Imperial College London

Imperial College London

Department of Mathematics

Evaluating approaches for estimating the Day-Ahead price and Water Value in hydro-dominated electricity market

Author: Thomas SIMON (CID: 02260739)

A thesis submitted for the degree of

MSc in Mathematics and Finance, 2022-2023

Declaration

The work contained in this thesis is my own work unless otherwise stated.
Signed : Thomas SIMON

Acknowledgements

I would first like to thank my supervisor and lecturer Dr. Cristopher Salvi for his support throughout my thesis and for the excellence of his lectures.

I am also grateful for the guidance that I have received from Marie Heintzmann during my internship who gave me the opportunity to do some interesting research and improve my knowledge of electricity markets.

Finally, I would like to thank my parents and my brother for their unwavering support and encouragement throughout my years of study. This accomplishment would not have been possible without you.

Abstract

Electricity pricing in a liberalised market is essentially based on matching supply and demand in a market clearing process. The supply curve, based on an increasing aggregation of the marginal costs of various means of electricity generation, induces a dynamism in prices that depends on the competition between these generation units. Therefore, how can the true marginal cost of production be defined in a system dominated solely by hydraulic generation? In other words, what value is given to the water stored and dedicated to electricity generation? In this thesis we elaborate on how a hydropower marginal value of water called Water Value is defined when considering no thermal units but availability to import or export power. Several general aspects of electricity markets and a possible stochastic modelling of the Day-Ahead price are given before defining Water Value and its impact on the functioning of the market in hydro-dominated countries. The results are tested numerically on data from Norway, a country that embodies this situation.

Contents

List of Figures

List of Tables

Introduction

Among the world of commodities, electricity remains relatively unique. Indeed, one of its fundamental characteristics is its inability to be stored in large quantities. This lack of storage has thus entailed typical behaviors in this market. Firstly, many traditional strategies and risk management techniques fail with this commodity due to lack of storage capacities of the underlying asset compared to stock or bond market. In addition, numerous physical factors such as energy transfer grid, transmission lines and availability of production sites have generated additional constraints that have a truly unique impact on prices. Furthermore, this market has developed three specific characteristics that many models attempt to reproduce: seasonality, that can be annual, weekly or even daily, which can be simulated by Wavelet-based modelling, see for example [2] and [3], the effects of spikes whose causes are multi- factorial such as congestion in transmission lines [4] and very sharp mean reversion behavior which appears generally after a sudden price peak [5]. Since the creation of liberal electricity trade, these parameters have been studied in numerous research articles, but those characteristics make all the analyses fairly complex since few techniques such as Black-Scholes modelling cannot be applied to this type of asset essentially because of the violation of the geometric Brownian motion hypothesis of the underlying.

Within this framework of scientific research on the modelling of electricity markets, there is a very specific type of market, only a handful of countries depend on today: namely hydro- dominated markets, i.e. where electricity production is essentially reliant on hydraulic units (run-of- river, hydro-reservoirs and Pumped Storage Power Station). In most countries, such as France, the United Kingdom and Spain, the coordination of hydro generation is often based on price indicators and frequently complements other means of generation. In these countries, hydro-assets do not drive prices, but adapt to them. On the other hand, in Norway, which is the reference country for this thesis, the electricity production is based on a strategy that relies heavily on water. The study of prices is thus more complex because additional meteorological factors, such as quantity of water that the country possesses and expects to receive over the study period are less foreseeable. To understand how water prices evolve, electricity producers need to take into account what the value of stored water against production strategy and the market is. Researchers and professionals refer to this concept more commonly as Water Value. The problem lies in the fact that there is no single interpretation of Water Value, and it all depends on the researcher's approach to the problem.

Nevertheless, two main definitions seem to stand out in the report [2]: profit-based water value, which corresponds to a variation in the company's profit for a marginal variation in its water resources, and cost-based water value, which is defined as the marginal cost of substituting a water resource with a thermal resource. The latter provides us with important information as it potentially links marginal production costs to prices observed on the markets. Several approaches have already been considered for calculating water value, as listed in the following document [6] based on estimation using a stochastic process, Linear Programming or even Mixed-Integer Linear Programming Algorithms.

Norway is divided into five zones Oslo (NO1), Kristiansand (NO2), Trondheim (NO3), Tromso (NO4) and Bergen (NO5). The market is handled by Nord Pool, a power exchange that facilitates the trading of electricity in the Nordic and Baltic regions. The Nord Pool Group was established in 1993 and has since become a key player in the energy industry, providing transparent and efficient electricity market solutions.

This paper is organised in the following manner. Chapter 1 evolves around a global description of the main features of electric market to give the reader an overview on how this market is organised. Since this thesis mainly focuses on hydraulic production systems, generation units are described and an analysis of Norway's situation is presented with quantitative statistics of the five

bidding zones. In Chapter 2, an initial price model is built, attempting to reproduce the behaviour of day-ahead prices directly from NordPool [7] data. The analysis is carried out more specifically for 2017, with a focus on the NO1 and NO4 zones, which globally reflect the two price behaviours in Norway. Thus, a deterministic part of the prices is first built with the analysis of a weekly trend and a theoretical part to finally add a stochastic part built from an ARMA process on the residuals. Finally, Chapter 3 focuses on a supply and demand model based on the concept of water value. After a general description of this type of model for prices and types of hydraulic production, a quantitative analysis is carried out to construct the demand curve using stochastic processes and a supply curve model is constructed based on influencing factors such as CO2 prices and carbon emission contracts to determine the factors influencing the formation of the day-ahead price. At the end, the concept of water value is analysed in more detail with a comparison of existing models, the situation in Norway, and modelling using optimisation solutions to market problems is carried out with the aim of recreating price dynamics in the NO1 zone.

Chapter 1

Electric spot prices

This part is dedicated to a global description of how electricity markets work. Besides, there are similarities between equities and bonds market, electricity is a special commodity whose interest has been gradually increasing since the Trade Liberalisation in the early nineties. For a better understanding on how pricing and trading are organised, behaviour of this commodity and how the market is built to answer requirements of consumers and producers must be analysed. That is why the first part presents some features of transport and storage and also the two main types of market, the Intraday and the Day-Ahead Market. Secondly, since this report focuses on electric hydrogeneration prices, a presentation of the hydro modelling supply plants is given and how they impact the behaviour of the market. Finally, a brief description of the case of Norway is underlined and statistical facts of Day-Ahead price models are presented.

1.1 Global Features and Market Microstructure

One key point when analysing the commodity sector is to take the physical constraints into account. Indeed, electricity is highly concerned because it has to be delivered at a specific time and given area, while one has to maintain a strict balance between supply and demand on the grid at all times. Some components, such as storage plants and transmission infrastructures, must be integrated into the market for the success of the consumer delivery contracts. Indeed, they ensure the foundations of well- functioning of the electricity market [8, Part 2].

1.1.1 Transmission

Like many raw materials, very often, electricity is not consumed where it is produced. A Transport network is needed to provide the link between the production facilities and the consumers [9]. One must bear in mind that electric current is a movement of electrons in a network, generally copper, between two areas experiencing different potential and the current takes all possible roads from one point to another. This phenomenon belongs to well-known physical laws, namely Kirchhoff's laws.

Definition 1.1.1 (Kirchhoff's laws)**.** Kirchhoff's first law or Kirchhoff's junction rule states that, for any node in an electricity grid, the sum of the currents flowing out of the node is equal to the sum of currents flowing into it. Note that the current is signed (positive or negative depending on the orientation rule) reflecting if it goes into or out of a node, we have the following mathematical formulation:

$$\sum_{n=1}^{N} I_n = 0 \tag{1.1.1}$$

where N is the number of branches connected to the node.

Transmission limits may be exceeded when the demand is very high and cannot be contained by supply leading to network breakdown. To enable liberalised electricity markets to open up, operators have had to define a number of rules both to prevent overloading and to ensure the stability of the network. According to [10][part 2.1.2], the market was organised by Transmission system operators (TSO) which by auction mechanism, had to buy a tax for transferring electricity

from one country to another, in the European area. Meanwhile, in the Northern countries (Norway, Finland, Sweden,..) members of NordPool marketplace, buyers and sellers send bids to the exchange market without concern for transport constraints and consequently the responsability is given to the market coordinator. From 2006 onwards, the last option has been mainly adopted across Europe to create a unification between all markets, a concept known as "Market coupling".

To ensure the stability of the electric networks, a smart grid concept has been developed in order to integrate new sources of energy produced by essentially renewable technologies. More flexibility is essential in the transmission grids to overcome spikes of supply and demands. In fact, researchers have shown tremendous interest in the resilience of electrical infrastructures by modelling them for example with a Bayesian network [4]. They have attempted to underline the advantages and disadvantages of current grids. Finally, more political decisions to put an end to consumption are settled to face incapacity to fulfil the demand when intermittent plants are no longer available to maintain flexibility on the market.

1.1.2 Storage

Electricity cannot be stored instantly because, by definition, in order to create an electric current, electrons have to move from point A to B. Without movement, there is no electric power. However, electricity can be stored by physical or chemical transformations to create a potential energy that can be restated in the future by inverted transformations. Through those technologies, electricity can be stored but transformed into another potential energy. This involves an optimisation issue between available generation plants and potential storage plants that can be used to answer the electricity consumption which is continuous over time. One type of storage that has really enabled the democratization of electricity storage is hydraulic reservoirs. Widely used in mountainous countries, where abundant rainfall and snowmelt enable dams to be filled with water, a potentially usable, long-term reserve of energy that can be converted into electricity, has made this type of production a widespread part of the energy mix in many countries.

1.1.3 Day-Ahead Market

Since the Liberalisation of Trade and the deregulation of electricity prices in the 1990s, two main types of market have emerged: Intraday market and Day-Ahead market. A Day-Ahead market enables the well-being of all electricity auctions [11, Part 5.2]. These two markets are complementary and have their own specific features that ensure the smooth running of electricity exchanges. One of the special features of the electricity market is the nature of the traded product. Unlike equity markets and even other commodity markets, electricity is a product that cannot be stored in quantity compared to the amount traded on the market. As a result, electricity cannot be traded at a given moment in time. Contracts have to be drawn up in advance, usually the day before, specifying the quantity to be delivered, the period during which the system operators must supply the electricity and, finally, the recipient customer. As its name suggests, the Day-Ahead market is a market where the spot price of electricity is traded for the hours of the following day. It is based on the principle of a blind auction that takes place once a day throughout the year. Market participants are entitled to submit bids and offers for the hours of day d until 12:00 CET on day d-1. After this time, orders are no longer accepted and the price is calculated using a supply and demand balancing algorithm.

There main type of order in the market is referred to as single hourly order. This type of order allows the trader to buy or sell a certain quantity of MW at a specific time and at an agreed price between the minimum and maximum price imposed by the market. Furthermore, there is more complexity because the buyer or the seller can choose to buy/sell a certain amount of electricity depending on the price. If the quantity does not depend on the price, we call this order price either independent order or a price dependent order. Let us consider the following example to further understand:

Example 1.1.1 (Dependent and independent orders)**.** Imagine one wants to buy 100 MW at any price between the minimum price and the maximum price, in the first two hours of the sale. The order will be described as the following:

Hour\Price	-200	-20	2	...	2000
1	100	100	100	...	100
2	100	100	100	...	100

Table 1.1: Example of price independent order

Let us now consider that the buyers care about the price. The following table shows an example of dependent price order:

Hour\Price	-200	-20	2	2.1	...	2000
1	100	80	20	18	...	-200
2	70	55	10	10	...	-180

Table 1.2: Example of price dependent order

According to NordPool website [7], NordPool interpolate volumes between adjacent pair of submitted price steps to find the correct traded volume.

1.1.4 Intraday Market

To complete Day-Ahead market, intraday market supported by balancing mechanisms and organised by TSO consists of a place that ensure exchanges between market players themselves to ensure the well balancing between consumption and generation. If this balance is not respected, balancing operators bid on the market to ensure the real-time assessment of the grid and send a financial penalty to market players. The time frame for trading is shorter than for the day-ahead market, which means that market operators can rebalance volumes very shortly before the delivery time of the day-ahead contract This market is outside the scope of this Master's thesis and so its operation will not be studied in detail. However, readers can refer to the following article [12] for further information.

1.2 Hydro modelling approaches

The different types of electricity production are important in defining the best approaches for modelling the market. Indeed, a country dependent on renewable energies is subject to constraints linked to the intermittent nature of energy production, due to variations in wind power , sunshine and water levels in rivers and reservoirs in dams for hydroelectric power. All these constraints are accompanied with additional costs that are reflected in the pricing model. In an approach based on hydro-production, it is relevant to establish the means of production available to generate electricity from water, in order to establish the cost constraints that will be reflected in the market prices. Hydro-power by reservoir storage is a controllable form of energy, that is that the quantity of electricity produced can be controlled and managed according to demand. This is essentially due to the capacity to store water in artificial dam lakes, which provides a reservoir of water available for generation. However, the amount of water available greatly depends on the inflow from rivers and rainfall.

One can consider water to be a renewable resource, as it follows a permanent regeneration cycle. In the current economic and climatic context, this resource is attracting attention because carbon dioxide emissions with this technology are virtually zero. As a result, many countries have reverted to using hydroelectric power generation systems, and a number of technologies have been developing in order to convert the potential energy of gravity contained in water, firstly into mechanical energy and then, with the help of turbines, into electrical energy.

Hydroelectric sources can be divided into two main categories:

1. Installations equipped with a reservoir, which can be further divided into:
    - Storage power plants;
    - Pumped-storage power plants;
2. Run-of-river power plants, which have a short residence time.

A storage power plant operates as a pivotal component of modern energy systems, adept at balancing electricity supply and demand fluctuations. These facilities store excess energy during periods of low demand and generate electricity when demand peaks. This process involves several stages: charging, storage, and discharging. During times of surplus electricity production, such as when renewable sources like solar and wind are abundant, the storage plant converts excess energy into a storable form, like chemical potential in batteries or gravitational potential in pumped hydro storage. The stored energy is then retained until it is needed, such as during peak demand hours or when renewable energy output is low. Upon demand surge, the stored energy is released by inverting the conversion process, converting stored potential back into electrical energy, which can be supplied to the grid. This dynamic interplay enables storage power plants to act as reliable buffers, enhancing grid stability, supporting renewable energy integration, and ensuring a continuous and responsive energy supply.

Pumped-storage power plants illustrate a specialized form of energy storage that utilize gravitational potential energy to store and release electricity. These facilities consist of two reservoirs located at various elevations. During periods of excess, the electricity generation, typically when renewable sources like solar and wind are abundant, surplus energy is used to pump water from the lower reservoir to the upper reservoir, thereby storing potential energy. When electricity demand increases, the stored water is released back to the lower reservoir, flowing through turbines to generate electricity. Pumped-storage power plants are highly efficient in terms of energy storage and release, and they provide rapid response capabilities to balance grid fluctuations.

Run-of-river hydropower is a sustainable energy technology that taps into the natural flow of rivers to generate electricity. Unlike traditional dam-based hydropower, run-of-river systems operate without the need for massive reservoirs, reducing environmental disruption. The process begins with diverting a portion of the river's flow into a channel or pipeline, which directs the water's energy towards a turbine. As the high-velocity water flows through the turbine, its kinetic energy turns the blades and generates mechanical energy. This energy is then converted into electricity through a connected generator. Since run-of-river plants work with the river's existing flow, they don't significantly alter the watercourse or ecosystem. They often employ weirs or structures to ensure a consistent water supply, especially during low-flow periods. The generated electricity is then fed into the grid, contributing to the overall energy supply. Run-of-river systems are particularly well-suited for rivers with steady flow rates, providing reliable renewable energy while minimizing the ecological impact often associated with conventional hydropower projects.

1.3 A state-of-the-art Norway

The case of Norway is presented in this section because this country can be considered as the subsequent applications of our models, since Norway is one of the first countries in the world to base its electricity production on hydro-based energy. This section therefore begins with an overall description of the situation in Norway, followed by a study of the dynamics governing its five bidding zones.

1.3.1 Quick facts

Norway is a unique country when it comes to electricity production. Because of its geographical and political commitment, around 90% of its electricity production is based on hydroelectric resources.

Norway remains one of the best examples in the global energy landscape thanks to its abundant and diverse energy resources, including hydropower, oil, and natural gas. The country's electricity situation is characterized by a robust energy mix, focusing on sustainability and the ongoing transition towards increased reliance on renewable sources. Hydropower has been the backbone of Norway's electricity production for over a century. The country's topography, with numerous rivers and waterfalls, has enabled the development of a vast hydropower infrastructure. This renewable resource accounts for the majority of Norway's electricity generation, offering a consistent and reliable source of energy. The ability to store energy in hydropower reservoirs allows Norway to manage electricity supply and demand efficiently, making it a key player in the Nordic and European electricity markets.

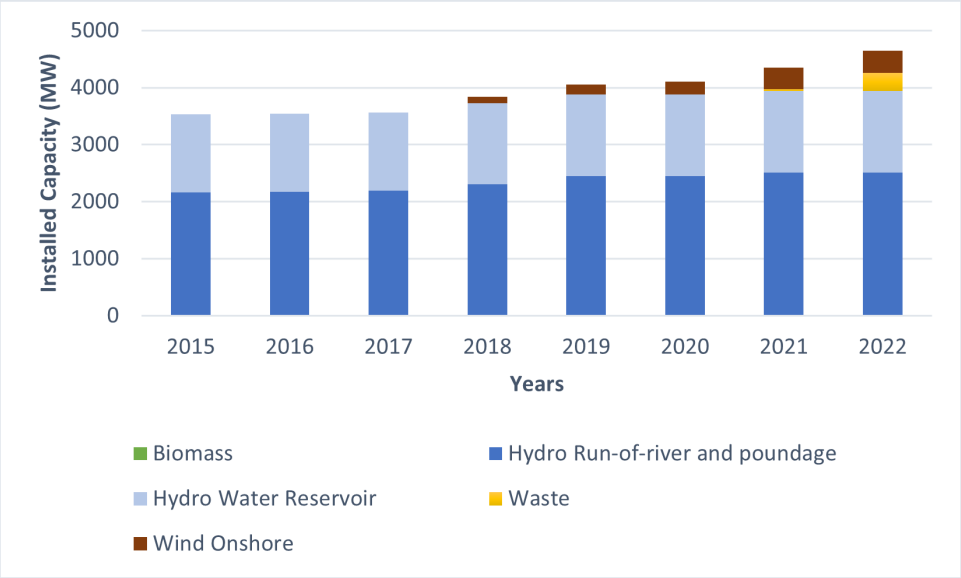


Figure 1.1: Installed production capacity Norway (2015-2022)

In Figure 1.1, Hydro power reservoir and Hydro Run-of-river and poundage are clearly the main sources of electric production. Besides the fact that Norway as a few thermal power capacity of production, thanks to network link transmission cables, electricity can be imported from bordering countries such as Sweden, Denmark, Finland and Germany in Norway's case. Inverse procedure of exporting electricity to neighbour countries is also possible and its mainly the case in Norway thanks to a low production costs by hydro generation. We can also note that the energy mix seems to be shifting slightly towards new sources such as onshore wind power, but also towards waste incineration and the use of biomass. This is fairly typical of European countries, which do not tend towards a monopoly of one type of production, but rather towards the broadest possible mix by integrating different types of production. The recent emergence of new production technologies plays an important role in price formation, as marginal production costs are not the same for each production category. New studies are being carried out into the implementation of new intermittent electricity generation techniques, with a particular focus on long-term prices. For example, in [13], some authors have studied how the future power market in Norway may be affected by risk factors using a probabilistic approach by incorporating forecast new share of renewable energies (Offshore Wind and PV).

The electric market in Norway is a dynamic and well-developed system that encompasses a range of participants and activities related to the generation and consumption of electricity. It operates under the principles of competition, transparency, and regulation to ensure efficient and reliable electricity supply across the country [14, Part 4]. At the core of the Norwegian electric market is the NordPool market, the leading power market in Europe, where electricity is traded in various markets, including the Day-Ahead, Intraday, and balancing markets. To complete section 1.1.3 and 1.1.4, balancing market ensures grid stability by addressing imbalances between scheduled and actual electricity flows.

Market participants include electricity producers, distributors, retailers, traders, and consumers. Producers generate electricity from various sources, with a significant emphasis on hydropower due to Norway's abundant water resources. Distributors manage the transmission of electricity through the grid, ensuring its safe and efficient delivery to consumers. Retailers provide electricity services to end-users, while traders engage in buying and selling electricity on the exchange to optimize their portfolios. The Norwegian electric market places strong emphasis on renewable energy and sustainability. Hydropower plays a central role, accounting for the majority of electricity generation. Additionally, Norway is engaged in exploring the integration of other renewable sources such as wind and solar powers. To ensure fair competition and market integrity, the electric market in Norway is regulated by government authorities such as the Norwegian Water Resources and Energy Directorate (NVE) and the Ministry of Petroleum and Energy. These entities oversee market operations, set regulatory frameworks, and monitor compliance with energy policies and standards.

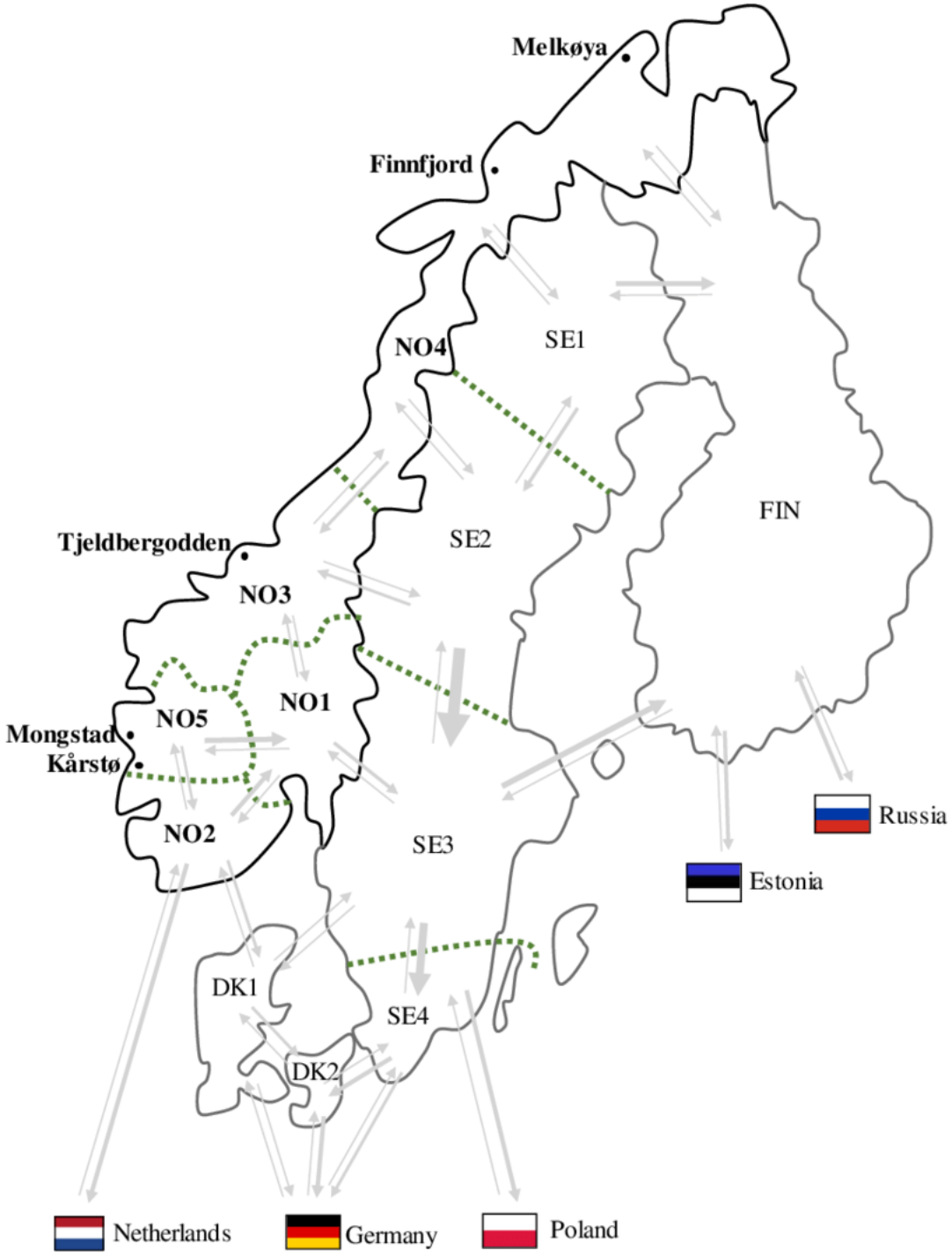


Figure 1.2: Map showing electrical interconnections in the Nordic zone [1]

1.3.2 Statistical Analysis of the five bidding zones in Norway

The purpose of this section is to investigate the main statistics of the five bidding zones. Data use for this are originally hourly day-ahead spot prices of the five bidding zones from January 2015 to December 2022.

