
14/09/2019	8	475.0	6 784.4	672.9	4 652.6	8 153.9	100.1	24.1	5 047.3
14/09/2019	9	475.0	6 807.4	1 145.0	5 532.4	8 030.0	100.4	231.5	5 082.3
14/09/2019	10	580.0	6 829.4	1 660.5	6 558.2	7 636.2	173.3	900.3	5 171.1
14/09/2019	11	580.0	6 852.4	1 668.0	6 441.4	7 218.1	251.6	1 613.5	5 287.9
14/09/2019	12	580.0	6 872.4	1 812.5	6 493.8	7 033.6	374.6	2 212.8	5 388.5
14/09/2019	13	580.0	6 891.4	1 730.3	6 625.1	6 857.6	423.4	2 607.2	5 449.8
14/09/2019	14	580.0	6 910.4	1 536.5	6 747.0	7 067.3	491.0	2 790.0	5 424.1
14/09/2019	15	580.0	6 929.4	1 020.2	6 304.9	7 457.6	531.9	2 776.1	5 406.6
14/09/2019	16	475.0	6 947.4	873.4	5 620.1	7 917.6	502.8	2 592.6	5 216.1
14/09/2019	17	475.0	6 962.4	1 072.5	5 048.3	8 056.6	531.3	2 291.2	5 159.8
14/09/2019	18	475.0	6 961.4	1 093.9	5 599.0	8 140.8	509.3	1 820.7	5 209.4
14/09/2019	19	580.0	6 963.4	2 284.2	6 806.1	8 283.5	532.3	1 125.4	5 249.0
14/09/2019	20	580.0	6 966.4	4 124.7	6 813.9	8 084.8	407.2	397.9	5 270.7
14/09/2019	21	580.0	6 970.4	5 807.2	6 821.2	7 708.3	303.1	55.8	5 379.7
14/09/2019	22	580.0	6 973.4	4 550.8	6 837.2	7 543.4	200.1	14.7	5 388.1
14/09/2019	23	580.0	6 978.4	2 722.3	6 662.9	7 639.0	184.1	66.9	5 330.0
14/09/2019	24	580.0	6 978.4	2 042.6	5 076.7	7 593.9	119.1	14.7	5 222.3

In table 3 another sample of a file containing historical bids for the first hour of the day 14/09/2019. The quantities of energy bided are expressed in MWh and the price is expressed in Euro/MWh. On a side note, for each day in the day-ahead market there is an average of 50,000 bids and offers.

Table 3 Sample of a file containing historical bids for day 14/09/2019

OMIE - Mercado de electricidad						
Mercado diario - Hora 1						
Hora	Fecha	Pais	Tipo Oferta	Energía Compra/Venta (MWh)	Precio Compra/Venta (EUR/MWh)	Ofertada (O)/Casada (C)
1	14/09/2019	MI	V	2.5	0	O
1	14/09/2019	MI	V	4.3	0	O
1	14/09/2019	MI	V	390.0	0.01	O
1	14/09/2019	MI	V	17.7	0.01	O
1	14/09/2019	MI	V	25.0	0.10	O
1	14/09/2019	MI	V	25.0	0.10	O
1	14/09/2019	MI	V	20.0	0.10	O
1	14/09/2019	MI	V	0.4	0.10	O
1	14/09/2019	MI	V	10.1	0.10	O
1	14/09/2019	MI	V	20.0	0.10	O
1	14/09/2019	MI	V	90.0	0.15	O
1	14/09/2019	MI	V	10.0	0.19	O
1	14/09/2019	MI	V	25.0	0.20	O
1	14/09/2019	MI	V	31.7	0.50	O

4.2 Data treatment and preparation

As mentioned in the previous section over 900 files containing around 81,000,000 data samples are collected for the purpose of this work. The majority of the files were in text format and there was a need to reformat all files to csv or excel format and clean all files from undesired data. Due to the great number of files and massive number of data samples a piece of code is developed by using python to handle these data efficiently and prepare it to be fed to the ANN model.

4.3 Future scenarios and projections

To be able to model the year 2030, a description of the future energy mix and production distribution must be provided. Three different projections were proposed by Pereira [59] that he developed by the help of EnergyPlan tool. A very brief review of the chosen projection is presented. In case of the need for more detailed information regarding the different projections the reader is advised to check Pereira's research [59]:

4.3.1 The Governmental Projection (RNC + PNIEC):

This projection consists of the Portuguese and the Spanish ministry's projections combined together. Table 4 represents the installed capacity distributions for years 2020, 2030 and 2040. And figure 9 represents a graphic though out the years.

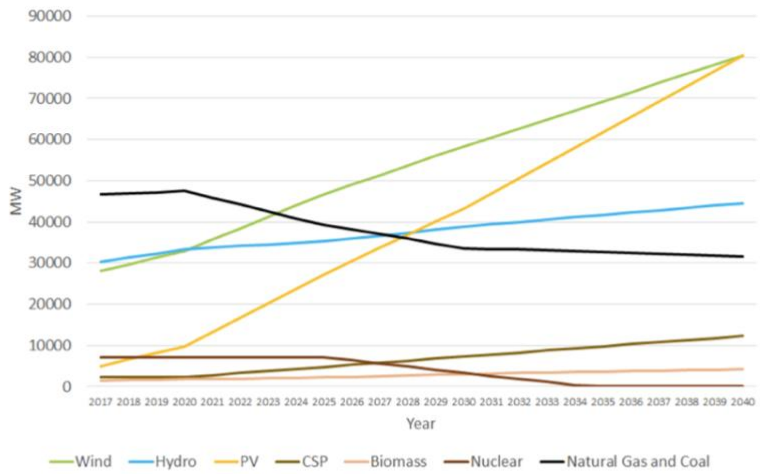


Figure 9 Technologies installed capacity for the (RNC + PNIEC) Projection, Pereira [59]

Table 4 Technologies' Capacity for Portugal and Spain for years 2020,2030 and 2040 based on (RNC + PNIEC)

(MW)	2020	2030	2040
Hydro	32,933	58,323	44,580
Wind	33,430	38,830	80,323
PV	9,804	43,277	80,333
CSP	2,303	7,303	12,303
Biomass	1,755	3,055	4,255
Nuclear	7,117	3,355	0
Natural Gas and Coal	47,493	33,655	31,709

In this study only one projection is taken into consideration as the future scenario for 2030. The chosen projection is the governmental projection (RNC + PNIEC). In Pereira's work there are 2 other projections which are the private entity projection (APREN) and the European entity projection (ENTSOE). The choice of (RNC + PNIEC) projection was due to the fact that the (RNC + PNIEC) is the only ambitious projection were nuclear and coal generation totally phase out. Coal and nuclear power plants are planned to be decommissioned by 2030 and 2035 respectively [59]. Moreover, data availability was a key point for the selection of this projection.

Table 5 represents the total annual electricity production in 2030 based on the (RNC + PNIEC) projection. And figure 10 represents the production though out the years till 2040.

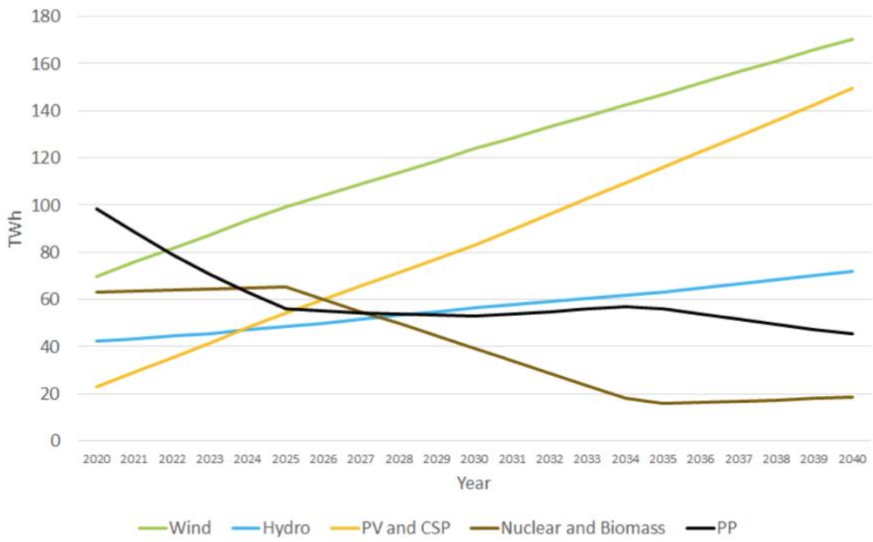


Figure 10 Total annual electricity production by technologies for the (RNC + PNIEC) Projection, Pereira [59]

Table 5 Portugal and Spain Electricity Production in 2030 (RNC + PNIEC)

Technology	2030 Production [TWh]
Hydro	56.49
Wind	123.81
Solar (PV+CSP)	82.91
Nuclear	26.21
Natural Gas	25.14
Coal	0
Other renewables (Biomass)	13.29

4.4 Artificial neural network model (ANN)

For the proposed methodology a computational intelligence technique is chosen to develop the practical component of this paper. this technique is limited to an artificial neural network algorithm used to forecast the quantities of electricity bid at different price segments for the DAEM mainly, zero and non-zero prices. Next chapter will focus more on the theoretical framework and implementation of the suggested neural network. In figure 11 a schematic view of the proposed neural network can be viewed.

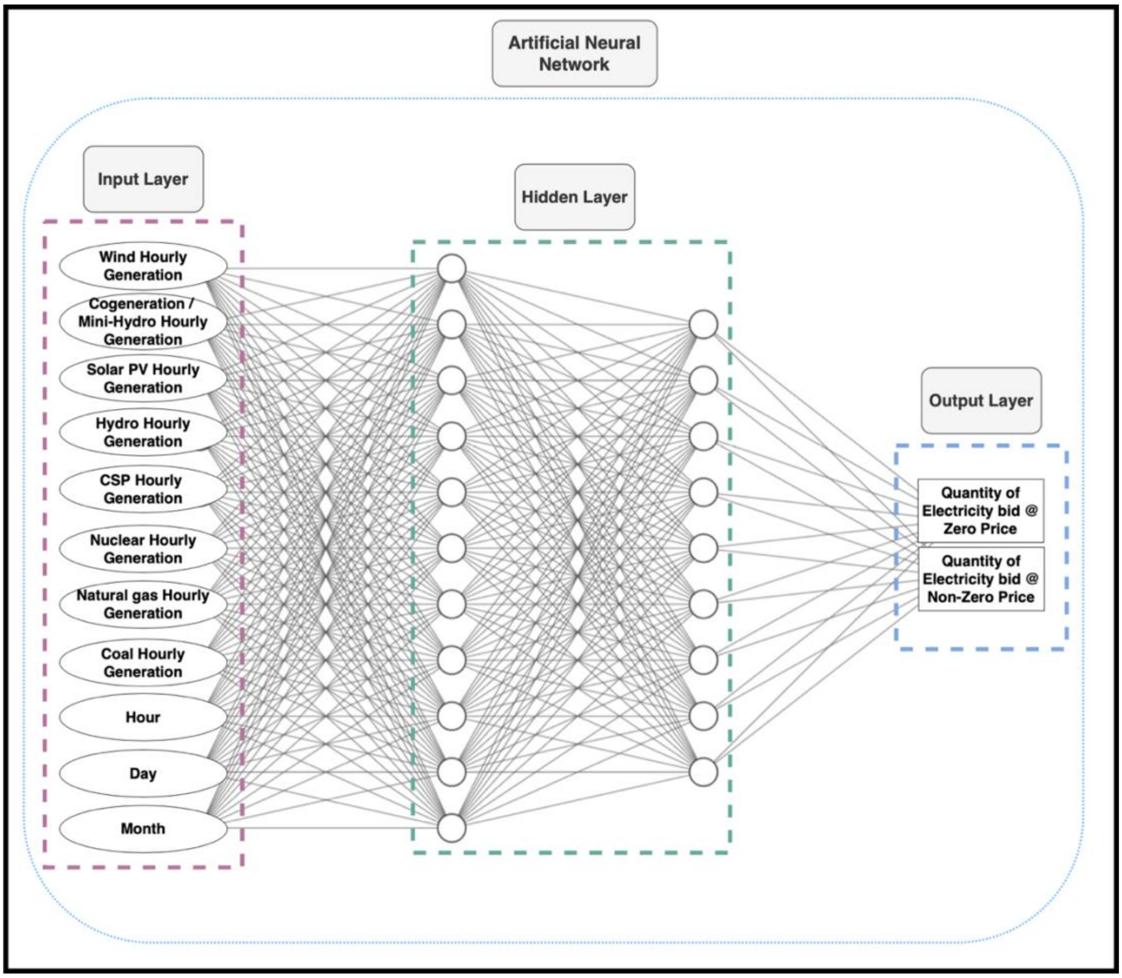


Figure 11 Schematic view of the proposed neural network

4.4.1 Model Description

In this section a brief summary of the model is presented to help the reader in gaining intuition about the model's main features. The model represented in the above figure has a main objective of predicting the total quantities of electricity bid at zero and non-zero price segments in the day ahead MIBEL market for year 2030. To achieve this goal, input variables are fed to the model to train, verify and predict for future scenarios. These input variables are discussed in detail in the next section.

4.4.2 Training and Validation Input Variables

The features shown below in figures 12 to 18, were used as the input variables to the ANN model for the training and validation sets. The training dataset is composed of years 2019 and 2020, on the other hand, the validation dataset is composed of the first 6 month of year 2021. More details regarding the validation process can be found in chapter 6. The reader should also note that all variables are with respect to the Portuguese and Spanish electricity wholesale market (MIBEL).

- Wind hourly generation (MWh):

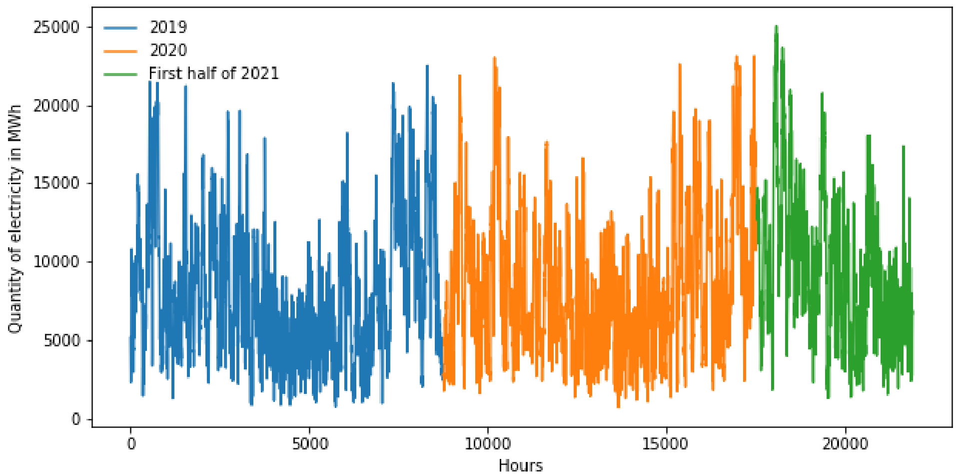


Figure 12 Wind hourly generation in (MWh)

- Solar hourly generation (MWh):

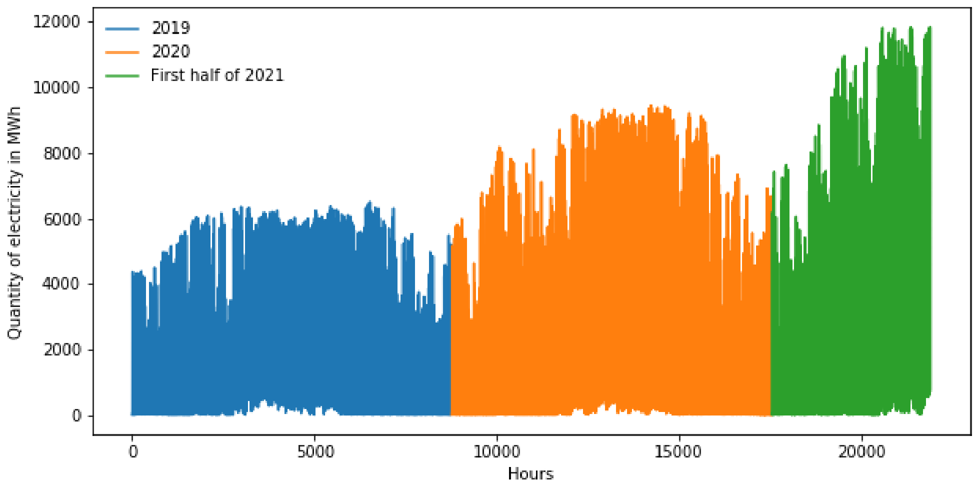


Figure 13 Solar hourly generation in (MWh)

- Hydro hourly generation (MWh):

