Intervention analysis of electricity markets

An analysis using the SARIMAX regression model of day-ahead and balancing markets in the Netherlands

Master thesis

Folkert Post

Delft University of Technology

Intervention analysis of electricity markets

An analysis using the SARIMAX regression model of day-ahead and balancing markets in the Netherlands

by

Folkert Post

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on October 2nd 2024, 16:00.

Student number:	5639891	
Thesis committee:	Prof. dr. ir. L. de Vries,	TU Delft
	Dr. ir. Ö. Okur,	TU Delft
	Dr. Dipl.-Ing. D.G. Ribó Pérez	Universitat Politècnica de València
	Dr. ir. H. van der Weijde	TNO

Cover: Image by freepik.com
Style: TU Delft Report Style, with modifications by Daan Zwaneveld

Preface

Dear reader,

In front of you is my final project I write as a student of Engineering and Policy Analysis (EPA) at TU Delft. EPA has allowed me to discover a world of interesting and complex problems and has given me a wide variety of tools to analyse, reflect on and tackle these problems. During my first year at EPA, just when exiting the final lockdown of the COVID-19 pandemic, Russia invaded Ukraine, causing energy prices to skyrocket in the Netherlands and the rest of Europe, showing the interconnectedness of our modern-day world. Personally, this got me interested in the topic of energy, markets and the future of our energy system which led me to pursue an internship at TNO. This, I found to be a very welcoming place, and here I got to meet lots of kind, interesting and very smart people who I got to be colleagues with for half a year.

This thesis has allowed me to combine many interests, particularly in modelling, data analysis and analysing complex systems. From problems with data sets, wrestling with the methodology and finally interpreting the meaning of all the results, this thesis process was full of challenges. Coincidentally, when researching ARIMA methods, I found out that George Box, after whom the Box-Jenkins approach is named, was also the one who coined the phrase: ”All models are wrong but some are useful.”. After having struggled endlessly to find the ”correct” model and sometimes forgetting about its usefulness and real-life applicability, I now agree with this maxim more than ever before.

Finalizing this master’s thesis also means that I will leave my time as a student at university behind me, a time I have greatly enjoyed. This has also meant that during the final stretch of my thesis I have had to figure out what goals to pursue outside university and who I wanted to be after this significant time in my life. This has meant that writing this thesis has taken longer than I initially expected but I’m glad for all the valuable lessons I learned, also during these last months. I’m very grateful to my supervisors, friends and family for their patience and for all the discussions we had about research, goals, ambitions and life in general.

Folkert Post
Delft, September 2024

Summary

Currently, a lot is changing on the electricity markets. New regulations and changing market conditions have put a renewed focus on these markets. The main problem looked at in this thesis is understanding how interventions quantitatively influence electricity markets. While some research has touched on this, there hasn't been much focus on explaining these effects through specific interventions or in the context of the Dutch market. This study focuses on two key electricity markets in the Netherlands: the Day-Ahead market and the Balancing market. We aimed to conduct an intervention study on these markets to better understand the impacts of various interventions.

To perform an intervention study, it's necessary to compare real-world situation with a counterfactual, essentially, what the market conditions would have been if the intervention hadn't occurred. To predict these counterfactuals, we developed models for both the DA and Balancing markets using the SARIMAX (Seasonal AutoRegressive Integrated Moving Average with eXogenous factors) method.

The SARIMAX model performed decently for the DA market but poorly for the balancing market. The main reason for this is the inherent unpredictability of the balancing market, which exists to address unpredictability in electricity supply and demand. Additionally, the local nature of the balancing market and the completeness of the dataset contributed to the difficulties in modelling it accurately. Balancing markets are inherently less predictable because they are the tool used to solve for the unpredictability itself. If their demand were completely predictable, they would not be necessary. However, predictability is not the same as explainability. Given the data we obtained, the models were able to explain, to a limited degree, a basic component of the balancing market behaviour, imbalances.

Given the need for a well-performing model to conduct intervention analysis, we used the DA model to study three interventions. Initial tests indicated promising outcomes for some interventions, showing significant impacts on DA prices. However, it must be noted that sensitivity analyses of the input data produced widely different effects of the interventions.

Firstly, SARIMAX modelling is fully dependent on the input data used for training and is essentially a black box model. We do not model the system described by the data but only the relationship between the dependent and independent variables directly. As a result, the model does not account for the different inputs or their relationships in the real world. SARIMAX models operate under the assumption that the system being modeled remains stable. However, some of the interventions tested in this study changed how the underlying system works, therefore making the SARIMAX model unsuitable in these cases. Thus, it is advised to use this modelling method only if the underlying process of the model, in this case, price formation in electricity markets, can be assumed to be stable.

In conclusion, while SARIMAX models can be useful for certain stable conditions, they might not be the best choice for analyzing systems where the underlying processes are significantly changing, such as in the case of market interventions on balancing markets.

Contents

List of Figures

List of Tables

List of abbreviations

Definitions

- **ANN**: Artificial Neural Network
- **ARIMA**: AutoRegressive Integrated Moving Average
- **aFRR**: Automatic Frequency Restoration Reserve
- **BRP**: Balance Responsible Party
- **BSP**: Balancing Service Provider
- **DA**: Day-Ahead
- **DSO**: Distribution System Operator
- **ENTSO-E**: European Network of Transmission System Operators for Electricity
- **FCR**: Frequency Containment Reserve
- **ITS**: Interrupted Time Series
- **mFRR**: Manual Frequency Restoration Reserve
- **OLS**: Ordinary Least Squares
- **RR**: Restoration Reserve
- **SARIMAX**: Seasonal AutoRegressive Integrated Moving Average with eXogenous regressors
- **SVR**: Support Vector Regression
- **TSO**: Transmission System Operator
- **VRES**: Variable Renewable Energy Sources

1 Problem Introduction

1.1. Problem Statement

Currently, electricity markets are shifting. Increasing deployment of renewables, the effects of the energy crisis in Europe and further electrification have put a large emphasis on the electricity grid and the price of electricity in our society (IEA, 2023) (CBS, 2023). This has further increased the importance of a reliable, affordable, and secure electricity grid (Eupean Parliament, 2023). Furthermore, in Europe, the markets for this electricity have undergone large transformations and have become more liberalized and interconnected in recent years.

While increasing variable renewable sources (VRES) of energy, such as wind and solar, helps reach climate goals, several new issues can occur during such a transition. The first is that this electricity is produced intermittently, which depends on weather conditions. Because of this, a lessening of traditional fossil generation has created additional challenges in regard to balancing the production and consumption of electricity as large-scale electricity storage is lacking. Additionally, traditional generation was very controllable which allowed for quick responses to grid instabilities, variable renewables currently do not match the balancing capabilities of traditional generation. This poses an issue as fossil generation is planned to be phased out. Therefore, large-scale integration of renewables could cause the grid to be more difficult to keep in balance (Hu, Harmsen, Crijns-Graus, Worrell, & van den Broek, 2018). This could mean that there will be an increase in the need for balancing capacity while the most common type of this capacity will be phased out. In the meantime, the underlying system determining the daily electricity price is shifting to more expensive fossil fuels and, at the same time, cheaper renewables. Figure 1.1 shows how VRES has increased in the Netherlands over the last years.

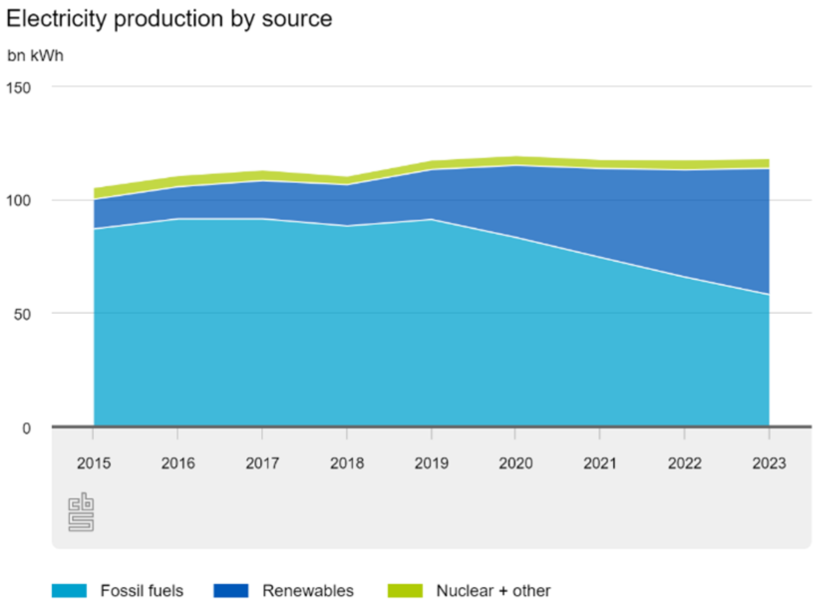

Figure 1.1: Electricity Generation Mix Netherlands (CBS, 2024)

Additionally, the last decades of the 20th century have seen European electricity markets become more established and integrated. During this time, integrating the electricity grids has brought issues that other markets have not. Due to electricity being very costly to store on a large scale, production and consumption have to match perfectly. Additionally, most electricity grids had been state-owned and centrally controlled, leaving little room for competition (Verbong & van der Vleuten, 2004) and making integration more difficult. A well-connected common market with fewer trade barriers would allow for more competition, lower prices, and higher reliability (Energy Regulatory Office, 2024). The transition from a state monopoly to a more liberalised market is illustrated in figure 1.2. To reach this, policies were made in the 1980s by the European Commission and are being implemented to this day (Meeus, 2020).