The data set is from NordPool website [7]. The day-ahead market price is given for each hour of the study period specified above. The first relevant observation is that electricity prices can be null or even negative. Indeed, it is remarkable that some hours of this study period have negative prices, which is very different from other markets found in finance. Thus, the returns and logarithms of

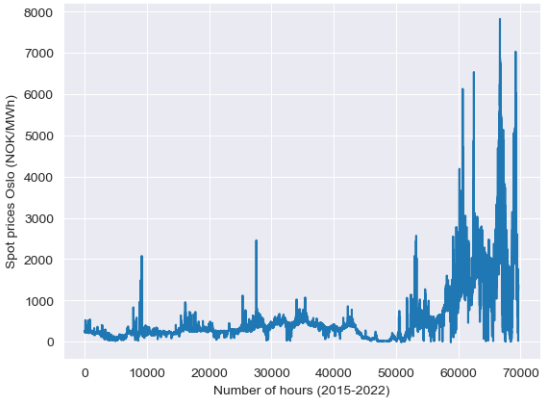
returns conventionally defined as follows are not defined at all points:

$$r_t = ln(\frac{P_t}{P_{t-1}}) \tag{1.3.1}$$

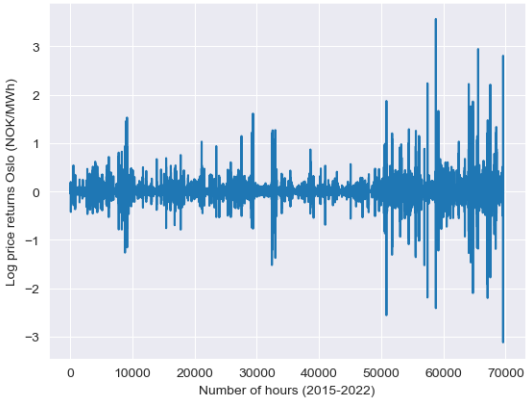
where :

1. P_t is the spot price at time t
2. P_{t-1} is the spot price at time t-1
3. r_t is the log-return at time t.

To overcome the problem of defining logarithmic returns, a constant will be added to each price to ensure that only positive values are used, so that this metric can be calculated. Therefore, we first plot Day-Ahead spot prices and returns for the five bidding zones : Oslo 1.3, Kristiansand 1.4, Trondheim 1.5, Tromso 1.6 and Bergen 1.7 from 2015 to 2022. Data are available thanks to Nord Pool AG [7].

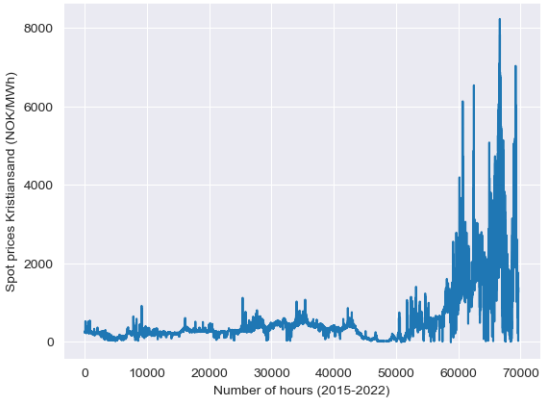


(a) Oslo spot prices (2015-2022)

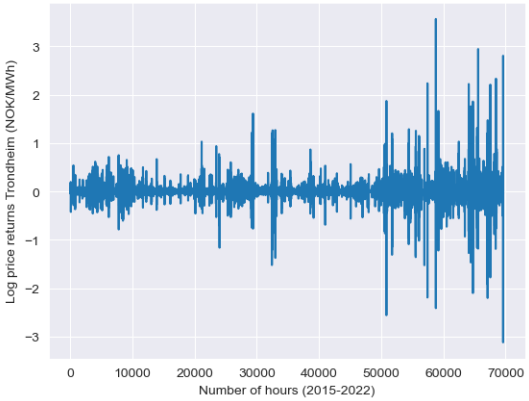


(b) Oslo log spot price returns (2015-2022)

Figure 1.3: Day-Ahead spot prices and returns Oslo N01 (2015-2022)

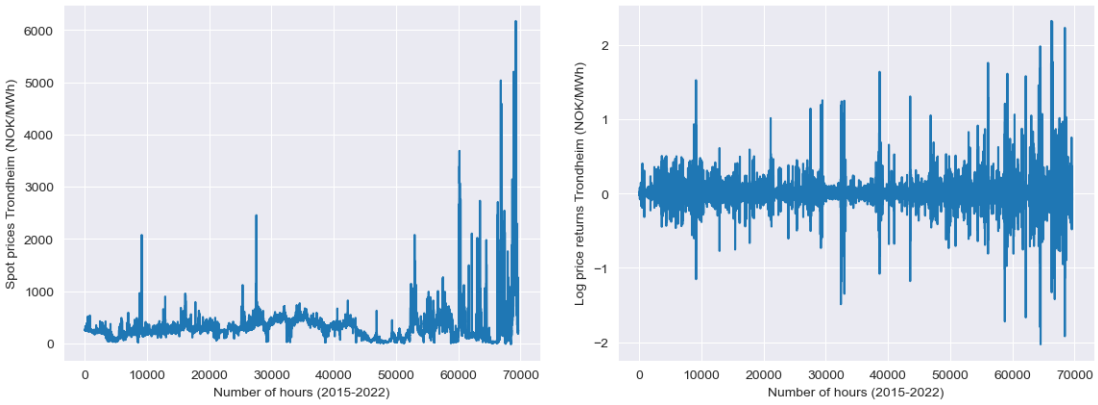


(a) Kristiansand spot prices (2015-2022)



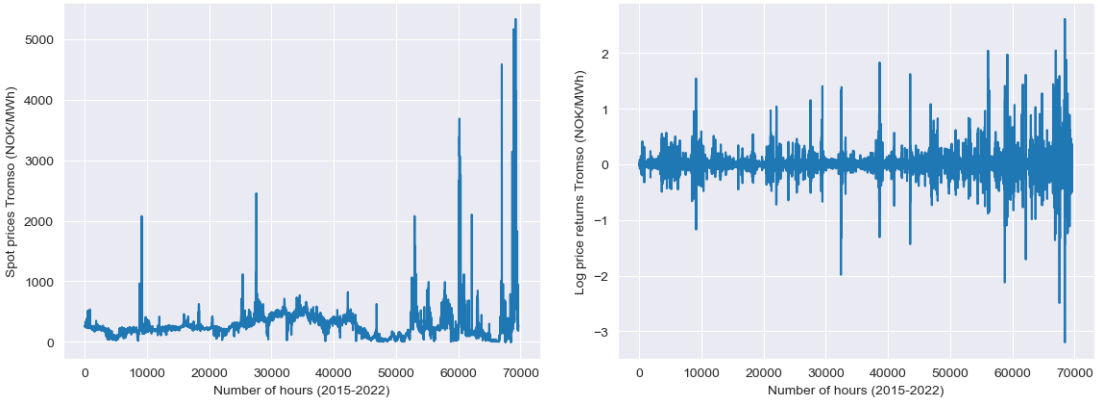
(b) Kristiansand log spot price returns (2015-2022)

Figure 1.4: Day-Ahead spot prices and returns Kristiansand N02 (2015-2022)



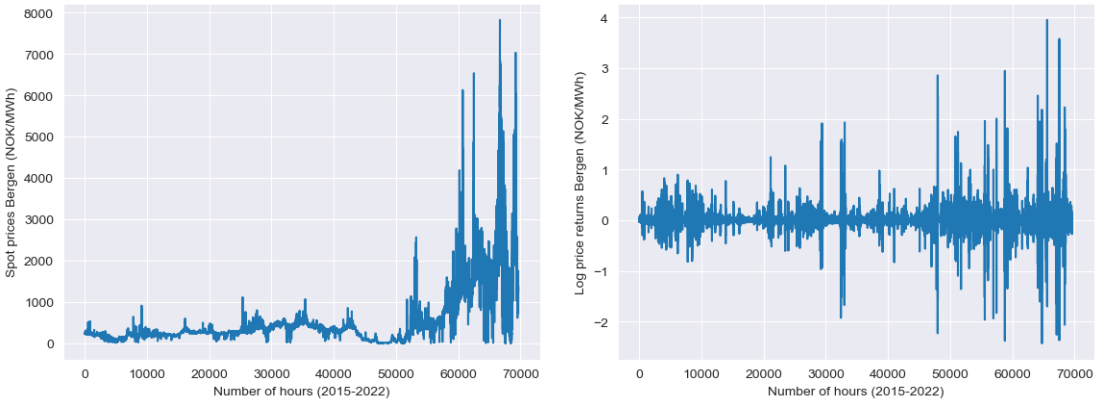
(a) Trondheim spot prices (2015-2022) (b) Trondheim log spot price returns (2015-2022)

Figure 1.5: Day-Ahead spot prices and returns Trondheim N03 (2015-2022)



(a) Tromso spot prices (2015-2022) (b) Tromso log spot price returns (2015-2022)

Figure 1.6: Day-Ahead spot prices and returns Tromso N04 (2015-2022)



(a) Bergen spot prices (2015-2022) (b) Bergen log spot price returns (2015-2022)

Figure 1.7: Day-Ahead spot prices and returns Bergen N05 (2015-2022)

The first element to understand about these graphs is that there are two distinct periods. The first runs from 2015 to the beginning of 2021 (around the 50,000th hour on plots). In fact, this period is marked by a certain periodicity, with classic behaviour including a seasonal trend and

extreme regimes (peaks) with a quick return to the mean. But since the end of 2020, the pattern has changed radically, with a drastic rise in prices and increased volatility. Price peaks are reaching levels almost 20 to 30 times higher than in the previous period, and the returns are much more pronounced. There are several explanations for these phenomena. The first one is the global pandemic caused by Covid-19, which has led to market disruptions and an economic recovery from 2021, boosting demand. In addition to this strong demand, soaring gas and oil prices are having a direct causal effect on electricity prices on the European market. There is also a purely political factor behind these prices. The European Union has embarked upon an energy transition based on alternative and intermittent means of production (solar, wind, to name a few) deemed to be "greener" for the environment. This decision has lead to some intermittency in the production systems and, consequently, to instantaneous falls in the supply curve, which necessarily implies an increase in prices. With interconnections between countries and zones within a country, prices can vary because of transmission capacity from one zone to another. For northern countries, prices depend on the ability to transmit power from one country to another. This, of course, applies to Norway, which exports a large proportion of its electricity, so its price is defined by the market and interconnection capacities.

If we focus on the first part of the graph until the end of 2019 (500, 000 hours approximately), all graphs globally show the same features, spikes that occurs at the same time on each zone which are really sharp and also return steadily to the mean. Moreover, a yearly periodicity is visible across the spot price graph with more volatility on winter periods than summer. Zooming on return plots, there is a significant correlation between all zones since returns have approximately the same amplitude at the same time. This is due to interconnections between Norway's areas that ensure a standardization of the market. NordPool considers that if there is sufficient interconnection capacity between the Scandinavian countries, then the price is defined for the region and then, depending on the constraints of each locality, the price may vary according market clearing process.

A final remark is that some zones are likely to share same features. Indeed, Oslo, Kristiansand and Bergen have the same behavior for electricity prices whereas Tromso and Trondheim have globally the same profile for returns and spot prices. Many factors like geography, types of production, interconnection with other countries play an important role to define the profile of curves. This point is discussed through next sections.

The following two tables table 1.3 and table 1.4 show a summary of Day-Ahead Spot prices and returns. The volatility value for spot prices is quite surprising. This can be explained by the recent market movements from 2020 onwards, which have seen hour-on-hour price variations of unprecedented amplitude.

	Oslo	Kristiansand	Trondheim	Tromso	Bergen
count	69648	69648	69648	69648	69648
mean	536.914	559.203	310.479	269.267	533.50
std	26.63	28.04	17.26	14.7	26.60
min	-19.72	-19.72	-21.370	-10.580	-0.960
max	7820.33	8224.63	6175.74	5335.090	7820.33

Table 1.3: Day-Ahead spot prices (NOK/MWh) of the five bidding zones (2015-2022)

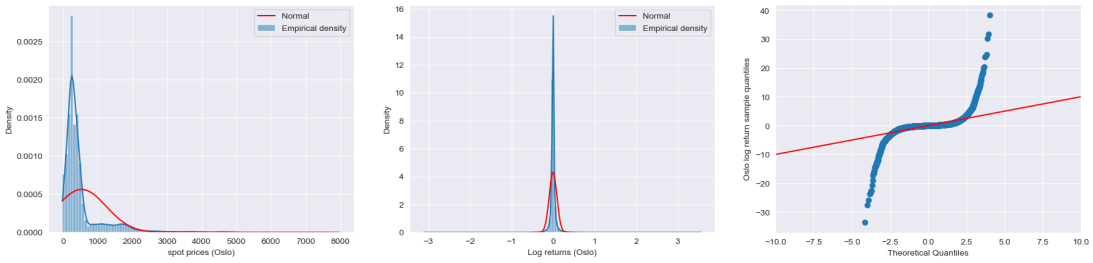
	Oslo	Kristiansand	Trondheim	Tromso	Bergen
count	69647	69647	69647	69647	69647
mean	2.2e-05	2.2e-05	-3.805e-07	-3.964e-07	2.3e-05
std	9.29e-02	9.20e-02	8.914e-02	8.758e-02	9.92e-02
min	-3.116	-3.116	-2.028	-3.197	-3.322
max	3.564	3.56	2.321	2.614	3.865

Table 1.4: Log returns of the five bidding zones (2015-2022)

The two types of zones are even more apparent when you look at the basic statistical data. Indeed, Oslo, Kristiansand and Bergen share same orders of magnitude whereas Trondheim and Tromso look similar. It is interesting to see why these zones behave in a same way. This basic data already represents a break with the traditional models used in the equity market. Indeed, it is remarkable that the price of electricity can be negative, whereas it is impossible to see a negative

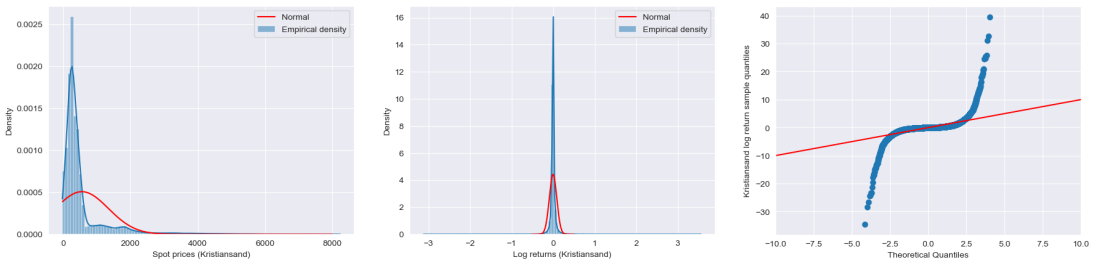
share price. This is because electricity is a commodity and its behaviour is also governed by physical laws that can lead to extreme price behaviour. Another observation concerns the difference between the average price and the maximum price, which in some cases is more than twenty times the average price. Here again, electricity is the only market where such a wide range of prices can be seen, and this is due to the difficulty of storing energy to smooth out extreme situations and maintain the smooth operation of the network, as explained in the section on 1.1.

After these initial descriptions, we now go into more detail about the structure of the distribution of prices and returns. To do this, histograms give us an initial impression of the law that can be associated with these distributions, whether the tail of the distribution is thicker or thinner, and whether the data are distributed symmetrically or with an imbalance. Another interest of these histograms is to see whether the Gaussian hypotheses can be applied to the Norwegian electricity market. The density curve of a Normal distribution calibrated to the data is also shown. The normality assumption is very important and essential in financial markets. Many simplifications flow from this assumption and it often provides explicit formulas for many problems. However, it is unlikely that we are in this situation, given the observations made on the graphs of returns.



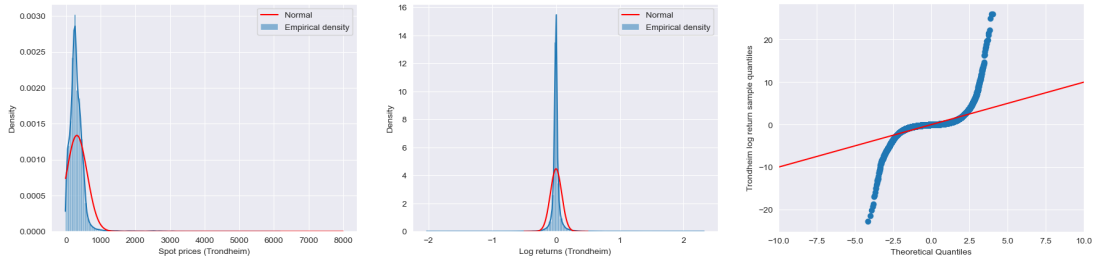
(a) Oslo spot price histogram (b) Oslo log spot price histogram (c) Q-Q plot of Oslo log spot price returns

Figure 1.8: Distribution of Day-Ahead prices and log returns Oslo N01 (2015-2022)



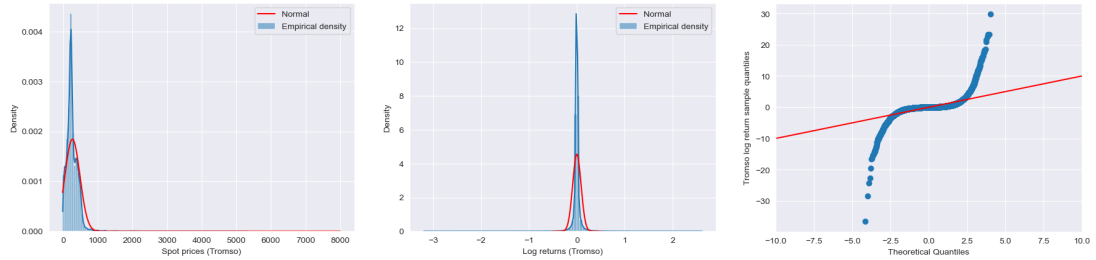
(a) Kristiansand spot price histogram (b) Kristiansand log spot price histogram (c) Q-Q plot of Kristiansand log spot price returns

Figure 1.9: Distribution of Day-Ahead prices and log returns Kristiansand N02 (2015-2022)



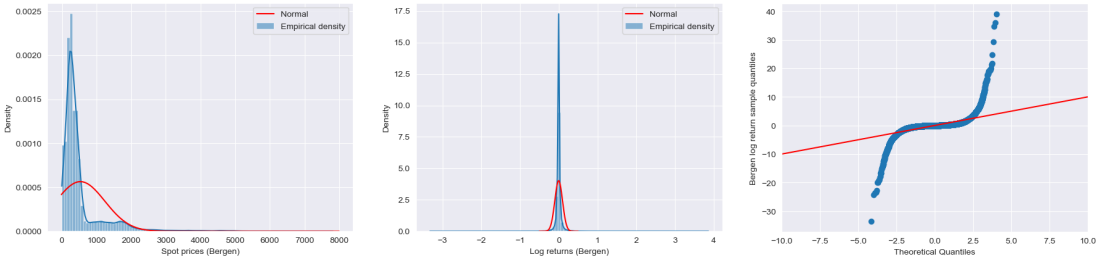
(a) Trondheim spot price histogram (b) Trondheim log spot price histogram (c) Q-Q plot of Tromso log spot price returns

Figure 1.10: Distribution of Day-Ahead prices and log returns Trondheim N03 (2015-2022)



(a) Tromso spot price histogram (b) Tromso log spot price histogram (c) Q-Q plot of Tromso log spot price returns

Figure 1.11: Distribution of Day-Ahead prices and log returns Tromso N04 (2015-2022)



(a) Bergen spot price histogram (b) Bergen log spot price histogram (c) Q-Q plot of Bergen log spot price returns

Figure 1.12: Distribution of Day-Ahead prices and log returns Bergen N05 (2015-2022)

In financial and in in other types of markets, the assumption of normal distribution is often used, although it is often rejected by the data. This description is well explained in [15] research paper which, using a multivariate approach to the risks associated with mutually generated portfolios, highlights the importance and credibility of the normal distribution hypothesis. For the remainder of the analysis, we can define the study hypothesis H0 as follows:

Definition 1.3.1 (H0)**.** H0: "The time series is normally distributed"

Two most common features used for comparing distribution to the normal one are called skewness and kurtosis. Skewness is a statistical measure that describes the asymmetry of the probability distribution of a real-valued random variable and Kurtosis is a statistical measure that quantifies the "tailedness" of the probability distribution of a real-valued random variable. Note that for normal distribution, the model should be 0 for skewness and 3 for kurtosis. As explained in the table 1.5 below, the results for these two characteristics are calculated for the five zones and for both Day-Ahead spot prices and log returns from 2015 to 2022.

The excess kurtosis is too high to consider that the returns follow a normal distribution. This first approach therefore concludes that we should abandon the idea of modelling returns using a

	Oslo	Kristiansand	Trondheim	Tromso	Bergen
skewness	1.75	1.89	1.08	1.08	2.67
kurtosis	190	214	94	133	215
H0 :	Rejected	Rejected	Rejected	Rejected	Rejected

Table 1.5: Skewness and Kurtosis of log returns (2015-2022)

normal distribution, and that a modelling approach such as the Black-Scholes model would not be a very accurate approximation for Day-Ahead prices. To be a little more precise and to definitively confirm the non-normality of the returns, two statistical tests inspired by [16, Chapter 3] are performed, the Jarque-Bera test and Shapiro-Wilk test.