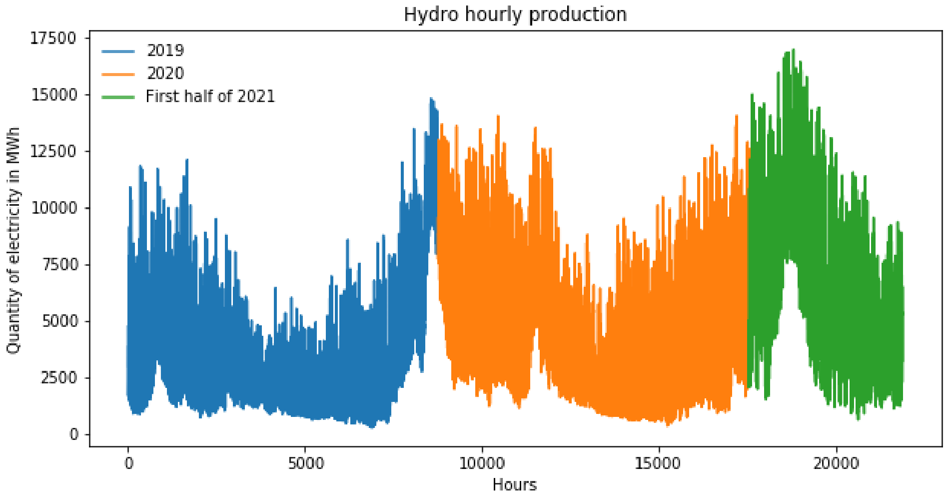


Figure 14 Hydro hourly generation in (MWh)

- Natural gas hourly generation (MWh):

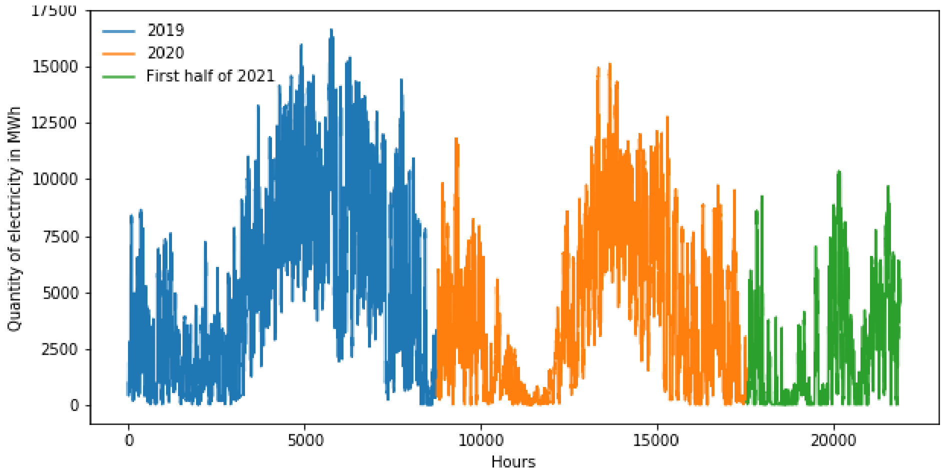


Figure 15 Natural gas hourly generation in (MWh)

- Coal hourly generation (MWh):

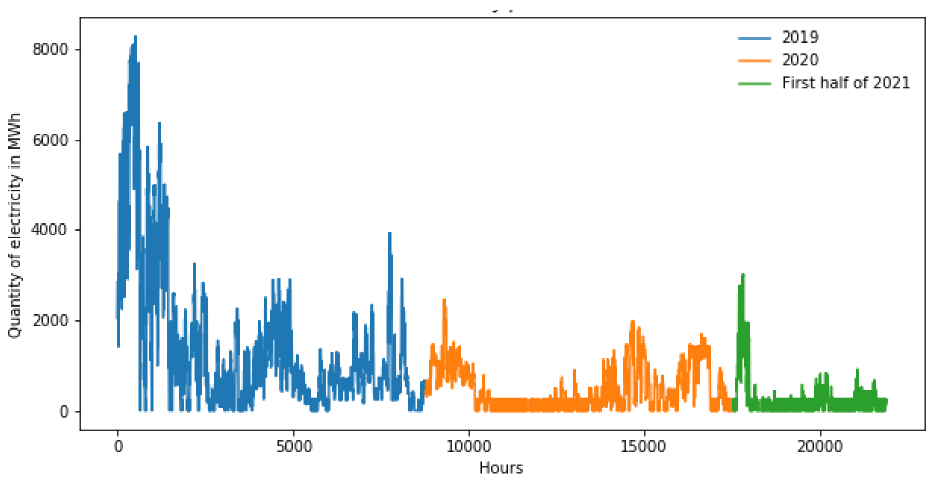


Figure 16 Coal hourly generation in (MWh)

- Nuclear hourly generation (MWh):

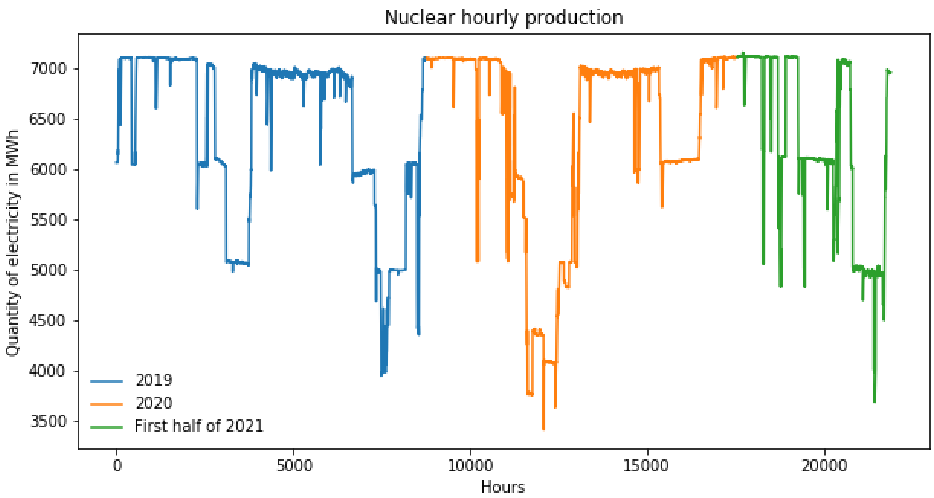


Figure 17 Nuclear hourly generation in (MWh)

- Other renewables hourly generation (MWh):

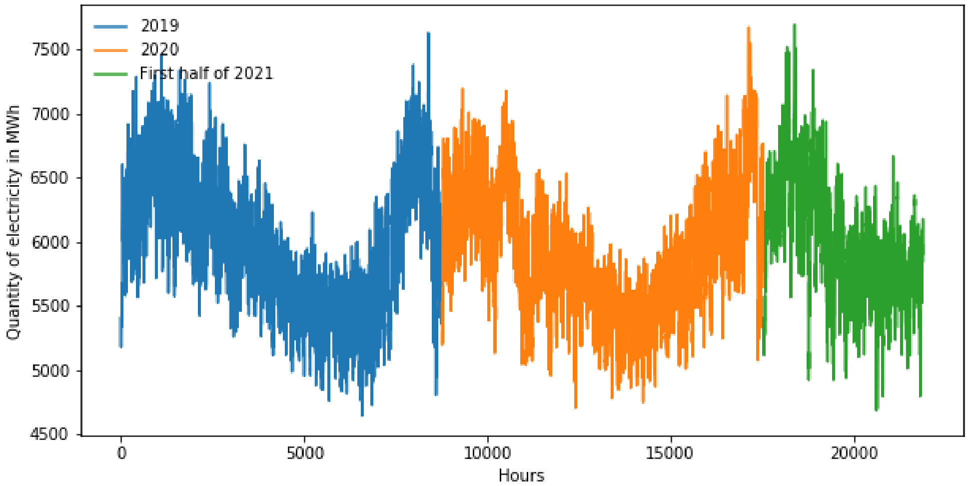


Figure 18 Other renewables hourly generation in (MWh)

- Hour
- Day
- Month

4.4.3 Training and Validation Output Variables

The features below were used as the output variables to train and validate the ANN. The model's output variables for the training and validation datasets are composed of the quantities of electricity bid at zero and non-zero price segments. And as explained before the training dataset included years 2019 and 2020, and the validation dataset is included the first 6 month of year 2021.

- Quantities of electricity bid at non-zero price:

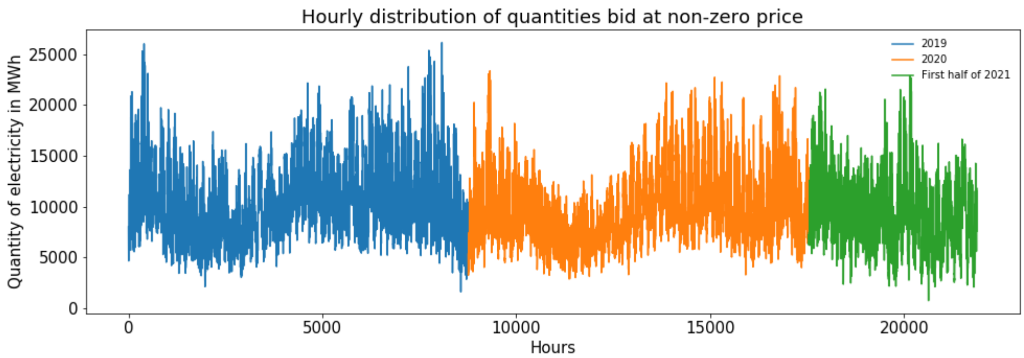


Figure 19 Distribution of quantities bid at non-zero price with respect to the MIBEL market

- Quantities of electricity bid at zero price:

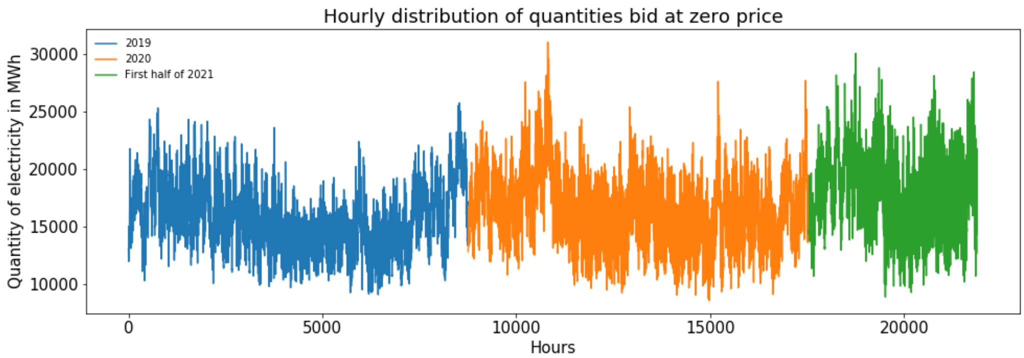


Figure 20 Distribution of quantities bid at zero price with respect to the MIBEL market

4.3.1 2030 Input Variables

Predictors, as defined above, are the variables (inputs) used to run the model, resulting in forecasting outcomes. The production data presented in Pereira's work [59] for 2030's MIBEL energy mix is presented on an annual scale as shown in Table 4, rather than on an hourly resolution as in the previous data used for model training and validation. It is necessary to convert these annual production values from a yearly basis to an hourly basis in order for the model to be able to forecast the quantities of electricity bid in the DAM for 2030. This section will cover this topic.

Production Patterns by Technology for 2030

Given the above explanation, production data for 2030 must be transformed from an annual (TWh/year) to an hourly (MWh/hour) basis. This means that a future hourly distribution of generation predictors must be constructed. The ideal approach was to collect past patterns of such variables and duplicate them for 2030, making an assumption that the distribution of today's inputs which are the hourly generation per technology will be similar in 2030. This assumption is mostly correct, because the most significant changes in the future will be in the total values of production, but not in their pattern. For instance, it is anticipated that in 2030, hydro generation will be at its peak during the winter season, while solar generation will be at its peak during the summer within the daylight hours, as it is now. As previously indicated, the difference will be only in the overall values of the generation, which are completely different. To acquire historical patterns, one can use a single year from the past years. In our case 2019 was the year used since only the data from the two years (2019 and 2020) were available from OMIE, and dude to COIVD-19, 2020 was excluded for the possibility of having data abnormalities.

After choosing year 2019 to resemble the production pattern, data normalization is applied to all data points, which entails dividing each predictor's hourly value by its yearly total. Wind normalization, for example, was carried out by dividing each hour of the year in terms of wind production by its entire yearly production. The same technique was used for the remaining predictors, yielding a normalized average year as a result. To determine the hourly distribution for 2030, simply multiply the normalized variables by the corresponding yearly value for 2030 obtained from the (RNC + PNIEC) projection, which is shown in Table 4.

Chapter 5

5 Theoretical Framework and Implementation

As mentioned before in chapter 4, an artificial neural network is developed for the practical component of this work. In this chapter the theory behind artificial neural networks is explained in a brief way to help the reader understand how an ANN algorithm operates.

5.1 Introduction to Artificial Neural Networks

Neural networks were discovered in the 1940s by the mathematician Walter Pitts and the neurophysiologist Warren McCulloch [60]. Despite the discovery of ANN, training these networks remained a mystery for twenty years. Later in the 1960s the concept of backward propagation was developed. It started to receive much attention in 2010 when the research community realized the potential of ANN and the great ability to solve problems that were previously unsolvable. To briefly describe what is a neural network made of the reader should simply start thinking of a neuron in a human brain as shown in figure 21.

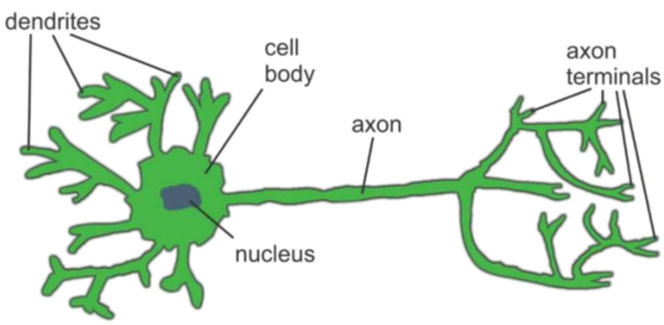


Figure 21 Human brain cell (neuron) , (Kinsley & Kukieła, 2020) [63]

Of course, it is not a perfect comparison, yet the mechanism of how they both work is pretty much the same. As shown in figure 22, the building block of an ANN is a neuron. Each neuron has an input (x_i), and each input is multiplied to a weight (w_i) then passed to a summation function. After summing, another adjustable parameter is added known as the bias (b). The purpose of this bias is to offset the output negatively or positively providing more degrees of freedom to the training procedure.

$$Net = \sum_{i=1}^{n} x_i w_i + b \qquad (1)$$

Following the summation and the bias addition as shown in equation 1, the (Net) value corresponding to the total input, is passed to an activation function resulting in the neuron's output (y). This output can be either the output of the network if the neuron is located in the last layer, or it can be the input to a neuron in another layer. Figure 22 represents a block diagram of a simple neural network.

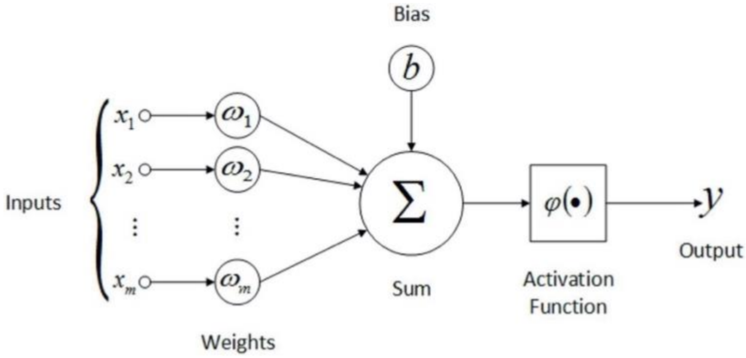


Figure 22 Block Diagram Representation of an Artificial Neuron

Now taking a step back to activation functions (AF), as explained above before the neuron can produce an output, the value of the equation $\sum_{i=1}^{n} x_i w_i + b$ is passed by an AF. This AF is usually chosen based on the complexity of the problem. For example, if the problem is non-linear, a linear activation function will not be adequate for solving the problem [46]. In general, for a neural network there are two types of AFs that can be delt with. The first is the AF used in hidden layers and the second is the AF used in the output layer. Usually, they are the same, but they can be also different.

After understanding how an artificial neuron works it is also important to understand how a group of neurons work together to form a network. Figure 23 illustrates a network formed of a number of artificial neurons. The network is composed of an input layer, n-hidden layers, and an output layer. The hidden layer is responsible for propagating and processing data that is passed by the input layer and eventually to the output layer.

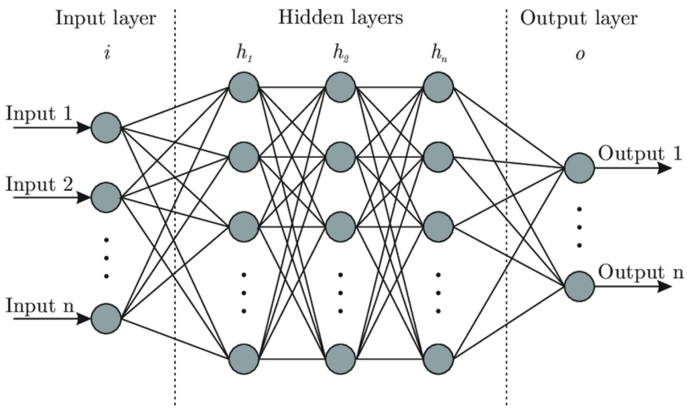


Figure 23 Representation of an ANN structure

5.2 Feed-forward neural network (FFNN)

A feed-forward algorithm is one of the most used in ANN and it is considered as the base of deep learning. it is also known for its simple yet very efficient architecture and being extensively used in supervised learning. From its name FFNN processes information in only one direction. Moving data from the input layer through the hidden layers and eventually reaching the output layer. This process is best described as forward direction. A representation of a FFNN can be also referred to in figure 11.