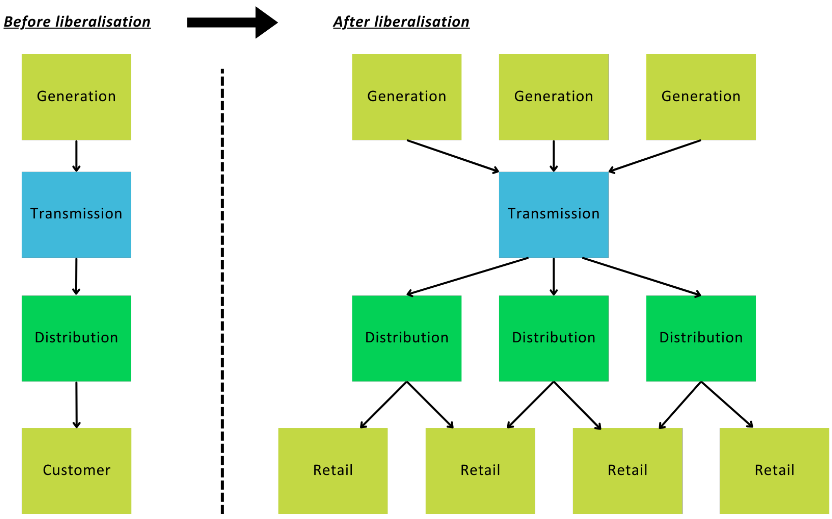

Figure 1.2: Overview of electricity generation and transmission (Pepermans, 2019)

The policies that stem from this approach are implemented in a step-wise manner. The whole system does not change at once; small steps are taken along the way. Each time a policy is implemented in markets, it has an expected effect on the market. Gaining an overview of large events, such as a new law being enacted or another large-scale event, can help understand how a complex system such as an electricity market behaves. This knowledge might prove interesting when deciding on future policies and aid in understanding the driving factors of the markets.

Therefore, studying the effects of market integration policies and large-scale events on electricity markets is the starting point for this research. We will take a high-level view of electricity markets and focus mainly on the Dutch Day-Ahead (DA) and balancing market. The DA market underwent large systemic changes, creating new scenarios with negative prices, large price fluctuations during the day, and an overall increase in price. Balancing markets are traditionally largely reliant on large generation units that provide inertia to the grid. With the trend of decreasing these large spinning reserves and the ever-more fluctuating supply and demand, it is expected that large changes in balancing market behaviour can be observed. These policy enactments and large-scale events that are assumed to have an impact on the system outcomes will, from now on, be referred to as interventions.

1.2. Research Objective

This research aims to study the effects of interventions on the Dutch electricity markets in the context of European integration and renewable production. The goal is to, with the help of statistical methods, analyze the Dutch DA and balancing markets during recent years. For this, the period of 2018 to 2023 has been chosen. This period contains significant changes to market design (TenneT, n.d.) as well as a large integration of renewables into the grid.

1.3. Research Gap

As stated, the goal of this research is to evaluate the impact interventions in electricity markets quantitatively have had in the Netherlands. This analysis will be done in a quantitative way where we look for a method that allows us to assess the impact of these interventions on the market.

As mentioned, the prices and volumes of electricity products have and are expected to change significantly in the coming years. The goal is to study if an effect on the prices and quantities of Dutch electricity market products can be observed due to the introduction of a market reform or other large scale event, therefore helping in the understanding of the effect of such interventions. A better understanding of the impact of these interventions could potentially help policymakers in the future when creating new policies for market reform.

1.4. Research Questions

Main research question: What has been the quantitative impact of market interventions on electricity markets in the Netherlands?”

From the main research question, it becomes clear that we need a modelling approach that allows us to consider external factors and account for market interventions over a long time frame.

To answer the main research question, the following sub-questions are made:

Sub-question 1: What have been the main developments in the Dutch day-ahead and balancing markets, and what interventions can be identified?

To measure the effects of market interventions, we need to first have an overview of relevant interventions that have taken place between 2018 and 2023. Answering this sub-question will give an overview and a sense of the direction the market is moving towards. The sub-question will be answered mainly based on (grey)literature research.

Sub-question 2: How can electricity market outcomes be modelled and explained?

Answering this sub-question will give insight into methods of modelling electricity markets and evaluation criteria for such models. It will be answered by comparing several modelling methods and applying one method to several Dutch markets.

Sub-question 3: How can impacts of market interventions and events be modelled and evaluated in electricity markets?

Having selected a method that allows for modelling electricity markets, we will use the model to study the effects of large-scale interventions such as increased collaboration and other large-scale events, thus answering this sub-question.

Sub-question 4: What are the policy implications of the outcomes of this analysis?

This question aims to identify if an intervention has had a quantifiable effect on balance markets and will be answered using the ITSA framework presented in the methodology. Using the modelling and the ITSA framework, the effects of the interventions might be seen.

These questions form the basis for the rest of this thesis and will be referred to in the following chapters.

1.5. Report Structure

This report will start with a more in-depth background chapter, introducing key concepts and developments. Then, an overview of the methodology will be given together with a proposed method of evaluation of market interventions for electricity markets in general. Afterwards, a chapter is dedicated to data gathering and analysis. Thereafter, a chapter is dedicated to modelling the electricity markets under study. Then, with the suitable models that were found, a chapter on intervention analysis shows the method of evaluation of large-scale market events. The thesis will be concluded by reflecting on the modelling process and intervention analysis. Finally, concluding remarks will be drawn, and further research will be proposed.

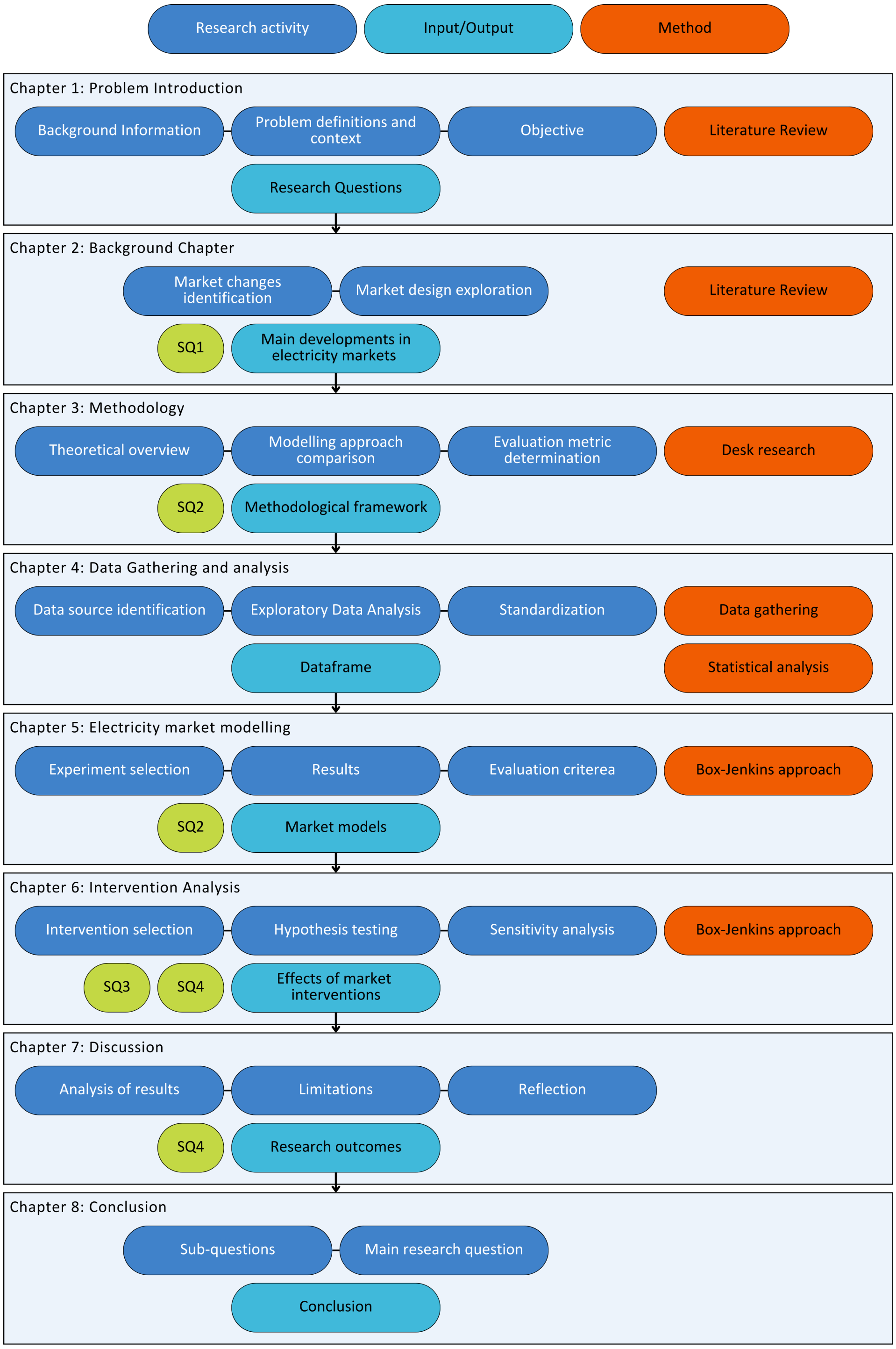

Figure 1.3: Research flow diagram

2
Background Chapter

This chapter will help set further context for the research and answer sub-question one by explaining the main developments on the DA and balancing markets. This section aims to explain the workings of electricity markets and balancing markets in particular, and the effects of renewable integration and market reform. Using recent (grey) literature and data, the section aims to explain this complex system clearly. Finally, this section will cover some large-scale events or market interventions that might be of interest.

2.1. Electricity Market Fundamentals and Day-Ahead Market

Between the production and the consumption of electricity is a large network of players, infrastructure and regulations that all have a particular role to fulfill. We will start by giving a small overview of the electricity network. Physically, Transmission System Operators (TSOs) control high-voltage transmission and balancing and Distribution System Operators (DSOs) manage the distribution to end users. National and international regulators determine the way in which all the players can act. Market participants can include large scale electricity generators, industries, and other consumers.

Historically, electricity systems were mostly a national or even regional affair. Markets for electricity have not always been there especially not on an international scale. Production and consumption was mostly done on a local level. There has always been trade in electricity even before markets were introduced in countries or coupled between countries but this was mainly done by large individual contracts between TSOs from different countries and was not as prevalent as it is nowadays.