Definition 1.3.2. The Jarque-Bera test statistic is calculated from:

$$JB = \frac{n}{6}(s^2 + \frac{(k-3)^2}{4}) \tag{1.3.2}$$

where n is the same size, s is the sample skewness and k is the sample kurtosis. The null hypothesis is rejected if the test statistic exceeds some predefined critical value, which is taken in the asymptotic limit from the χ^2_2.

Definition 1.3.3. The Shapiro-Wilk test statistic is defined as:

$$W = \frac{\sum_{i=1}^{n} a_i x_i)^2}{(\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{1.3.3}$$

where x_i are the order statistics from the empirical sample, $\bar{x}$ is the mean and a_i are appropriate constant values attained from the means and covariance matrix of the order statistics. Again, the null hypothesis is rejected if the test statistic exceeds some predefined critical value.

The results of these two tests are shown in the table 1.6 below:

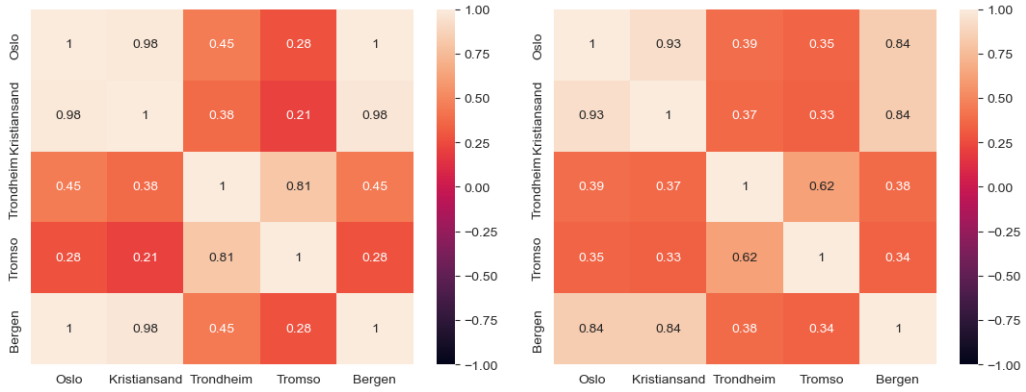
areas	Jarque-Bera test		Shapiro-Wilk test	
Log Returns	stat	p-value	stat	p-value
Oslo	105327420	0.0	0.478	0.0
Kristiansand	133218022	0.0	0.473	0.0
Trondheim	25664905	0.0	0.576	0.0
Tromso	51458630	0.0	0.488	0.0
Bergen	134487777	0.0	0.455	0.0
H0 :	Rejected		Rejected	

Table 1.6: Log returns normality tests

The results are indisputable. All the p-values are 0, which means that we have to reject the normality hypothesis 1.3.1. This confirms the hypothesis put forward by the graphical analysis that it is inappropriate to use a normal model for the distribution of returns.

Finally, Heating maps for correlation make no doubt about the dependency between zones. High correlation come out for Oslo, Kristansand and Bergen according to 1.13(a) (superior to 0.95). For returns it's practically the same 1.3.2. On the same hand, Tromso and Trondheim have a high correlation for spot prices and returns.

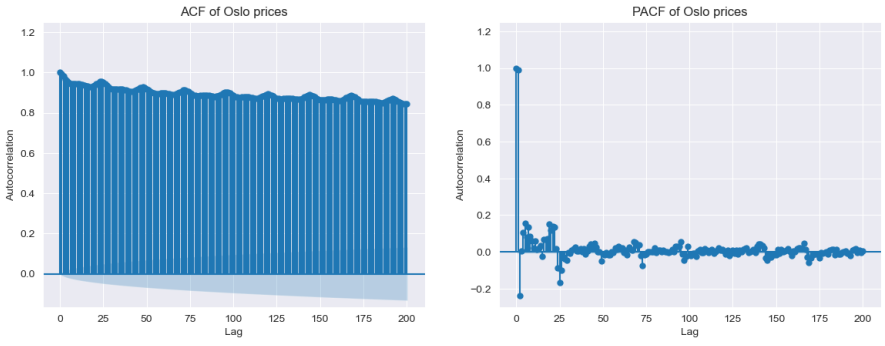
A second point of interest when studying a time series is the auto-correlation between observations from a same time series. To study this phenomenon, Autocorrelation functions (ACF) and Partial autocorrelation functions (PACF) are two features that underline much information about the structure of our time series. We can expect same remarks from the precedent part such as similar shape of curve for Oslo, Kristiansand and Bergen zones and on the other hand between Trondheim and Tromso and also natural correlation corresponding to the structure of a day and week. Indeed, electricity spot prices as explained before are mainly described with intern seasonality in days, weeks and years. Therefore, it is more likely to observe spikes of ACF at lag 24, 48, 72. Several points seem to emerge from these graphs. The first is, as expected, a very high degree of similarity in the graphs between Oslo, Kristiansand and Bergen ACF and PACF for spot prices



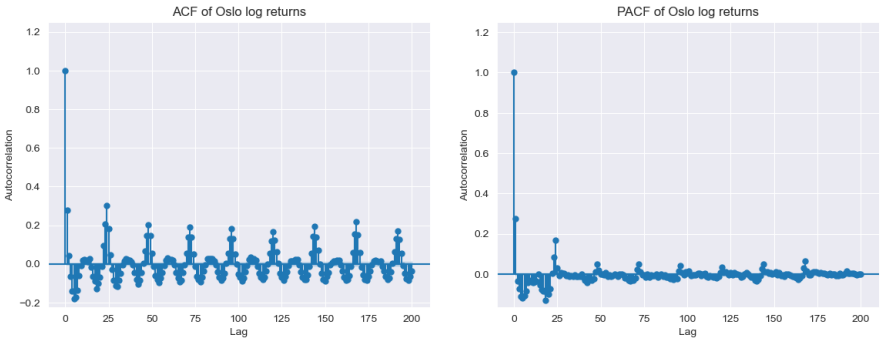
(a) Spot price correlations between bidding zones (b) Log return correlations between bidding zones

Figure 1.13: Heating map of spot prices and returns (2015-2022)

and returns 1.14, A.1, A.3 and also between Trondheim and Tromso A.2 and 1.15. Moreover, we observe in each ACF plots two main features. The first one is a smooth decrease when the lag increases and the second one is small peaks at lag 24,48,and other multiples of 24. It should be noted, however, that the decline is much more marked in the Tromso and Trondheim regions, as the graphs 1.15(a) and A.2(a) show. PACF curves for each zones confirm these observations, with a slow decay and increasingly small peaks in terms of amplitude, reflecting a reduction in the correlation of the series as the lag increases.

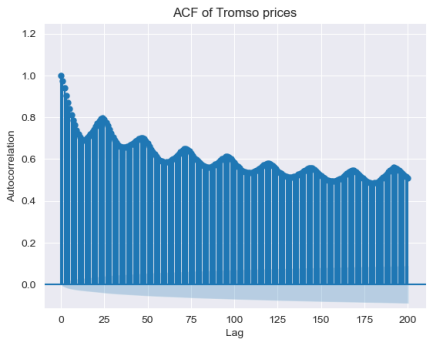


(a) ACF Oslo prices NO1 (2015-2022) (b) PACF Oslo prices NO1 (2015-2022)

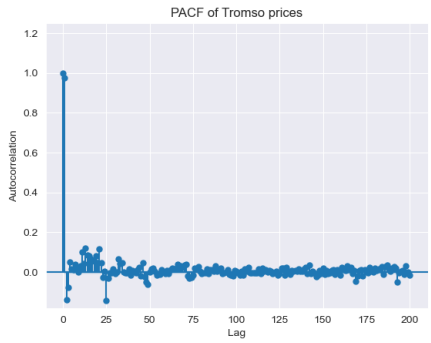


(c) ACF Oslo log returns NO1 (2015-2022) (d) PACF Oslo log returns NO1 (2015-2022)

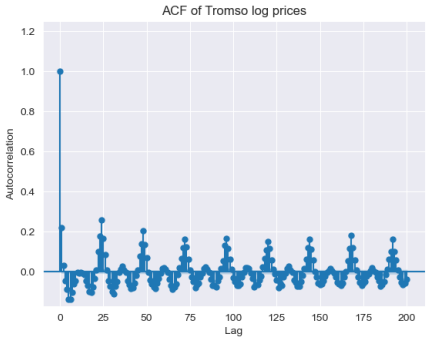
Figure 1.14: ACF and PACF Day-Ahead prices and returns Oslo NO1 (2015-2022)



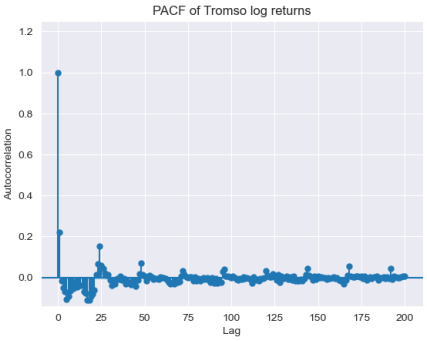
(a) ACF Tromso prices NO4 (2015-2022)



(b) PACF Tromso prices NO4 (2015-2022)



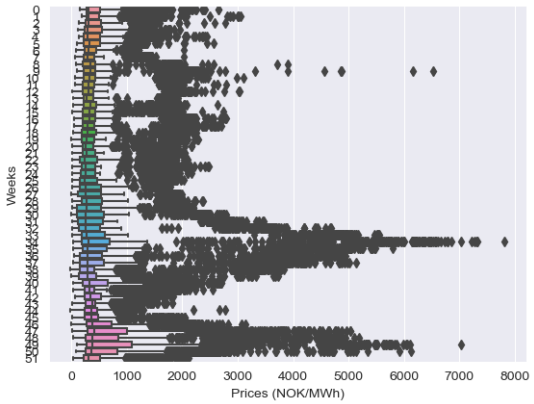
(c) ACF Tromso log returns NO4 (2015-2022)



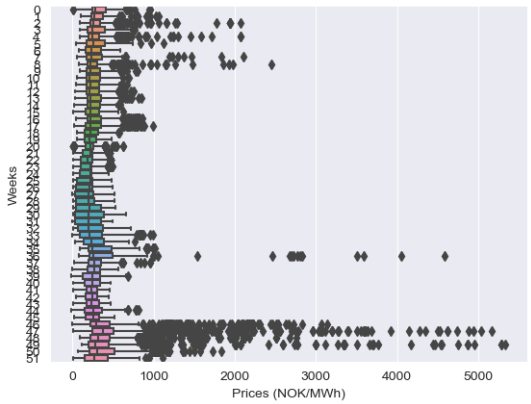
(d) PACF Tromso log returns NO4 (2015-2022)

Figure 1.15: ACF and PACF Day-Ahead prices and returns Tromso NO4 (2015-2022)

As the structure of our time series is globally analysed, we can now focus on a distinction between months in each year from 2017 to 2022. Indeed, electricity price is under physical constraint and obeys to many dynamics driven by the market but also by natural elements such as temperature, wind, melting of snow,etc. Therefore, seasons and more precisely months and weeks in each year reflect a change of those constraints on electricity prices. To expand on this point and to go deeper into our analysis, Box plots of each areas (NO1,NO2,NO3,NO4 and NO5) are presented to give to the reader a viewpoint about sparsity of Day-Ahead spot prices each week.



(a) Box plot Oslo NO1 (2017-2022)



(b) Box plot Tromso NO4 (2017-2022)

Figure 1.16: Box plots NO1 and NO4 (2017-2022)

A remark would be to notice is that the behavior of N01,NO2 and NO3 are roughly the same with somehow a high volatility with many values "outside the box" according to figures 1.16(a), A.4 and A.6. As explained in this section, this phenomenon is due to the fact that those areas have

to import electricity from other zones to fulfil the demand and are more exposed to fluctuations of market due to transmission costs and power reserve management. On the other hand NO3,NO4 are less volatile as show Figures A.5 and 1.16(b) which is explained by the fact that hydro-capacities of production in this two zones are superior to the demand. So less electricity has to be imported and hydro plants are used to generate more electricity power than the local demand to improve profit by exporting overproduction and when prices are high.

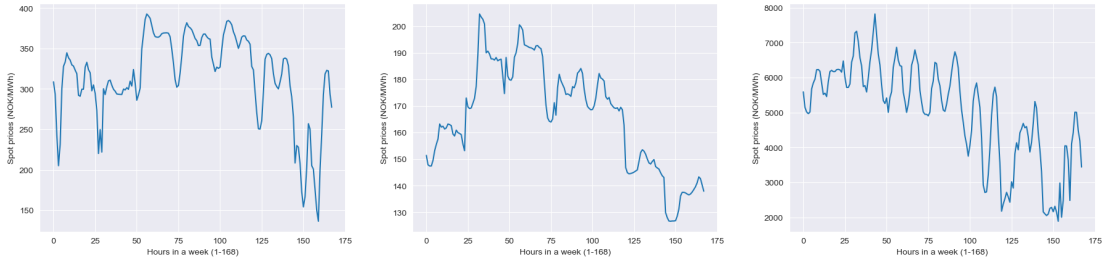
Chapter 2

Stochastic dynamic model

In this part, we attempt to use an approach based on stochastic calculus. The idea is to derive directly the spot price based on historical data instead of modelling the market. Indeed, after closer examination, especially on previous history of electric spot price, it seems that one can model the dynamic of the price with some stochastic dynamics. To be able to better fit the curve price, some features have to be transcribed in our model. More precisely, three components of the dynamic must be incorporated: the seasonality over days, weeks and years, sharp spikes due to inelastic demand and tendency over weeks. Those spikes lead to another interesting feature, which is the mean-reversion of the price. After substantial increases in price, one can observe that the price goes back quickly to the mean.

2.1 Observations on data

As explained in the previous section, some features of our time series have to be quantified. When we are interested in modelling, the first step is to observe the dynamics that we are trying to quantify. By way of example, the weeks 18 (2018), 36 (2020) and 35 (2022) are shown in the three figures 2.1.



(a) NO1(OSLO)-2018-Week18 (b) NO1(OSLO)-2020-Week36 (c) NO1(OSLO)-2022-Week35

Figure 2.1: Weekly spot price change over hours NO1

A number of factors stand out in these graphs 2.1. The first is that the general dynamics of each week are broadly the same. In fact, there is a temporal periodicity of daily frequency throughout the week, but also, more subtly, a periodicity within the day. What's more, each week seems to be driven by an increasing or decreasing weekly trend. This trend must be taken into account, as it often leads to problems of non-stationary in the time series. This point is highlighted in many electricity markets in Europe, particularly in the Iberian market [17].

2.2 Weekly trend

Thanks to the first observations made in part 2.1, this section focuses on estimating a trend in our data. Since mainly the five bidding zones have a behavior of two distinct dynamics, we only compute studies on NO1 and NO4 areas which embody the best these two behaviors.

2.2.1 Linear Trend estimation

Identifying linear trend in our hourly data in each week have two main goals. The first one is to differentiate each week by the coefficient of the trend and secondly by subtracting this trend to the original data, the time series of Day-Ahead prices is more stationary. Last remark is useful for next sections focusing on periodicity and fitting the residuals.

To compute trends, linear regression is performed through each week. Recall that this estimator is given by the following definition:

Definition 2.2.1. We assume a dependence of the form

$$y_i = f(x_i) + \epsilon_i, \text{ for i = 1,...,n} \tag{2.2.1}$$

where f(x) $\equiv \alpha + \beta x$ is a linear function, and the sequence $(\epsilon_i)_{i=1,\dots,n}$ are centered independent random noises with constant variance σ^2.

The least-square estimator $(\hat{\alpha}, \hat{\beta})$ is the solution to the minimisation problem

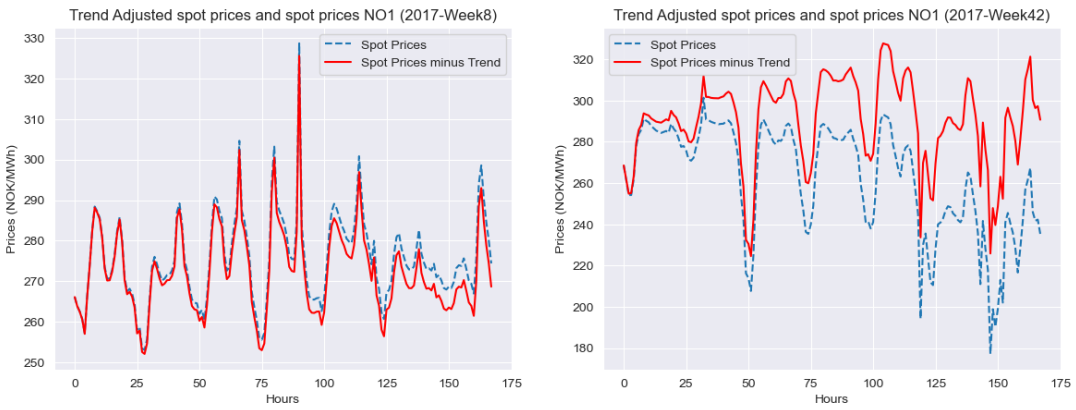
$$(\hat{\alpha}, \hat{\beta}) := \arg\min_{((\alpha,\beta)} \mathcal{L}(\alpha, \beta) = \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2 \tag{2.2.2}$$

Theorem 2.2.1. The solution to (2.2.2) reads

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \text{ and } \hat{\beta} = \frac{\sum_{i=1}^{n} x_i(y_i - \bar{y})}{\sum_{i=1}^{n} x_i(x_i - \bar{x})} \tag{2.2.3}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$.

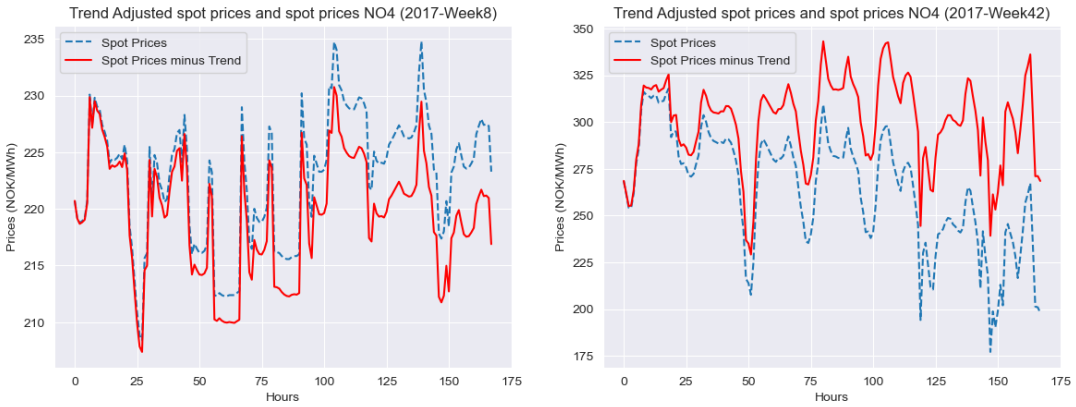
Since linear trend coefficient is compute, we forecast the trend over each week. New time series are computed by subtracting linear trends to original Day-Ahead spot prices. Results for weeks 8 and 42 in 2O17 for Oslo (NO1) and Tromso (NO4) are displayed in Figures 2.2 and 2.3. In our study, the x_i are the hours of the week and the y_i are the Day-Ahead prices associated with these hours. By convention, the first hour is priced at index 0 and is incremented by one until the end of the week. Thus, each week 168 observations are considered.



(a) Weekly trend adjusted Day-Ahead spot prices NO1 (2017-Weeks8)

(b) Weekly trend adjusted Day-Ahead spot prices NO1 (2017-Weeks42)

Figure 2.2: Weekly trend adjusted Day-Ahead spot prices NO1 (2017-Weeks8-42)



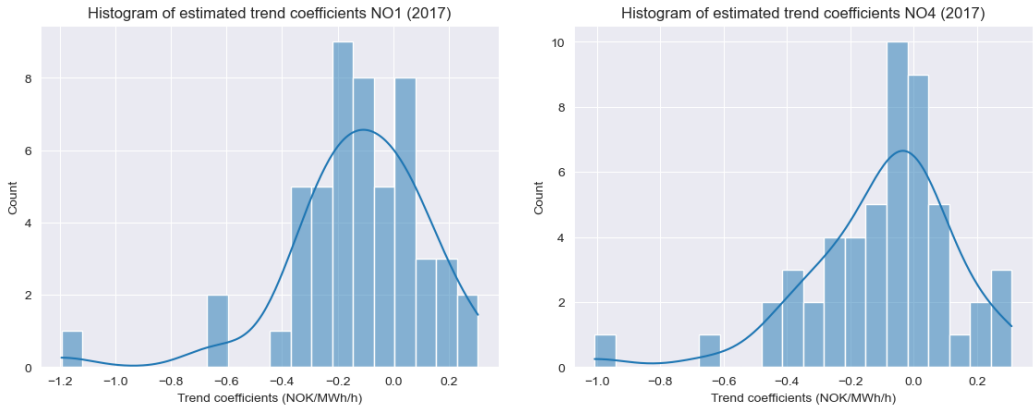
(a) Weekly trend adjusted Day-Ahead spot prices NO4 (2017-Weeks8)

(b) Weekly trend adjusted Day-Ahead spot prices NO4 (2017-Weeks42)

Figure 2.3: Weekly trend adjusted Day-Ahead spot prices NO4 (2017-Weeks8-42)

Remark 2.2.2. Hypothesis of stationary is likely to be respected after linear trend adjustment as Figures 2.2 and 2.3 shows in red. Indeed, a drift that the reader can observe in blue curves seems to disappear in red curve. More careful and quantitative facts are given in the next section 2.2.2 focusing on stationary.

To finally give a global view of the behavior of trend coefficients in 2017, the 52 linear trend coefficients of each zones are reassembled together in histograms 2.4.



(a) Histogram of estimated trend coefficients NO1 (2017)

(b) Histogram of estimated trend coefficients NO4 (2017))

Figure 2.4: Histogram of estimated trend coefficients NO1-NO4 (2017)

The distribution of coefficients is not the same between the different zones. In fact, the histogram 2.4(a) shows a higher concentration of trend coefficients distributed between [-0.4,0.2] than on the histogram 2.4(b). Recall that Oslo NO1 zone is an import zone for electricity whereas Tromso NO4 zone exports electricity mostly every day. This difference is quite remarkable, and reflects the dynamics of prices in relation to power generation capacity in the zone.

2.2.2 Test of stationarity

Stationarity is a key factor to describe a time series.

Definition 2.2.3. (Stationary) A time series $(X_t)_{t\in\mathcal{Z}}$ is said to be stationary if

$$(X_{t_1}, ..., X_{t_n}) \overset{d}{=} (X_{t_{1+k}}, ..., X_{t_{n+k}}) \tag{2.2.4}$$

where $\overset{d}{=}$ is equality between joint distributions.

The Augmented Dicky Fuller test (ADF test) is a way for checking stationary in time series. This test is classified as a unit root test that determines how a time series is defined by a trend. Null and alternate Hypothesis are defined by the following definition 2.2.4 and 2.2.5 :

Definition 2.2.4. Null Hypothesis (H0): Time series has a unit root, meaning it is non-stationary. It has some time dependent structure.