5.3 Practical implementation and architecture

To construct and implement the code for this ANN model, Jupiter was used for the development of the work for this thesis. Jupiter is an open-source web application that enables the user to create and share files that contain live code, equations, visualizations, and narrative text. This software package serves well in applications like data cleaning and transformation, numerical simulation, statistical modelling, data visualization, machine learning. With the help of machine learning libraries like Keras, a functional model was developed. This library contains the fundamentals for building Machine learning, such as optimizers, activation functions, weights initializers and so on. However, the entire compilation and code writing process was completed throughout the practical development of this thesis. In this section the reader should get a better understanding of the model architecture.

5.3.1 Input nodes

The constructed ANN model aims to predict the hourly quantities of electricity bid at the day-ahead MIBEl market categorized into zero and non-zero bids. As mentioned before in chapter 4, the number of variables used to forecast the quantities of electricity for both price segments are 10. Therefore, for each variable a corresponding input neuron was considered in the input layer. These input neurons are fed forward to the hidden layers.

5.3.2 Output nodes

The output nodes are also related to the size of the output vector. In our case the outputs of the model are the quantities of electricity categorized in two price segments: Zero and Non-zero. Therefore, the number of output nodes are two. One for each price category.

5.3.3 Hidden layers

Identifying the number of hidden layers is usually a tricky task since there is no rule of thumb to follow but. In this work the followed methodology of identifying the number of hidden layers was by trial and error. Different numbers of hidden layers and nodes were used and the combination with the best results was eventually chosen. Two hidden layers were used for the network. The first one is composed of 10 nodes and the second one is composed of 9 as seen previously in figure 11.

5.4 Training hyperparameters

After setting the model architecture and preparing the data to be fed into the model, the training phase starts. In order to reach the most optimum solution, tuning of hyperparameters is a must. To do this, different combinations of hyperparameters are tested and evaluated for the best accuracy. These hyperparameters are:

5.4.1 Batch size

Batch size is usually what controls the number of training samples that are propagated through the model before the internal parameters are updated. Generally training data is often huge and contains a very large amount of data points that goes through the training process. Therefore, a data set with millions of data point can be very time consuming and computational demanding if each data point is used to calculate the cost function and update the internal parameters. For this reason, mini-batches came to use. Mini batches simply solve the problem by dividing the data sets in batches, each batch contains a number of data points, and the internal parameters are only updated after the propagation of each batch instead of each single data point.

The batch size is also important and influences the model accuracy since a very big batch size can cause underfitting, and very small batch sizes can cause overfitting leading to a biased model. For this thesis the optimal batch size was found to be 7. In another way the model calculates the internal parameters after processing 7 of training points.

5.4.2 Epochs

Epoch is the parameter that controls the number of times the training dataset is passed through the algorithm. In our case the epoch was set to 30.

5.4.3 Optimizer

Optimizers are algorithms that are used to adjust the elements of a neural network such as learning rates and weights for the aim of minimizing the loss function. This optimization process is important for providing the most accurate result as possible of the ANN model. For this work, Adam optimizer was selected to be the one in use.

Adaptive Moment (Adam) Estimation is a technique for optimizing gradient descent algorithms. When dealing with big problems involving a huge number of data or parameters, the method is extremely efficient. It also uses less memory compared to other methods. Adam computes adaptive learning rates and momentums for each parameter, which means that parameters that have a large influence on the cost function are assigned lower momentums and learning rates, and vice versa. This results in a more balanced approach between the parameters, resulting in a smoother algorithm convergence. Adam is considered a hybrid of the gradient descent with momentum and the RMSP algorithms.

- Momentum

  This approach is used to accelerate the gradient descent algorithm by taking the exponentially weighted average of the gradients into account. Using averages enables the algorithm to converge to the minima more quickly.

- Root Mean Square Propagation (RMSP)

  RMSprop, or root mean square prop, is an adaptive learning technique that uses the exponential moving average rather than the cumulative sum of squared gradients.

The magnitude of each step in the optimisation process is described as the learning rate; it is usually higher in the early steps, providing a faster convergence rate, but its value begins to decrease as one gets closer to the solution point, providing a smoother convergence to the minimum.

5.4.4 Activation Function

As explained before the activation function is what transforms the weighted sum of the inputs to an output. For the FFNN developed in this work ReLU activation function was used.

The ReLU is one of the most widely used activation functions. It has been utilized in almost all convolutional neural networks and deep learning algorithms. As seen in the figure below, the ReLU is only half-rectified (from bottom). When z is less than zero, R(z) is zero, and when z is more than or equal to zero, R(z) is equal to z.

Figure 24 provides a visual representation of the ReLU function.

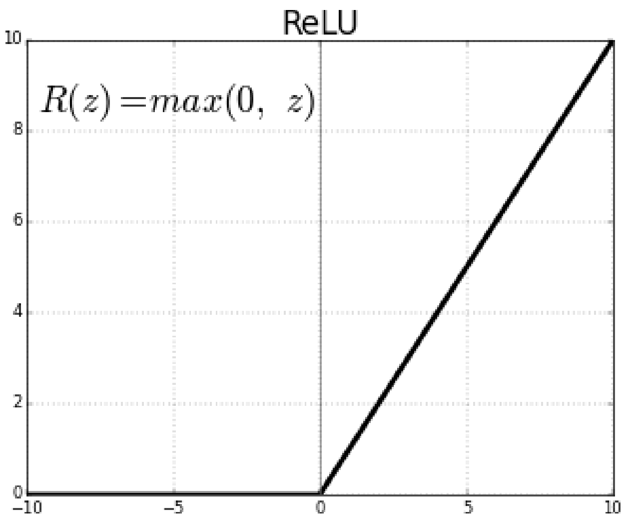


Figure 24 ReLU function

5.4.5 Cost Function

A cost function is a single value that measures how accurate the neural network as a whole did with respect to the fed training data. In this case, the used cost functions that best fit our model is the mean squared error (MSE). Equation 2 represents the mathematical formula for MSE.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \qquad (2)$$

MSE = Mean Squared Error
n = Number of data points
Yi = Observed values
$\hat{Y}$ = *Predicted* values

5.5 Model Training

The process of determining the value of the hyperparameters that minimize error is known as network training. To accomplish this, the model must be fed with specific data, known as training data, which contains predictors and solutions resembling historical inputs and output, ensuring the model's ability to learn, adapt, and minimize error. There are numerous algorithms, known as optimizers, that are used to determine the value of each hyperparameter that minimizes the global error, with the majority of them relying on gradient descent techniques.

Gradient descent is an optimization approach for determining the parameters of a function that minimizes a cost function. The procedure begins with initial values for the function's coefficients. These could be 0 or a small random number. By entering the coefficients into the function, the cost is calculated. Then the cost's derivative is computed. The derivative is a mathematics concept that relates to the slope of a function at a specific position. It is important to know the slope so that we can shift the coefficient values in the right direction on the next iteration to get a reduced cost. This procedure is done until the cost of the coefficients is 0.0 or near to zero.

Figure 25 provides a visual representation of the gradient descent principle, where the weight value (W) is being tweaked in order to minimize the cost function [61].

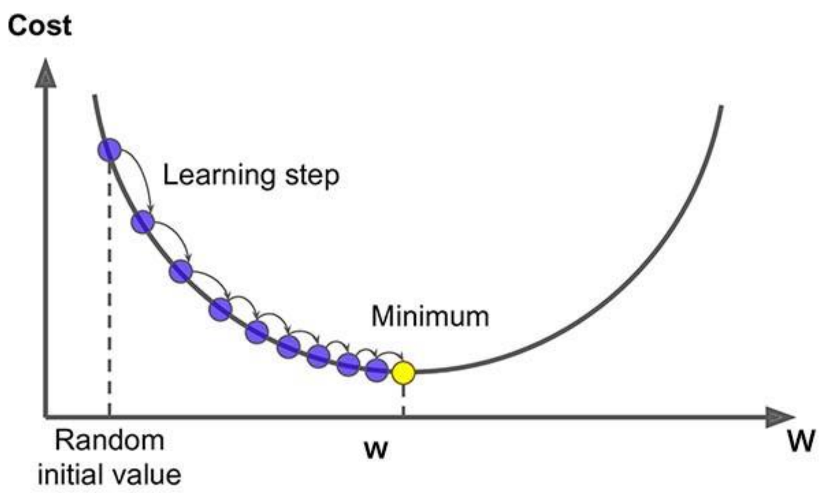


Figure 25 Visual Representation of a Cost Function and the Gradient Descent Optimization Process.

5.6 Validation Matrices

In this thesis some of the classical error measures are used to evaluate and compare the performance of the forecasting model due to their simplicity and effectiveness in describing the accuracy of the models under study. For instance, the Mean Absolute Error (MAE) allows the comparison of the real value of error. The Mean Absolute Percentage Error (MAPE) is one of the measures based on percentage errors and it is a good method to compare the relative error between different forecasts.

Also being scale independent enables MAPE to compare forecasts with datatypes on different scales. The Root Mean Squared Error (RMSE) is a useful metric when comparing datatypes having the same scale. Last but not least, the Coefficient of Variation of Root-Mean Squared Error (CV-RMSE) is an efficient way of estimating the predictive capability of a model.
A brief review on the used Key Performance Indicators (KPIs) is performed below:

5.6.1 Mean Absolute Percentage Error (MAPE)

MAPE is the sum of the individual absolute errors divided by the number of fitted points (n). It is the average of the percentage errors. It measures the accuracy in terms of a percentage as follows:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| * 100\%$$

- n is the number of fitted points.
- A_t is the actual value.
- F_t is the forecast value.

5.6.2 Mean Absolute Error (MAE)

MAE measures the average magnitude of the errors in a set of forecasts by calculating the difference between the forecasted value and real value. The MAE equation can be expressed as follows:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|F_t - A_t|$$

5.6.3 Root Mean Squared Error (RMSE)

RMSE is defined as the square root of the average squared error. It is a measure of how far from the regression line data points are. In other words, it measures how spread out the prediction errors. The following equation represents the RMSE calculation:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(F_t - A_t)^2}{n}}$$

5.6.4 Coefficient of Variation of Root-Mean Squared Error - CV(RMSE)

The CV-RMSE is simply RMSE normalized by the mean value. Usually a CV-RMSE below 25% indicates a good model fit with acceptable predictive capabilities. This KPI is also expressed in percentage and is calculated as follows:

$$CV(RMSE) = \frac{1}{Y}\sqrt{\frac{\sum_{t=1}^{n}(F_t - A_t)^2}{n}}$$

- Y is the mean value.

5.7 Conclusion

The Fifth chapter offers insight into the theory and practical use of neural network algorithms. The processes for initialization and training were handled, and a final model was built. The next phase, model validation, will be covered in Chapter 6, where we will evaluate the model's performance, accuracy, and capacity to extend to new cases.

In table 6 a summary of the used hyperparameters are represented.

Table 6 Hyperparameters Optimization values

Hyperparameter	Model
Batch size	7
Epochs	30
Optimiser	Adam
Cost function	RMS
Activation function	ReLU

Chapter 6

6 Model validation

After the construction of the model's algorithm, setting all the hyperparameters and training the model, it is very important to evaluate its accuracy before using it to forecast. To perform the validation, process the idea is to have historical data for both the inputs, which are equivalent to the model predictors, and their corresponding outputs, as model targets. This enables us to compare the model with real data, thus, allowing to quantify the error. For this work, the validation process will be based on feeding the model with historical data of the previously explained inputs for year 2021 and comparing the output values of the hourly quantities of electricity bid in the day-ahead MIBEL market with the true historical values. This chapter is composed of four sections. The first section represents a visualization of the real validation data, the second section focuses on the visual analysis of the validation results, the third section evaluates of the model accuracy, error analysis and compares the model to a similar study. Finally, the fourth section is the conclusion.

6.1 Validation Data

As mentioned before, the model's goal is to forecast the hourly quantities of electricity bid in the day-ahead MIBEL market. These quantities are categorized into two price segments, the zero and non-zero prices. The validation procedure is conducted using the historical hourly data from the first half of 2021. Only the first 6 month corresponding to 4320 data points of year 2021 were used in the validation dataset due to data limitations. Figures 26 and 27 Present the real values of the hourly quantities of electricity bid in the DAM for the two price segments. Each data point resembles the total quantity of electricity bid into the DAM at that specific hour.

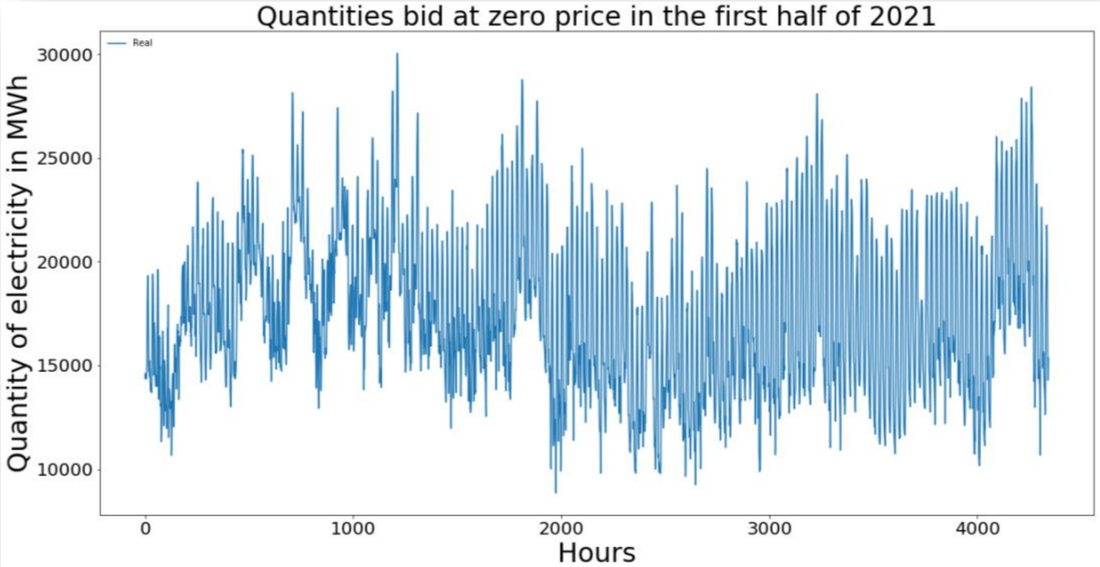


Figure 26 Quantities of electricity bided at zero price for the first half of 2021 in Portugal and Spain

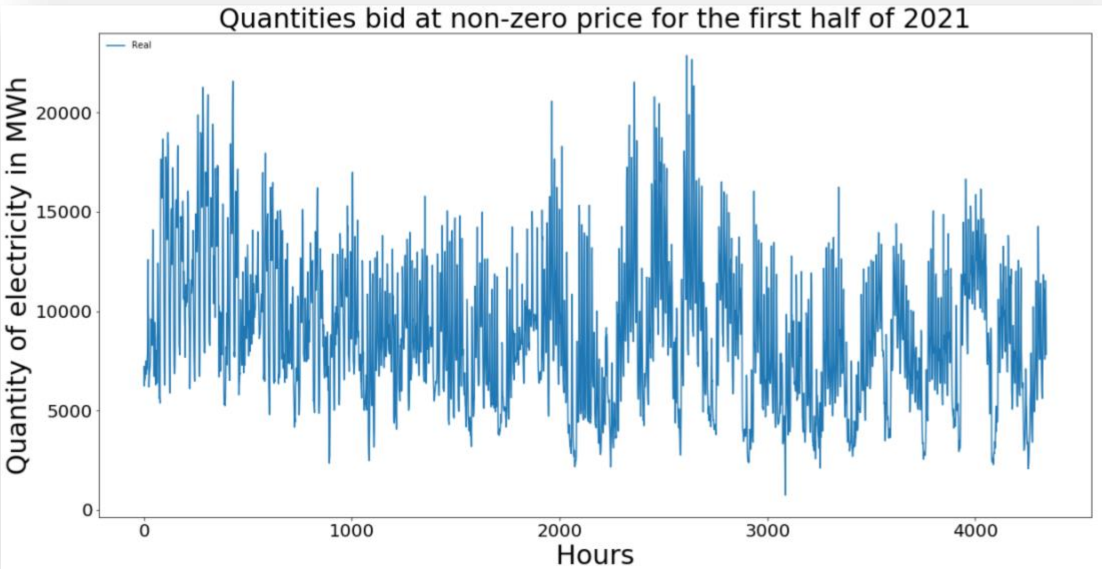


Figure 27 Quantities of electricity bided at non-zero price for the first half of 2021 in Portugal and Spain

6.2 Validation forecasting results and visual analysis

With the information gathered from OMIE regarding the hourly generation categorized by technology and using the hour, day, and month as inputs to take seasonality into consideration, it was possible to run the model and forecast the quantities for the first half of 2021 as it was the only data available at the time of the study. During the training phase it was made sure that the model doesn't have access to the validation data. This is a crucial step to obtain a good model validation.