Since the 1990s, several regulatory frameworks have been implemented to work towards a common market for electricity in Europe. These packages included many regulations to streamline national electricity regulation and created new institutions such as the European Network of Transmission System Operators for Electricity (ENTSO-E) and the Agency for the Cooperation of Energy Regulators (ACER). These regulations included several network codes that form the basis for market rules. With these rules, European countries are currently working towards a common market for electricity with the aim of creating a more efficient and reliable system.

Power exchanges are where the market parties come together. They allow consumers and suppliers to trade electricity, for example, on day-ahead and intraday markets. The day-ahead market (DAM) works with bids and an algorithm to optimize the prices everywhere in Europe in the Single Day-Ahead Coupling (SDAC) for all the bidding zones.

Figure 2.1: Single Day-Ahead Coupling

This integration also occurred for balancing reserves in balancing markets. Since Balancing Responsible Parties (BRPs) do not always have their portfolios perfectly balanced. Under new regulations, TSOs have the role of balancing the grid. As such, TSOs procure balancing resources in balancing markets as they normally cannot own their own resources. Previously, most of these resources were procured nationally, but this is also shifting more and more towards an international market.

Most markets work with a merit order. Market participants can place bids, which are then combined and optimized for constraints from which a merit order is created. The merit order works in a way that prioritizes the dispatch of the cheapest supplier first. The demand drives dispatch, which determines the price. An example of a merit order for a day-ahead market is given in Figure 2.2.

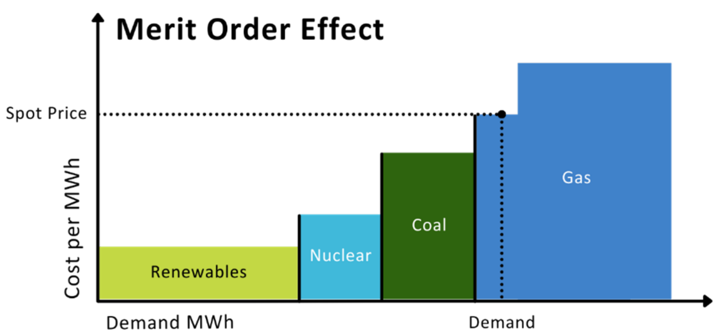

Figure 2.2: Merit Order effect

While this research focuses on the DA and balancing market, they are not the only markets where electricity gets traded. These electricity markets operate on several different timeframes to ensure a reliable balance between supply and demand. These timeframes include:

- **Long-Term Markets**: These markets involve contracts spanning several years. They allow producers and consumers to hedge against future price volatility and secure long-term supply.
- **Day-Ahead Markets**: Here, electricity trading occurs one day before the actual delivery. Market participants submit their bids and offers, and the market clears based on these, establishing prices for the following day.
- **Intraday Markets**: These markets enable the trading of electricity within the same day of delivery. They provide flexibility for participants to make adjustments due to unexpected changes in demand or generation availability.
- **Real-Time Markets**: Also known as balancing markets to address immediate imbalances between electricity supply and demand. TSOs manage these markets to ensure the grid remains stable and reliable, activating reserves as needed.

The time between demand and supply mostly determines the type of contract that is used, and thus, this falls on a continuum, as shown in Figure 2.3.

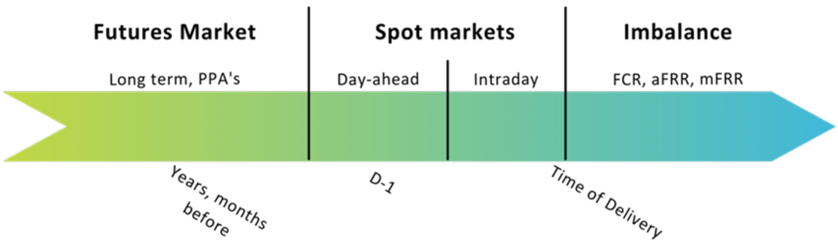

Figure 2.3: Electricity Markets Timeline (Moitroux, 2020)

2.2. Balancing markets fundamentals

In order to better understand the implications of the research, we continue by giving a more detailed explanation of balancing and balancing markets.

2.2.1. Solving Imbalances

Imbalances in the electricity grid happen all the time. The five main drivers of system imbalances are unplanned outages of interconnectors, forecasting errors for VRE generation and load, scheduling leaps and unplanned outages for thermal or hydro plants (**hirth_balancing_2015**). Imbalances can be solved automatically by Balance Responsible Parties (BRPs). BRPs are market parties who are responsible for the balance of their portfolio of generation and consumption in the electricity market. To use the electricity grid, a BRP needs to either solve imbalances within its own portfolio or trade away the imbalance in the markets mentioned above. BRPs can be producers, large-scale consumers, energy suppliers or traders and need to be registered with the TSO (“Balansverantwoordelijken (BRPs)”, n.d.). There are several reasons why a BRP cannot balance its portfolio, leading to a system imbalance. Once the TSO registers a system imbalance, it will dispatch balancing capacity from a Balancing Service Provider (BSP) to restore the balance.

As TSOs are not allowed to own energy-generating or consuming infrastructure, they will use BSPs to do the actual balancing. These BSPs must also register with the TSO and will, based on the type of reserve offered, receive an automated signal to power up or power down equipment from TenneT. To maintain frequency in the Dutch grid, TenneT needs to balance the grid by either up-regulating or down-regulating generation so the frequency stays within the 50 +/- 0,2 Hz range. The severity of the deviation will determine the actions a TSO takes and how the balance is restored. When the balance gets lost in the grid, there are several steps and reserves that are used:

1. Frequency Containment Reserves (FCR)
2. Frequency Restoration Reserves (aFRR)
3. Manual Frequency Restoration Reserves (mFRR)
4. Restoration Reserve (RR)

Depending on the severity of the imbalance and its duration, the different reserves are activated. A detailed overview of the design parameters of these reserves can be found in the appendix.

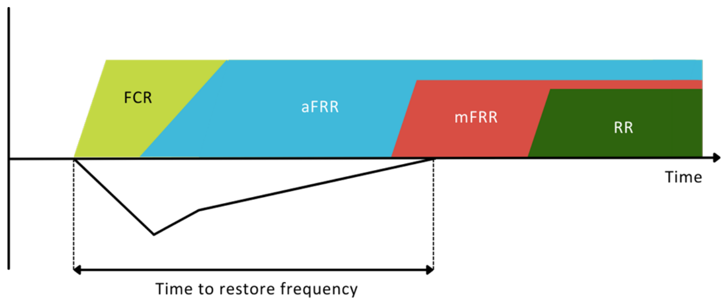

Figure 2.4: Frequency reserve activation schematic (Roumkos, Biskas, & Marneris, 2021))

Since BRPs have to pay for the imbalances they create, there is an incentive to self-balance and thus trade in the markets. The figure below illustrates BRPs trading away their excess consumption or generation on the spot markets and how a BSP is activated to solve an imbalance in real time. The BRP has an overproduction that could not be accounted for by the spot markets, thus an imbalance is created, the TSO will activated capacity to restore the system balance and the BRP has to remunerate the TSO for the capacity activated at the time of the caused imbalance.

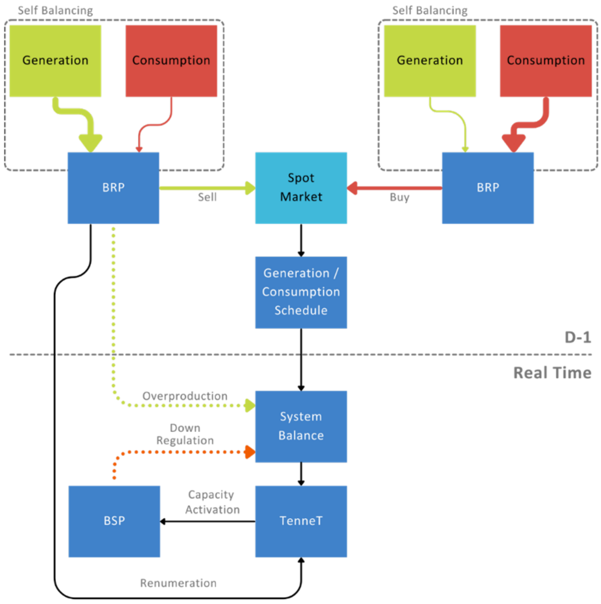

Figure 2.5: Imbalance Restoration

2.2.2. Types of Balancing products

Different types of balancing products have already been introduced. In the context of the balancing market, two main directives are important. The main directive is EU2017/2195, which frames the European balancing markets and is still being implemented today. Furthermore, EU2017/1485 is an important directive. European laws provide a directive for the countries to implement themselves. The development of regulation is crucial to get an overview of market developments. In this case, the EU has put out documents (European Commission, n.d., 2023) indicating the new changes that include the aim to create stable prices and avoid volatility in electricity prices. This section will provide more details on the types of balancing products, their roles and the degree of integration with other countries in procurement or activation.

2.2.3. Frequency Containment Reserves (FCR)

The FCR is the capacity for up or down-regulating the electricity grid and is applied automatically. Procurement and exchange of FCR capacity is done on a common market, the FCR Cooperation. This includes Austrian, Belgian, Czech, Danish, Dutch, French, German, Slovenian and Swiss TSOs (ENTSO-E, n.d.).

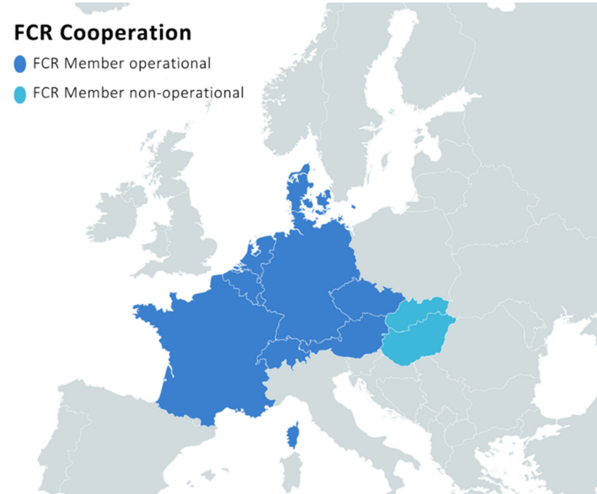

Figure 2.6: FCR Cooperation members (Regelleistung.net, n.d.)