Definition 2.2.5. Alternate Hypothesis (H1): Time series does not have a unit root, meaning it is stationary. It does not have time-dependent structure.

The satistics use to reject or not the null hypothesis is p-value. A p-value below a threshold (5% or 1%) reject the null hypothesis (stationary), otherwise the test fails to reject the null hypothesis.
The results of the ADF test are shown in the tables 2.1 and 2.2 below:

	Statistic	p-value	CV 1%	CV 5%	H0 1%	H0 5%
NO1 Week8	-5.823	0	-3.471	-2.879	Rejected	Rejected
NO1 Week8 (LA)	-5.862	0	-3.471	-2.879	Rejected	Rejected
NO1 Week42	-3.670	0.004	-3.471	-2.879	Rejected	Rejected
NO1 Week42 (LA)	-5.298	0.000	-3.471	-2.879	Rejected	Rejected

Table 2.1: ADF Test results NO1 (2017-Week8-42)

	Statistic	p-value	CV 1%	CV 5%	H0 1%	H0 5%
NO4 Week8	-3.781	0.003	-3.471	-2.879	Rejected	Rejected
NO4 Week8 (LA)	-4.008	0.001	-3.471	-2.879	Rejected	Rejected
NO4 Week42	-3.273	0.016	-3.471	-2.879	Non-Rejected	Rejected
NO4 Week42 (LA)	-5.473	0.002	-3.471	-2.879	Rejected	Rejected

Table 2.2: ADF Test results NO4 (2017-Week8-42)

Remark 2.2.6. (LA) stands for Linear Adjustment. This corresponds to the new time series obtained once the linear trend has been removed.

Once the linear trend has been removed from the Day-Ahead price time series, the HO hypothesis is more likely to be rejected. Indeed, ADF statistics and p-values are lower after the linear adjustment, that reject more likely the hypothesis of tendency in our time series.

In this first stage, the new time series has gained in stationarity, which makes it possible to eliminate a spurious trend for the study of periodicity and the calibration of a stochastic model on the residuals.

2.3 Periodic deterministic function

One of the main feature of electric market is the periodicity in prices. Indeed, daily, weekly and yearly seasonality play a key role on observed prices. When one think of modeling electricity prices, it's an important part to integrate into the model. The main question is how to fit seasonality with price.

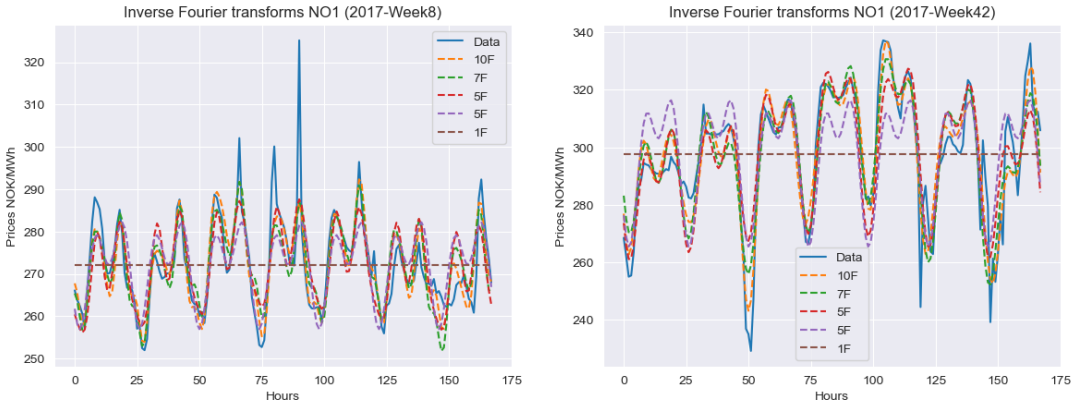
After a closer look in the previous section, it seems that some periodicity existed in the market. Indeed, after an overview on graphs and deeper on the data, we remark that a daily and an in-daily shape composed each week. In order to reconstruct and forecast the weekly electricity spot price, a mathematical tool, Fourier series, is really useful to this goal. A Fourier series is an expansion of a periodic function f(x) in terms of an infinite sum of sines and cosines. The study of Fourier series is known as harmonic analysis and is useful as a way to reconstruct a periodic function into a set of simple terms called harmonics that can be plugged in, and then recombined to an approximation to it to whatever accuracy is desired [18, Chapter 3].

Concerning our times series for weekly prices, the idea is to decompose weekly data. We apply Fourier analysis to the hourly-resolved profiles of electricity price over a weekly time horizon. Fourier analysis allows to represent or approximate any generic function through a Fourier series, i.e., the sum of trigonometric functions with different frequencies and amplitudes. In the case of the electricity price profile, S_t, the Fourier series representation is given by :

$$S_t = a_0 + \frac{2}{N}\sum_{n=1}^{N/2} a_n cos(\frac{2n\pi t}{N}) + b_n sin(\frac{2n\pi t}{N}) \tag{2.3.1}$$

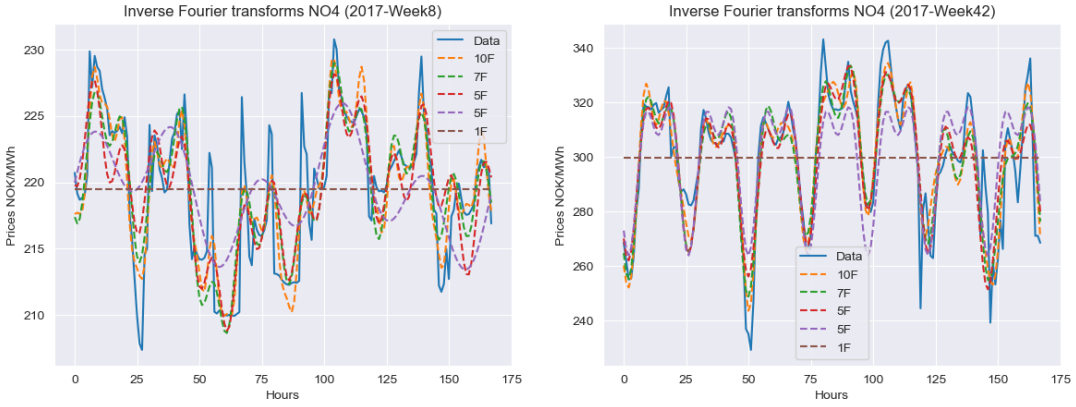
where a_0 corresponds to the mean value of S_t, a_n and b_n are the Fourier coefficients defining the shape of the periodic functions at frequency n, and N is the length of the original time series.

Fourier Transform is applied to the time series that has become stationary by subtracting the linear trend in the previous part 2.2. The inverse transform is then calculated to recover the original series, but filtering the frequencies by order of importance. To assess the impact of keeping more frequencies or not, RMSE is performed to quantify the accuracy of our model.



(a) Inverse Fourier transforms NO1 (2017-Week8) (b) Inverse Fourier transforms NO1 (2017-Week42)

Figure 2.5: Weekly trend adjusted Day-Ahead spot prices NO1 (2017-Weeks8-42)



(a) Inverse Fourier transforms NO4 (2017-Week8) (b) Inverse Fourier transforms NO4 (2017-Week42)

Figure 2.6: Weekly trend adjusted Day-Ahead spot prices NO4 (2017-Weeks8-42)

	1F	3F	5F	7F	10F
Oslo NO1 Week8	10.64	7.37	6.53	5.91	5.20
Tromso N04 Week8	5.30	4.20	3.16	2.77	2.36
Oslo NO1 Week42	21.54	13.93	11.40	9.75	8.14
Tromso NO4 Week42	23.09	14.03	11.14	10.10	8.88

Table 2.3: RMSE of Inverse Fourier transforms NO1-NO4 (2017-Week8-42)

An improvement is observable when referring to the table 2.3 when the number of frequencies retained for our model increases. Keeping only the continuous component and the two most

important significant frequencies clearly improves the RMSE. For the sake of simplicity and to avoid overfitting, only the constant component and the first two most important frequencies are kept for the rest of this section. The periods selected are 7, 14 hours and the 0 for the constant component.

Once the inverse Fourier transform has been performed, the periodic component is subtracted from the stationary time series, leaving only the stochastic residual component.

2.4 ARMA Process

According to [19, Part 3.2] ARIMA process is one of the principal choices of modelling electricity. Autoregressive Integrated Moving Average models was developed by Box and Jenkins in 1976 [20]. Main reason is the mean-reverting behavior of electricity price. This model requires the time series to be stationary for meaningful results to be obtained. Another advantage of this process is the convergence of this process for long-term forecasts of the series to the unconditional mean of the series. Since a previous work was done to make time series stationary, the model uses in this section in only an ARMA process thanks to the non-necessity of differentiating the series to make it stationary.

To define a mathematical context for this process, several definitions are recalled in order to justify the use of the ARMA model.

Definition 2.4.1. (Covariance Stationary Process) A process $(X_t)_{t\in\mathcal{Z}}$ is covariance-stationary if

- $(X_t)_{t\in\mathcal{Z}}$ is square-integrable.
- For any s,t ∈ Z and k ∈ Z, the mean function μ and autocovariance function γ of $(X_t)_{t\in Z}$ satisfy $\mu(t) = \mu(t+k), \gamma(s,t) = \gamma(s+k, t+k)$.

Definition 2.4.2. (White Noise) A covariance stationary process $(X_t)_{t\in Z}$ is called a white noise if $\rho(h) = \begin{cases} 1 & h = 0, \\ 0 & h > 0 \end{cases}$ If $(X_t)_{t\in Z}$ is a white noise with $\mu(t) = 0$ and $\gamma(0) = \sigma^2$, then we write $(X_t)_{t\in Z}$ $\sim$ WN(0,σ^2).

Definition 2.4.3. (ARMA Process) The process $(X_t)_{t\in Z}$ is a zero-mean ARMA(p,q) process, where p,q = 0,1,..., if it is covariance-stationary and satisfies

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} \ \forall t \in \mathbb{Z}, \tag{2.4.1}$$

where $\phi_1, ..., \phi_p$ and $\theta_1, ..., \theta_q$ are the parameters of the process and $(\epsilon_t)_{t\in Z} \sim$ WN(0,σ^2).

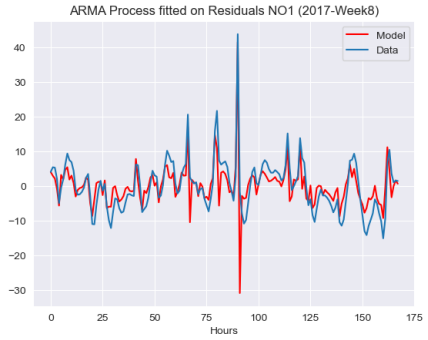
The order of our ARMA process is chosen on the basis of the ACF and PACF figures defined in section 1.3.2. A good approximation for both zones (NO1 and NO4) is p=1 and q=1. The ARMA process is calibrated using the ARIMA module in the "statsmodels" library in the Python language. The results of both the model statistics and the graphs for the NO1 and NO4 zones are presented to give the reader an idea of the coherence of the choice of this model. Remember that the model is defined on day-Ahead price data by subtracting the linear trend and the periodic function.

SARIMAX Results

Dep. Variable:	y	No. Observations:	
Model:	ARIMA(1, 0, 1)	Log Likelihood	
Date:	Sat, 26 Aug 2023	AIC	
Time:	10:05:10	BIC	
Sample:	0	HQIC	
	- 168		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025
const	0.0349	1.530	0.023	0.982	-2.963
ar.L1	0.5002	0.167	2.990	0.003	0.172
ma.L1	0.1633	0.178	0.919	0.358	-0.185
sigma2	33.8701	1.603	21.135	0.000	30.729

Ljung-Box (L1) (Q): 0.01 Jarque-Bera (JB):



(a) ARMA process statisitcs NO1 (2017-Week8)

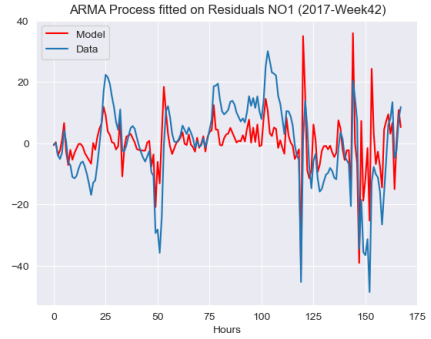
(b) ARMA process fitted on Residuals NO1 (2017-Week8)

Figure 2.7: ARMA process statisitcs and chart NO1 (2017-Weeks8)

SARIMAX Results

Dep. Variable:	y	No. Observations:	
Model:	ARIMA(3, 0, 1)	Log Likelihood	
Date:	Sat, 26 Aug 2023	AIC	
Time:	10:30:55	BIC	
Sample:	0	HQIC	
	- 168		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025
const	0.2194	3.937	0.056	0.956	-7.497
ar.L1	0.3605	0.696	0.518	0.604	-1.004
ar.L2	0.3694	0.503	0.734	0.463	-0.617
ar.L3	-0.0553	0.087	-0.637	0.524	-0.226
ma.L1	0.3940	0.687	0.573	0.566	-0.953
sigma2	75.5262	4.619	16.351	0.000	66.473



(a) ARMA process statisitcs NO1 (2017-Week42)

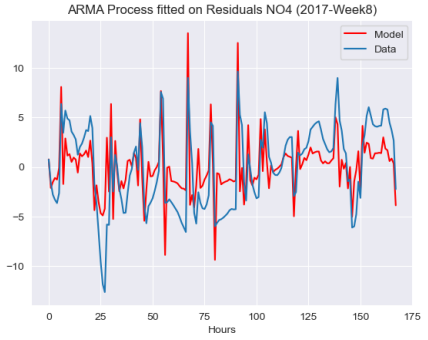
(b) ARMA process fitted on Residuals NO1 (2017-Week42)

Figure 2.8: ARMA process statisitcs and chart NO1 (2017-Week42)

SARIMAX Results

Dep. Variable:	y	No. Observations:	
Model:	ARIMA(3, 0, 1)	Log Likelihood	
Date:	Sat, 26 Aug 2023	AIC	
Time:	10:39:46	BIC	
Sample:	0	HQIC	
	- 168		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025
const	-0.0210	0.837	-0.025	0.980	-1.662
ar.L1	0.7104	0.986	0.720	0.471	-1.223
ar.L2	0.0360	0.758	0.047	0.962	-1.449
ar.L3	-0.0907	0.098	-0.925	0.355	-0.283
ma.L1	0.0524	1.003	0.052	0.958	-1.914
sigma2	8.3711	0.723	11.578	0.000	6.954



(a) ARMA process statisitcs NO4 (2017-Week8)

(b) ARMA process fitted on Residuals NO4 (2017-Week8)

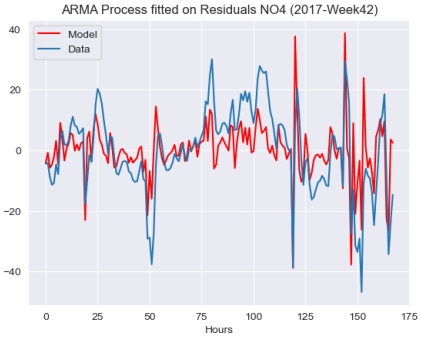
Figure 2.9: ARMA process statisitcs and chart NO4 (2017-Week8)

SARIMAX Results

Dep. Variable:	y	No. Observations:	
Model:	ARIMA(3, 0, 1)	Log Likelihood	
Date:	Sat, 26 Aug 2023	AIC	
Time:	10:45:02	BIC	
Sample:	0	HQIC	
	- 168		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025
const	-0.2418	3.174	-0.076	0.939	-6.463
ar.L1	0.3581	1.287	0.278	0.781	-2.164
ar.L2	0.3104	0.960	0.323	0.746	-1.571
ar.L3	-0.0547	0.064	-0.850	0.395	-0.181
ma.L1	0.4126	1.280	0.322	0.747	-2.096
sigma2	87.1349	5.791	15.046	0.000	75.785

(a) ARMA process statisitcs NO4 (2017-Week42)



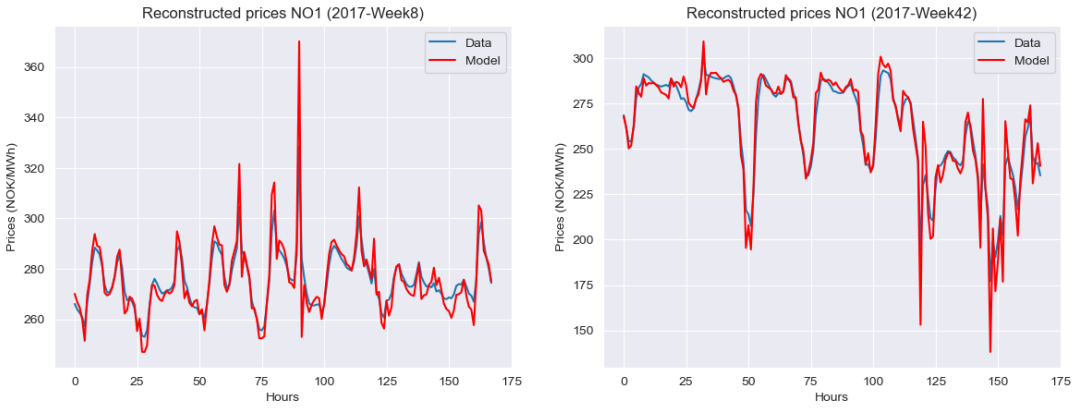
(b) ARMA process fitted on Residuals NO4 (2017-Week42)

Figure 2.10: ARMA process statisitcs and chart NO4 (2017-Week42)

ARMA models provide a powerful framework for modeling mean-reverting processes by incorporating both autoregressive and moving average components. The results of the processes calibrated on historical data reflect the overall behaviour of our residuals. Sudden price movements leading to mean reversion are well captured as Figures 2.11 and 2.12 show.

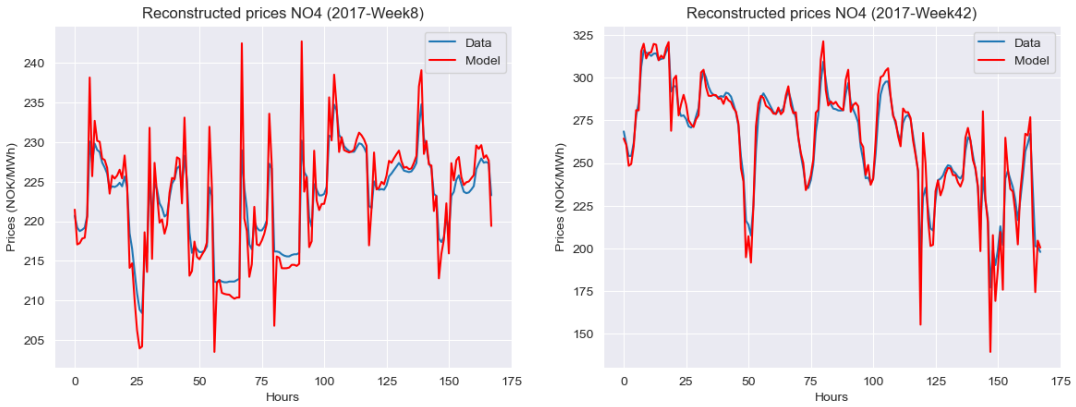
2.5 Reconstructed Day-Ahead Price

To conclude this section, the day-Ahead price for weeks 8 and 42 in the NO1 and NO4 zones is reconstructed from the results obtained in previous section. For each week, the weekly trend, the periodic function and the ARMA process are added to create the day-ahead price estimator. Results are shown graphically (Figures 2.11 and 2.12) and RMSE and MAE are computed.



(a) Reconstructed Day-Ahead Price NO1 (2017-Week8) (b) Reconstructed Day-Ahead Price NO1 (2017-Week42)

Figure 2.11: Reconstructed Day-Ahead Price NO1 (2017-Weeks8-42)



(a) Reconstructed Day-Ahead Price NO4 (2017-Week8) (b) Reconstructed Day-Ahead Price NO4 (2017-Week42)

Figure 2.12: Reconstructed Day-Ahead Price Day-Ahead spot prices NO4 (2017-Weeks8-42)

	NO1 Week8	NO1 Week42	NO4 Week8	NO4 Week42
RMSE	6.60	4.68	8.65	4.35
MAE	4.13	3.36	5.48	3.25

Table 2.4: RMSE and MAE reconstructed Day-Ahead price NO1-NO4 (2017-Week8-42)

The results presented in the table 2.5 are very encouraging for an approach to price with this breakdown into three parts. The main dynamics of the day-ahead price, i.e. periodicity, sudden price movements and the weekly trend, seem to be well integrated into this model. The rather low values of the RMSE and MAE over these two weeks of study reflect the effectiveness of this model for considering a future price projection with these parameters, which are outside the scope of this master's thesis.

Chapter 3

Market clearing model, a water value approach

We saw earlier that it is possible to approximate the Day-Ahead price. But is it possible to reconstruct the dynamic market processes that lead to the formation of these prices? This is the aim of this section that takes a different approach to that used previously, to approximate as closely as possible the supply curve based on hydraulic generation in order to deduce the resulting price.

The market clearing price is the linchpin of efficient market operation, shaping the equilibrium between supply and demand across various economic sectors, most notably in electricity markets. Within this context, the market clearing price serves as the pivotal point where the quantity of goods or services offered equals the quantity coveted by consumers, ensuring optimal resource allocation and fair value exchange. Market clearing process is based on supply and demand curve but for hydro dominated market where mainly all production is done by hydro generation plants, the supply curve is linked to the water stored in reservoirs which influences the marginal cost function. In this part, the Water Value concept attempts to throw new light on this particular point by providing a new interpretation of the marginal costs in hydro-dominated electricity production.

3.1 A global description

Market participants, ranging from electricity producers and consumers to traders, submit bids and offers detailing the quantity of electricity they are willing to supply or purchase at varying price levels for a specific period, often an hour. These bids and offers are subsequently sorted in ascending order, forming a supply curve representing the quantity of electricity available at each price point, while the corresponding demand curve embodies the quantity consumers seek at each price level. The point of intersection between these curves embodies the potential market clearing price [21, Part 3.3].

To pinpoint the precise market clearing price, the grid operator or market platform identifies the highest-priced bid whose quantity matches or exceeds the quantity sought at the intersection. This highest-priced bid becomes the marginal offer and establishes the market clearing price, applicable to all accepted trades. Bids and offers below the market clearing price are approved, with trades executed at this equilibrium rate. The market clearing price is not only an economic benchmark but also a dynamic indicator [22, Part3,4,5]. It mirrors the cost of production for the marginal producer - the participant whose electricity offering sets the market clearing price - highlighting the price at which the last unit of electricity supplied matches the last unit demanded.

The significance of the market clearing price reverberates throughout the energy landscape. Consumers make consumption choices as long as the price remains beneath their valuation. This delicate balance aids grid operators in sustaining stability by ensuring a harmonious match between electricity supply and demand. Furthermore, the market clearing price provides critical insights to market participants, governments, and regulators. Producers adjust their output and investment strategies based on the prevailing clearing price, while consumers adapt their energy consumption habits. Regulatory bodies, equipped with these pricing signals, can shape policies to enhance market competitiveness, incentivize renewable energy integration, and mitigate imbalances.