6.2.1 Zero price segment

As seen in figure 28, A visual representation between the real and forecasted data is presented for the first 6 month of year 2021 in an hourly interval resulting in 4320 hourly data points starting January. Also, a zoom-in of the real and forecasted values for March is represented to give a closer comparison of the real and forecasted curves. It is clear that the model is able to forecast the main pattern of the quantities bided in the DAM. It can be also observed that there are some limitations forecasting sharp peaks and dips.

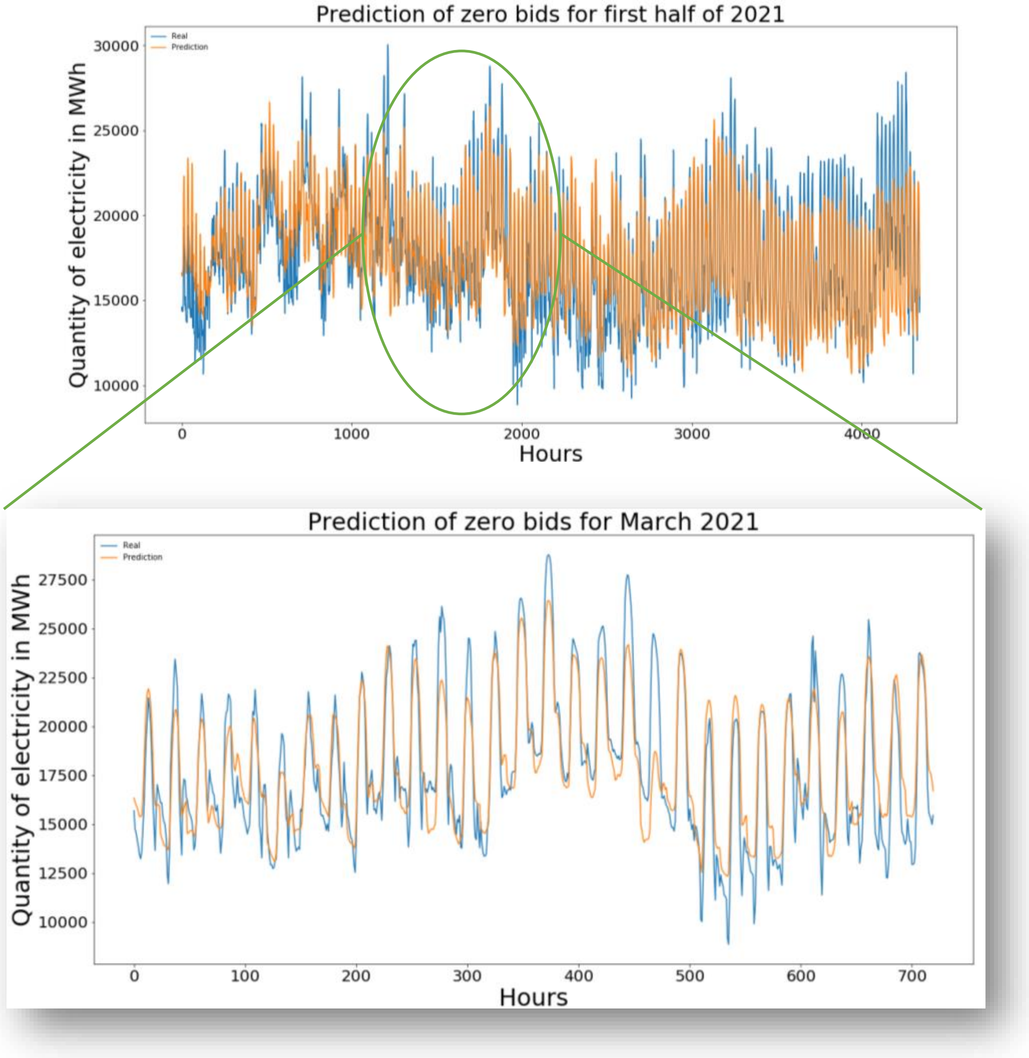


Figure 28 Real and forecasted quantities for zero price segment

6.2.2 Non-Zero price segment

For the non-zero price segment, the validation procedure is the same. Figure 29 presents the real and forecasted values of the hourly quantities of electricity bided in the DAM for the non-zero price segment. It can be observed that the model is able to forecast the main pattern of the quantities yet there are also some limitations forecasting sharp peaks and dips. A zoom-in of the real and forecasted values for March is represented to give a closer comparison of the real and forecasted curves.

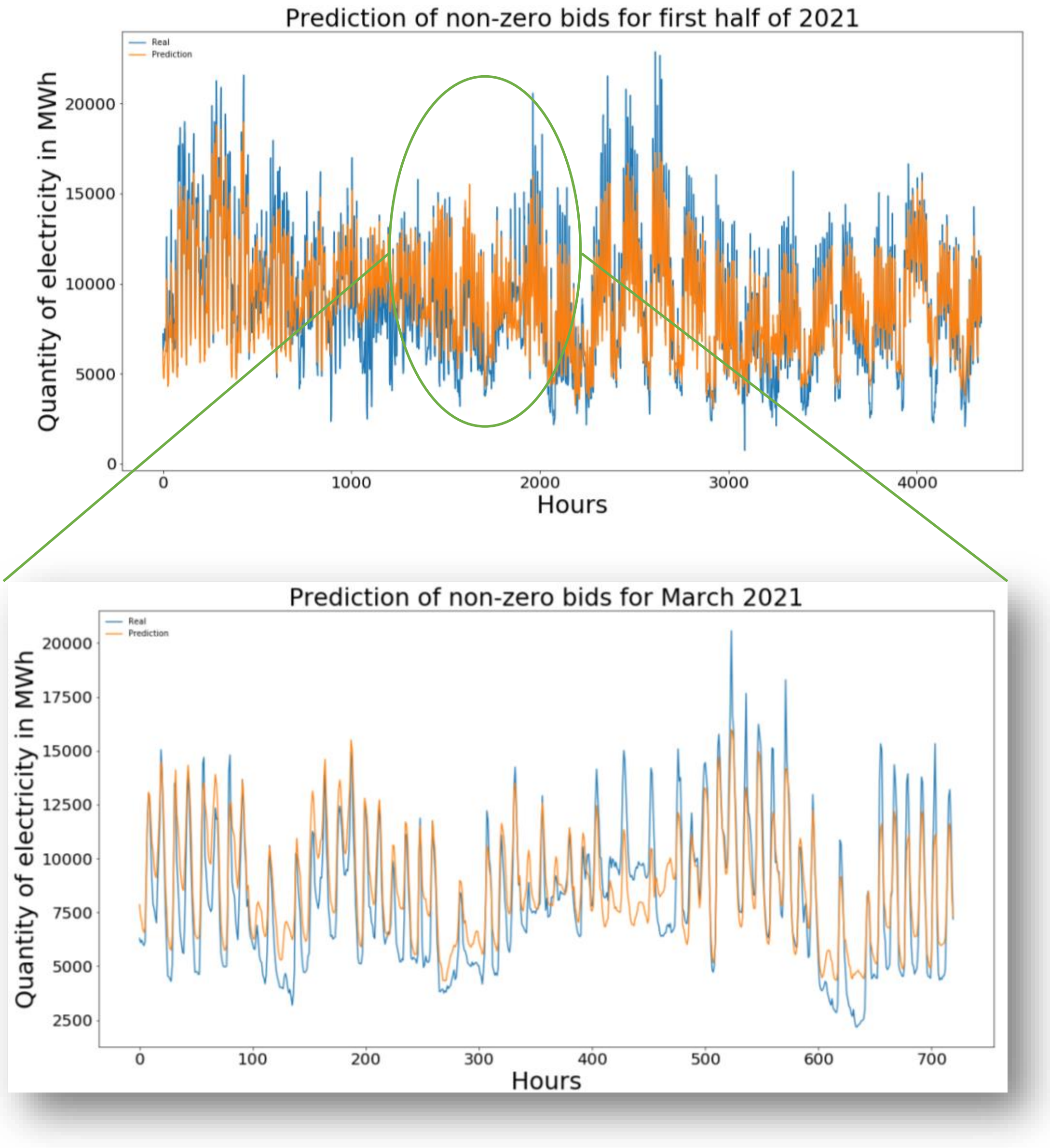


Figure 29 Real and forecasted quantities for non-zero price segment

6.3 Error Analysis

Following the prior visual study, a numerical evaluation is critical. Furthermore, it is also important to compare the models to similar studies as it provides an indication on how the model performs compared to similar models. In this section some forecasting key performance indicators (KPIs) are used to measure the model's accuracy (or error), utilizing the MAPE, MAE, RMSE and CV-RMSE indicators.

Results of the previously explained KPIs are presented in the following table to evaluate the model accuracy for both price segments.

Table 7 KPIs of the developed model for the two price segments

Model	MAPE (%)	MAE (MWh)	RMSE (MWh)	CV-RMSE (%)
Zero price segment	8.7	1534.9	1950.3	11.1
Non-zero price segment	13.6	1148.8	1490.53	16.3

As mentioned before, the model can clearly predict the main pattern of the forecasted quantities, yet it suffers some limitations when forecasting sharp peaks and dips. Such limitation can be justified by the ANN input variables. The fact that the model is provided only 10 variables mainly divided into hourly production of each technology and time, to fully describe the supplier's behaviour, made it hard for the model to detect such steep movements. An explanation of such limitation is that the model is forecasting quantities of electricity bid in the DAM for the two price segments. these quantities mainly depend on the production from each technology and the production also depend on weather conditions specially regarding renewable resources (represented by zero price segment), therefore a sudden change in weather conditions will eventually cause a change in the renewable production which will induce a change in the production of non-renewable resources to be able to supply the demand. A way of tackling this limitation would be by introducing hourly variables that represents weather conditions along with the proposed inputs, which can be very challenging and data availability is not guaranteed.

Comparison with the Literature

It is important to compare the results with similar studies having the same approach as this work. As explained in chapter 3, literature about this topic is extremely limited, and studies with similar approaches as the developed model in this dissertation, are very scarce.

One of the papers mentioned in the literature review developed a forecasting method to predict the bidding curve of generation players in the Iberian electricity market MIBEL, which is the same market under study. As explained before this forecasting model is constructed as a two-step artificial neural network (ANN) prediction model. The first step model works on the prediction of the amount of energy to be bid at zero price for a certain hour. The second step model involves the prediction of rest of the bidding curve. In this section, only the first step model will be relevant for the comparison with our developed model. The reported results showed that the first model that forecast the amount of electricity bid at zero price achieved a MAPE value of 17.5 %. On the other hand, the developed model in this thesis achieved MAPE value of 8.7% for the zero-price segment. Indicating a better forecasting performance in favour to the developed model.

6.4 Conclusions

Throughout this chapter, a detailed error analysis was performed, resulting in the conclusion that the built model, can be used to forecast new data samples. Despite the model’s limitations predicting the sharp peaks and dips, it is a feasible way to predict the bidding behavior of the electricity market in 2030. Finally, the model's accuracy was validated with similar error values by comparing it to other studies.

Following model validation, the next stage is to use it to forecast for 2030 with confidence that it will give trustworthy results. This subject will be covered in the next chapter.

Chapter 7

7 Results and Discussion

Following creation, training, and validation, the model is ready to be utilized for predicting, with a high degree of confidence and reliability in the outcomes. The following chapter is broken into two sections: the first previews the models' inputs for 2030, and the second offers the simulated future outcomes for 2030 with a detailed discussion about the findings.

7.1 2030 Model Inputs

In order to forecast for 2030, these inputs had to be estimated to be fed into the model. This was achieved by selecting a reference year which is 2019 and assuming that the production distribution will be similar in 2030. The suggested assumption is found to be reasonable since the most significant changes in the future will be in the total values of production but not in the production pattern which follows seasonality. Figures 30 to 37 show the hourly distribution of generation variables for each technology with respect to year 2030 based on the methodology explained in chapter 4.

As seen in figure 30 the hourly nuclear generation pattern is mostly stable during the whole year. Which is expected since nuclear power plants take a long time to start up and they usually operate in a steady mode to provide the base load to the grid.

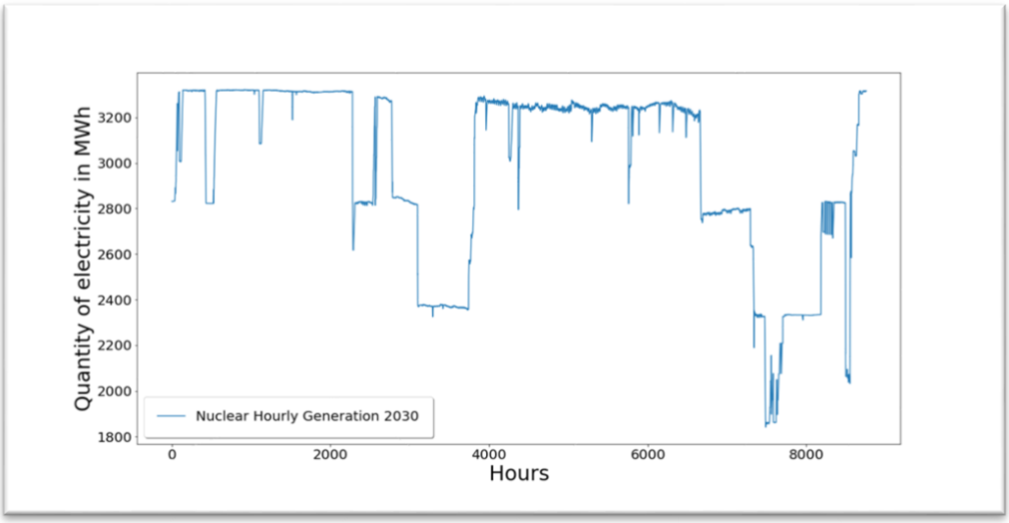


Figure 30 Predicted 2030's Hourly Nuclear Generation for Portugal and Spain.

In figure 31, it can be observed that hydro generation produces the most during the winter, as well as in the early spring days, when reservoirs have stored water from the rainy season and rivers still have significant amounts of water. During the summer, hydro power is reduced to a minimal level, due to the shortage of river water and rain. Hydro output begins to increase in between October and November, when the rainy season begins.

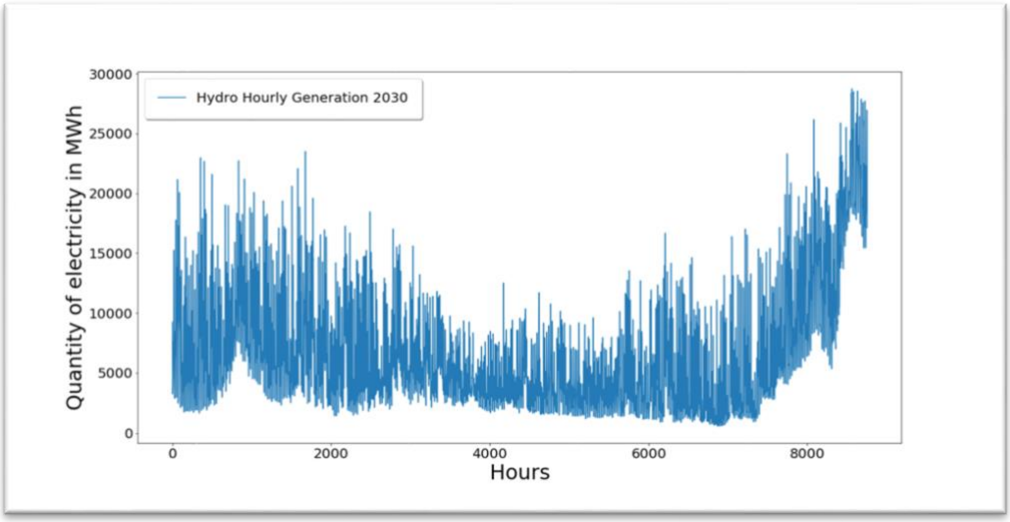


Figure 31 Predicted 2030's Hourly Hydro Generation for Portugal and Spain.

On a more focused scale over only one-hour is presented in figure 32, hydro production is noticed to be at its maximum starting from the afternoon hours till around 9 pm. These hours are called peak hours. It is known that nuclear and fossil fuel plants are not very efficient for generating power for short periods of increasing demand during the mentioned peak hours, and this is due to the long start up time they need to be in operation, making them more efficient for supplying the base load. On the other hand, hydroelectric generators can be switched ON/OFF almost instantaneously which made hydropower more responsive to peak demand than most other energy sources. Water can be stored over the night in a reservoir until it is needed during the day, then released through turbines to produce power to help meet the peak load.