In general, the bids of BSPs are awarded by the use of a general optimization principle that first awards the core share and then optimizes based on export limits while keeping other constraints in mind (ENTSO-E, n.d.). Prices are determined by marginal pricing and bidding happens on daily auctions for six 4-hour blocks.

2.2.4. Frequency Restoration Reserves (FRR)

In the Netherlands, FRR is split into two components: aFRR and mFRR. aFRR is the secondary reserve, and mFRR is the tertiary emergency reserve.

2.2.5. Automated Frequency Restoration Reserve

The frequency restoration reserves are the secondary layer of protection that allowed the frequency to be restored. Currently, in the Netherlands, aFRR is procured in its own operating area. The plan to incorporate aFRR into a common market under the PICASSO program but this has not yet taken place. This means that TenneT is currently solely responsible for the all the FRR processes in the Netherlands. Currently the integration of the Dutch aFRR with the European PICASSO program is delayed (Autoriteit Consument en Markt (ACM), 2022).

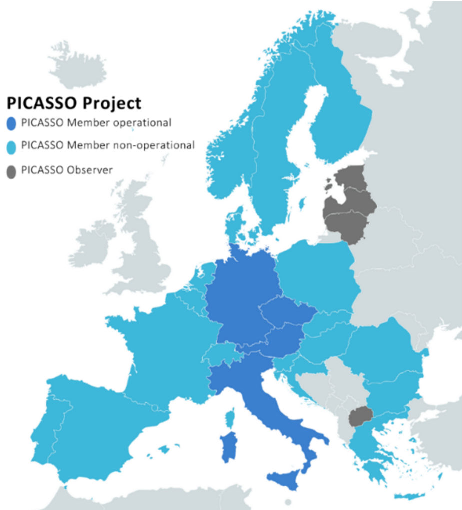

Figure 2.7: PICASSO Project members

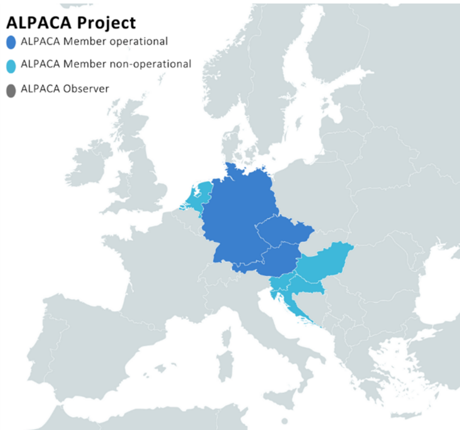

Figure 2.8: ALPACA project members

As aFRR is split into both an up and downwards and an energy and capacity component, this means that there are multiple markets. As the Netherlands is currently not operational on the PICASSO platform, the market is limited to the Netherlands Dutch only. Regulations require a minimal amount of aFRR capacity to be available and tendered by the TSOs as all times this capacity is thus its own market. Furthermore, there is also the possibility of voluntarily supplying energy to the aFRR system/market this is the energy market. In general the capacity bids are for a longer time frame than the energy bids and are traded via the ALPACA Project. Both capacity bids and energy bids determine the merit order list.

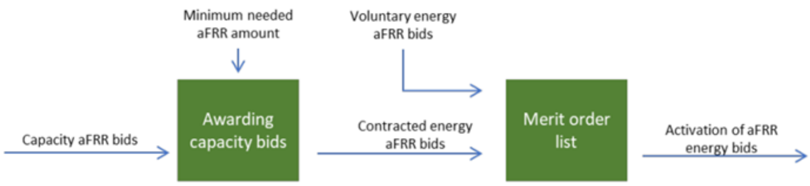

Figure 2.9: Market clearing aFRR capacity and energy bids (TenneT, n.d.)

2.2.6. Imbalance netting

Imbalance netting is done by TSO on FRR activation via the International Grid Control Cooperation (IGCC). Via this cooperation, TSOs know of the others TSOs balancing states and can coordinate their activation of FRR making it possible to reduce the total amount of activation through cooperation.

Figure 2.10: IGCC project member

2.2.7. Manual Frequency Restoration Reserve (mFRR)

Finally, as the name indicates, the Manual Frequency Restoration Reserve or mFRR is a reserve of balancing power that can be switched on manually, this reserve has a relatively large size and is to be activated within 15 minutes after the signal given by the TSO. Normally, the mFRR capacity is reserved but rarely activated.

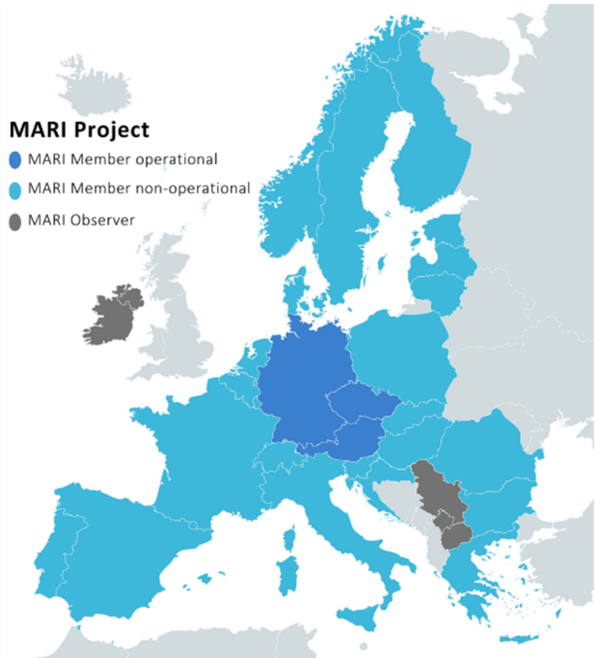

Figure 2.11: MARI project members

2.2.8. Balancing Market developments

First, van der Veen and Hakvoort (2016) state that there is an uncertainty about the effects of design option in the balancing market and propose a framework for policy makers to face this challenge. Hu et al. (2018) also states the importance of market redesign and proposed changes to the European electricity markets that would allow for integration of large scale vRES, additionally indicating the inability of some market designs to facilitate this. Additionally, (Poplavskaya et al., 2021) have studied the effects of market design changes on outcomes of the balancing market with the help of agent based modelling and propose for the further research to study the implications of balancing market integration and reinforcement learning. Newbery, Pollitt, Ritz, and Strielkowski (2018) state that the integration of renewables will have significant impact on the markets need to be restructured.

German Paradox

The German Paradox was observed where higher integration of renewables coincided with the demand and price for balancing services decreasing (Hirth et al., 2015). Ocker and Ehrhart (2017) have studied the large scale integration of renewables on balancing markets in Germany in 2017 and found that while renewable generation grew, the demand for balancing did not. They described this as the German paradox and accredited it mainly to more efficient markets for balancing power.

2.2.9. Changes in regulation balancing markets

Currently, in the scope of the FCR market, 3 major changes to the FCR market have been identified since Jan 1st 2020. Mainly, a shortening of the FCR product from 1 day to 4 hours, a shortening gate opening time of 14 days to 7 and three parties joining the cooperation. The FCR market has shown large changes in recent years in around the time of these market design changes. Currently, these market designs have not been quantitatively reviewed on a longer timescale which makes it interesting to study the FCR market. Additionally, the Dutch aFRR market has seen large price increases in the last years which have been displayed on this market, aFRR balancing energy can traded internationally in contract to the current national setup. In previous years, the aFRR market has seen several changes in it regulation as well, mainly focused on the preparation for integration of the national market into PICASSO. This integration has been done by harmonizing the national aFRR markets so that their designs match and thus can create a common market. There have been several large market design changes that might have an influence on the outcome of the Dutch aFRR market but research on this is currently lacking. Previous research indicated that there is a relation between the outcome of electricity markets and new market design for Greek markets (Papaioannou, Dikaiakos, Dagoumas, Dramountanis, & Papaioannou, 2018) but these have not been studied for Dutch markets yet. As the FCR and aFRR market are the most active, we will limit our scope to these two markets. The timeline shows there has a rapid introduction of new policies in both the FCR and aFRR markets in the period between 2019 and 2021.

2.3. Changes in day-ahead markets

Similarly to balancing markets, day ahead markets have undergone several large scale changes in past years. Firstly, there has been a large number of additions to the SDAC (ENTSO-E, 2022). As the SDAC is the main place where electricity is traded, it is also expected that this leads to more competition and higher overall reliability (Newbery, Strbac, & Viehoff, 2016). Additionally, Great Britain exited the SDAC due to Brexit. There have been two additional changes in the form of the adaptation of the Clean Energy Package by the EU which allowed for more active consumer participation in markets and increased focus on decarbonisation (Anchustegui, 2020). Additionally, the rules for international electricity flow have been altered in the 2021 which could have an impact on the imports and exports and therefore on the final electricity price (ENTSO-E, 2022).

2.4. Market development overview

To conclude this background chapter, an overview of the main developments in the balancing and day-ahead market is given. Firstly, it must be noted that we looked at market design changes for the FCR, aFRR and the coupled day-ahead market. Additionally, there have been several large shocks such as COVID-19, Brexit and the Russian invasion of Ukraine which have had an impact on the price of electricity and market behaviour (European Central Bank, 2022)(Do et. al., 2024).