3.2 Supply Curve

The supply curve in a market illustrates the relationship between the quantity of a good or service that producers are willing to offer for sale and the corresponding prices at which they are willing to sell. Based on merit order principle, sources of energy are based on ascending order of marginal price of production. This competition between electricity products creates the supply curve which described generation constraints reflected in marginal costs. Various factors can shift the supply curve, such as changes in production costs, technology, input prices, or government policies, ultimately impacting the equilibrium price and quantity in the market.

3.2.1 Main Features

The supply curve is one of the main elements for elaborating the electric spot price in a market clearing process. Indeed, by matching with the demand curve, one can find the market clearing price associated to the equilibrium quantity. At all time, the supply curve is constructed by adding together the bid offers ordered by increasing marginal costs of production. A new supply curve is thus created by adding next to each other all types of production of an area.

To model this supply curve, it remains necessary to find a function that reproduces its shape. Thus, the function must be increasing, continuous and must adopt a behaviour reproducing an exponential character towards high quantities of electricity produced. One of the primary factors is the cost of production, which encompasses various expenses incurred during electricity generation. These costs include capital expenditures for building power plants, operational costs such as fuel and maintenance, and regulatory compliance expenses. As the price of electricity increases, it becomes more profitable for producers to generate electricity, resulting in a greater quantity being supplied. Additionally, the type of energy sources used for electricity generation plays a pivotal role in shaping the supply curve. Different energy sources have varying production costs, efficiency levels, and environmental impacts. For instance, renewable sources like solar and wind have lower marginal costs. On the other hand, fossil fuels such as coal and natural gas have variable costs associated with fuel prices and emissions regulations. As these costs change, the supply curve adjusts accordingly. Technological breakthroughs and innovations also influence the supply curve by altering the efficiency and cost-effectiveness of electricity generation methods. Breakthroughs in energy storage, grid management, and generation technologies can lower production costs. Conversely, disruptions in supply chains or technological constraints can reduce supply and therefore cause an increase in price.

3.2.2 Literature review for fitting supply curve with linear and logistic curves

If we attempt at approximating the supply curve with mathematical function, two types of models are mainly quoted in the literature, the linear 3.2.1 and the logistic 3.2.2 supply curve. For example, [23, Part 3] considers a model for forecasting based on an approximation of supply and demand curves compile with bid data of auctions.

$$\text{For linear curve: } P_t = \begin{cases} 0 & \text{if } Q_t < Q^{min} \\ a + b * Q_t & \text{if } Q_t \in (Q^{min}, Q^{max}) \end{cases} \tag{3.2.1}$$

$$\text{For logistic curve: } P_t(Q_t) = \begin{cases} 0 & \text{if } Q_t < Q^{min} \\ a + \frac{b_1}{1+e^{-b2(Q_t - b3)}} & \text{if } Q_t \in (Q^{min}, Q^{max}) \end{cases} \tag{3.2.2}$$

where Q_t is the demand at time t.

This short presentation allows us to better understand the first techniques of modelling the supply curve. Subsequently another model is chosen for section 3.4 but it is interesting to present different existent approaches.

3.3 Demand curve

We now need to define the demand curve. In order to achieve this, several approaches can be considered. As we have already illustrated the importance of the Ornstein-Uhlenbeck (OU) process

in modelling the market, and the fact that it lends itself particularly well to modelling the demand curve, we are going to use it to complete a periodic trend, which we will model using a period function whose period frequencies is be determined by the Fourier transform as done in section 2.3.

Definition 3.3.1. For the following section, we assume that the demand curve is built from two process, one is a deterministic and periodic function and the second is an Ornstein-Uhlenbeck process. Therefore, the demand curve is defined as the following :

$$D_t = P_t + S_t, \forall t \in [0, T] \tag{3.3.1}$$

where T is the studied period, $(P_t)_{t\in[0,T]}$ is a periodic function and $(S_t)_{t\in[0,T]}$ is an Ornstein-Uhlenbeck process.

3.3.1 Ornstein-Uhlenbeck process

As stated by definition 3.3.1, residuals defined by $\forall t\ S_t = D_t - P_t$ are realisations of an Ornstein-Uhlenbeck process.

Remark 3.3.2. The function P_t can be modelled using the same study carried out in part 2.3.

Definition 3.3.3. We define the Ornstein-Uhlenbeck process as followed :

$$\begin{cases} dS_t = -\theta S_t dt + \sigma dW_t \\ S_0 = S_o \end{cases} \tag{3.3.2}$$

where θ is the mean reversion rate and σ is the volatility.

Theorem 3.3.1. The explicit solution of an Ornstein-Uhlenbeck process $(S_t)_{t\in[0,T]}$ with $S_0 = S_o$ is:

$$S_t = S_o \exp(-\theta t) + \sigma \int_0^t \exp(-\theta(t-s))\, \mathrm{d}W_s \tag{3.3.3}$$

where W_s is the standard Wiener process.

Proof. By using variation of parameters, defining $X_t = S_t \exp(\theta t)$ and by Itô's lemma:

$$\begin{aligned} dX_t &= \theta S_t \exp(\theta t)dt + \exp(\theta t)dS_t \\ &= \sigma \exp(\theta t)dW_t \\ S_t \exp(\theta t) - S_o &= \int_0^t \exp(\theta s)\mathrm{d}W_s \text{ ,(By Integrating from 0 to t)} \\ S_t &= S_o \exp(-\theta t) + \sigma \int_0^t \exp(-\theta(t-s))\, \mathrm{d}W_s \end{aligned}$$

□

The following simulation equation 3.3.4 is used for generating paths of Figure 3.1. The equation is obtained by splitting time into time step with the same length.

$$S_{t+dt} = S_t \exp^{-\theta dt} + \sigma \sqrt{\frac{1 - \exp^{-2\theta dt}}{2\theta}} \mathcal{N}_{0,1},\ \forall t \in [0, T - dt] \tag{3.3.4}$$

where θ is the mean reversion rate and σ is the volatility and $\mathcal{N}_{0,1}$ is the Normal law.

Figure 3.3.1 shows the result of 10 paths simulated according to the parameters defined in the description.

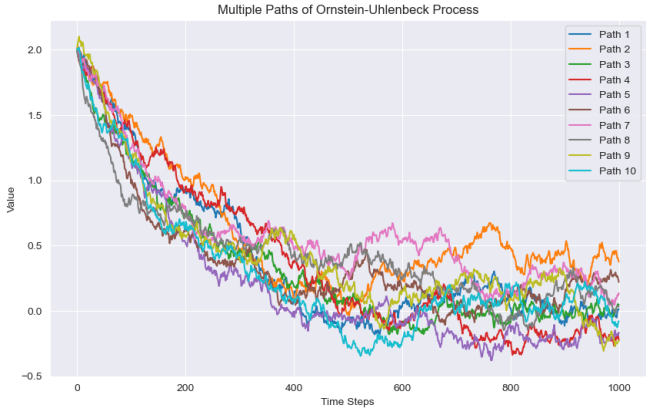


Figure 3.1: 10 sample paths from the following Ornstein-Uhlenbeck process (x0=2.0, $\theta = 0.2$, $\sigma = 0.2$, dt=0.01, number of steps=1000, number of paths=10)

The OU process is used for its mean-reverting nature. To illustrate the properties of parameter-dependent dynamics θ and σ, a sensitiveness analysis is presented in Figures 3.2 to understand how these parameters impact the behaviour of the curve.

(a) Theta sensitiveness of OU process (x0=2.0, sigma =0.2, dt=0.01, number of steps=1000, number of paths=5)

(b) Sigma sensitiveness of OU process (x0=2.0, Theta =1, dt=0.01, number of steps=1000, number of paths=5

Figure 3.2: Theta and Sigma sensitiveness of OU process

The θ parameter is responsible for the mean reversion characteristic. The greater the coefficient is, the greater the mean reversion is. The figure 3.2(a) clearly shows this phenomenon, with a return to the mean 0 in less than 200 steps for $\theta = 2$ and a very slow decrease for $\theta = 0.001$.

The σ parameter is responsible for the amplitude of volatility. For a fixed mean reversion coefficient, each amplitude depends on σ, which can lead to a deviation from the mean, resulting in sudden jumps in amplitude as Figure 3.2(b) shows.

3.3.2 Calibration of Ornstein-Uhlenbeck process to demand curve

In this section, two tools are presented to fit the Ornstein-Uhlenbeck process to the demand curve. First, we present a calibration by least squares regression and secondly by Maximum Likelihood estimation.

Least squares regression needs linear relationship between explained variable S_{t+1} and the observable variable S_t. It's the case for simulation 3.3.4 where the relation between two consecutive observations is linear with iid Normal term ϵ such as

$$S_{t+\Delta} = \alpha + \beta S_t + \epsilon \tag{3.3.5}$$

where $\alpha = 0$, $\beta = \exp(-\theta\Delta)$ and $sd(\epsilon) = \sigma\sqrt{\frac{1-\exp^{-2\theta\Delta}}{2\theta}}$

By rewriting these equations:

$$\begin{cases} \theta &= -\frac{\log(\beta)}{\Delta} \\ \sigma &= sd(\epsilon)\sqrt{\frac{2\theta}{1-\exp^{-2\theta\Delta}}} \end{cases}$$

Theorem 3.3.2. (β and $sd(\epsilon)$ estimators) Estimators for β and $sd(\epsilon)$ are given by the following formula:

$$\beta = \frac{n\sum_{i=0}^{n-1} S_{i+1}S_i - \sum_{i=0}^{n-1} S_i \sum_{i=0}^{n-1} S_{i+1}}{n\sum_{i=0}^{n-1} S_i^2 - (\sum_{i=0}^{n-1} S_i)^2} \tag{3.3.6}$$

$$sd(\epsilon) = \sqrt{\frac{n\sum_{i=0}^{n-1} S_{i+1}^2 - (\sum_{i=0}^{n-1} S_{i+1})^2 - \beta(n\sum_{i=0}^{n-1} S_{i+1}S_i - \sum_{i=0}^{n-1} S_i \sum_{i=0}^{n-1} S_{i+1})}{n(n-2)}} \tag{3.3.7}$$

Once an estimate has been made of β and $sd(\epsilon)$, θ and σ can be easily deduced.

Maximum Likelihood estimation requires conditional distribution of S_{t+1} given S_t. The conditional probability density of an observation given a previous observation (with a time step of Δ) is normally distributed and given by 3.3.8

$$f(S_{i+1}|S_i, \theta, \sigma) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp(-\frac{(S_{i+1} - S_i \exp(-\theta\Delta))^2}{2\sigma_1^2}) \tag{3.3.8}$$

where $\sigma_1^2 = \sigma^2 \frac{1-\exp(-2\theta\Delta)}{2\theta}$.

By asymptotic analysis, since Δ is small compared to θ, we can assume $\sigma_1^2 \sim \sigma^2\Delta$.

We derive from the conditional density function the log-likelihood function:

$$\begin{aligned} \mathcal{L}(\theta, \sigma_1) &= \sum_{i=0}^{n-1} \log f(S_{i+1}|S_i, \theta, \sigma_1) \\ &= -\frac{n}{2}\log(2\pi) - n\log(\sigma_1) \\ &- \frac{1}{2\sigma_1^2}\sum_{i=0}^{n-1}[S_{i+1} - S_i\exp(-\theta\Delta)]^2 \end{aligned} \tag{3.3.9}$$

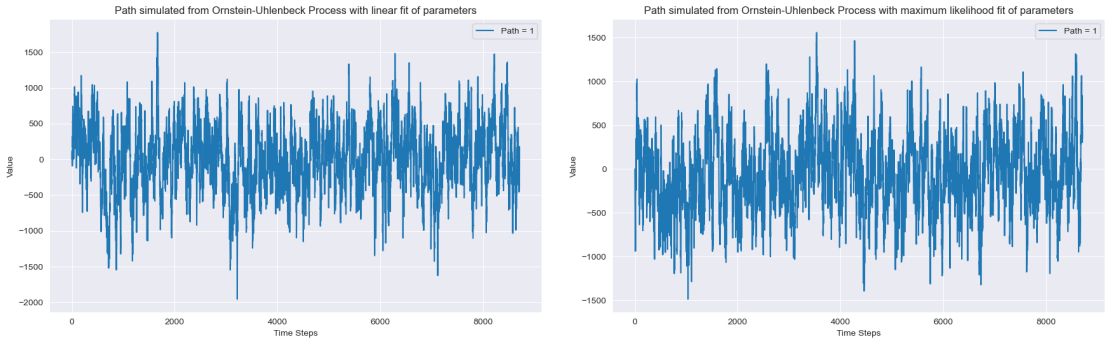
To find the maximum of this log-likelihood curve, all partial derivatives have to be zero. Thanks to equation 3.3.9, we derive partial derivatives associated to θ and σ_1.

$$\begin{aligned} \frac{\partial\mathcal{L}(S_i, \theta, \sigma_1)}{\partial\theta} = 0 \iff \\ -\frac{1}{2\sigma_1^2}\sum_{i=0}^{n-1} 2[S_{i+1} - S_i\exp(-\theta\Delta)]S_i\Delta\exp(-\theta\Delta) = 0 \iff \\ -\frac{\Delta\exp(-\theta\Delta)}{\sigma_1^2}\sum_{i=0}^{n-1}[S_{i+1}S_i - \exp(-\theta\Delta)S_i^2] = 0 \iff \\ \theta = -\frac{1}{\Delta}\log(\frac{\sum_{i=0}^{n-1} S_{i+1}S_i}{\sum_{i=0}^{n-1} S_i^2}) \end{aligned} \tag{3.3.10}$$

$$\begin{aligned} \frac{\partial\mathcal{L}(S_i, \theta, \sigma_1)}{\partial\sigma_1} = 0 \iff \\ -\frac{1}{\sigma_1} + \frac{1}{\sigma_1^3}\sum_{i=0}^{n-1}[S_{i+1} - S_i\exp(-\theta d\Delta)]^2 = 0 \iff \\ \sigma_1^2 = \frac{1}{n}\sum_{i=0}^{n-1}[S_{i+1} - S_i\exp(-\theta\Delta)]^2 \\ \sigma_1^2 \sim \frac{1}{n\Delta}\sum_{i=0}^{n-1}[S_{i+1} - S_i\exp(-\theta\Delta)]^2 \end{aligned} \tag{3.3.11}$$

The solution for σ^2 depends on θ. However θ can be estimated only with data. Therefore, we can find optimal solution for σ and θ by solving equations 3.3.10 and 3.3.11 by using historical data of demand.

A simulation of the same process but with the two different ways of calibration is shown in Figures 3.3

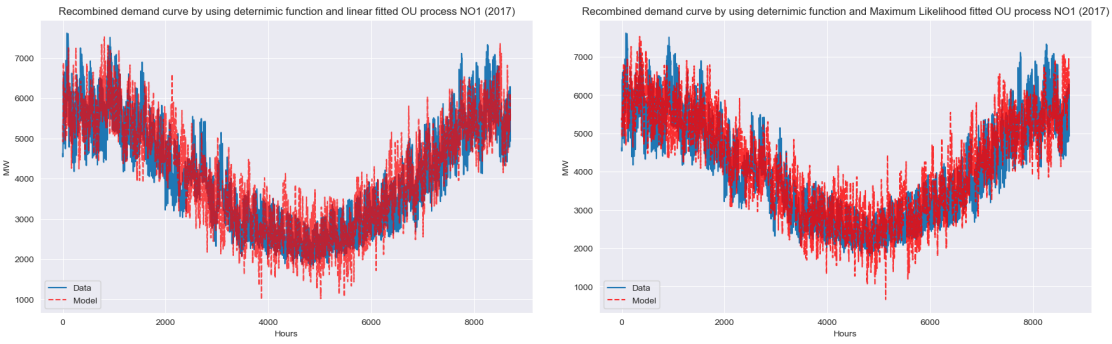


(a) Simulated path from Ornstein-Uhlenbeck Process with linear fit of parameters (sigma = 14493, theta=466)

(b) Simulated path from Ornstein-Uhlenbeck Process with maximum likelihood fit of parameters (sigma = 14492, theta=466)

Figure 3.3: Simulated path from Ornstein-Uhlenbeck Process with linear and maximum likelihood fit

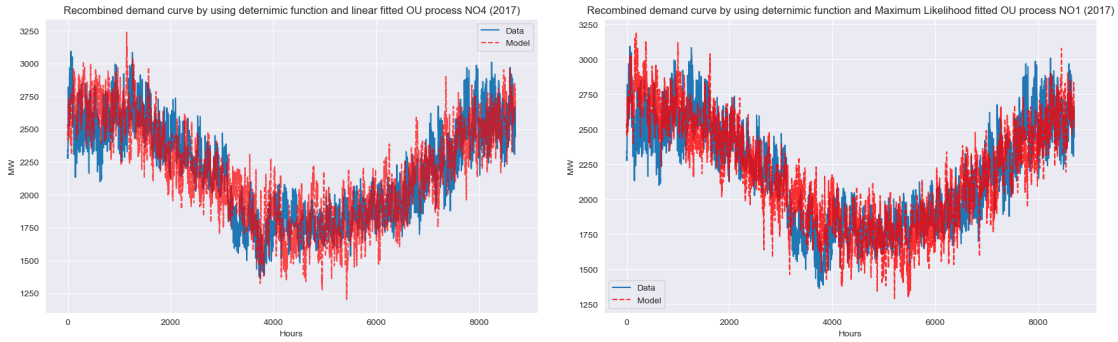
Using preciously calibrated OU processes, we can predict annual demand and thus forecast its evolution in our market clearing model. The reconstructed demand for the NO1 and NO4 zones are shown in the figures 3.4 and 3.5. The part of the modelling of the deterministic function is not described in this section because the same study was carried out in advance in section 2.3. It can be seen that the model follows the annual demand profile well, with peak amplitudes corresponding to the data. Nevertheless, the model could be improved with the possibility of adding a stochastic process for volatility that can be found, for example, in the Heston model.



(a) Path simulated from Ornstein-Uhlenbeck Process with linear fit of parameters(sigma = 14493, theta=466)

(b) Path simulated from Ornstein-Uhlenbeck Process with maximum likelihood fit of parameters (sigma = 14492, theta=466)

Figure 3.4: Reconstructed demand NO1 (2017)



(a) Path simulated from Ornstein-Uhlenbeck Process with linear fit of parameters(sigma = 4450, theta=469)

(b) Path simulated from Ornstein-Uhlenbeck Process with maximum likelihood fit of parameters (sigma = 4411, theta=469)

Figure 3.5: Reconstructed demand NO4 (2017)

3.3.3 Integration of run-of-river technologies in the demand curve

. One of the main types of hydro-generator used for electricity production is the "run-of-river". Compared to dams, "run-of-river" (ROR) don't store water above the gate but use the flow of the river to produce power. In this context, Norway has considerably installed many of them along rivers to produce low-price electricity in part to develop aluminum industry. Quantity in MW produce by those technologies cannot be negligible against the hydro reservoir plants. In addition, this means of production is considered as fatal production, i.e. the operator cannot manage the quantity of water to optimise production; he must necessarily produce instantaneously. However, the marginal production cost of these systems is considered to be zero. This is why, in order to integrate these means of production into our modelling, it is better to integrate them as a reduction in demand. The amount of power available from run-of-river units will be subtracted from demand to create an adjusted demand that must be met solely by generation from hydroelectric reservoirs.

In order to quantify electricity production from this energy source and integrate it in a market clearing model, an idea is to link the production to water flow. Indeed, water flow can be considered as a stochastic process following a stochastic differential equation. Consider a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a Brownian motion $(W_t)_{t\geq 0}$. According to [24, Part 3, page 1713], we can express the water flow by the following SDE :

$$\begin{aligned} &d\mathcal{Q}_t = b_s(\mathcal{Q}_t, t)dt + \sigma_s(\mathcal{Q}_t, t)dW_t, \text{ if t} \in [\text{s,T}], \\ &\mathcal{Q}_s = q \end{aligned} \tag{3.3.12}$$

Since, a seasonality trend is identifiable in data of water flows, the idea is to divide the production into two parts, once for the deterministic component and the second for the stochastic part. This stochastic process can be model by a mean-reverting process with mean 0.

Let assume a process with stochastic process but minus the seasonal deterministic function, $S_t = R_t - r(t)$. Furthermore, we assume that this process follows an Orstein-Uhlenbeck process reverting towards 0, i.e.

$$dS_t = -\kappa S_t dt + \sigma dW_t, \text{ where } \kappa \geq 0, \sigma \geq 0. \tag{3.3.13}$$

Theorem 3.3.3. Water flow process $\mathcal{Q}_t = \exp(r(t)+S_t)$ follows the following stochastic differential equation

$$d\mathcal{Q}_t = (\frac{dr(t)}{dt} + \frac{1}{2}\sigma^2 - \kappa(\log(\mathcal{Q}_t) - r(t)))\mathcal{Q}_t dt + \sigma \mathcal{Q}_t dW_t \tag{3.3.14}$$

Proof. By using Itô's lemma :

$$
\begin{aligned}
d\mathcal{Q}_t &= d(exp(r(t) + S_t)) \\
&= (\frac{dr(t)}{dt} e^{(r(t)+S_t)} dt + dS_t e^{(r(t)+S_t)} + \frac{1}{2}\sigma^2 e^{r(t)} d[e^S, e^S]_t \\
&= (\frac{dr(t)}{dt} \mathcal{Q}_t dt - \kappa S_t dt + \sigma dW_t)\mathcal{Q}_t + \frac{1}{2}\sigma^2 \mathcal{Q}_t dt \\
&= (\frac{dr(t)}{dt} + \frac{1}{2}\sigma^2 - \kappa S_t)\mathcal{Q}_t dt + \sigma \mathcal{Q}_t dW_t
\end{aligned}
$$

□

Run-of-river unit generally operates in On/Off mode. As a result, the few operating costs are generally start-up and shut-down costs. Electricity is generated by a turbine powered by the movement of the watercourse, which drives an alternator. The physics of electricity production using this method are shown in the following system 3.3.15 :

$$
\mathcal{P}(Q) = \begin{cases} 0 & \text{if } \mathcal{Q} \leq \mathcal{Q}_{min} \\ c \times \eta(\mathcal{Q}) \times \mathcal{Q} & \text{if } \mathcal{Q}_{min} \leq \mathcal{Q} \leq \mathcal{Q}_{max} \\ c \times \eta(\mathcal{Q}_{max}) \times \mathcal{Q}_{max} & \text{if } \mathcal{Q}_{max} \leq \mathcal{Q} \end{cases} \tag{3.3.15}
$$

where $c = \rho g h$ with ρ density of water, g acceleration of gravity and h the level of water and where η is an efficiency function depending on river flows.

For the following sections, this model will not be used since the generation data for run-of-river is directly available on the Entsoe website [25]. However, an approach based on the model explained above could be envisaged, provided that access to historical river data is available.

3.4 Basic Model of market clearing based on exponential supply curve

To complete the section 3.2 on the supply curve, this section looks in more detail at a dynamic approach based on an exponential function.

3.4.1 Description of the model

It is clear that the supply curve behaves dynamically over time. Supply is not constant for every hour of every day of the year. As a result, the model needs to incorporate a time expenditure to reflect hour-by-hour variations. To do this, the values of the potential electricity reserves contained in the hydraulic reservoirs is integrated hour by hour, as are all the other variables in our system. Unlike Barlow [26], who considers the structure of supply to be fixed, we vary the structure of supply and demand over time.