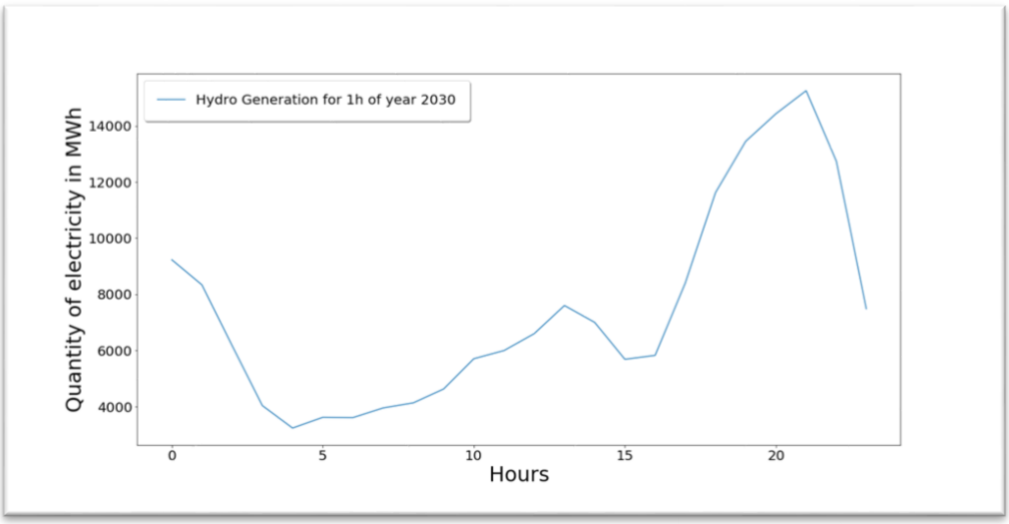


Figure 32 Predicted Hydro Generation for the first day of year 2030 for Portugal and Spain.

Natural gas hourly distribution shown in figure 33, is not consistent throughout the year. It is also the only fossil fuel technology in the market since coal is planned to phase out in 2030. Natural gas is highly dependent on the renewable power generation, and usually natural gas production is low during periods of high renewable production and vice-versa.

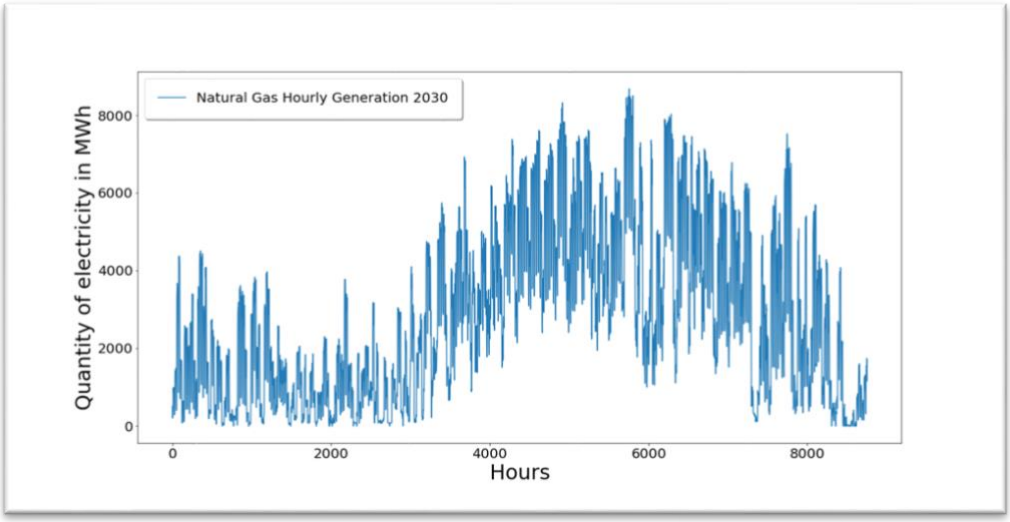


Figure 33 Predicted 2030's Hourly Natural Gas Generation for Portugal and Spain.

In figure 34, it can be observed that the hourly wind generation exhibits a more random pattern, with consecutive hours exhibiting extremely varying levels of production. At the same time, higher production can be observed throughout the winter and autumn seasons, when higher wind speeds are observed for longer hours.

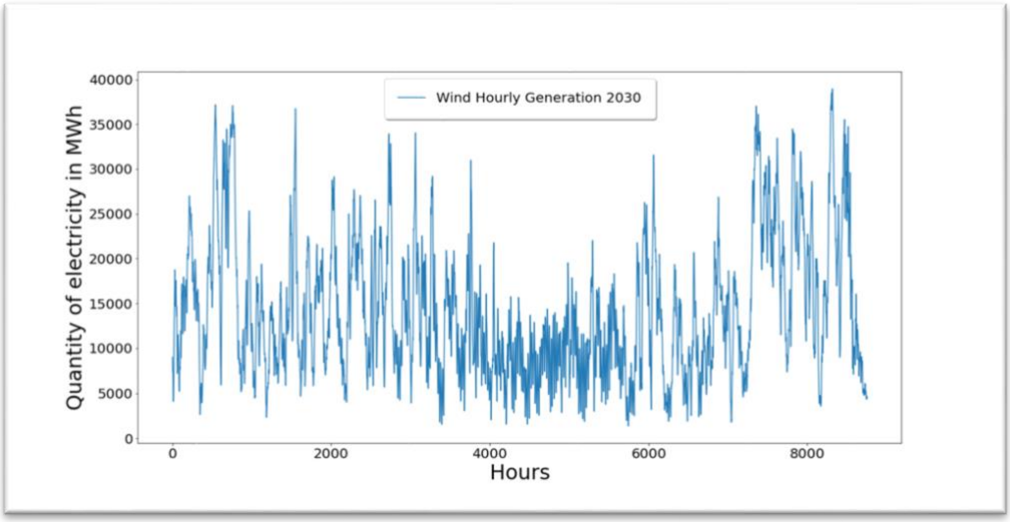


Figure 34 Predicted 2030's Hourly Wind Generation for Portugal and Spain.

Solar generation, figure 35, follows a constant hourly pattern. The production peaks in the summer season, and slightly decreases starting from the end of summer until winter. This occurs for obvious reasons, during the summer season the sky is clearer, irradiation is greater, and the maximum number of sun light hours is achieved. During wintertime the opposite happens. Therefor the production decreases. Figure 36 shows the typical hourly behaviour for the first day of year 2030. During the first and last hours of the day we can observe that usually there is no solar production, and the peak occurs during the afternoon. For obvious reasons, this happens because solar only produces electricity from day light.

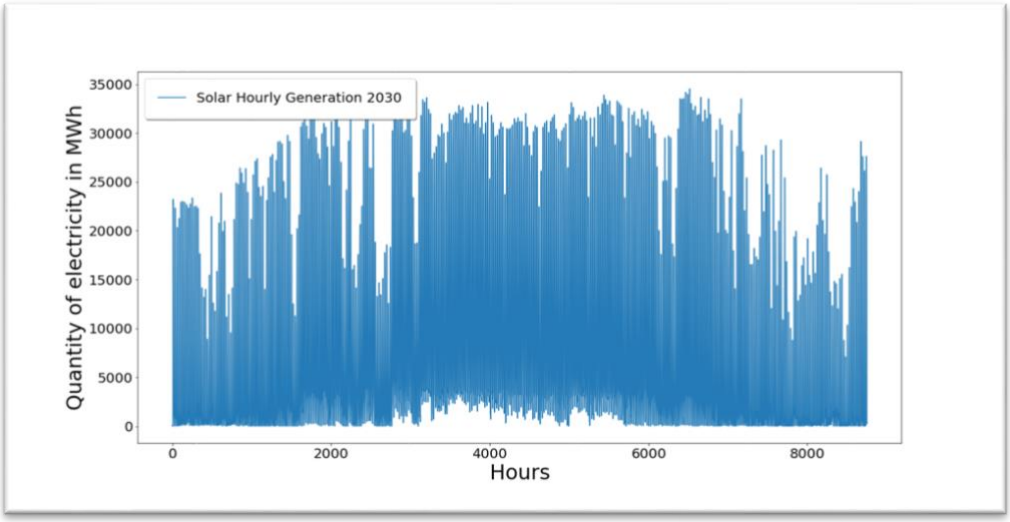


Figure 35 Predicted 2030's Hourly Solar Generation for Portugal and Spain.

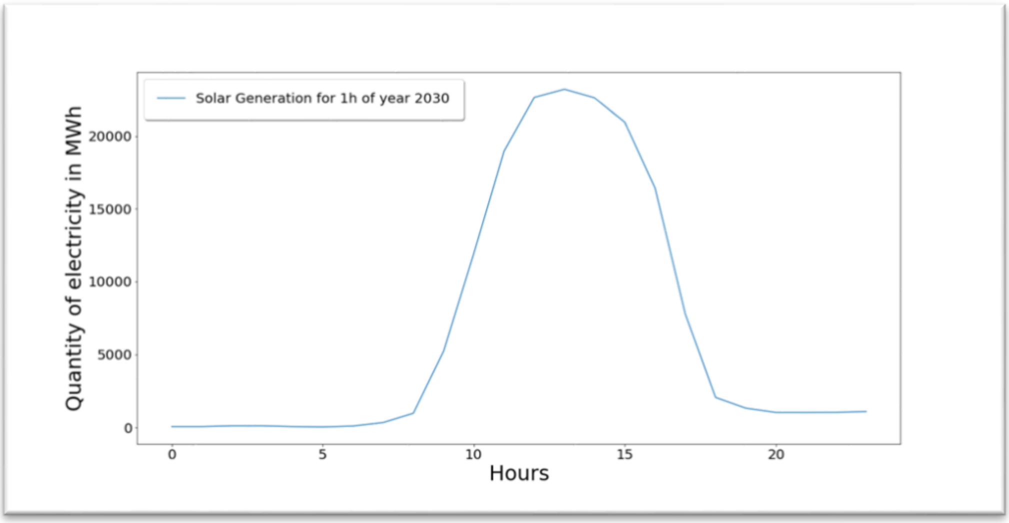


Figure 36 Predicted Solar Generation for the first day of year 2030

In figure 37 the pattern for other renewable generation is shown below. Other renewables contain cogeneration and mini-hydro production. The distribution decreases in the summer season and peaks in the winter season as seen.

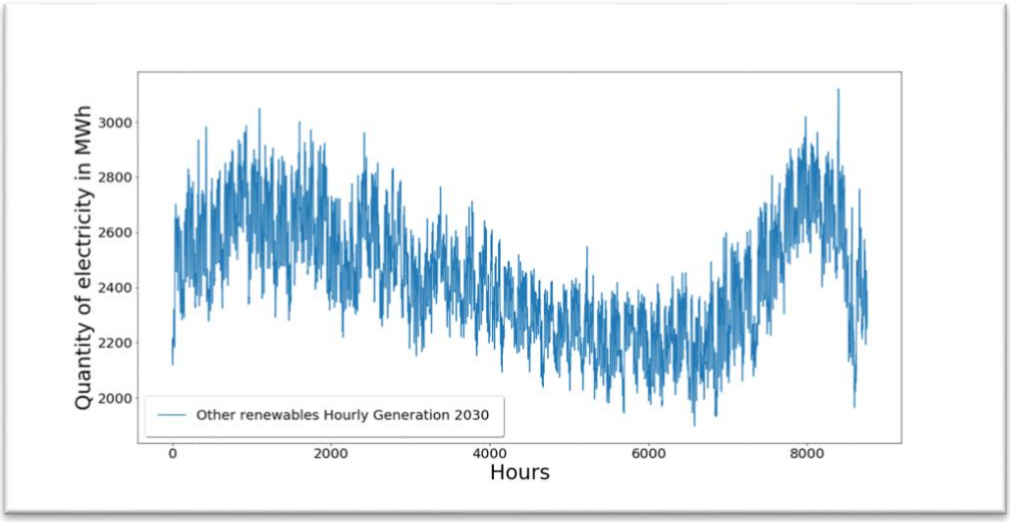


Figure 37 Predicted 2030's Hourly Other renewables Generation for Portugal and Spain.

7.2 2030 Model Results

Having described on an hourly basis, all the models' inputs for 2030, it is possible to feed the model to forecast for year 2030 and obtain the quantities of electricity categorized into zero and non-zero price segments, that are going to be bid in the day ahead MIBEL market for Portugal and Spain. Figures 9 & 10 provide an overview of the forecasted hourly quantities distribution for 2030.

7.2.1 Zero Price Segment

Looking at the simulated results in figure 38, it can be seen that the quantities of electricity bid at zero price will increase for year 2030 with a mean value of the distribution of 21463.86 MWh compared to a mean value of 15666.92 MWh for 2019. This observed increase was excepted and can be justified by the high increase in renewable production and the decrease in fossil fuel generation for 2030.

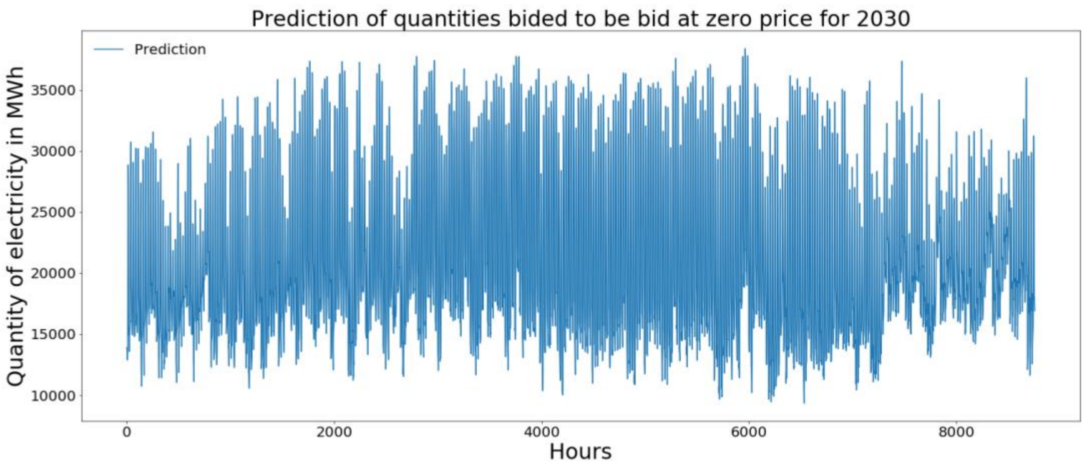


Figure 38 Forecasted hourly quantities bid at zero price for 2030, MIBEL

As explained in previous chapters, it is clear that the share of renewable production will highly increase in the MIBEL market, and fossil fuel production will decrease considerably, with coal powerplants phasing out the energy mix by 2030.

In year 2019 the renewable electricity had a share of 39% of the total production, on the other hand for 2030 the planned share of renewable production should rise to 86%. The greatest investment will go to solar technologies, which is intended to raise production by 707 % over 2019 in Portugal and Spain. The overall increase in the quantities of electricity bid at zero price in the day ahead market reflects the impact of such a substantial change in the energy mix.

To closely analyse the results obtained above, the forecasted hourly quantities to be bid at zero price segment are divided into 4 sections. Each section represents a season of the year as follows:

7.2.1.1 Summer:

As it can be observed in figure 39, the summer season show the highest quantities bid at zero price over the 4 seasons. This observation supports the speculation that solar technologies are going to contribute with a big share in the future energy mix.

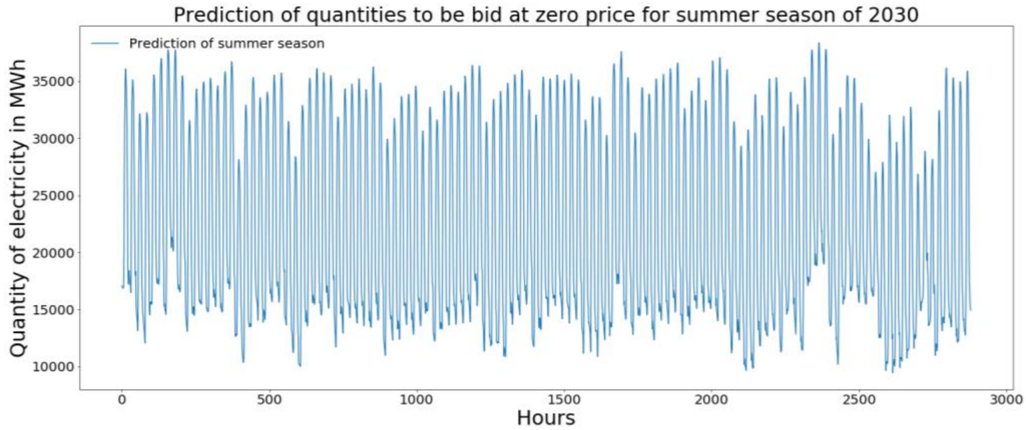


Figure 39 Forecasted hourly quantities bid at zero price during summer 2030, MIBEL

Figure 40 presents on an hourly scale the distribution of the average quantities over the whole season. It can be observed that at night the quantities are the lowest, mostly coming from wind. The quantities then start to rise during the day reaching the peak at afternoon as a result of the solar contribution.