Having identified the main regulatory changes on the markets for balancing electricity and day-ahead electricity as well as several large scale shocks to the electricity system, we gained insight into the main interventions that have taken place on these markets. With this knowledge, the first sub-question can be answered. A completed timeline of the largest market design changes can be found in figure 2.12

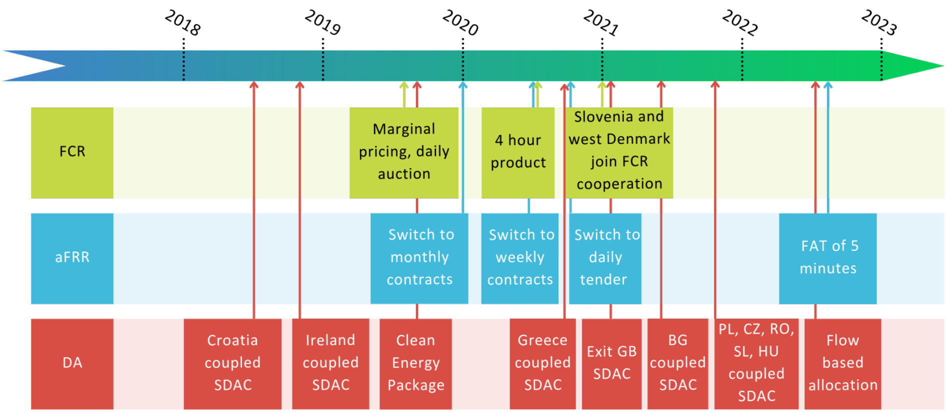

Figure 2.12: Market policy implementation

3
Methodology

This chapter explains how sub-questions 2 and 3 will be studied and answered and provide context on electricity market modelling. Having obtained a clearer overview of balancing markets and electricity markets in general in the background chapter and an overview of the interventions being performed in the market, we will now look towards finding a method that sufficiently helps answer these research questions.

3.1. Methodological Overview

To answer the sub-questions proposed in the introduction and with the background chapter in mind, figure 3.1 illustrates a methodological overview. This section, the methodology section, will substantiate this figure and explain the considerations for answering the sub-questions.

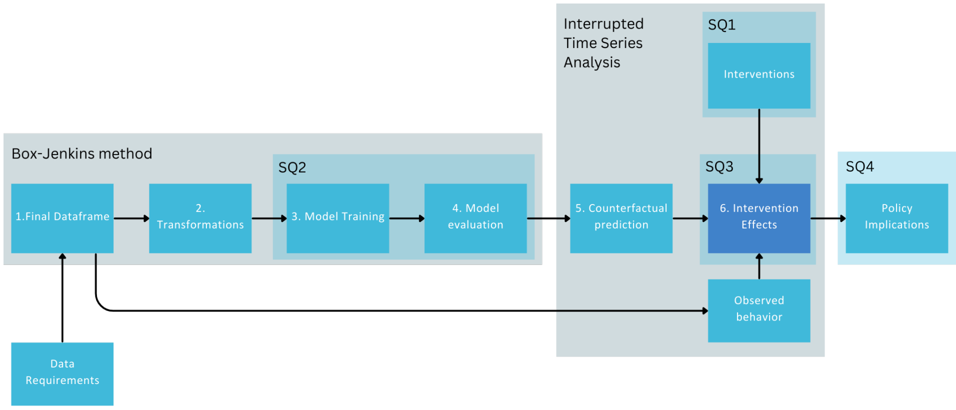

Figure 3.1: Methodological framework

First, data will be collected and transformed, after which modelling can take place. These outcomes will be compared with the real-world scenario from which policy implications can be obtained

3.2. Interrupted Time Series Analysis

To answer the main research question, ideally, one would be able to do an experiment on the response of the market to a large intervention and being able to determine causality. Shadish, Cook, and Campbell (2002) mention several ways of trying to determine causality with differing degrees of reliability. A best-case scenario would be where we can confidently say that a market intervention has had a clear impact on market outcomes. In this case, we can attribute the market outcomes to the intervention.

However, due to the complexity of market behaviour, especially in interconnected electricity markets, it is hard to assign a certain outcome purely to a market design change. We can still attempt to best account for all other factors that have driven these market outcomes and try to gain insight into the effects of a market design change, albeit with a lower degree of confidence. Additionally, there is complexity in the interventions; the market design changes are implemented at different times for different TSOs, and thus, there is complexity in both the system under study and the intervention itself. This complexity will limit the power of a study that foregoes accounting for it. In formulating an approach to answer the main research question, there were two paths to go down. Scaling down the system such that the complexity could be appropriately captured with a relatively large degree of certainty or giving a higher level overview with less power. In this case, the latter has been chosen.

To perform such an analysis, we can look at methods that are being used in other complex systems where we want to observe the impact of an intervention. Preferably, we would be able to have more than one identical system so that we could perform an experiment and directly compare outcomes. However, since there is only one market for each electricity product, the methodology is not a true experiment but a quasi-experiment. In the current case, it is not possible to perform an experiment on multiple markets as the intervention has already taken place, and there is only one unique market to which the intervention can be applied. The main methodology applied to study whether a market design change has had an impact on the outcome variables is that of an interrupted time series analysis (ITSA). An interrupted time series analysis consists of a dependent variable that is measured pre- and post-intervention and a counterfactual. The counterfactual is the outcome in the case the intervention would not have taken place. In our case, we do not know the counterfactual and thus need to predict the counterfactual to compare it with the post-intervention outcome. The assumption is that the intervention has had an impact on the dependent variable and that shift in level or behaviour of the variable post can be attributed to the intervention.

To answer the main research question, several sub-questions have been proposed in the previous section. The nature of these sub-questions differs, which leads to different methods being used to answer them. The first two sub-questions focus mainly on desk research into academic and grey literature on the Dutch imbalance market and on electricity pricing developments in general. From these first two sub-questions, a list of parameters can be identified that can be used to answer the subsequent sub-questions. Sub-question three attempts to quantitatively assess the impact of market design change by making use of statistical methods for the analysis of time series. The aim is to identify if a policy change has made a significant impact on the outcomes of the balancing market and to explain these. The final sub-question will integrate the previous sub-questions and looks at the main takeaways for policy makers from the research.

Thus, a method is needed to help explain the effect of such market reforms or interventions on prices and volumes in that market and what has driven these factors in the first place. First, the definition of a time series: “A time series is a sequence of data at equally spaced points in time” Examples of time series are the population of the Netherlands from 1950 to 2023 per year or the financial performance of a stock measured per day. A time series has three components: Trend, Seasonality, and Residual. The trend captures the general direction that the time series is moving in. Seasonality captures effects that happen on a specific frequency, and the residual is what is left after both trend and seasonality have been taken out is random. If it still has a trend or seasonality, it is not random, and thus, seasonality or trend has been missed during decomposition.

3.2.1. Interrupted Time Series Analysis evaluation metrics

Papaioannou et al. (2018), evaluate the effects of interventions on their models by adding dummy variables. When a dummy variable is added to the regression model, it will allow us to see if the variable is significant and learn about the coefficient of the variable. The dummy variable is a stand-in for the intervention and will, at the moment of the intervention, produce a step signal that indicates that the intervention is active. The intervention dummy variable can be modelled as other signals as well such as a step function with a decay or an increasing function with time, however, a single step change is easiest to model and easiest to interpret.

3.3. Counterfactual Prediction Models

Several methods exist to study the statistical effects of interventions in a system. The most common approaches to these analyses are 1) Difference-in-Difference (DiD) models, 2) Segmented Interrupted Time Series Analysis or 3) Intervention analysis using autoregressive integrated moving average models(ARIMA) (Li et al., 2021).

To select the correct model for counterfactual prediction, we must consider some requirements for the method. The first requirement is that the model be explainable. This stems from the research questions and aids in understanding what factors drive the market outcomes. Secondly, we want to use a model that can account for external variables and seasonal data. Lastly, we want the model creation to be replicable, thus not using randomness, or at least replicable randomness. Methods of counterfactual prediction:

- OLS
- ANN
- SVR
- ARIMA Models

Turner et al. (2021), have compared six statistical methods and have found that although results differ depending on the method chosen, due to the nature of interrupted time series analysis, one should take a cautious approach to interpreting the results of an interrupted time series analysis.

3.4. Model Conceptualization

The purpose of the model will be to evaluate interventions. From modelling, it should become clear what the impact has been on market outcomes. In an ideal situation, we would see a clear impact of the intervention if one is present.

There are several ways of studying how a market's price and volume reflect external variable changes. In our case, the method should be explainable, and thus, we are not looking for a black box but for what factors have influenced the final outcome. To do this, we will study potential market drivers and policy changes and create a regression analysis. Brijs, De Vos, De Jonghe, and Belmans (2015) have used statistical regression methods to study the impact of renewable generation capacity on balancing market prices in Europe. However, Brijs et al. (2015) did not review the effect of market design. Studying the effect of external variables on electricity markets using the SARIMAX method is not new and has been performed by Kraft, Keles, and Fichtner (2020). This study focused on using the SARIMAX method to explain the price drivers of the FCR price between 2014 and 2018. Several studies have researched the impact of market design changes on the price of electricity in the wholesale market in combination with external variables or fundamentals. Both Papaioannou et al. (2018), and Petrella and Sapio (2012) have used ARMAX and SARMAX methods in analysing the impact of market design changes, or as they state it changes in market architecture on the price of wholesale markets in Greece and Italy. The approach used in these papers most closely resembles the approach in this research but differs in that both concern the wholesale market and not the Dutch balancing market. Similar external variables such as day of the week, energy prices, temperature prices and load/demand variables can be used. Tarsitano and Amerise (2017) have used the SARIMAX method before and combined it with dummy variables and have shown relation between day of the week and electricity demand in Italy. The impact of policy on electricity DA pricing in combination with regression has been studied before (Papaioannou et al., 2018). A regression of external variables on the price of balancing markets has also been studied before by Bunn, Inekwe, and MacGeehan (2021), Lago, Marcjasz, De Schutter, and Weron (2021), and Lucas, Pegios, Kotsakis, and Clarke (2020).

3.5. SARIMAX Regression

The SARIMAX regression model is an extension of the ARIMA model used often in econometrics and economics research. SARIMAX stands for Seasonal, AutoRegressive, Integrative, Moving Average and eXogeneous. These terms are linearly combined in the model to describe the underlying process while accounting for seasonality and previous values. Firstly let us look at the ARIMA component:

- **AutoRegressive**: This is the component of the function that predicts based on past values
- **Integrative**: This is the component that differences the series, which achieves stationarity
- **Moving Average**: This models the relation between the average of past values and the current timestep

This is extended with a seasonal and exogenous component.