The choice of supply curve model has evolved over time. A relatively simplistic model is the linear model as presented in 3.2.2, which is rarely used because it does not reflect the exponential rise in prices when demand soars. Research has therefore focused on functions that translate this exponential rise in prices, hence the use of exponential function modelling of the supply curve, which can be found in [27]. However, it is not an exponential function that is used in this section. If we refer to [26] and [28], a power function better reflects the dynamics of the supply market and allows sudden price spikes to be modelled more accurately than an exponential function.

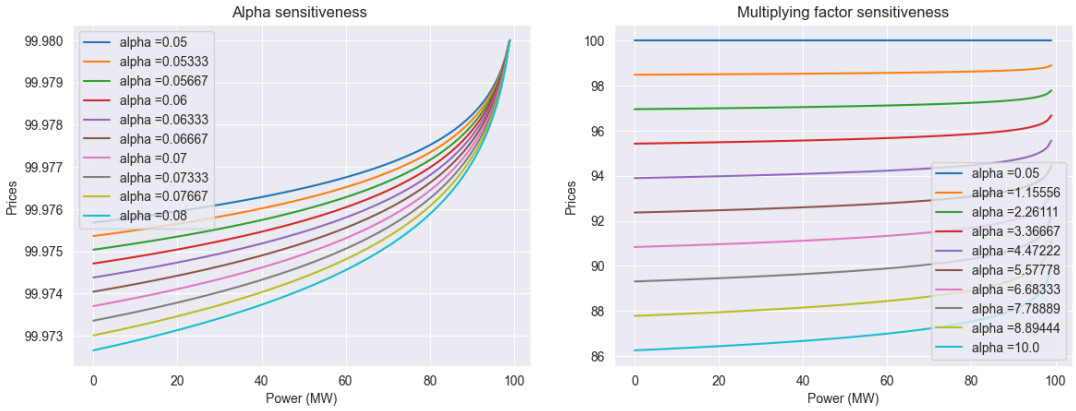
Furthermore, the supply curve has to fit well with physical conditions that the market has set for the modelling. Indeed, Nordpool imposes a maximal price to limit outrageous prices if a major problem should occurs on the network. For the Day-Ahead spot price on the Norwegian market, this price is settled by NordPool at 50000 NOK/MWh. In addition, market operators cannot send an infinite quantity of electricity to the market because infrastructure limits transmission. This is why a maximum supply is added to the model to take account of this behaviour.

Therefore based on [28], the same model is used for the supply curve:

$$
P_t(S_t) = \bar{P} - e^{a_t}(\bar{S} - S_t)^{\alpha} \tag{3.4.1}
$$

where S_t is the supply at time t, $\bar{S}$ is the maximal installed supply capacity, $\bar{P}$ is the maximal price and α and a_t are coefficients which can be constant or time-dependent of physical parameters.

To understand why this modelling is interesting for our model, two coefficient sensitivity analysis for α and a_t are presented in 3.6.



(a) α sensitiveness of the supply curve

(b) a_t sensitiveness of the supply curve

Figure 3.6: α and a_t sensitiveness of the supply curve

Figure 3.6(a) shows how the parameter α change the shape of the supply curve especially when the power approaching the limit supply value. When α increases, the curve becomes steeper when power increases. For small values of α, the curve gets a linear shape with a constant coefficient of growth. This graph accurately reflects a quality that our model must have, which is to adapt the growth of the curve according to the available supply and in a supply and demand model that must therefore be equal to demand.Figure 3.6(b) shows the impact of the multiplier coefficient a_t on the supply curve. We can see that it translates and modulates supply. Subsequently, when the coefficient is determined by concrete variables, it can be used to simulate an adjustment of supply according to the quantity of electricity that can be loaded onto the network, and thus reflect the daily modulations due to the means of production.

Based on previous remarks and analyses, supply curve has to be adapted to macro data which by causal effect influence shape of our curve. Three specific energy data are used to model multiplying factor a_t which are:

- Reservoir levels (MWh) (named as r_t)
- Carbon emission right prices (NOK/MWh) (named as P_t^{EUA})
- Coal prices (NOK/MWh) (named as P_t^{Coal})

Therefore, four models are defined, each of those incorporates a new variable in addition to others. M1 Model 3.4.2 is built with only constant for a_t, M2 Model adds reservoir levels variable, M3 Model incorporates carbon emission right prices and finally coal prices are added to model M4. All equations are listed in 3.4.2:

$$P_t^1 = \bar{P} - \exp(a)(\bar{S} - S_t^1)^\alpha \qquad (M1) \qquad (3.4.2a)$$

$$P_t^2 = \bar{P} - \exp(a + b_r^2 \times r_t)(\bar{S} - S_t^2)^\alpha \qquad (M2) \qquad (3.4.2b)$$

$$P_t^3 = \bar{P} - \exp(a + b_r^3 \times r_t + b_p^3 \times P_t^{EUA})(\bar{S} - S_t^3)^\alpha \qquad (M3) \qquad (3.4.2c)$$

$$P_t^4 = \bar{P} - \exp(a + b_r^4 \times r_t + b_p^4 \times P_t^{EUA} + b_c^4 \times P_t^{Coal})(\bar{S} - S_t^4)^\alpha \qquad (M4) \qquad (3.4.2d)$$

One reason why this structure of supply curve is useful and wide spread is because of exponential behaviour and the recovering of linear structure after a log transformation. After readjusting terms and taking log in both sides, equations 3.4.2 becomes:

$$log(\bar{P} - P_t^1) = a + \alpha \times log(\bar{S} - S_t^1) \quad (M1) \quad (3.4.3a)$$
$$log(\bar{P} - P_t^2) = a + b_r^2 \times r_t + \alpha \times log(\bar{S} - S_t^2) \quad (M2) \quad (3.4.3b)$$
$$log(\bar{P} - P_t^3) = a + b_r^3 \times r_t + b_p^3 \times P_t^{EUA} + \alpha \times log(\bar{S} - S_t^3) \quad (M3) \quad (3.4.3c)$$
$$log(\bar{P} - P_t^4) = a + b_r^4 \times r_t + b_p^4 \times P_t^{EUA} + b_c^4 \times P_t^{Coal} + \alpha \times log(\bar{S} - S_t^4) \quad (M4) \quad (3.4.3d)$$

3.4.2 Principles of linear regression

In order to find the most accurate parameters for solving 3.4.3, we are going to apply principles of linear regression and how we can compute the Ordinary least squares estimator (OLS).

Suppose we have independent observations of our target variable and also our regressor variables and we assume that target variables and regressor variable are linked by a linear model described as follows:

$$\forall i \in [1, n],\ y_i = \theta_0 + \theta_1 x_{i,1} + ... + \theta_p x_{i,p} + \epsilon_i \quad (3.4.4)$$

where p represents the number of parameters and n the number of observations. We can describe this setup with a matrix modelling setup :

$$\mathbf{X} = \begin{bmatrix} 1 & x_{1,1} & \ddots & x_{1,p} \\ 1 & x_{2,1} & \ddots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & \ddots & x_{n,p} \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},\ \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix}$$

.

When one performs an ordinary least squares regression, the main goal is to minimise the following Lost function by choosing the proper vector θ :

$$L(\theta) = \|Y - X\theta\|_2^2 \quad (3.4.5)$$

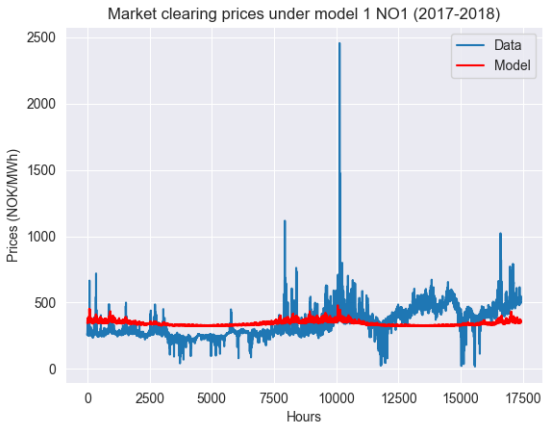
where $\|.\|$ is the $\mathbf{L}^2$ norm. By differentiating and taking the gradient equals to 0, we obtain an explicit formula for OLS estimator : $\hat{\theta} = (X^T X)^{-1} X^T Y$.

3.4.3 Results

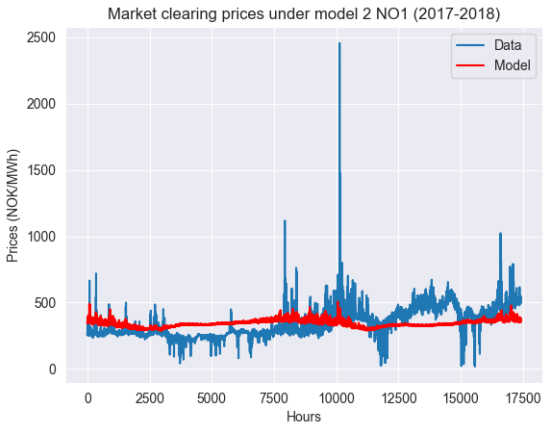
Section 3.4.2 enables to set up a clear mathematical model to fit our models. The idea of this study is to highlight the macro data that most influences the price of electricity. It is not designed to accurately reflect the sudden changes that the market can undergo on an hour-by-hour basis. In fact, our data is mainly expressed in daily periods and interpolated in a constant way over the whole day. As a result, any sudden movements cannot be explained using linear regression models alone. An approach based on a more precise definition of the market structure is presented in the following section 3.5.

As described in the section 3.4.1, four models are studied, with a variation in the variables considered between each model. The variables considered are firstly a model without any physical macros, then the second only includes the level of reservoirs in the study area, then we add the prices of carbon contracts issued by Europe and finally the price of coal is added to the last model. Undeniably other parameters can be considered, such as temperature, but this is outside the scope of this section. The analysis will cover a two-year period from the beginning of 2017 to the end of 2018, starting with Norway's NO1 zone. The prices are into Norwegian kroner per MWh (NOK/MWh), as are the prices of emissions contracts and coal.

The results are presented in two different ways. The first way chosen is a graph, to enable the reader to see the impact of the macro considered on the accuracy of the model's calibration on prices.

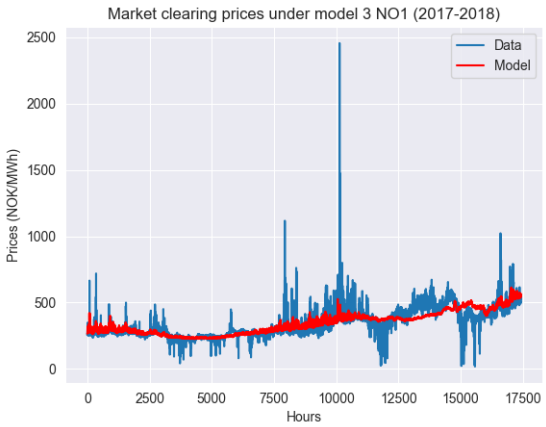


(a) Market clearing prices Model 1 NO1 (2017-2018)

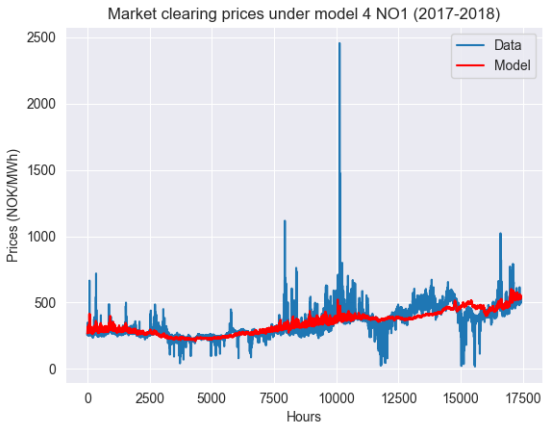


(b) Market clearing prices Model 2 NO1 (2017-2018)

Figure 3.7: Market clearing prices Model 1-2 NO1 (2017-2018)



(a) Market clearing prices Model 3 NO1 (2017-2018)



(b) Market clearing prices Model 4 NO1 (2017-2018)

Figure 3.8: Market clearing prices Model 3-4 NO1 (2017-2018)

Models \ coefficients	a	b_r	b_{eua}	b_{coal}	α
1	10.807	–	–	–	$7.50e^{-4}$
2	10.805	$-2.68e^{-10}$	–	–	$1.050e^{-3}$
3	10.807	$1.18e^{-10}$	$-2.46e^{-5}$	–	$1.056e^{-3}$
4	10.809	$1.72e^{-10}$	$-2.37e^{-5}$	$-1.96e^{-6}$	$1.045e^{-3}$

Table 3.1: Linear regression coefficients

To access the quality of the performance, two well-known indicators are used: the mean absolute estimator and the root mean squared error (RMSE) :

$$MAE = \frac{1}{N}\sum_{n=1}^{N}|(\hat{y_n} - y_n)| \tag{3.4.6}$$

$$RMSE = \frac{1}{\sqrt{N}}\sqrt{\sum_{n=1}^{N}(\hat{y_n} - y_n)^2} \tag{3.4.7}$$

Model	1	2	3	4
R Score	0.029	0.062	0.6073	0.6102
RMSE	106.70	104.86	67.83	67.59
MAE	88.62	87.27	40.8	40.69

Table 3.2: Statistics for evaluating approximation errors

Table 3.2 presents a few statistics to quantify the accuracy. R score, RMSE and MAE have been computed and acquaint in this table. To access precision of the fitting of a linear regression, Rscore gives a first approach. In our results, model 1 and 2 have a really low Rscore compared to model 3 and 4. Recall that higher the R score is, closer the model is to data. A sharp increase is observed between model 2 and model 3 when prices of carbon emission is added to the model. This is confirmed by observing the difference between plot 3.7 and plot 3.8 because Models 3-4 represented by a red trend in our model fit better variations of Day-Ahead spot prices. This observation is confirmed by statistics. For example, RMSE drops significantly from approximately 100 to 68 between model 2 and 3. The same is observed with MAE estimator which is divided by 2 between model 2 and 3. It seems that the price of carbon emission permits is an essential feature for establishing prices in Norway.

3.5 The water value

In this final part, the concept of water value is extended to the case of Norway where traditional definitions of water value for electricity production aren't well adapted due to the lack of production by thermal plants. The main goal of this section is to review traditional approaches for defining water value and trying to create a new understanding of this concept.

3.5.1 Description

Day-Ahead electricity prices has a dynamic based on price clearing which is a result of a balance between supply and demand every hour. Variations of supply and demand from one hour to the next one can appears involving high volatility in the market. Thanks to the flexibility of hydro-power production, allocation of water to be converted into electricity to maximise profit, is managed with influence of water stocks, Day-Ahead prices and installed power. The operators face the decision whether to use the water in the hydro reservoirs now or later. Therefore, the relevant costs are the opportunity costs (water value) of using the water in the future [2, Part 2, 2.1].

When we look at hydro-power generation, one of the key steps in understanding price formation is the coordination of hydro-power units between other generation systems. Globally, the decision-making of any production system is based on minimising the overall production costs under the various constraints that apply to our system. With the liberalisation of the electricity markets, the competition between producers has taken hold, creating a decentralized decision-making, traditionally carried out by a national and central company towards private companies. As a result, a company owning hydraulic assets has to plan a production schedule that generally covers an entire year, taking into account its production strategy as well as the strategy of its competitors, in order to maximise the benefits of its decisions. For a company to manage its production, access must be granted to both hydraulic and thermal production units, and its aim will be to decide when to activate one in favour of the other. If we refer to the article [29, Part 2], energy systems can interact with each other in two ways, either temporally or geographically. To create a link between two markets, one needs a storage unit capable of retaining energy from one market in order to transfer it to another. In the case of electricity markets, reservoirs are an ideal tool because water enables electricity to be stored in the form of potential energy that can be released again at selected times. In addition, hydraulic reservoirs provide a link between long-term programming and short-term operations, as it is very easy and inexpensive for a dam to switch from an operational state to a shutdown state, which ensures great operational flexibility and a means of optimising its profitability in relation to the market.

This is precisely the objective of a hydro producer company, to optimise its water stock with the aim of using it when sufficient profitability signals appear in order to cover short-term demand while maintaining a long-term management strategy.

A hydroelectric generating company will therefore be interested in trying to assess the value of its water stock in relation to its planned medium and long-term schedule. The aim is to find a way of comparing the potential value of the reserve with the market price in order to obtain a price signal to trigger production. This is the purpose of water value, which is an interesting indicator for getting an idea of how much marginal cost it would be worth bidding hydro production on the market to optimise our profitability.

3.5.2 Model

For considering a model in Norway in order to compute water value at each time step, a framework is needed to better fit the situation in this country. Therefore, thanks to the description in part 1.3, one can consider the market in Norway as a centralized environment dominated by a central company which is responsible of the exploitation of hydro storage plants, run-of-river hydro generation plants. The quantity of electric imports in Norway can be considered as electricity produced by thermal plants because traditionally electricity bought on the market is at the most expensive of bidding price and this price are reached by thermal plants. This last remark is essential because it is the central point of the redefinition of Water Value for countries whose thermal production is insufficient to cover demand. Indeed, thereafter, many changes in the modelling of constraints appear and is presented in the course of this section.

In this framework, the objective function is to minimize the total generation costs by thermal units or in the view point of Norway to minimise imports of overseas electricity and promote the quantity of electricity exported in line with its available generation resources. To simplify the formulation of the problem, a unique hydro storage plant is consider as well as an inelastic demand. This situation matches very well with the description of [30, part 3.1].

Traditionally for solving this cost minimization problem, one can try to solve the following problem 3.5.1. It is exactly the first approach we can find in [31, part 3].

$$\begin{aligned} \underset{T_t, H_t}{min} \quad & \sum_{t=1}^{t_{max}} C_t(T_t) \\ \text{subject to} \quad & T_t + H_t = D_t \text{ for all t} \\ & \underline{H} \leq H_t \leq \overline{H} \text{ for all t} \\ & 0 \leq T_t \leq \overline{T} \text{ for all t} \\ & \sum_{t=1}^{t_{max}} H_t \times l_t \leq R_{max} \end{aligned} \tag{3.5.1}$$

where T_t is the thermal power (MW) at time t, H_t is the hydro power (MW) at time t, D_t is the demand (MW) at time t, $\underline{H}$ is the minimal hydro power (MW), $\overline{H}$ is the maximal hydro power (MW), $\overline{T}$ is the maximal thermal power (MW), R_{max} is maximal total hydro production (MWh), l_t is the duration of each period of production (h) and $C_t(T_t)$ is the cost function.

The problem 3.5.1 is the basic framework. Equation $T_t + H_t = D_t$ represents the supply and demand balance equation at each t with inelastic price. Inequalities are described physical dynamics of generation constraints by settling power limits and maximum capacities of production at all t.

Now that the basic framework has been defined by 3.5.1, it needs to be adapted to best model the situation in Norway. As mentioned at the beginning of this section, Norway has very little thermal production capacity. In addition, as the study focuses on the NO1 and NO4 zones, thermal production for these zones is non-existent. As a result, thermal production T_t is replaced by imported or exported production from adjacent zones. Thus, thermal generation becomes imported or exported power (positive sign for power imports and negative sign for power exports), which has consequently changed the limits of generation by no longer reducing by 0 but by considering the transmission capacities between the study area and its neighbours. Furthermore, the previous model only included a general constraint on hydro generation, which must not exceed maximal reserves. By refining the model, a daily time step constraint (or even hourly if the needs are on this scale) can be added by inserting a dynamic for emptying and filling reservoirs according to water inflows and hydraulic generation for the period [30, Part 3]. Finally, the new mathematical formu-

lation for minimising the costs over the all period of one year with daily steps under constraints is stated as follows:

$$\begin{aligned}
\min_{T_{w,d},H_{w,d},R_{w,d}} \quad & \sum_{w=1}^{52}\sum_{d=1}^{7} C_{w,d}(T_{w,d}) \\
\text{subject to} \quad & T_{w,d} + H_{w,d} = D_{w,d}\ \forall(w,d), \quad (\eta_{w,d}) \\
& R_{w,d} - R_{w,(d-1)} + H_{w,d} \times 24 = I_{w,d}\ \forall(w,d), \quad (\mu_{w,d}) \\
& 0 \leq H_{w,d} \leq \overline{H_{w,d}}\ \forall(w,d), \\
& \underline{R_{w,d}} \leq R_{w,d} \leq \overline{R_{w,d}}\ \forall(w,d), \\
& \underline{T_{w,d}} \leq T_{w,d} \leq \overline{T_{w,d}}\ \forall(w,d), \\
& R_{1,0} = R_{in} \\
& R_{52,7} = R_{end}
\end{aligned} \tag{3.5.2}$$

where $T_{w,d}$ is the imported/exported power (MW) for each day of the year, H_t is the average hydro power (MW) for each day of the year, D_t is the demand (MW) for each day of each week, $R_{w,d}$ is the reservoir level in energy (MWh) for each day of the year, $I_{w,d}$ is the water inflow (MWh) for each day of each week, $\overline{H}$ is the maximal hydro power (MW), $\underline{T_{w,d}}$ is the minimal exported power (MW), $\overline{T}$ is the maximal exported power (MW), $\underline{R_{w,d}}$ is the minimal level of reservoir (MWh), $\overline{R_{w,d}}$ is the maximal level of reservoir, R_{in} and R_{end} are initial and final condition for reservoir level and $C_{w,d}(T_{w,d})$ is the cost function.

The overall structure of the algorithm 3.5.1 has been retained, but a number of constraints have been added to provide the best possible model of production dynamics throughout the study period. Equation $(\eta_{w,d})$ represents the equality between supply and demand in the same way as the previous algorithm 3.5.2, except that this time it is discretized for each day of each week and $T_{w,d}$ is imported power. The most important change in the description of the dynamics is probably the addition of the equation $(\mu_{w,d})$ which describes the evolution of the energy level contained in the reservoirs subject to the hydraulic production and the quantity of water entering the system. The variation in energy of the water stored in the reservoir between two consecutive days translated by $R_{w,d} - R_{w,(d-1)}$ is the result of the amount of energy entering during the day minus the average production of the day by the hydraulic system multiplied by 24 hours.

To find optimal solutions, Lagrange multipliers are applied to each of these constraints. The most important Lagrange multipliers are $\eta_{w,d}$ and $\mu_{w,d}$ which represents the marginal costs for each time period and the value of interest : the Water Value according to [30, part 3.1]. A similar approach to define those coefficients is also used by [32, part 3.1] but Water Value associated to the additional profit that would arise if an unit of water could be used for hydro generation. When thermal generation is replaced by the possibility of importing or exporting, as in our case study, the link between the two definitions is established. Indeed, if the electricity producer is in a situation where it is considering importing electricity, then its decision relates to cost-based water value as described of the cost by subsisting hydro generation by imported electricity which extends the definition given by [30, part 4.1]. On the other hand, if the electricity producer is in a situation where it is considering exporting electricity, then this Water Value is associated to the additional profit for producer that would arise if an unit of water could be used for hydro exportation generation, which is consistent with [32] definition.