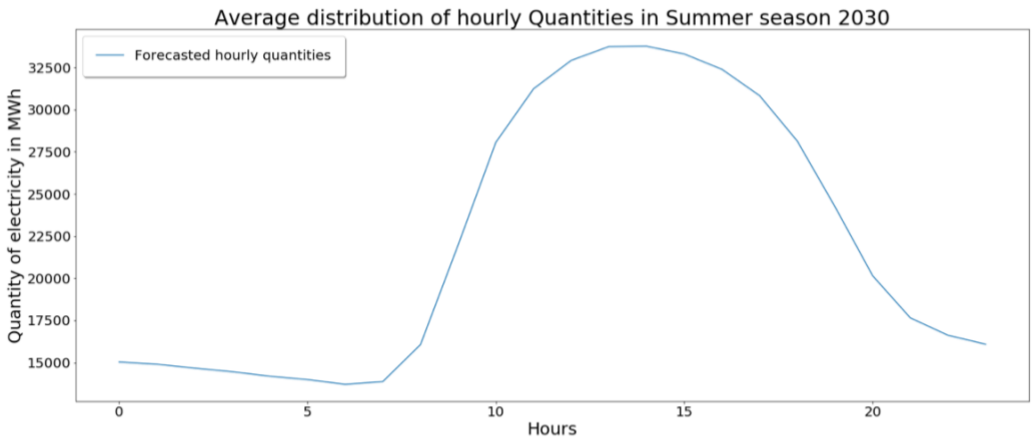


Figure 40 Forecasted average distribution of hourly quantities during summer 2030, MIBEL

7.2.1.2 Autumn:

In figure 41 it can be noticed that the quantities start to decrease compared to the summer season. Also, some fluctuations are observed. This can be a result of the decrease of solar production due to cloudy autumn weather conditions.

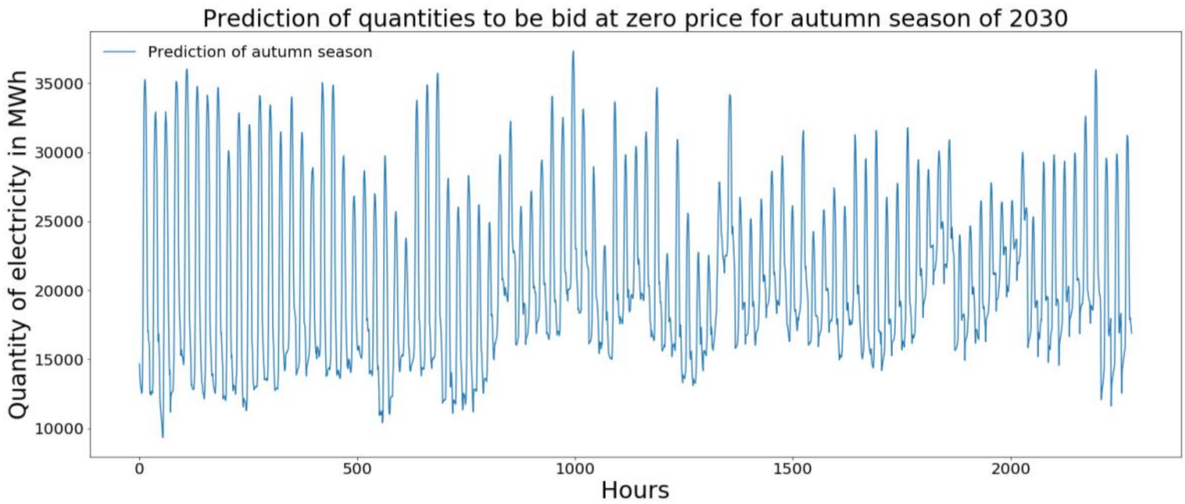


Figure 41 Forecasted hourly quantities bid at zero price during autumn 2030, MIBEL

From figure 42 the same pattern for the average hourly distribution as the summer season is observed but with a lower peak.

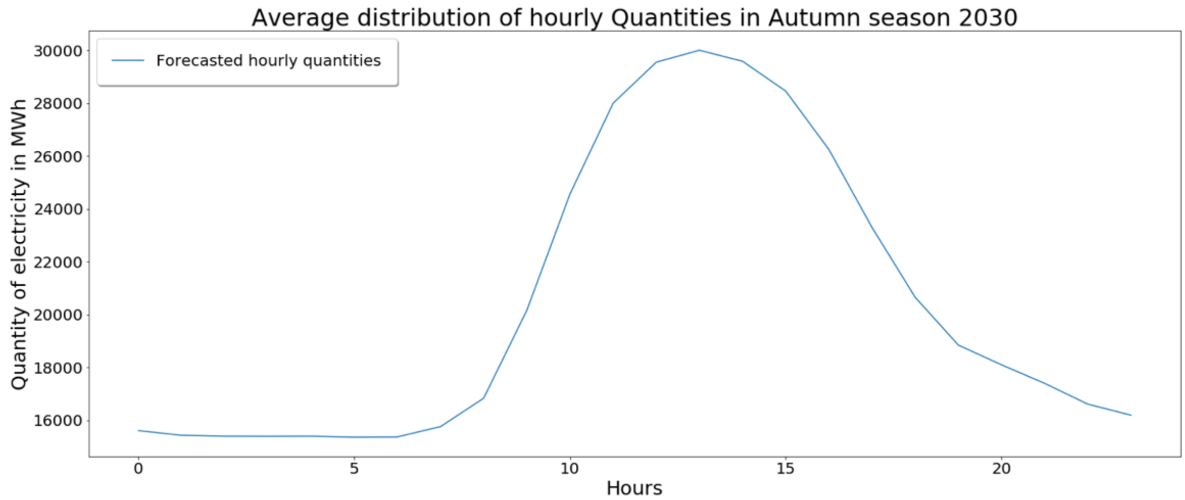


Figure 42 Forecasted average distribution of hourly quantities during autumn 2030, MIBEL

7.2.1.3 Winter:

For the winter season it can be seen in figure 43 that the quantiles bid at zero price decreases even more for some days than the autumn season. Which is again due to the decrease of solar production as a result of the winter weather conditions.

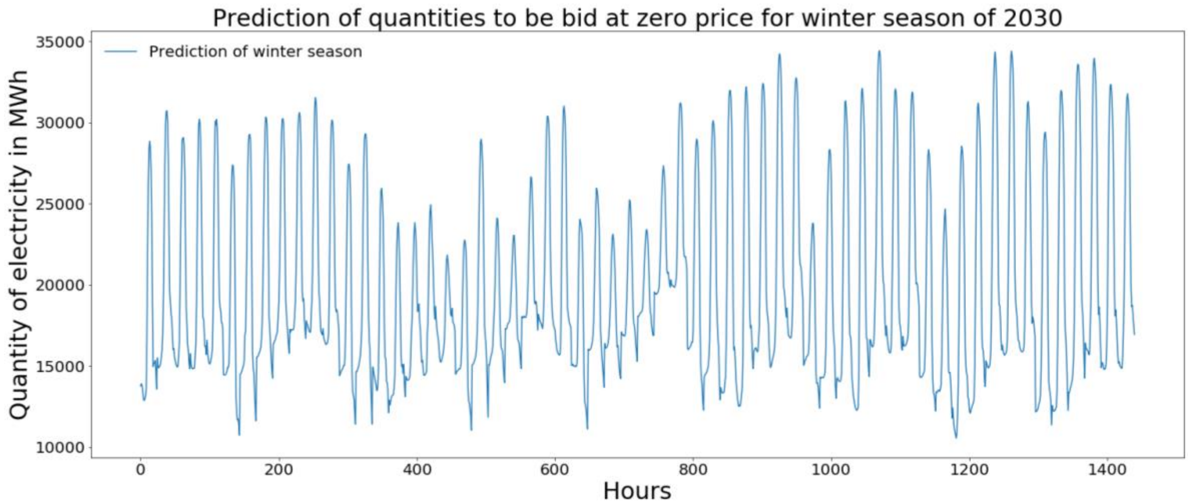


Figure 43 Forecasted hourly quantities bid at zero price during winter 2030, MIBEL

Again, the average distribution seen in figure 44 has the same pattern, yet there is a noticeable decrease of quantities.

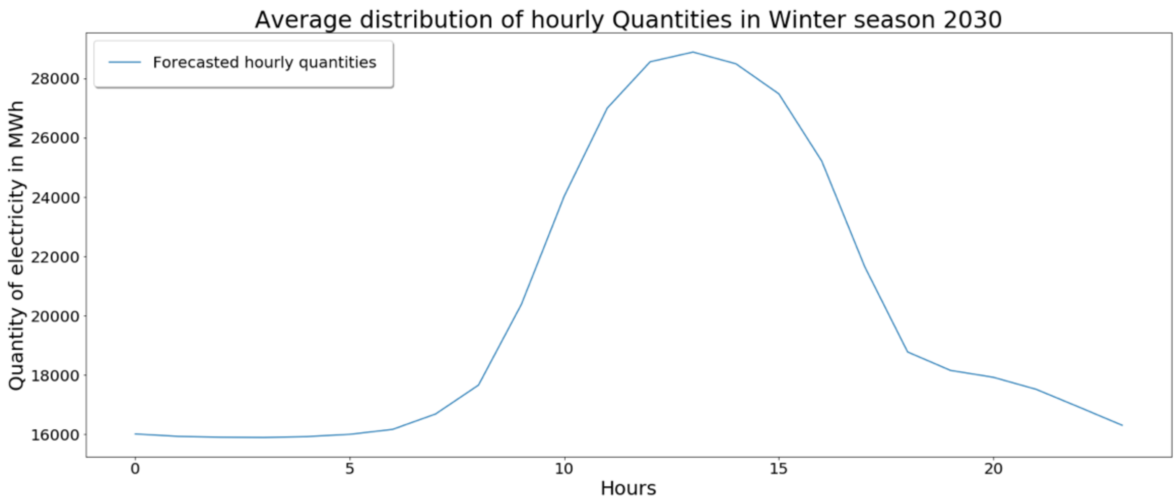


Figure 44 Forecasted average distribution of hourly quantities during winter 2030, MIBEL

7.2.1.4 Spring:

In figure 45 it can be noticed that quantities start to increase again heading towards the summer season where big quantities of renewable electricity is bid in the market.

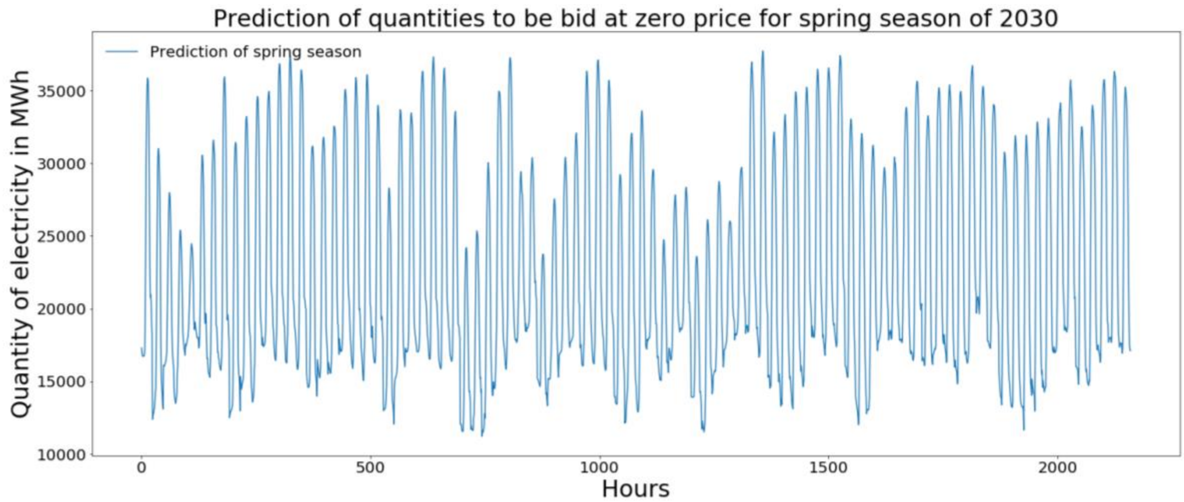


Figure 45 Forecasted hourly quantities bid at zero price during spring 2030, MIBEL

The pattern of the average hourly distribution is shown in figure 46. The pattern is still the same and the peak is increasing again near the summer levels.

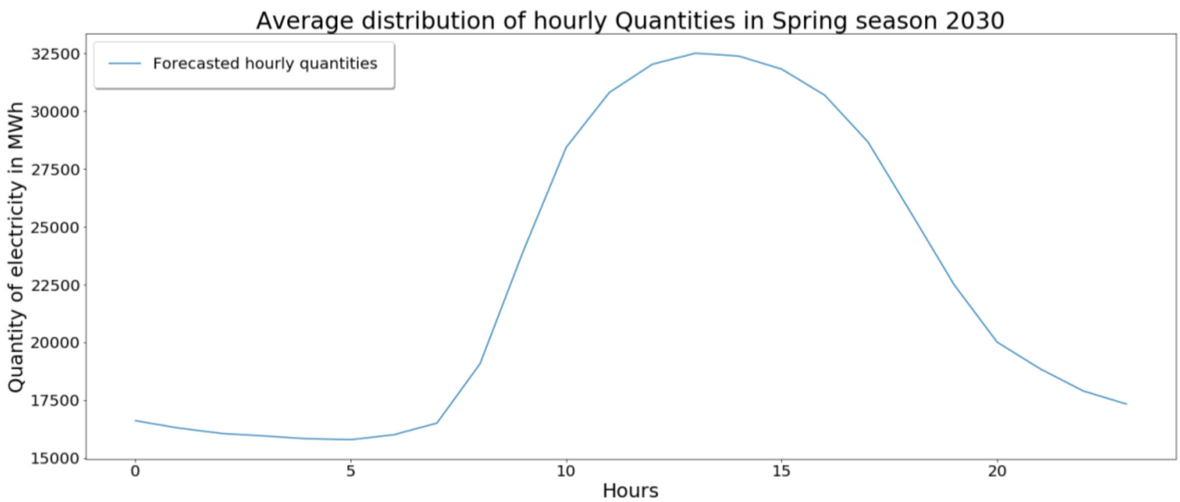


Figure 46 Forecasted average distribution of hourly quantities during spring 2030, MIBEL

7.2.2 Non-Zero Price Segment

Looking at the simulated results in the figure 47, it can be observed that the quantities of electricity bid at non-zero price will not be affected much, only a slight increase from a mean value of 10606.17 MWh in 2019 to 10980.58 MWh for year 2030. This can cause some confusion as nuclear and natural gas power plants will decrease their production and coal will phase out, supposedly leading to less quantities bid at non-zero price. But This can be justified by the fact that hydro generation is set to increase by 2030, and hydro generators usually bid their quantities of electricity at a non-zero prices to maximize plant profitability and also for their pump and storage capabilities. Therefore, the quantiles bid at non-zero price will hold ground and will not be affected much by the planned change of the 2030 energy mix.

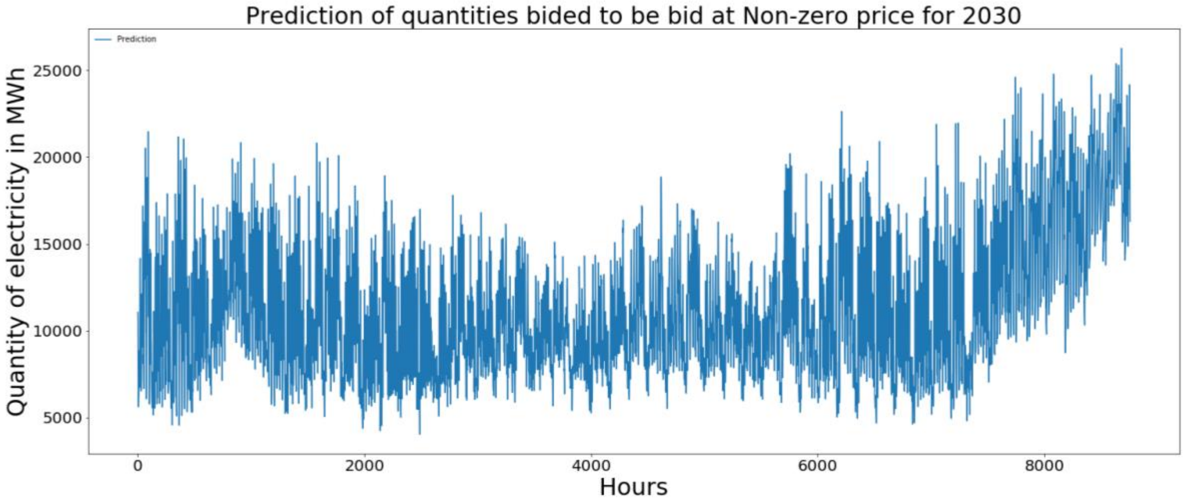


Figure 47 Forecasted hourly quantities bid at non-zero price for 2030, MIBEL

Again, in order to closely analyse the results obtained above the forecasted hourly quantities to be bid at non-zero price segment are divided into 4 seasons following the same procedure with the zero-price segment.

7.2.2.1 Summer:

In figure 48 the quantities of non-zero bids are represented for the summer season. First thing that can be noticed that is different from the quantities bid at zero price is the fluctuations. This is caused as the technologies that bid at non-zero price usually adjust their production to meet the demand requirements. It can be also observed that the quantities bid at non-zero price is low throughout the summer days and only starts to increase heading towards autumn. Which fits well with the observations made for zero bids in the previous section, showing that they have an inversely proportional relation.