- **Seasonal**: This component accounts for seasonally repeating patterns
- **Exogeneous**: This component accounts for the influence of exogenous variables on the prediction

In this case, the SARIMAX model is trained on data from before the intervention, and a counterfactual time series is created after the policy intervention. The counterfactual is the best possible prediction of the SARIMAX model of the situation where the policy intervention would not have occurred.

Using the SARIMAX regression to create we need to take several assumptions into account. Namely, ARIMA methods assume that data is stationary which needs to be accounted for before applying the method. The method of implementing ARIMA regressions is often done by making use of the Box-Jenkins methodology that outlines how to prepare, tune and evaluate an ARIMA model. In this analysis, the ‘pmdarima’ model is used for the fit of the SARIMAX model.

3.6. Box-Jenkins methodology

The Box-Jenkins is a commonly used modelling process for time series analysis initially developed in the 1970s to create models that perform forecasting tasks. The aim of the process is to structure the approach to statistical modelling. The original process is divided into three steps but more modern approaches add two additional steps namely, data preparation and model application. The Box-Jenkins approach is as follows:

1. **Data Preparation** means to create a dataset that can be used in the process of modelling. Not all data can be used with all modelling techniques directly and will need to be cleaned, transformed, and/or differenced before use.
2. **Model Selection** is used to select an appropriate model to be used by graphing the data and interpreting the graphs. AIC (Akaike’s Information Criterion) is one of the methods of selecting a model that is used often.
3. **Parameter Estimation** means finding the right parameter that fits the model best to the data that goes in. There are several ways to find the model parameters but one option is to perform a grid search to find the model with the lowest AIC score.
4. **Model Checking** is a step that involves checking the performance of the model against a benchmark and can be done by applying a train-test split to the data and using evaluation criteria such as RMSE or MAPE to evaluate model performance. If the model is found to be inadequate, step two needs to be performed again.
5. **Model Application** is the final point of the Box-Jenkins methodology, meaning the model can perform its tasks.

This approach forms the basis for many modelling procedures and has been used and updated by Peixeiro (2022) and is extended for the SARIMAX application.

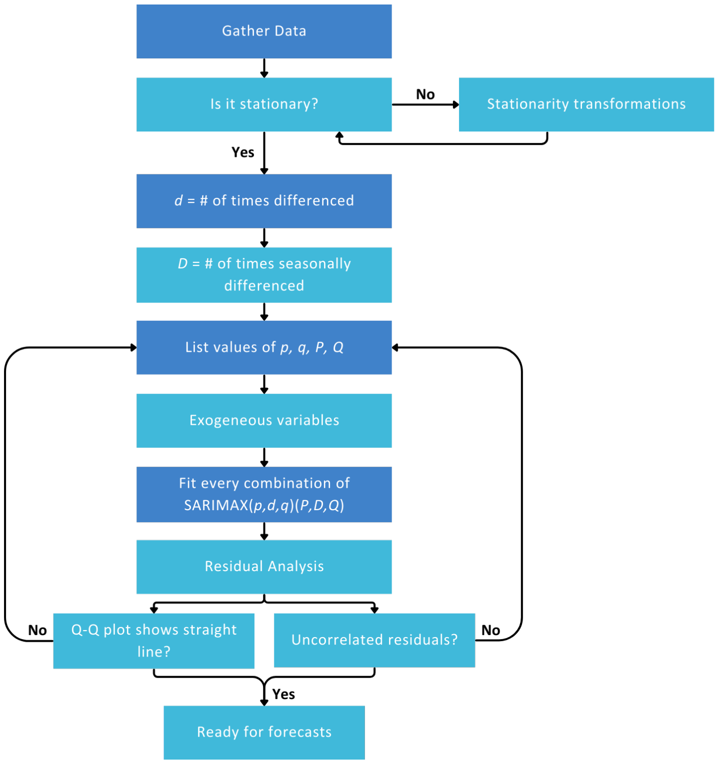

Figure 3.2: Box-Jenkins Methodology (Peixeiro, 2022)

3.6.1. Data Gathering

The first step of the Box-Jenkins approach is data gathering. It is important in time series analysis that the dataset is clean and comprehensive. There are requirements for a dataset. The first Is that all time series have the same resolution. Since data sources differ in their resolution, first, they must be aligned. An example of different data resolution is a daily gas price data point versus a minute imbalance data point. There exist several ways of dealing with this issue, but one common approach is reducing and averaging data to a less refined level of detail. The level of detail is determined by several things, such as the purpose of the analysis, computational power and data availability. As it is our goal to study the long-term effects of market intervention, we have chosen to average the data to a daily resolution. More details about this process can be found in the following chapter.

3.6.2. Stationarity Testing

The first step of the Box-Jenkins approach is to see if the time series is stationary. Non-stationary time series should be differenced to make them stationary.

Stationarity has three main characteristics. A stationary time series should have a:

1. Constant mean with time
2. Constant variance with time
3. Constant autocorrelation with time

Several tests exist to check if a time series is stationary. The most commonly used tests are the Augmented Dicky Fuller (ADF) test and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. If the time series are non-stationary, they can be transformed by differencing. Differencing is that, instead of their absolute values, the differences from time step to time are taken. How often this differencing is done determines the d (and D for seasonality) hyperparameter for the SARIMAX model.

3.6.3. Hyperparameter estimation

To obtain the best score, the SARIMAX model should be tuned correctly, meaning the order of each component should be optimal to yield the lowest error. While the *d* and *D* variables are found by differencing, the best fitting order for other variables is found by trial and error. Here, we perform a method called grid search. In grid search, for every order in a user-set range, a model is trained and evaluated, after which the score is stored. The best score is associated with the best hyperparameters to use.

Since we will use Python for the SARIMAX implementation, the 'pmdarima' package is used. This package automatically accounts for the selection of hyperparameters by performing a grid search for the model parameters associated with the lowest AIC score.

3.6.4. Training and evaluation

After hyperparameters have been selected using the lowest AIC score, we can test for model validity. To test this, the dataset is split into a testing and a training set. Under the assumption that the dependent variable behaves similarly to the effects of external factors in the pre-intervention time period, the data can be split into two sets: a testing and a training set. In accordance with other literature, we will use an 80/20 split. The training set is used to find the most suitable model. This model is then evaluated against the testing dataset, allowing for a test of the model's validity. To quantify the model performance, in general, with SARIMAX applications, the mean absolute percentage error (MAPE) and root mean square error (RMSE) terms are used to evaluate model performance.

3.6.5. Residual analysis

The final step of the Box-Jenkins approach is to check the model's residual. If these residuals are normally distributed and do not have clear auto correlation, they can be assumed to be white noise, which is a strong indication of a working model. Thus, residual checking is done by looking for autocorrelation in the residual using ACF and PACF plots. Furthermore, the normality of the residuals is checked by a quantile-quantile Plot (Q-Q Plot). If this plot shows a straight diagonal line, we can assume the residuals to be normally distributed. Once the residual analysis is done, the model can be used to for predictions and estimation.

4

Data Gathering and Analysis

The main aim of this chapter is to explain where the data was gathered, if it is complete, and to analyze the basic components of the data. Since model performance is largely dependent on the quality of the data, it helps to form the basis on which to answer sub-questions 2 and 3. This means that most data analysis will be done before the electricity market modelling chapter. This chapter thus starts by explaining the approach to data gathering and will end with a final data frame.

4.1. Data gathering approach

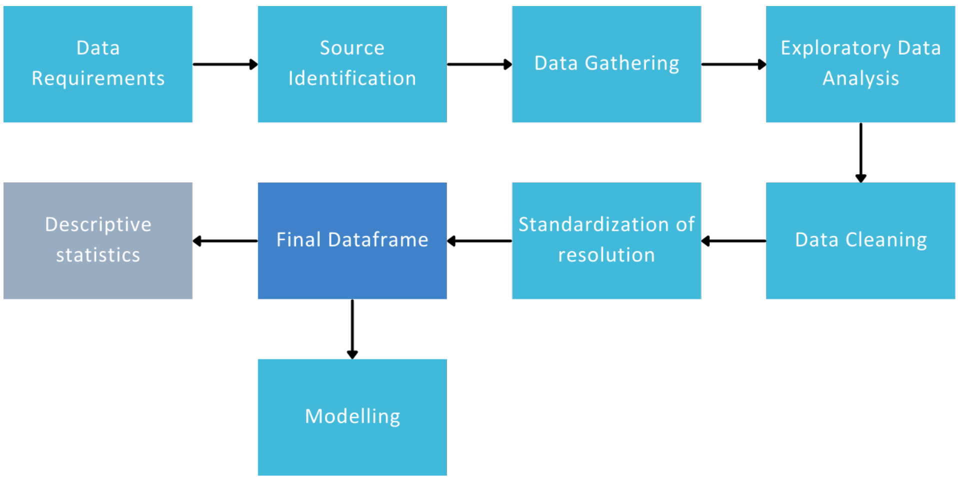

Figure 4.1: Data gathering and analysis approach

4.2. Data requirements

Literature has already indicated what sources of data are correlated with electricity market prices and volumes. Required data for day-ahead market modelling was found from Papaioannou et al. (2018) and Lago, De Ridder, and De Schutter (2018) and includes generation, consumption, temperature and several other variables. Kraft et al. (2020) and Hirth et al. (2015) have identified several variables for imbalance modelling such as forecasting errors, outages, temperature and whether there is a holiday or not. Therefore, we will start by gathering data for these relevant variables.