The Lagrangian expression of optimisation problem 3.5.2 is given by:

$$\begin{aligned}
\mathcal{L} &= \mathcal{L}(T_{w,d}, H_{w,d}, R_{w,d}, \eta_{w,d}, \mu_{w,d}, \gamma^{down}_{w,d}, \gamma^{up}_{w,d}, \delta^{down}_{w,d}, \delta^{down}_{w,d}, \alpha_{w,d}, \beta_{w,d}) \\
&= \min_{T_{w,d},H_{w,d},R_{w,d}} \sum_{w=1}^{52}\sum_{d=1}^{7} C_{w,d}(T_{w,d}) + \eta_{w,d}(T_{w,d} + H_{w,d} - D_{w,d} \\
&+ \mu_{w,d}(R_{w,d} - R_{w,(d-1)} + H_{w,d} - I_{w,d}) - \gamma^{down}_{w,d} * H_{w,d} + \gamma^{up}_{w,d}(H_{w,d} - \overline{H_{w,d}}) \\
&+ \delta^{down}_{w,d}(\underline{R_{w,d}} - R_{w,d}) + \delta^{up}_{w,d}(R_{w,d} - \overline{R_{w,d}}) + +\zeta^{down}_{w,d}(\underline{T_{w,d}} - T_{w,d}) + \zeta^{up}_{w,d}(T_{w,d} - \overline{T_{w,d}}) \\
&+ \alpha_{w,d}(R_{1,0} - R_{in}) + \beta_{w,d}(R_{52,7} - R_{end}), \forall(w,d)
\end{aligned} \tag{3.5.3}$$

The optimal solution has to fulfill first order conditions :

$$\frac{\partial \mathcal{L}}{\partial T_{w,d}} = \frac{\partial C_{w,d}(T_{w,d})}{\partial T_{w,d}} + \eta_{w,d} - \zeta_{w,d}^{down} + \zeta_{w,d}^{up} = 0 \tag{3.5.4a}$$

$$\frac{\partial \mathcal{L}}{\partial H_{w,d}} = \eta_{w,d} + \mu_{w,d} - \gamma_{w,d}^{down} + \gamma_{w,d}^{up} = 0 \tag{3.5.4b}$$

$$\frac{\partial \mathcal{L}}{\partial R_{w,d}} = \mu_{w,d} - \delta_{w,d}^{down} + \delta_{w,d}^{up} = 0 \tag{3.5.4c}$$

$$\eta_{w,d}(T_{w,d} + H_{w,d} - D_{w,d} + B_{w,d}) = 0 \tag{3.5.4d}$$

$$\mu_{w,d}(R_{w,d} - R_{w,(d-1)} + H_{w,d} - I_{w,d}) = 0 \tag{3.5.4e}$$

$$\gamma_{w,d}^{down} * H_{w,d} = 0 \tag{3.5.4f}$$

$$\gamma_{w,d}^{up}(H_{w,d} - \overline{H_{w,d}}) = 0 \tag{3.5.4g}$$

$$\delta_{w,d}^{down}(\underline{R_{w,d}} - R_{w,d}) = 0 \tag{3.5.4h}$$

$$\delta_{w,d}^{up}(R_{w,d} - \overline{R_{w,d}}) = 0 \tag{3.5.4i}$$

$$\zeta_{w,d}^{down}(\underline{T_{w,d}} - T_{w,d}) = 0 \tag{3.5.4j}$$

$$\zeta_{w,d}^{up}(T_{w,d} - \overline{T_{w,d}}) = 0 \tag{3.5.4k}$$

$$\alpha_{w,d}(R_{1,0} - R_{in}) = 0 \tag{3.5.4l}$$

$$\beta_{w,d}(R_{52,7} - R_{end}) = 0 \tag{3.5.4m}$$

If we assume ∀(w,d), when the maximum and minimum hydro power generation constraints are not binding, that involves $\gamma_{w,d}^{up} =$ and $\gamma_{w,d}^{down} =$. The same reasoning is applied to import generation constraints that involve $\zeta_{w,d}^{down} = 0$ and $\zeta_{w,d}^{up} = 0$. Water value is finally given by $\eta_{w,d}$ because its embodies cost by subsisting hydro generation by imported electricity $\frac{\partial C_{w,d}(T_{w,d})}{\partial T_{w,d}}$. Then, the first condition of 3.5.4 gives us directly a formula for the water Value (WV):

$$\eta_{w,d} = -\frac{\partial C_{w,d}(T_{w,d})}{\partial T_{w,d}} \tag{3.5.5}$$

Hence, fixing a function for the costs $C_{w,d}$ which is differentiable ∀(w,d), gives us directly the Water value for each time.

For the marginal cost function, based on the work made by [32, Part 5.1], we can globally approximated marginal imported production costs by a linear function. This is true if the imported quantity does not exceed a limit of transmission.

In the following, we define $c_{w,d} = \frac{\partial C_{w,d}(T_{w,d})}{\partial T_{w,d}}$ as a function only dependant of $T_{w,d}$:

$$c_{w,d} = \frac{\partial C_{w,d}(T_{w,d})}{\partial T_{w,d}} = a_{th} + b_{th} * T_{w,d} \tag{3.5.6}$$

where a_{th} and b_{th} are constants.

3.5.3 Data

Data	Unit	Time Frame	Source
Norway spot prices NO1	NOK/MWh	01/2017-01/2018 (Hourly)	(Nord Pool Spot AS)
Water inflows	MWh	01/2017-01/2018 (Monthly)	(NVE)
Load	MW	01/2017-01/2018 (Hourly)	(Nord Pool Spot AS)
Generation per production Type	MW	01/2017-01/2018 (Hourly)	(Entso-e Transparency)
Water Reservoirs	MWh	01/2017-01/2018 (Weekly)	(Entso-e Transparency)

Table 3.3: Data used in simulation

Remark 3.5.1 (Data sources)**.** As usual, when data come from different sources, a reshape and reformatting work has to be done in order to match the right time frame. In our case, two main

sources are used for the model in 3.5.2, Entso-e transparency which is a main data base of different countries, technologies and information for European electricity market whose the quality has been certified in many references such as in [33] and NordPool Spot AS which is the leading power market in Europe, offering day-ahead and intraday markets and data base for Northern countries.

One of the main issues for reprocessing data is the time scale variations between different sources. Indeed, for electricity markets, most of data such as load, prices are given in hourly. However, more macroeconomic and statistic data are often found in daily or weekly scale. In addition, our model defined in 3.5.2 takes in input daily data. Thus, we have to choose a strategy to interpolate data for re-scaling them to the correct time scale.

Water Reservoirs are given in week steps. Since daily data are key to our analysis, we interpolate the missing values by a linear Lagrange interpolation.

For load quantity, since we need the daily load of each weeks of the all year, we have to choose on how approximate daily load with hourly data. To keep it simple, we just take the mean of each hour of each day.

Finally, Water inflows are quite difficult to predict and different approaches can be considered to approximate inflows. For example [34] using a probability viewpoint by estimating the distribution with a modified Maximum Likelihood estimation and Bayesian inference. In this paper, we try another approach only based on the curve of the inflows during the year. Then, we estimate the profile of quantity of MWh falls during the year by taking the generation of MWh by hydroelectricity during the entire year on the interest area.

For the simulation, $R_{in} = 3157000MWh$, $R_{end} = 3471000MWh$, and $\forall$(w,d), $\underline{T_{w,d}}$=-7800MW, $\overline{T_{w,d}}$=7800MW, $\overline{H_{w,d}} = 1365MW$ based on ENTSO-E NO1 data for Norway.

3.5.4 Results

Model 3.5.2 runs are performed using data from the year 2017 as explained in 3.5.3.

To test the explanation, we apply model 3.5.2 to Norway's NO1 zone, where a weekly curve showing the evolution of the hydraulic energy reserve is available on the Entso-e Transparency website [25]. Similarly, run-of-river production is estimated using the Entso-e transparency platform. The values used are estimated for the periods of each day of the year 2017 summarised in the following table 3.4:

Days	1-130	131-200	201-300	301-363
Daily production (MWh)	19200	28800	24000	21600

Table 3.4: Run-of-river generation (MWh) NO1 2017

These daily values for run-of-river generation are subtracted from the demand observed each day to obtain an adjusted demand corresponding to the netting demand defined part 3.3.3. The inflow profile used for modelling with demand adjusted to production by run-of-river technologies is presented in the appendix A.7. To certify the consistency of our results, we need a control variable that we can easily compare with the data. As the evolution of energy in the reservoirs is an accessible variable, it serves as a benchmark for the accuracy of our simulation based on optimisation problem 3.5.2.

The following graph 3.5.4 shows the result of the simulation for the control value with the data explained above.

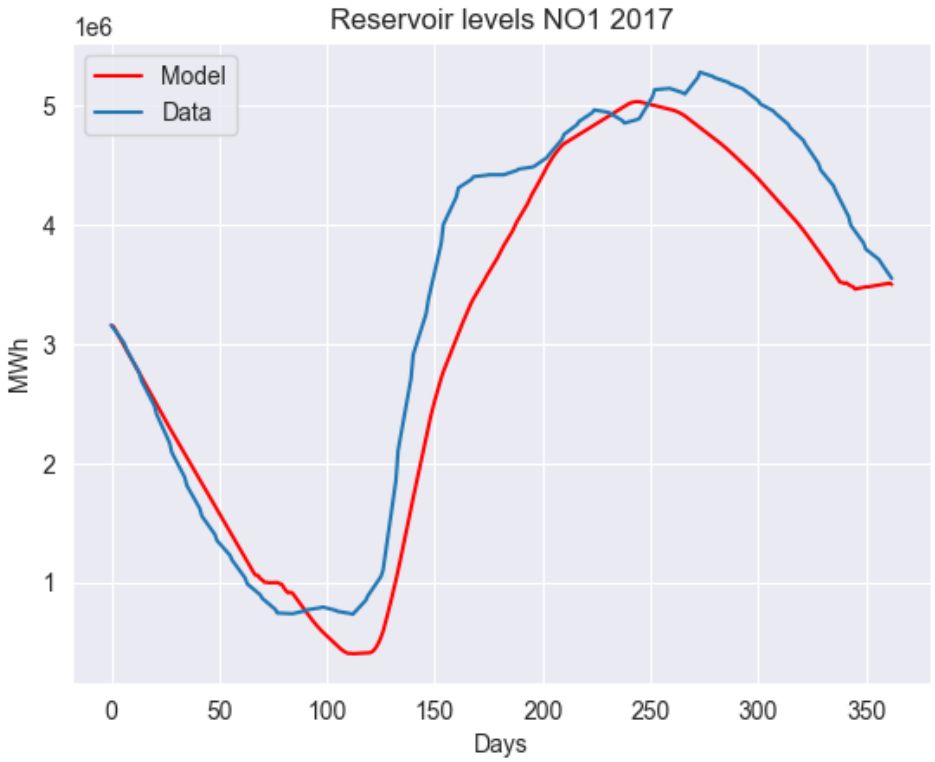
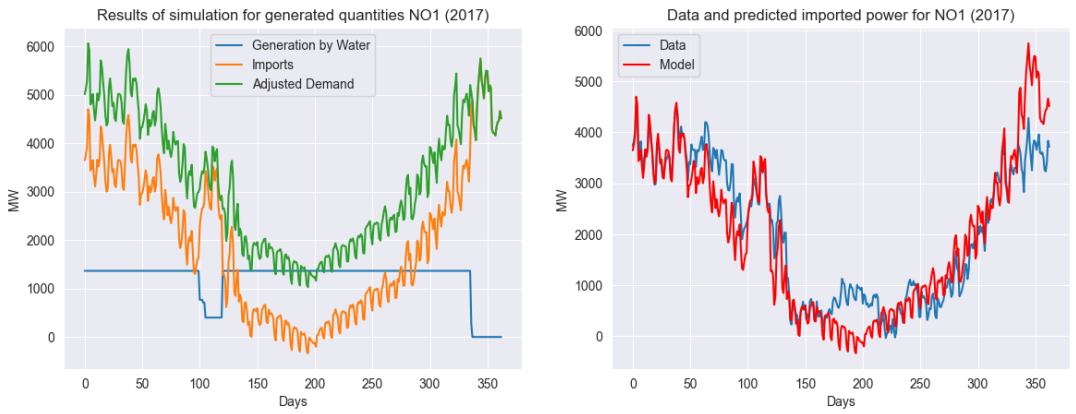


Figure 3.9: Data and predicted shape of level in reservoirs in NO1 (2017)

Remark 3.5.2. (Optimisation algorithm) The optimisation problem is calculated using Python language and the "scipy.optimise" library. The function minimize with method='SLSQP' enables minimizing a scalar function of one or more variables using Sequential Least Squares Programming.

The overall behaviour of the evolution of energy in the reservoirs follows the evolution of the data. Nevertheless, there are a few notable points of divergence, particularly at the very end of the year. In fact, the decrease in the reservoir in the simulation is faster and more abrupt than the data. The final level of the reservoir is reached before the end of the year. There are two reasons for this. The first is simply that the evolution of reservoirs depends on the inflow curve. The amount of water imported into the reservoir plays an important role and it is just estimated from historical hydraulic production quantity. As a result, it is difficult to predict the actual quantity of water available throughout the year. The second is the profile of these inflows over the year. In fact, just like the total quantity of water to be estimated, the inflow curve depends on a daily interpolation made from monthly data, which gives rise to numerous imprecisions and can therefore explain these discrepancies. Nevertheless, the model transcribes the general evolution of the reservoir level, which guarantees that our model is a good approximation to reality. The effects of filling and emptying are well reproduced at the corresponding times.

The model 3.5.2 has several degrees of freedom to find the optimum. It can play on the values of hydraulic production but also on the quantity imported or exported of electricity with the other zones. Finally, the level of the reservoirs is also linked to these production values. Thanks to NordPool [7], it is possible to access the quantity of imported or exported power (MW) each day in the NO1 zone. However, hydraulic production data are not available. Thus, another proof of good behaviour of our model is to compare the quantities of imported and exported power (MW) predicted by the model with the real data. The figures 3.11 present these comparisons. Remember that at each point in time, hydraulic generation and imported power (counted positively) or exported power (counted negatively) must be equal to demand adjusted by run-of-river generation.



(a) Generated consumption of each type in NO1 (2017) (b) Daily Day-Ahead spot prices and Water Value NO1 (2017)

Figure 3.10: Data and predicted power NO1 (2017)

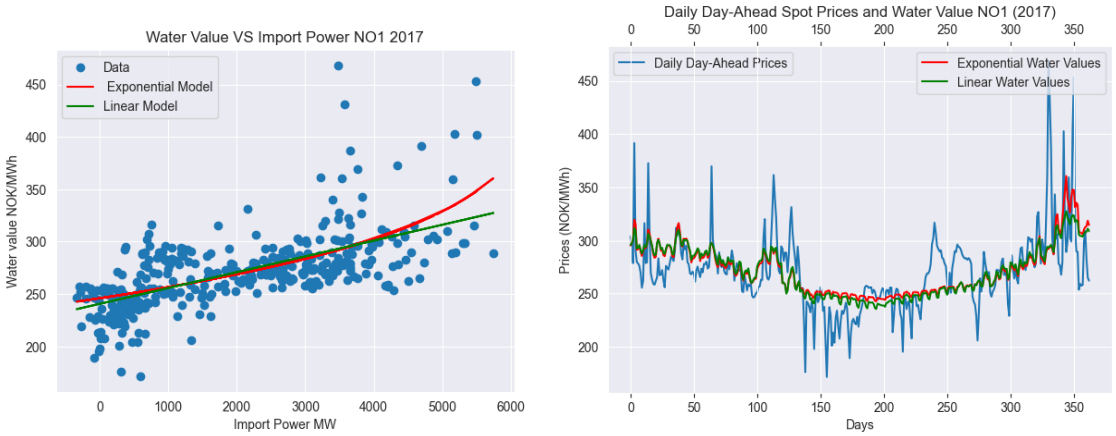
The graph shows changes in hydraulic generation and the quantity imported in 2017. It can be seen that the understanding of hydraulic production is fairly binary, i.e. the operation either favours maximum production from the reservoirs or turns it off. Over most of the year, production is at its maximum, except at two periods, around the hundredth day and from the three hundred and fortieth. The second period is due to the shutdown of the reservoirs, as the overall level has reached its final level. On the other hand, for the first period, it is more difficult to provide a justification for this shutdown behaviour, which is due to the inflow profile imposed as a hypothesis for our system. However, one reason for this is that the level of the reservoirs (MWh) has reached its low limit.

Since we have compute imported production, we can calculate the Water Value from observed market prices using the formula 3.5.6. After displaying the day-ahead price observed on the market against imported production, we note that the linear modelling seems to be improved by a change to exponential function for marginal imported costs. Based on work done part 3.4.1, we define the exponential function for Water Values as followed:

$$WV_t = \frac{\partial C_{w,d}(T_{w,d})}{\partial T_{w,d}} = \bar{P} - \exp(a)(\bar{S} - T_t)^{\alpha} \tag{3.5.7}$$

where $\bar{P} = 50000 NOK/MWh$, $\bar{S} = 6500 MW$.

Models 3.5.7 and 3.5.6 are calibrated by using linear regression on historical prices in NO1 in 2017. Results are shown in Figures 3.11.



(a) Water Value VS Imported Power NO1 (2017) (b) Daily Day-Ahead spot prices and Water Value NO1 (2017)

Figure 3.11: Daily Day-Ahead spot prices and Water Value NO1 (2017)

Models	R-score	MAE	RMSE
Exponential Water Values	0.48	19.61	28.22
Linear Water Values	0.56	19.55	28.15

Table 3.5: Results for both linear and exponential Water Values NO1 (MWH)

The results are quite surprising, as it seems that the linear model performs better than the exponential model. Indeed, the table 3.5 shows better results for the R-score and the MAE for linear model and similar result for RMSE. However, the difference between these two models is not marked, we have no model that outperforms another among these two.

The overall trend in prices is well respected. Winter periods are well represented, with higher volatility than during the summer. The linear model is better able to capture periods when prices are low, whereas the exponential model is better able to capture periods of high demand and therefore higher day-ahead prices.

However the daily comparison reveals that also both models are not fully capable to capture the jumps and troughs of the Norwegian electricity prices in 2017. We find the same results in the following paper [35, Part 5.2], which considers another approach to quantify Water Value. There are several reasons for this phenomenon. The first is that the jumps appear at specific times of the day, usually in the late afternoon, suddenly. Since,daily Day-Ahead prices and the demand are averaged for each day, we loose the effect of the increase in demand on an hourly basis. Secondly, the reader should remember that this model is made up of a single reservoir which covers all the production in the NO1 zone in 2017. This means that only one associated Water Value is defined. In reality, each reservoir has its own supply curve and therefore its own associated Water Value, which is then aggregated to build the supply curve. This extension of the Water Value calculation model to a competitive multi-agent environment can be found in the paper [30, Part 4.2]. Finally, this model predicts the best daily price that a reservoir operator can hope to obtain for his water reserve. The use of electricity production is then adapted to this value because it allows us to know at what time it will be interesting to produce in order to optimise the operation of the dam. During the day, demand varies from hour to hour, which leads to price movements over the day. So, with the Water Value defined for the day, the operator can choose the most interesting hours of production when prices are highest and import electricity when the price falls below the Water Value; his stock of water is saved for future hours when the price will be higher. This is mostly the case in summer, as can be seen in the figure 3.11(b). From day 140 to day 230, the price falls sharply and it becomes very rare to have production opportunities at Water Value prices. Hydro generation is therefore slowed down to be used when prices become attractive again.

Many constraints can be added to our model. We have only considered hydraulic constraints and network constraints at the interconnection capacity level. But environmental constraints are increasingly being imposed on electricity producers, leading to a change in production optimisation policy [36].

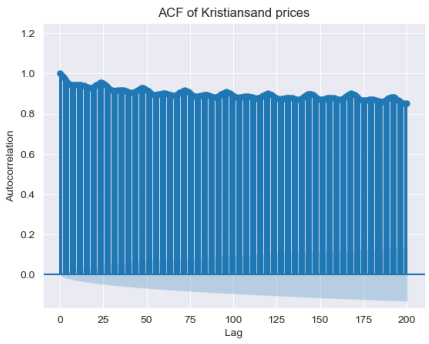
Conclusion

In this thesis, we have highlighted the non-normality of electricity market distributions and presented several studies and graphs on the correlation and auto-correlation of Day-Ahead prices for the five bidding zones in Norway. We have concluded that these five zones are divided into two groups because of their hydropower generation capacity and interconnection capacity. Finally, having identified a weekly trend in spot prices, adapted a deterministic function to transcribe the periodicity of the Day-Ahead price time series, an ARMA stochastic process was used to best transcribe the residuals. Despite the good performance of this approach on the Norwegian data, the internal dynamics of the market structure are not reflected in this model. Thus, a more detailed modelling of the supply curve and the redefinition of the Water Value for hydro-dominated zones with import and export capacity with adjacent zones was detailed to re-transcribe the market clearing price process. A stochastic modelling approach using an Ornstein-Uhlenbeck process is described for the demand curve. Thus, by imposing dynamic constraints on reservoir levels and on the possibility of importing or exporting, without forgetting the principle of equalising supply and demand at every moment, an optimisation problem led us to redefine Water Value and to recover the trend in Day-Ahead prices observed in Norway in 2017. The results show a good approximation of the general dynamics of the curve, despite an inability to capture price peaks. For future studies, models with a spike component to reflect these sudden dynamics of price increases or decreases could be considered.