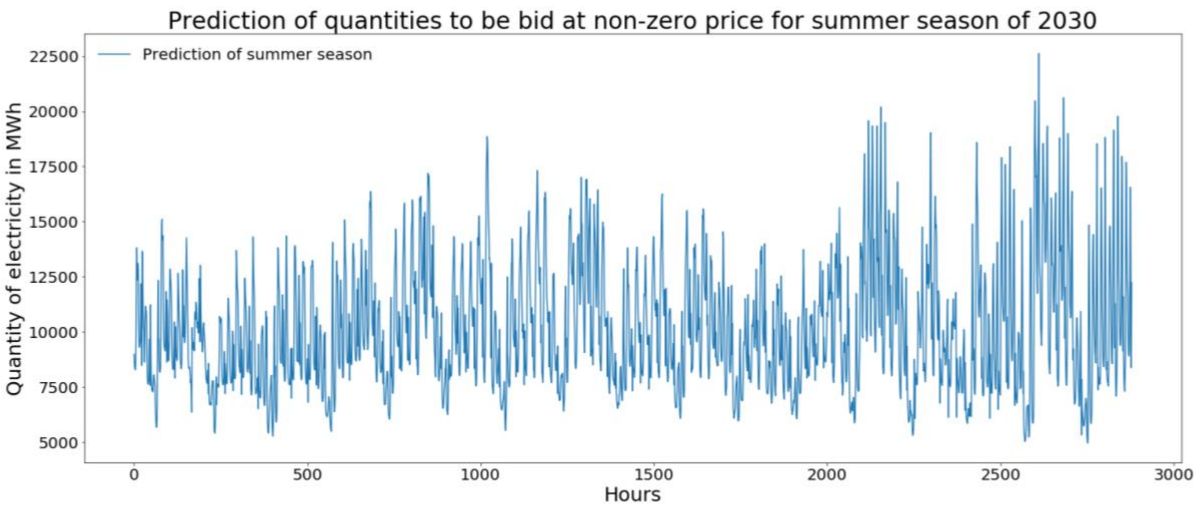


Figure 48 Forecasted hourly quantities bid at non-zero price during summer 2030, MIBEL

The pattern of the average hourly distribution for the summer season is shown in figure 49. The pattern hear is different from the zero-price segment. Two peaks are seen. A Moring peak to supply the morning demand that can't be achieved by only renewable sources. And the evening peak where technologies like solar is not producing and the demand should be supplied by other generators.

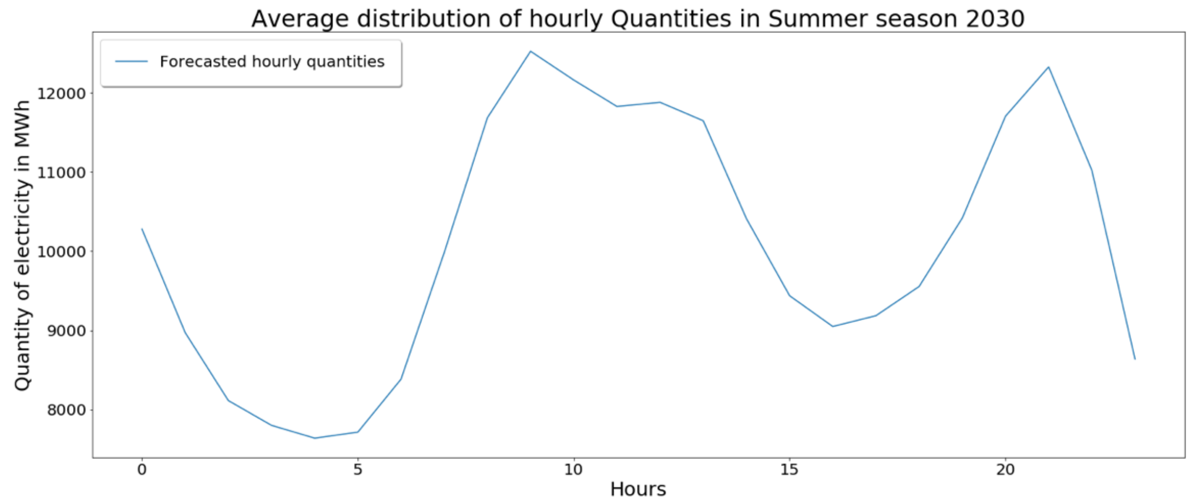


Figure 49 Forecasted average distribution of hourly non-zero price quantities during summer 2030, MIBEL

7.2.2.2 Autumn:

In figure 50 the quantities bid at nonzero prices show an increase from the beginning of the autumn season and towards the winter season. This is a result of the decrease of renewable quantities as discussed before, and the need to compensate this decrease in quantities to supply the demand.

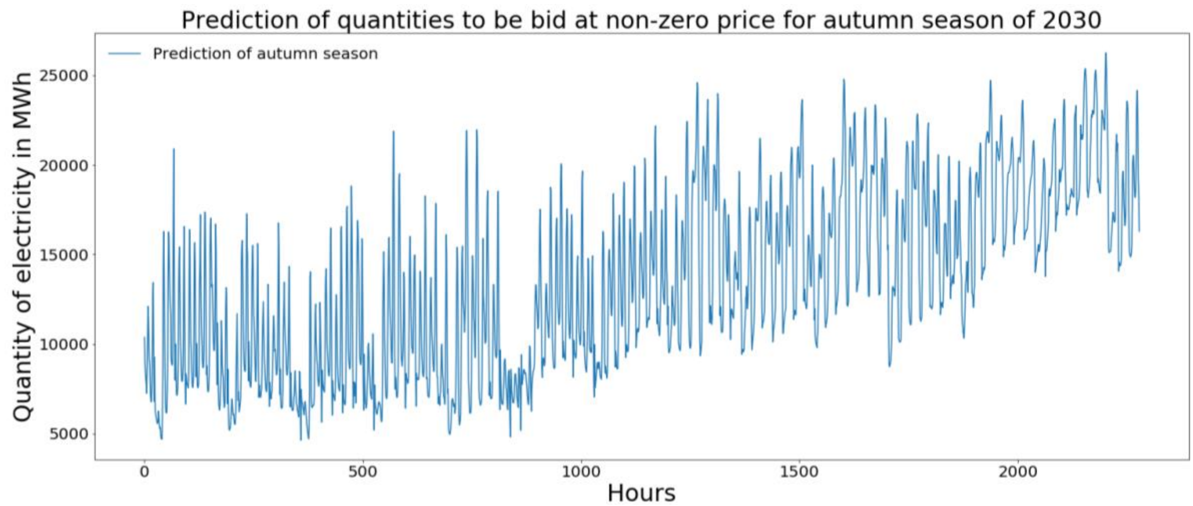


Figure 50 Forecasted hourly quantities bid at non-zero price during autumn 2030, MIBEL

In figure 51 the average hourly distribution during the autumn season has a similar pattern yet there is a noticeable increase of quantities compared to the summer season.

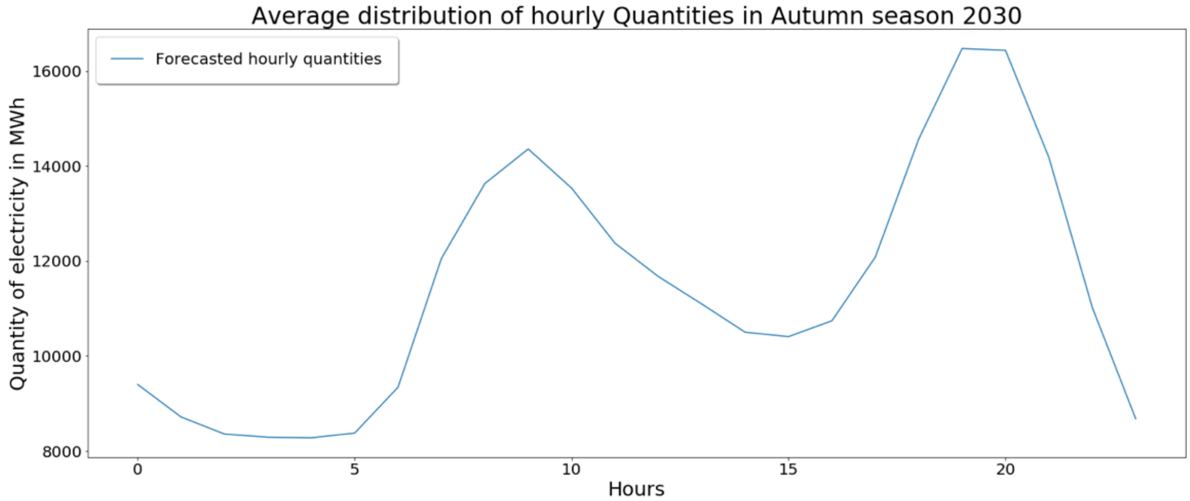


Figure 51 Forecasted average distribution of hourly non-zero price quantities during autumn 2030, MIBEL

7.2.2.3 Winter:

In figure 52 the quantities bid at non-zero prices during the winter season are represented. The fluctuations can be still noticed for the same reason explained before.

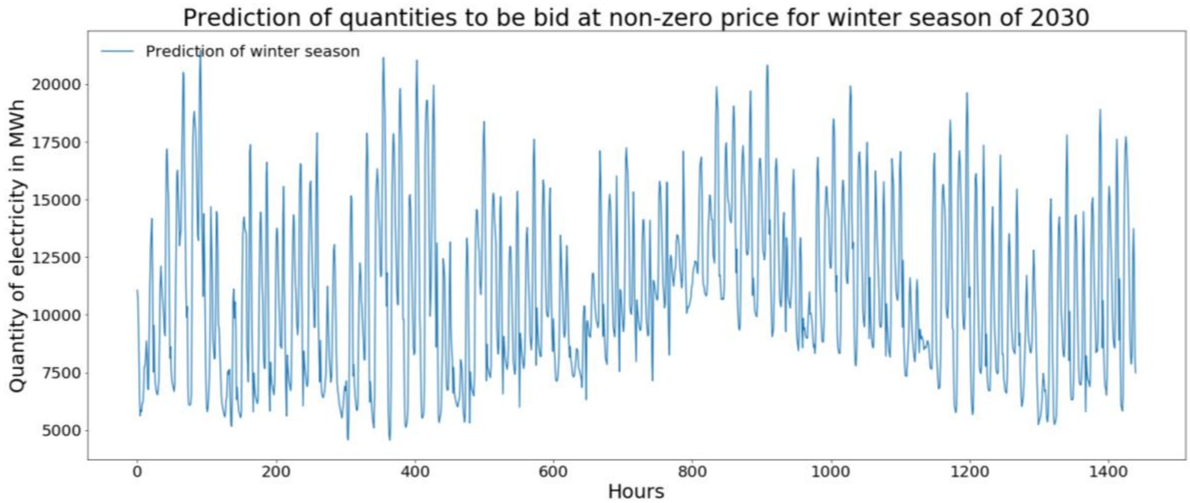


Figure 52 Forecasted hourly quantities bid at non-zero price during winter 2030, MIBEL

In figure 53 the average hourly distribution during the winter season maintains a similar pattern. But the average quantities increase compared to the previous season.

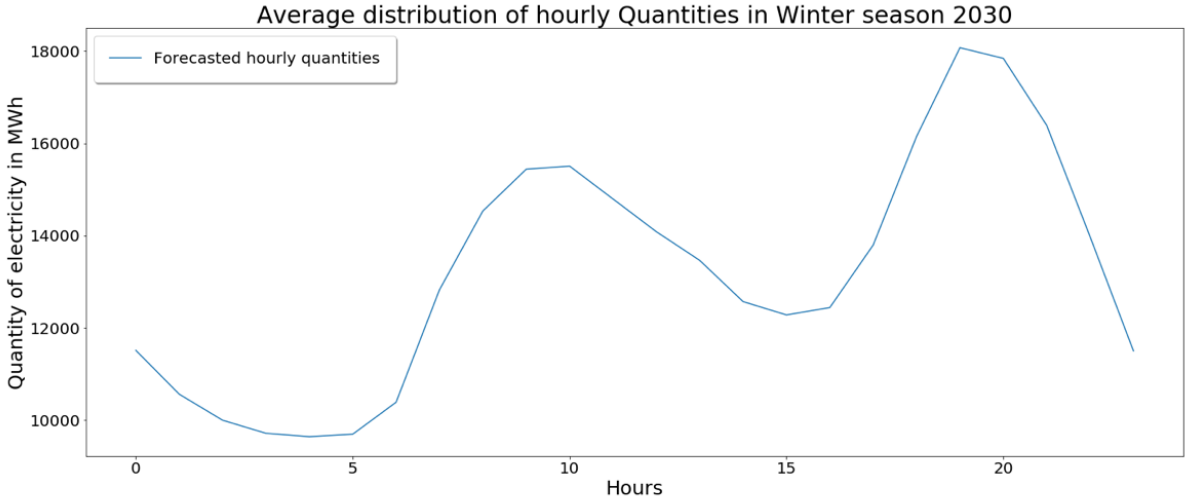


Figure 53 Forecasted average distribution of hourly non-zero price quantities winter 2030, MIBEL

7.2.2.4 Spring:

In figure 54 it can be noticed that quantities start to decrease again as we go towards the summer season. As a result of the increasing of renewable production.

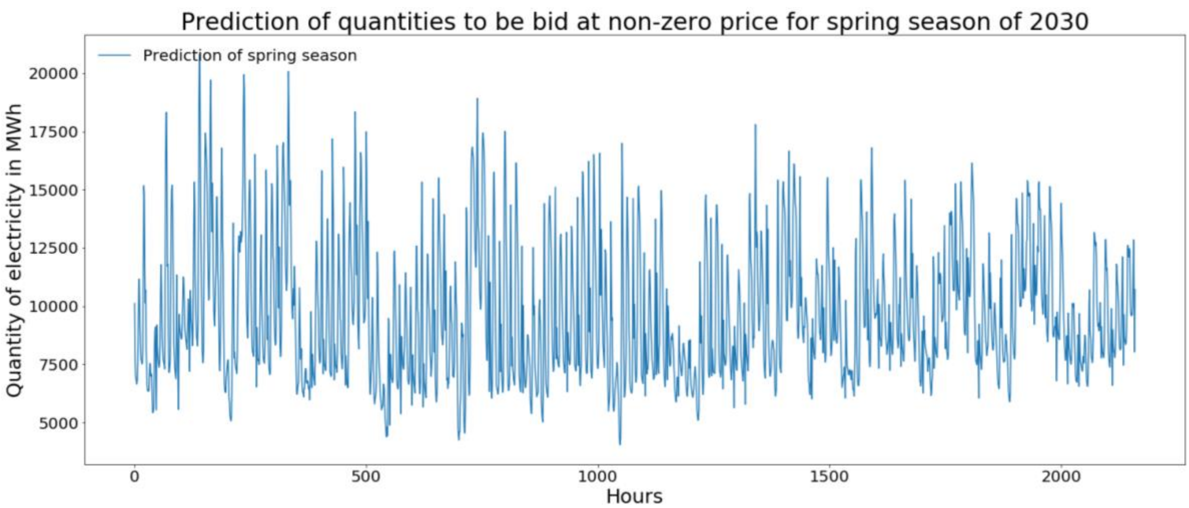


Figure 54 Forecasted hourly quantities bid at non-zero price during spring 2030, MIBEL

The pattern of the average hourly distribution is shown in figure 55. As observed the pattern is still the same showing two peaks and is decreasing quantities heading towards the summer levels.

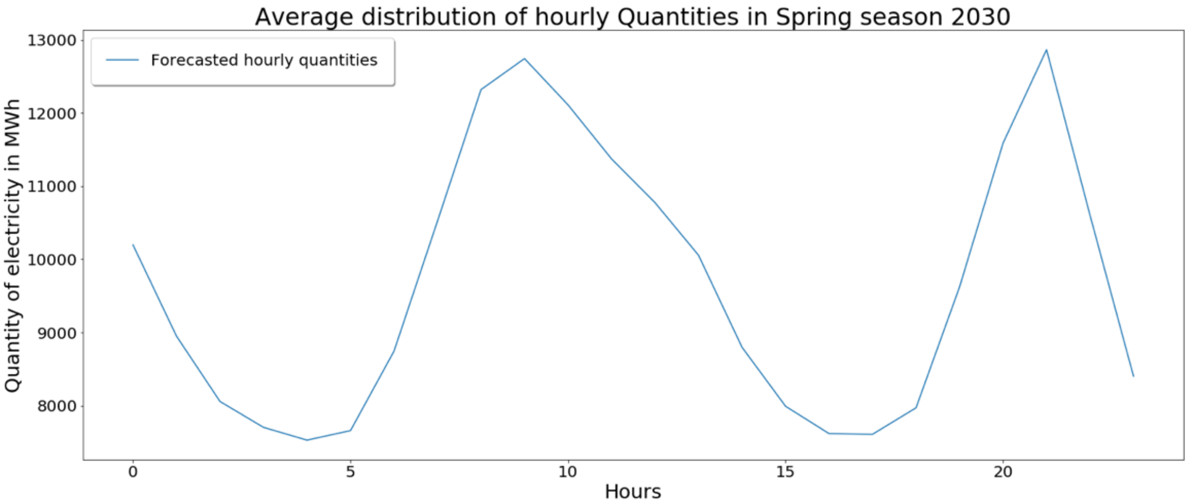


Figure 55 Forecasted average distribution of hourly non-zero price quantities spring autumn 2030, MIBEL

7.1 Conclusions

Form this chapter is possible to conclude that in 2030 an overall increase of quantities bid at zero price will occur. Such increase is caused by the enormous amount of renewable electricity generation, that increases from 39% in 2019 to 86% in 2030, specially from solar energy. On the other hand, the quantities bid at non-zero price will not be affected as explained in the section above.