4.3. Source identification

4.3.1. ENTSO-E

Under EU regulation, TSOs are required to publish data on their networks to the open platform of ENTSO-E, the ENTSO-E Transparency platform. This contains a wealth of data on the generation and consumption of electricity in the European Union. The platform includes data on load, such as total load and lord forecasts. Furthermore, many different types of energy are available, such as coal, gas, biomass, nuclear, wind, solar, and hydro. Both actual production numbers and TSO forecasts for the following day are available. Additionally, the platform contains data on cross-border flows, day-ahead prices and several balancing datasets such as capacity prices, imbalances and activated capacity. Although the dataset here can be accessed freely, not all of them are complete, as every TSO is responsible for publishing their own data to the platform. Even though the regulation calls for all TSOs to publish their data, many differences between datasets describing the same variable still exist, which will require further investigation before usage in modelling. Additionally, different definitions of variables and sparse documentation of the platform require a thorough approach to exploratory data analysis.

4.3.2. KNMI

It has been shown that the weather can have an impact on the demand and production of electricity. Therefore, for our dataset, we will take temperature data from the Dutch Meteorological Institute KNMI. This data is openly available.

4.3.3. EEX

Data from the EEX exchange has been obtained with permission from the platform for the purpose of this research. Data on TTF Price and CO2 price are not publically available in the resolution that would be required for this study and therefore commercial data has been used.

4.3.4. Nationaal Energie Dashboard

The Nationaal Energy Dashboard (National Energy Dashboard) is an initiative from TenneT, Gasunie, and several other data partners. The platform provides real-time estimates of the sources of electricity infeed into the Dutch grid. Since Tennet has indicated that the data on ENTSO-E may not be representative, this source was used as a backup. Other renewable data is also available for the Netherlands on this platform. Still, ENTSO-E data is the preferred source because it comes directly from TenneT instead of a third party.

4.4. Data gathering

Having identified the sources of data for this research, the following step was to obtain the data. For both the KNMI data and the Nationaal Energie Dashboard, data can be easily downloaded from their public websites. Obtaining ENTSO-E data required the use of an FTP client before obtaining .csv files. Lastly, data from EEX was obtained via email. Once the data was collected, it was interpreted into Pandas dataframes for a first analysis. The data has been gathered from the following sources:

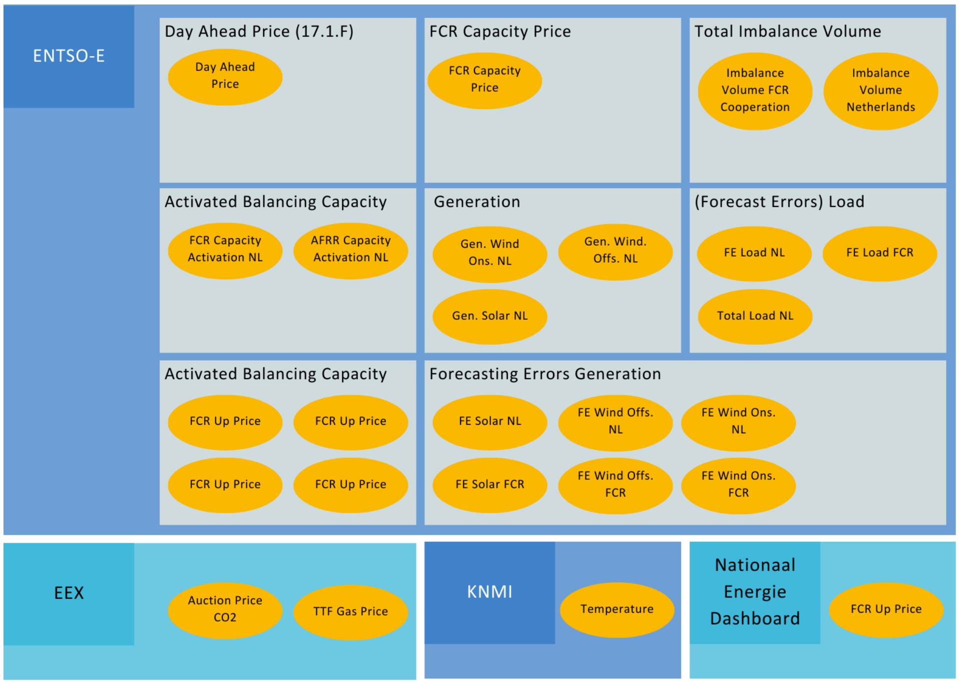

Figure 4.2: Timeseries and Sources

4.5. Exploratory Data Analysis

Since several different types and sources of data describe the electricity market, we first need to ensure that all the data aligns and can be used for time series modelling. Therefore, we will give an overview of all the raw data used in the table below, together with its resolution, source, data type, and the time frame for which it is available.

Table 4.1: Data sources and available date range

Variable name	Source	Start Date	End Date
	ENTSO-E		
ForecastError Load FCR Cooperation	6.1.B - 6.1.A	1-1-2018	31-12-2022
ForecastError Load NL	6.1.B - 6.1.A	1-1-2018	31-12-2022
Total Load Value	6.1.A	1-1-2018	31-12-2022
ForecastError Solar FCR Cooperation	14.1.D - 16.1.B/C	1-1-2018	31-12-2022
ForecastError Solar NL	14.1.D - 16.1.B/C	1-1-2018	31-12-2022
ForecastError Wind.offs. NL	14.1.D - 16.1.B/C	1-1-2018	31-12-2022
ForecastError Wind.offs. FCR Coop.	14.1.D - 16.1.B/C	1-1-2018	31-12-2022
ForecastError Wind.ons. FCR Coop.	14.1.D - 16.1.B/C	1-1-2018	31-12-2022
ForecastError Wind.ons. NL	14.1.D - 16.1.B/C	1-1-2018	31-12-2022
FCR Up Price	17.1.F	2020-05	2023-01
Imbalance total NL	17.1.H	1-1-2018	17-12-2019
Imbalance total FCR Cooperation	17.1.H	1-1-2018	1-1-2023
Not specified down activation volume (NL aFRR activation)	17.1.E	27-5-2020	1-1-2023
DA Price	12.1.D	2018-01	2023-01
Wind ons. NL Output	16.1.B/C	1-1-2018	31-12-2022
Wind offs. NL Output	16.1.B/C	1-1-2018	31-12-2022
	EEX		
Auction Price CO2	EEX Emissions Market	18-1-2018	19-12-2022
TTF Price	EEX Spot Market	21-5-2019	3-1-2023
	KNMI		
Temperature	Royal Netherlands Meteorological Institute (KNMI)	2018-01	2023-01
	Nationaal Energie Dashboard		
NL Solar Output	Nationaal Energie Dashboard	1-1-2018	31-12-2022
Holiday	Python	1-1-2018	31-12-2022

During the exploration of the gathered time series, we found that some had missing values. Therefore, we will need to do some data cleaning

4.5.1. Data Cleaning

Some timeseries had data missing, which have been addressed by omitting them from the analysis or by forward filling in the case of a small gap in the data. This becomes clear when performing the analysis on the dataframe at the end of this chapter.

4.5.2. Standardization of Resolution

All data is first averaged to hourly data before exporting to the final dataframe. In the final regression, the data is once again averaged to daily resolution to make training the SARIMAX model quicker and more effective. Since the choice was made for a longer time frame approach, the daily averaging is justified but does decrease resolution and potential effectiveness.

4.6. Final dataframe

The final full dataframe for the modelling section contains the following variables:

Table 4.2: Descriptive statistics Final Dataframe

variable name	mean	min	max	std	# observations
DateTime	2020-07-02	2018-01-01	2023-01-01	-	1827
FCR_UpPrice	23.70	0.00	260.26	16.68	987.00
DayAheadPrice	94.07	-3.85	699.75	98.08	1827.00
Temperature	11.26	-6.54	29.75	6.13	1827.00
Auction Price €/tCO2	39.79	7.68	97.51	24.92	1820.00
TTF_Price	50.80	3.02	312.74	56.82	1322.00
Full_TotalImbalanceVolume	89.47	-966.34	1334.21	275.60	1826.00
fcrcoop_imbalance_total_NL	2.66	-11.78	21.04	4.63	716.00
NotSpecifiedDownActivatedVolume	14.43	0.00	21.21	3.60	952.00
TotalActivatedVolumeAFRR	17.28	0.00	54.08	7.20	1827.00
Holiday	0.03	0.00	1.00	0.16	1827.00
TenetTotalImbalance	5.31	-134.01	166.60	41.96	1827.00
TotalImbalanceVolumeNL	1.36	-19.79	13.52	5.77	716.00
ForecastErrorTotalSolar	610.10	-1641.46	3844.30	655.35	1826.00
ForecastErrorNLSolar	675.21	-0.70	2264.27	534.58	1826.00
ForecastErrorTotalWindOffshore	-43.20	-2870.23	2294.60	515.83	1826.00
ForecastErrorNLWindOffshore	-169.48	-1352.62	379.69	281.72	1826.00
ForecastErrorTotalWindOnshore	466.36	-4725.49	8344.27	1290.52	1826.00
ForecastErrorNLWindOnshore	563.77	-108.09	3390.87	562.23	1826.00
LoadForecastErrorFCR	-2424.12	-66601.16	14842.39	8478.04	1827.00
LoadForecastErrorNL	-326.78	-4166.29	4472.01	1279.54	1827.00
TotalLoadValue	12400.52	8691.25	15443.67	1418.75	1827.00
NL_WindOnOutput	508.65	8.36	2193.71	429.52	1826.00
NL_WindOffOutput	623.13	1.92	2359.92	514.91	1826.00
NL_SolarOutput	27.26	0.00	134.93	32.63	1826.00
NED_NedSolarActualized	1118.42	29.51	4930.32	1012.65	1826.00

4.7. Limitations to final dataframe

The final dataframe, used as the foundation for analysis, primarily consists of market data, including generation numbers and electricity prices, but does not include direct information on market interventions or policy changes. The data was aggregated and scaled to uniform 1-hour blocks, with missing data points from several countries being estimated or filled in. While this process may have introduced some errors that could propagate through the dataset, the resulting dataframe still provides a solid starting point for analysis. It offers valuable insights into market activities and serves as a robust foundation for further investigation into the impacts of market dynamics and interventions.