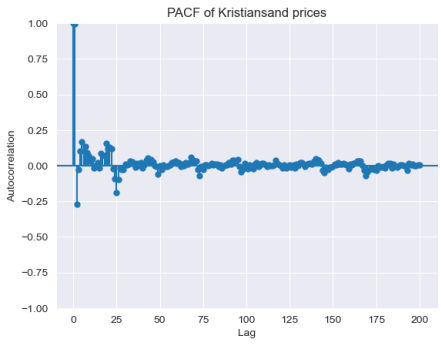
Appendix A

Additional graphs

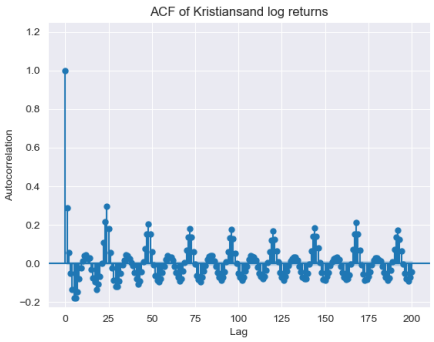
A.1 Additional graphs of the five bidding zones in Norway



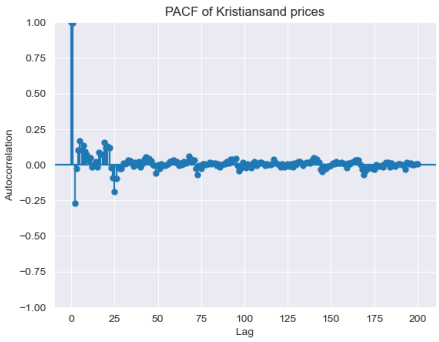
(a) ACF Kristiansand prices NO2 (2015-2022)



(b) PACF Kristiansand prices NO2 (2015-2022)

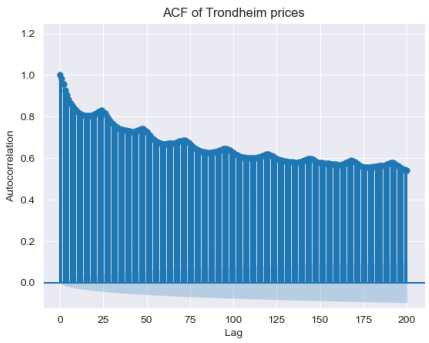


(c) ACF Kristiansand log returns NO2 (2015-2022)

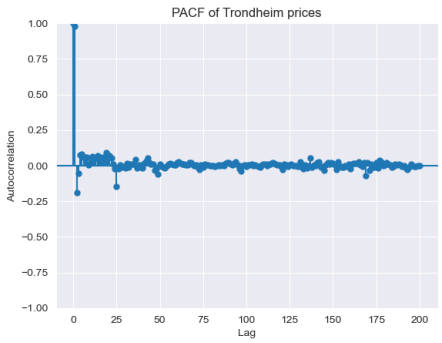


(d) PACF Kristiansand log returns NO2 (2015-2022)

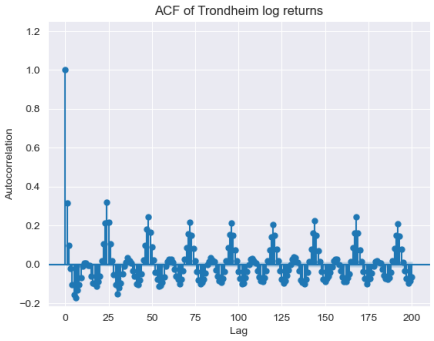
Figure A.1: ACF and PACF Day-Ahead prices and returns Kristiansand NO2 (2015-2022)



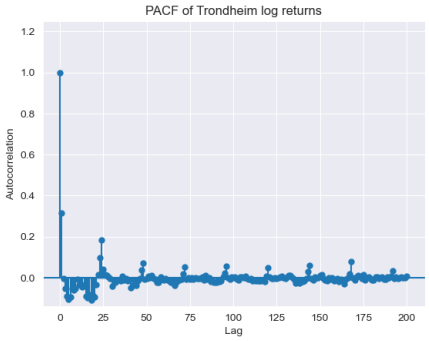
(a) ACF Trondheim prices NO3 (2015-2022)



(b) PACF Trondheim prices NO3 (2015-2022)

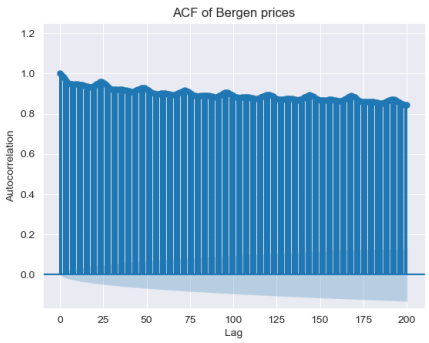


(c) ACF Trondheim log returns NO3 (2015-2022)

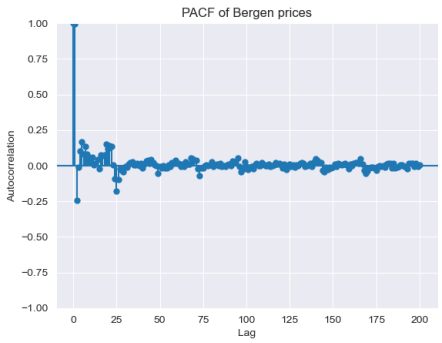


(d) PACF Trondheim log returns NO3 (2015-2022)

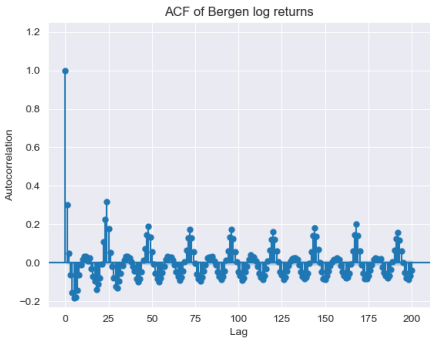
Figure A.2: ACF and PACF Day-Ahead prices and returns Trondheim NO3 (2015-2022)



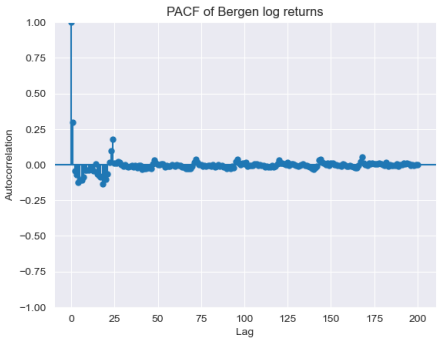
(a) ACF Bergen prices NO5 (2015-2022)



(b) PACF Bergen prices NO5 (2015-2022)



(c) ACF Bergen log returns NO5 (2015-2022)



(d) PACF Bergen log returns NO5 (2015-2022)

Figure A.3: ACF and PACF Day-Ahead prices and returns Bergen NO5 (2015-2022)

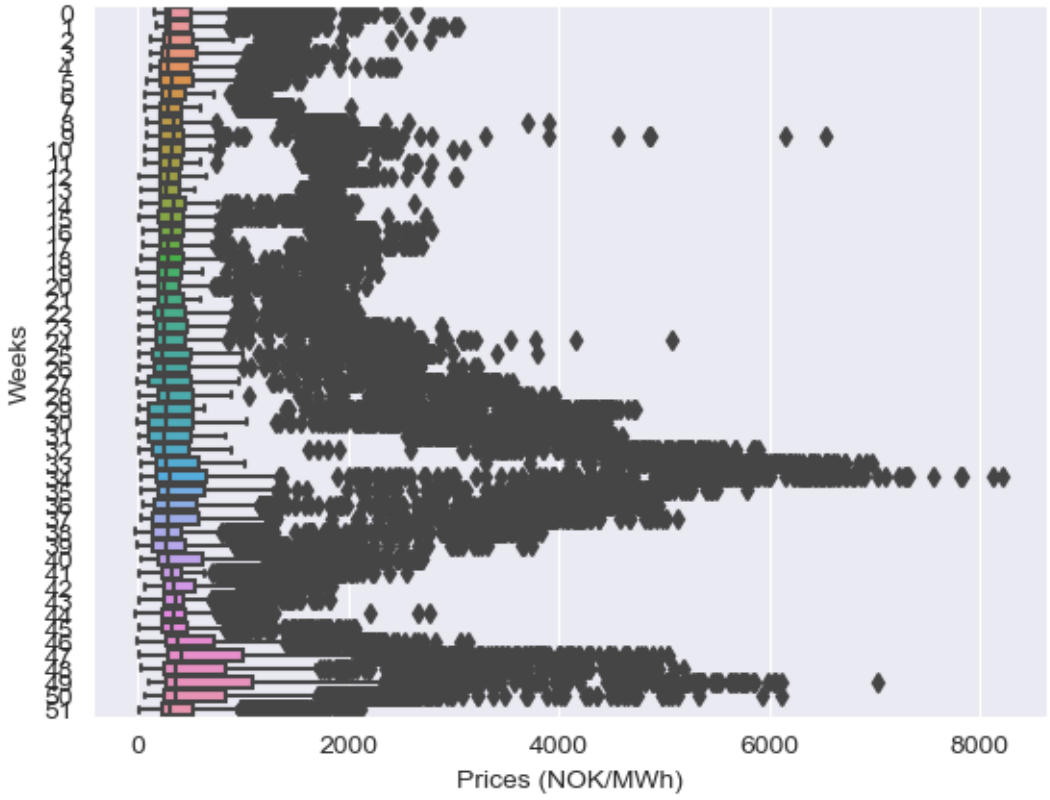


Figure A.4: Box plot Oslo NO2 (2017-2022)

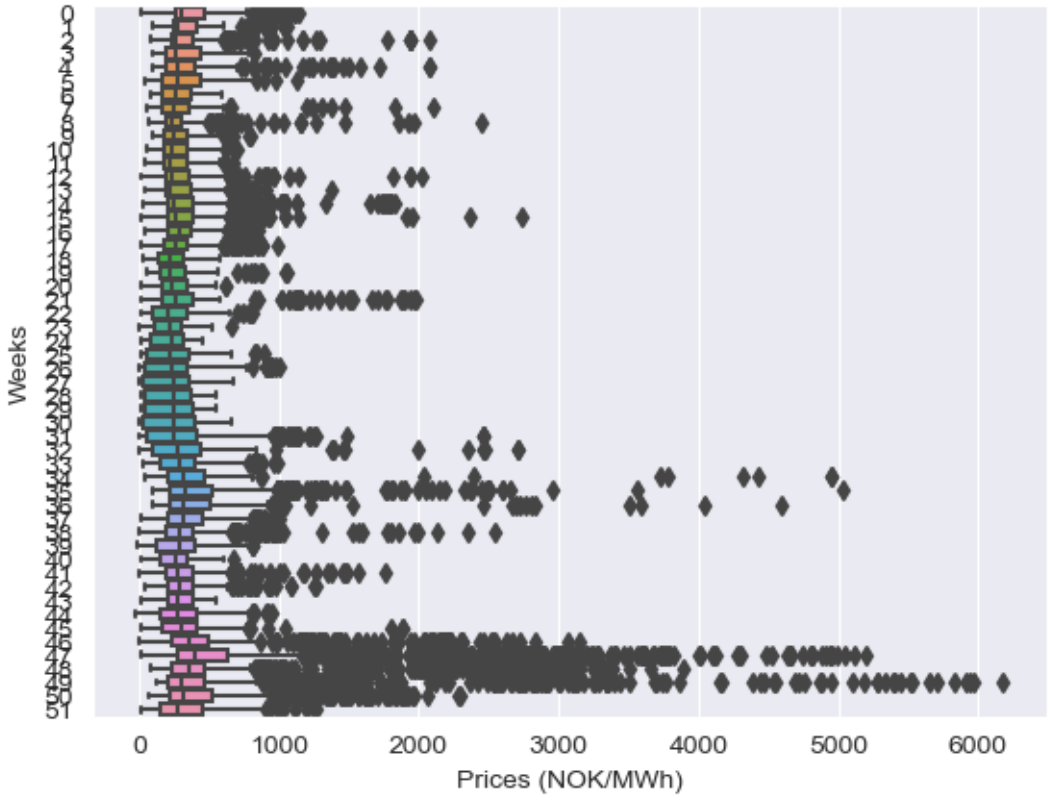


Figure A.5: Box plot Trondheim NO3 (2017-2022)

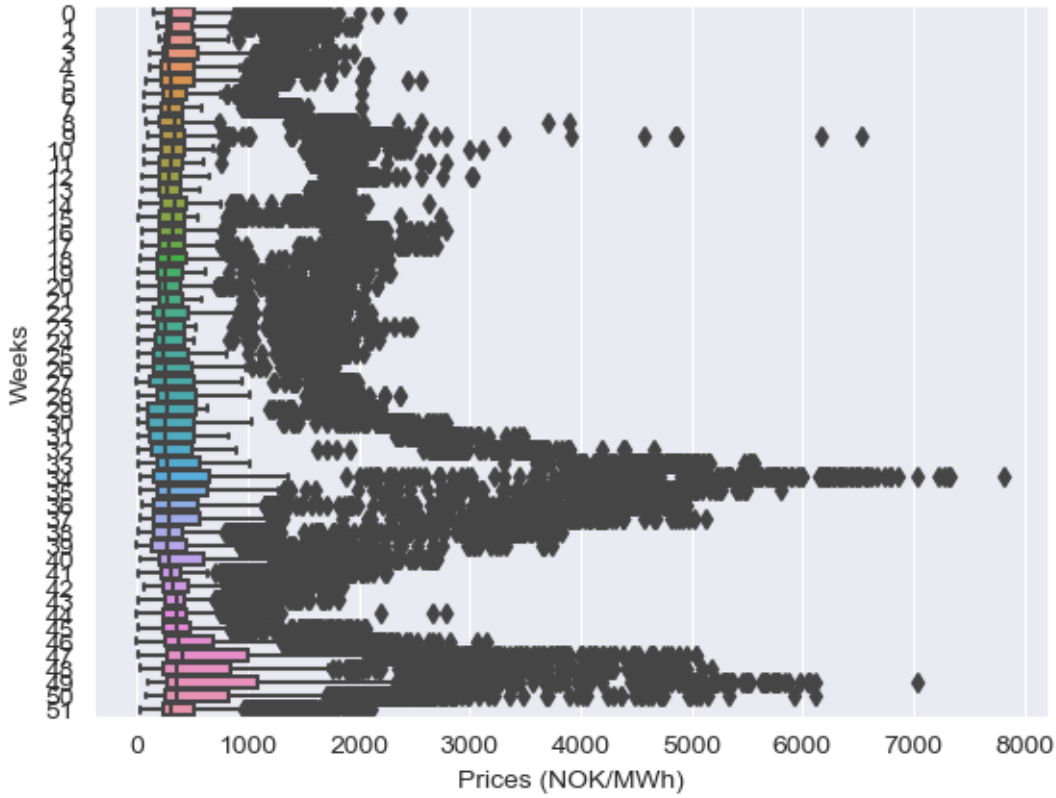


Figure A.6: Box plot Bergen NO5 (2017-2022)

A.2 Market clearing graphs

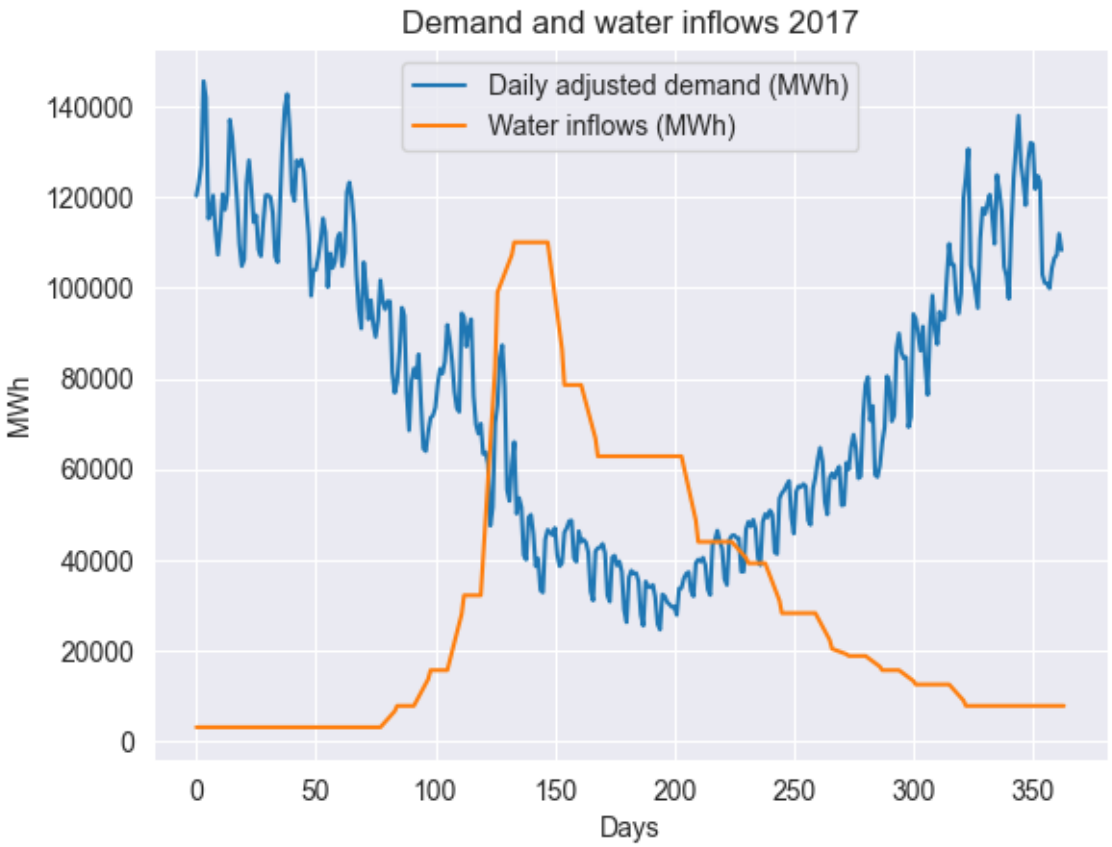


Figure A.7: Water inflows and Daily adjusted demand (MWh) NO1 (2017)

Bibliography

[1] John Clauß, S Stinner, C Solli, Karen B Lindberg, Henrik Madsen, and Laurent Georges. A generic methodology to evaluate hourly average co2eq. intensities of the electricity mix to deploy the energy flexibility potential of norwegian buildings. In *Proceedings of the 10th International Conference on System Simulation in Buildings, Liege, Belgium*, pages 10–12, 2018.

[2] Maria Sandsmark and Berit Tennbakk. Ex post monitoring of market power in hydro dominated electricity markets. *Energy Policy*, 38(3):1500–1509, 2010.

[3] Weibiao Qiao and Zhe Yang. Forecast the electricity price of us using a wavelet transform-based hybrid model. *Energy*, 193:116704, 2020.

[4] Niamat Ullah Ibne Hossain, Raed Jaradat, Seyedmohsen Hosseini, Mohammad Marufuzzaman, and Randy K Buchanan. A framework for modeling and assessing system resilience using a bayesian network: A case study of an interdependent electrical infrastructure system. *International Journal of Critical Infrastructure Protection*, 25:62–83, 2019.

[5] Stefano Zedda, Gb Masala, et al. Price spikes in the electricity markets how and why. In *Energy Transition: European and Global Perspectives*. Hellenic Association for Energy Economics (HAEE), 2018.

[6] Ignacio Guisández and Juan Ignacio Pérez-Díaz. Evaluating approaches for estimating the water value of a hydropower plant in the day-ahead electricity market. In *The International Workshop on Hydro Scheduling in Competitive Markets*, pages 8–15. Springer, 2018.

[7] Nord pool group website, 2023.

[8] Peter Cramton. Electricity market design. *Oxford Review of Economic Policy*, 33(4):589–612, 2017.

[9] Paul L Joskow and Jean Tirole. Transmission rights and market power on electric power networks. *The Rand Journal of Economics*, pages 450–487, 2000.

[10] René Aïd. *Electricity derivatives*. Springer, 2015.

[11] Udi Helman, Benjamin F Hobbs, and RICHARD P O'NEILL. The design of us wholesale energy and ancillary service auction markets: Theory and practice. In *Competitive electricity markets*, pages 179–243. Elsevier, 2008.

[12] Priyanka Shinde and Mikael Amelin. A literature review of intraday electricity markets and prices. *2019 IEEE Milan PowerTech*, pages 1–6, 2019.

[13] Eirik Ogner Jåstad, Ian M Trotter, and Torjus Folsland Bolkesjø. Long term power prices and renewable energy market values in norway–a probabilistic approach. *Energy Economics*, 112:106182, 2022.

[14] Torstein Bye and Einar Hope. Deregulation of electricity markets: the norwegian experience. *Economic and Political Weekly*, pages 5269–5278, 2005.

[15] Michele Costa, Giuseppe Cavaliere, and Stefano Iezzi. The role of the normal distribution in financial markets. In *New Developments in Classification and Data Analysis: Proceedings of the Meeting of the Classification and Data Analysis Group (CLADAG) of the Italian Statistical Society, University of Bologna, September 22–24, 2003*, pages 343–350. Springer, 2005.

[16] Leo Quigley and David Ramsey. Statistical analysis of the log returns of financial assets. *BSc in Financial Mathematics. University of Muenster, Germany*, 2008.

[17] Ângela Paula Ferreira, Jenice Gonçalves Ramos, and Paula Odete Fernandes. A linear regression pattern for electricity price forecasting in the iberian electricity market. *Revista Facultad de Ingeniería Universidad de Antioquia*, (93):117–127, 2019.

[18] John Semmlow. *Signals and systems for bioengineers: a MATLAB-based introduction*. Academic press, 2011.

[19] Dominik Möst and Dogan Keles. A survey of stochastic modelling approaches for liberalised electricity markets. *European Journal of Operational Research*, 207(2):543–556, 2010.

[20] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[21] Juan M Morales, Antonio J Conejo, Henrik Madsen, Pierre Pinson, Marco Zugno, Juan M Morales, Antonio J Conejo, Henrik Madsen, Pierre Pinson, and Marco Zugno. Clearing the day-ahead market with a high penetration of stochastic production. *Integrating Renewables in Electricity Markets: Operational Problems*, pages 57–100, 2014.

[22] William W Hogan. Multiple market-clearing prices, electricity market design and price manipulation. *The Electricity Journal*, 25(4):18–32, 2012.

[23] Aitor Ciarreta, Blanca Martinez, and Shahriyar Nasirov. Forecasting electricity prices using bid data. *International Journal of Forecasting*, 39(3):1253–1271, 2023.

[24] Marcus Olofsson, Thomas Önskog, and Niklas LP Lundström. Management strategies for run-of-river hydropower plants: an optimal switching approach. *Optimization and Engineering*, pages 1–25, 2021.

[25] Entso-e transparency platform. `https://transparency.entsoe.eu`. Accessed: 2023-08-31.

[26] Martin T Barlow. A diffusion model for electricity prices. *Mathematical finance*, 12(4):287–298, 2002.

[27] Manuela Buzoianu, Anthony Brockwell, and Duane Seppi. A dynamic supply-demand model for electricity prices. 2005.

[28] Ronald Huisman, David Michels, and Sjur Westgaard. Hydro reservoir levels and power price dynamics: Empirical insight on the nonlinear influence of fuel and emission cost on nord pool day-ahead electricity prices. *J. Energy & Dev.*, 40:149, 2014.

[29] Lennart Söder. Analysis of electricity markets, 2011.

[30] Javier Reneses, Julian Barquín, Javier García-González, and Efraim Centeno. Water value in electricity markets. *International Transactions on Electrical Energy Systems*, 26(3):655–670, 2016.

[31] Elis Nycander and Lennart Söder. Modelling prices in hydro dominated electricity markets. In *2022 18th International Conference on the European Energy Market (EEM)*, pages 1–6. IEEE, 2022.

[32] James Busby Bushnell. *Water and power: Hydroelectric resources in the era of competition in the western US*. Program on Workable Energy Regulation, 1998.

[33] Lion Hirth, Jonathan Mühlenpfordt, and Marisa Bulkeley. The entso-e transparency platform–a review of europe's most ambitious electricity data platform. *Applied energy*, 225:1054–1067, 2018.

[34] Julia Lutz, Lars Grinde, and Anita Verpe Dyrrdal. Estimating rainfall design values for the city of oslo, norway—comparison of methods and quantification of uncertainty. *Water*, 12(6):1735, 2020.

[35] Christopher Jahns, Caroline Podewski, and Christoph Weber. Supply curves for hydro reservoirs–estimation and usage in large-scale electricity market models. *Energy Economics*, 87:104696, 2020.

[36] Linn Emelie Schäffer, Magnus Korpås, and Tor Haakon Bakken. Implications of environmental constraints in hydropower scheduling for a power system with limited grid and reserve capacity. 2023.