Chapter 8

8. Conclusion

Several subjects were discussed during the development of this dissertation in order to fulfill the objectives of this work. This chapter is broken into two sections that summarize the key findings and conclusions of this research and recommendations for future work. The first section summarizes the most important conclusions and findings from the entire research process. The second section discusses future work and enhancements.

8.1 Findings and Conclusion

The main objective of this paper is to forecast the quantities of electricity bid in the 2030 day-ahead market, to study the Iberian market behaviour to the increasing renewable penetration, and its effect on the quantities that are bid at the day ahead market. The quantities of electricity bid are based on two price segment, zero and non-zero prices. The zero-price segment is the first part of the supply curve and usually expresses the quantity of renewable generation. The model that was used for the forecasting purpose is an artificial neural network model that uses input data like; hourly generation by technology, date and time, to forecast the quantities of electricity for the mentioned price segments. This would be the first step towards forecasting the whole supply curve.

By the end of this dissertation, and after reviewing all of the developed work, it is necessary to highlight some of the most important findings and conclusions reached throughout the process.

First, in order to have a model, data had to be collected. And since the MIBEL market is operated by Portugal and Spain. Therefore, it is mandatory to gather information with reference to both countries. Historical hourly bids and historical hourly electricity production per technology were collected from OMIE, which is the Spanish market operator. OMIE only provided two years and half of hourly data. From 2019 till June 2021 from which the model was fed.

It was very important to have a description of the future energy mix and production distribution of the year under study. For this work only one projection is considered for the future scenario for 2030. The chosen projection is the governmental projection (RNC + PNIEC). The mentioned projection was selected due to the fact that it is the only ambitious projection where nuclear and coal generation totally phase out. Coal and nuclear power plants are planned to be decommissioned by 2030 and 2035 respectively. Moreover, data availability was a key point for the selection of this projection.

In order to forecast for 2030, the model inputs resembled in the hourly production of the mentioned year had to be developed. The production data presented in the (RNC + PNIEC) projection for 2030's MIBEL energy mix is presented on an annual scale, rather than on an hourly resolution. Therefore, it was necessary to convert these annual production values from a yearly basis to an hourly basis in order for the model to be able to forecast the quantities of electricity bid in the DAM for 2030. This was accomplished by collecting past patterns of such variables and duplicate them for 2030, assuming that the distribution of today's inputs which are the hourly generation per technology will be similar in 2030. This assumption is fairly correct, because the most significant change in the future will be in the total values of production, not in their distribution.

The model’s training dataset was composed of years 2019 and 2020. After the model was trained, validation was performed to evaluate the model’s accuracy. The idea behind validating is to have historical (real) data for both the inputs, which are equivalent to the model predictors, and their corresponding outputs, as model targets. This allows us to compare the model’s outputs with real data and compute the error. The validation dataset used was the first half of year 2021. A part of the validation process the developed model is also evaluated by some error metrics to provide an overview of the models' capabilities to forecast, resulting in MAPE values of 8.7% and 13.6%, MAE values of 1534.9 and 1148.8 MWh, RMSE values of 1950.3 and 1490.53 MWh and finally CV-RMSE values of 11.6% and 16.7%. These values are respectively for zero and non-zero quantities.

According to the model findings, the expected increase in renewable penetration by 2030 tends to increase the quantities of electricity bid at zero price and hence lower the average hourly electricity market price as a result of shifting the supply curve against the demand. For year 2030 the increase of renewable generation is the main trend therefore, it is concluded that the quantities bid at zero price will significantly increase and this will cause spot prices to decrease at hours with high renewable generation.

Finally, it is clear that this study was created utilizing public data from the Iberian Electricity Market. However, the suggested approach may be carried out with different electricity markets if the variables of the prediction model are modified, and the weights of the associated ANN's are retrained.

8.2 Recommendation of Future Work

With all of the previously mentioned findings and information presented in this thesis, it is obvious that the developed model can be enhanced to provide a more comprehensive knowledge of the topic. Therefore, the following proposals for future development are recommended:

- Acquiring the missing generation information:

  In order to achieve the end goal of modeling the supply curve information regarding the type of technology associated with each bid should be known, to be able to categorize bids by their technology type and build a simple aggregated supply curve by calculating the average price of each technology. A starting point for obtaining this information would be the TSO of both Portugal and Spain REE and REN.

- Involving different input Variables:

  Hourly generation, Month, Day and hour, were the chosen explanatory variables to model electricity prices in the long run. Other variables like; hourly load, hourly weather forecasts, future fuel prices and carbon emission costs, can be found useful and could add effectiveness to the model for the proposed problem.

- Considering other future scenarios:

  Test different 2030's scenarios, different considerations about future energy mix can influence the model output. In this work the most optimistic scenario was used to model the future energy mix. Therefore, other future scenarios can be also simulated.

Bibliography

[1] E. Commission. 2020 climate and energy package, 2019. URL https://ec.europa.eu/clima/policies/strategies/2020_en. Accessed: 2-07-2021.

[2] E. Commission. 2030 climate and energy framework, 2019. URL https://ec.europa.eu/clima/policies/strategies/2030_en. Accessed: 2-7-2021.

[3] E. Comission. 2050 long-term strategy, 2019. URL https://ec.europa.eu/energy/en/topics/energy-strategy-and-energy-union/2050-long-term-strategy. Accessed: 02-07-2021.

[4] Ritchie, H., & Roser, M. (2021). CO_2 and Greenhouse Gas Emissions. URL https://ourworldindata.org/emissions-by-sector. Accessed: 02-07-2021.

[5] DESCRIPTION OF THE OPERATION OF THE MIBEL. (2009). Retrieved 1 May 2021, from https://www.mibel.com/wp-content/uploads/2018/08/Estudio_MIBEL_EN_v2.pdf

[6] O. de los Mercados Diario (OMIE). Omie, 2019. URL http://www.omie.es/inicio. Accessed: 21-10-2021.

[7] R. E. N. (REN). Redes energéticas nacionais, 2019. URL https://www.ren.pt. Accessed: 21- 10-2021.

[8] R.E.de España (REE). Red eléctrica de España, 2019.URLhttps://www.ree.es/es/. Accessed: 21-10-2021.

[9] Aggregate supply and demand curves MIBEL | OMIE. (2021). Retrieved 1 August 2021, from https://www.omie.es/en/market-results/daily/daily-market/aggragate-suply-curves?scope=daily&date=2021-01-15&hour=1

[10] Technical Data. (2020). Retrieved 5 July 2021, from https://datahub.ren.pt/media/udphiynt/dados-técnicos-2020.pdf

[11] APREN - Production. (2021). Retrieved 5 July 2021, from https://www.apren.pt/en/renewable-energies/production

[12] R. E. de España (REE). The Spanish Electricity System 2020, 2021. URL https://www.ree.es/sites/default/files/publication/2021/07/downloadable/inf_sis_elec_ree_2020_0EN_0.pdf. Accessed: 11-8-2021.

[13] 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. doi: 10.1016/j.ijforecast.2014.08.008

[14] Liu, Z., Yan, J., Shi, Y., Zhu, K., & Pu, G. (2012). Multi-agent based experimental analysis on bidding mechanism in electricity auction markets. International Journal Of Electrical Power & Energy Systems, 43(1), 696-702. doi: 10.1016/j.ijepes.2012.05.056

[15] Ventosa, M., Baíllo, Á., Ramos, A., & Rivier, M. (2005). Electricity market modeling trends. Energy Policy, 33(7), 897-913. doi: 10.1016/j.enpol.2003.10.013

[16] Day, C., Hobbs, B., & Pang, J. (2002). Oligopolistic Competition in Power Networks: A Conjectured Supply Function Approach. IEEE Power Engineering Review, 22(5), 68-68. doi: 10.1109/mper.2002.4312211

[17] Batlle, C., & Barquin, J. (2005). A Strategic Production Costing Model for Electricity Market Price Analysis. IEEE Transactions On Power Systems, 20(1), 67-74. doi: 10.1109/tpwrs.2004.831266

[18] Bower, J., & Bunn, D. (2000). Model-Based Comparisons of Pool and Bilateral Markets for Electricity. The Energy Journal, 21(3). doi: 10.5547/issn0195-6574-ej-vol21-no3-1

[19] Santos, G., Pinto, T., Praça, I., & Vale, Z. (2016). MASCEM: Optimizing the performance of a multi-agent system. Energy, 111, 513-524. doi:10.1016/j.energy.2016.05.127

[20] Gao, C., Bompard, E., Napoli, R., & Zhou, J. (2008). Design of the electricity market monitoring system. Proceedings of DRPT 2008 (pp. 99–106), art. no. 4523386.

[21] Liebl, D. (2013). Modeling and forecasting electricity spot prices: A functional data perspective. The Annals Of Applied Statistics, 7(3). doi: 10.1214/13-aoas652

[22] Gonzalez, V., Contreras, J., & Bunn, D. (2012). Forecasting Power Prices Using a Hybrid Fundamental-Econometric Model. IEEE Transactions On Power Systems, 27(1), 363-372. doi: 10.1109/tpwrs.2011.2167689

[23] Johnsen, T. (2001). Demand, generation and price in the Norwegian market for electric power. Energy Economics, 23(3), 227-251. doi: 10.1016/s0140-9883(00)00052-9

[24] Barlow, M. (2002). A DIFFUSION MODEL FOR ELECTRICITY PRICES. Mathematical Finance, 12(4), 287-298. doi: 10.1111/j.1467-9965.2002.tb00125.x

[25] Coulon, M., & Howison, S. (2009). Stochastic behavior of the electricity bid stack: from fundamental drivers to power prices. The Journal Of Energy Markets, 2(1), 29-69. doi: 10.21314/jem.2009.032

[26] Hamilton, J. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. Econometrica, 57(2), 357. doi: 10.2307/1912559

[27] Shahidehpour, M., Yamin, H., & Li, Z. (2003). Market operations in electric power systems: forecasting, scheduling, and risk management. John Wiley & Sons.

[28] Taylor, J. (2010). Triple seasonal methods for short-term electricity demand forecasting. European Journal Of Operational Research, 204(1), 139-152. doi: 10.1016/j.ejor.2009.10.003

[29] Tong, H., & Lim, K. (1980). Threshold Autoregression, Limit Cycles and Cyclical Data. Journal Of The Royal Statistical Society: Series B (Methodological), 42(3), 245-268. doi: 10.1111/j.2517-6161.1980.tb01126.x

[30] Weron, R., & Misiorek, A. (2006). Short-term electricity price forecasting with time series models: A review and evaluation. In W. Mielczarski (Ed.), Complex electricity markets (pp. 231–254). Łódź: IEPŁ& SEP.

[31] Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50(4), 987–1007. https://doi.org/10.2307/1912773

[32] B. F. Hobbs, U. Helman, S. Jitprapaikulsarn, S. Konda and D. Maratukulam, “Artificial neural networks for short-term energy forecasting: Accuracy and economic value,” Neurocomputing, vol. 23, no 1-3, pp. 71-84, 1998.

[33] N. Kohzadi, M. S. Boyd, B. Kermanshahi and I. Kaastra, “A comparison of artificial neural network and time series models for forecasting commodity prices,” Neurocomputing, vol. 10, no 2, pp. 169-181, 1996.

[34] H. Z. Zou, G. P. Xia, F. T. Yang and H. Y. Wang, “An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting,” Neurocomputing, vol. 70, no 16-18, pp. 2913-2923, 2007.

[35] Buckley, J., & Yoichi, H. (1995). Neural nets for fuzzy systems. Fuzzy Sets And Systems, 71(3), 265-276. doi: 10.1016/0165-0114(94)00282-c

[36] Nauck, D., & Kruse, R. (1999). Neuro-fuzzy systems for function approximation. Fuzzy Sets And Systems, 101(2), 261-271. doi: 10.1016/s0165-0114(98)00169-9

[37] Hong, Y., & Hsiao, C. (2002). Locational marginal price forecasting in deregulated electricity markets using artificial intelligence. IEE Proceedings - Generation, Transmission And Distribution, 149(5), 621. doi: 10.1049/ip-gtd:20020371

[38] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273-297. doi: 10.1007/bf00994018

[39] Čižek, P., Härdle, W., & Weron, R. (Eds.) (2011). Statistical tools for finance and insurance (2nd ed.). Berlin: Springer.

[40] Wang, L., & Fu, X. (2005). Data mining with computational intelligence. Springer.

[41] Sansom, D. C., Downs, T., & Saha, T. K. (2002). Evaluation of support vector machine based forecasting tool in electricity price forecasting for Australian national electricity market participants. Journal of Electrical and Electronics Engineering, Australia, 22(3), 227–233.

[42] F. Ziel and R. Steinert, “Probabilistic mid- and long-term electricity price forecasting,” Renewable and Sustainable Energy Reviews, vol. 94, pp. 251-266, 2018.

[43] J. P. S. Catalão, S. J. P. S. Mariano, V. M. F. Mendes and L. A. F. M. Ferreira, “Short- term electricity prices forecasting in a competitive market: A neural network approach,” Electric Power Systems Research, vol. 77, no 10, pp. 1297-1304, 2007.

[44] U. Ugurlu, I. Oksuz and O. Tas, “Electricity Price Forecasting Using Recurrent Neural Networks,” Energies, no 11.5, p. 1255, 2018 .

[45] Pardo, A., Meneu, V., & Valor, E. (2002). Temperature and seasonality influences on Spanish electricity load. Energy Economics, 24(1), 55-70. doi: 10.1016/s0140-9883(01)00082-2

[46] MIRASGEDIS, S., SARAFIDIS, Y., GEORGOPOULOU, E., LALAS, D., MOSCHOVITS, M., KARAGIANNIS, F., & PAPAKONSTANTINOU, D. (2006). Models for mid-term electricity demand forecasting incorporating weather influences. Energy, 31(2-3), 208-227. doi: 10.1016/j.energy.2005.02.016

[47] I. Vehviläinen, T. Pyykkönen, “Stochastic factor model for electricity spot price- the case of the Nordic market”, Energy Economics 27 (2005) 351-367.

[48] H. Mohammadi, “Electricity prices and fuel costs: Long-run relations and short-run dynamics,” Energy Economics, vol. 31, no 3, pp. 503-509, 2009.

[49] D. Kotur and Ž. Mileta , “Neural network models for electricity prices and loads short and long-term prediction,” em 4th International Symposium on Environmental Friendly Energies and Applications (EFEA), 2016.

[50] R. K. Agrawal, F. Muchahary, and M. M. Tripathi, "Long term load forecasting with hourly predictions based on long-short-term- memory networks," in 2018 IEEE Texas Power and Energy Conference (TPEC), 2018, pp. 1-6.

[51] H. Sangrody, N. Zhou, S. Tutun, B. Khorramdel, M. Motalleb, and M. Sarailoo, "Long term forecasting using machine learning methods," in 2018 IEEE Power and Energy Conference at Illinois (PECI), 2018, pp. 1-5.

[52] L. Jiang and G. Hu, "Day-Ahead Price Forecasting for Electricity Market using Long-Short Term Memory Recurrent Neural Network," in 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV), 2018, pp. 949-954.

[53] Barlow, M. (2002). A DIFFUSION MODEL FOR ELECTRICITY PRICES. Mathematical Finance, 12(4), 287-298. doi: 10.1111/j.1467-9965.2002.tb00125.x

[54] Buzoianu, M., Brockwell, A., and Seppi, D. J. (2005). A dynamic supply-demand model for electricity prices. Technical report, Carnegie Mellon University.

[55] R. A. Soares, J. T. Saraiva, J. N. Fidalgo and B. C. Martins, "Forecast of the bidding curve of generation players in the Iberian electricity market," 2015 12th International Conference on the European Energy Market (EEM), Lisbon, Portugal, 2015, pp. 1-5, doi: 10.1109/EEM.2015.7216659.

[56] Shah, I., & Lisi, F. (2019). Forecasting of electricity price through a functional prediction of sale and purchase curves. Journal Of Forecasting, 39(2), 242-259. doi: 10.1002/for.2624

[57] Flammini, M., Prettico, G., Mazza, A., & Chicco, G. (2021). Reducing fossil fuel- based generation: Impact on wholesale electricity market prices in the North-Italy bidding zone. Electric Power Systems Research, 194, 107095. doi: 10.1016/j.epsr.2021.107095

[58] Ziel, F., & Steinert, R. (2016). Electricity price forecasting using sale and purchase curves: The X-Model. Energy Economics, 59, 435-454. doi: 10.1016/j.eneco.2016.08.008

[59] G. M. Pereira, Master Thesis: " Simulation of the future Iberian power system till 2040," IST, 2020.

[60] Piccinini, G. (2004). The First Computational Theory of Mind and Brain: A Close Look at Mcculloch and Pitts's "Logical Calculus of Ideas Immanent in Nervous Activity". Synthese, 141(2), 175-215. doi: 10.1023/b:synt.0000043018.52445.3e

[61] A. Géron, Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow: Concepts,tools and techniques to build

[62] World Greenhouse Gas Emissions: 2016. (2021). Retrieved 1 November 2021, from https://www.wri.org/data/world-greenhouse-gas-emissions-2016

[63] Kinsley, H., & Kukieła, D. Neural networks from scratch in Python.