5

Electricity Market Modelling

This chapter will study the possibility of modelling electricity markets and thus helps to answer sub-question 2: *"How can electricity market outcomes be modelled and explained?"*. This will be done by applying the Box-Jenkins modelling method from Chapter 3 and the data gathered in Chapter 4.

As we will study several electricity markets in the Dutch context and we have already found some limitations in the data available, this chapter has been split into several experiments. Every experiment will be explained in more detail including an overview of the expected system, model specifications and model evaluation. Modelling outcomes will determine what models are suitable for testing the interventions on in chapter 6.

To answer sub-question 2, we'll start by trying to model day-ahead price behaviour and run intervention experiments. Afterwards, we will model imbalances and the associated activation and price. This order is chosen as the systems become more complex to model with each further step. For every experiment, an assumption of the system and, thus, an overview of the complexity and modelling considerations is given.

5.1. Experiment Selection

First, we will give a short overview of the experiments that are to be run. This is done to find scenarios where sufficient data is available to have a functioning model, which can then be used to do intervention testing. We will increase the complexity of the models by seeing if more complex behaviour can be estimated with the variables present. Therefore, we'll start with the following scenarios.

We will test for different models:

1. Day-ahead price modelling
    (a) Netherlands
2. Imbalance volumes modelling
    (a) Netherlands

   (b) FCR Cooperation

3. Activated Reserves modelling

   (a) FCR Netherlands

It must be noted that the original dataframe also contains the possibility to model other types of balancing variables but that these models did not fit the required criteria and thus have been omitted from the final report.

Comparing the different scenarios and accounting for their differences will give insight into the differences in model prediction accuracy and might also help gain insight into the differences when modelling these electricity market variables.

Table 5.1: Description of various capacity and price metrics

Type	Relevance	Format	Resolution	Source
Day Ahead Prices	The price of electricity for the next day is crucial for market participants to plan their operations.	€/MWh	Daily	ENTSO-E
Imbalance volume Netherlands	Total imbalance in the Netherlands	MW	Daily	ENTSO-E
Imbalance volume FCR Cooperation	Total imbalance in the whole FCR Cooperation	MW	Daily	ENTSO-E
Activated FCR demand	The activated FCR capacity is a measure of activated MWh of capacity for each 15-minute block.	MWh	Daily	ENTSO-E

5.2. Implementation of Box-Jenkins model

To apply the Box-Jenkins modelling approach, all the data was first collected using a Python Dataframe. This process was explained in chapter 4. The following step is to check the data for stationarity, and if the dependent time series is non-stationary, transformations are applied. These transformations are done by using differencing, which can also determine the d and D hyper-parameters for the model using the KPSS test. After d and D have been found, the other model hyper-parameters were found using the pm.autoarima function. The final model was selected for the model that has the lowest AIC score. Afterwards, the residuals of the model were analysed alongside the RMSE and MAPE scores to evaluate model performance. Every model was trained until an intervention date, meaning that all the data before was part of the training set, and all the data after was part of the testing set. The RMSE and MAPE were determined by comparing the model performance with real-world observations of the dependent variable. The RMSE and MAPE scores, along with visual analysis, will determine whether the model is suitable for intervention analysis.

5.3. Day-ahead price modelling

5.3.1. Introduction

As mentioned, DA markets have been modelled before using the SARIMAX approach but not during the timeframe that we are studying with large underlying systemic changes. Infeed of renewables and load has been determined as having a significant correlation with day-ahead prices Mosquera-López and Nursimulu, 2019 and, therefore, is expected to play a large role in price determination. Furthermore, temperature, price of gas and the price of CO2 will be used in the model Gabrielli, Wüthrich, Blume, and Sansavini, 2022. Due to data availability limitations, we have ignored imports and exports.

Table 5.2: Description of Day-ahead prices model scope

Type	Relevance	Format	Resolution	Source	Period
Day-ahead prices	Widely studied and modelled, used for validation of methodology.	EUR/MWh	Daily	ENTSO-E	24/05/2019 - 01/12/2022

5.3.2. System assumptions

As the price for day-ahead electricity is set by the point where the load meets the merit order that comes from bids from large-scale producers and consumers, we assume the underlying system to be as follows:

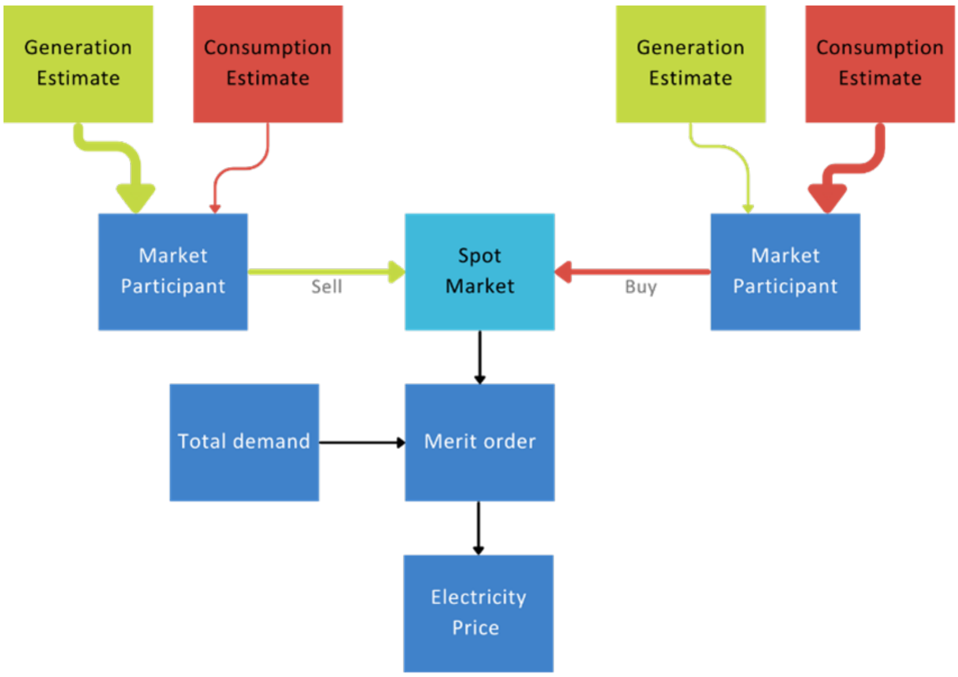

Figure 5.1: Methodological framework

5.3.3. Data selection and descriptive statistics

The following table indicates the variables used for this analysis and their unit, source and resolution. The timeseries are aggregated to a daily level, and the aggregation method is described in the appendix.

Table 5.3: Description of variables and their sources, units, and resolutions

Variable	Source	Unit	Resolution
Day-ahead Price	ENTSO-E	EUR/MWh	Daily
Wind Onshore Generation	ENTSO-E	MW	Daily
Wind Offshore Generation	ENTSO-E	MW	Daily
Solar Generation	ENTSO-E	MW	Daily
Load	ENTSO-E	MW	Daily
TTF Price	EPEX Spot	EU/MW	Daily
CO2 Price	EPEX Spot	EU/Tonne	Daily
Temperature	KNMI	Celsius	Daily
Holiday	Python	DAY	Daily

Table 5.4: Statistical summary of variables

Variable name	Mean	Min	Max	Std. Dev.	#observations
DayAheadPrice	107.6	-3.9	699.7	108.1	1287.0
NL_WindOffOutput	698.8	3.7	2218.7	554.8	1287.0
NL_WindOnOutput	585.6	18.9	2156.7	451.0	1287.0
Holiday	0.0	0.0	1.0	0.2	1287.0
Auction Price Euro/tCO2	47.2	14.6	97.5	24.0	1287.0
TTF_Price	49.4	3.0	312.7	56.6	1287.0
NED_NedSolarActualized	1402.2	51.5	4930.3	1069.6	1287.0
Temperature	11.8	-6.0	29.0	5.9	1287.0
TotalLoadValue	12085.1	8691.3	15068.0	1362.8	1287.0

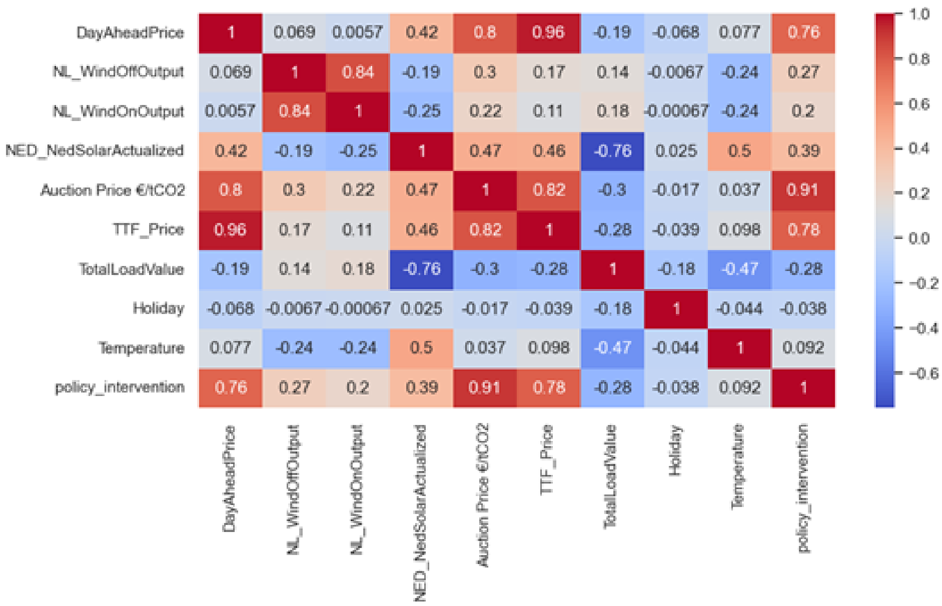

Figure 5.2: Correlation Matrix Day-ahead price

5.3.4. Limitations

Following the literature, we can expect that day-ahead prices can be explained with a relatively high degree of certainty. However, it must be noted that all variables are aggregated to a daily mean and that only data for the Netherlands is used. Furthermore, data concerning solar generation can be seen to be relatively low in value compared to actual installed capacity, which is due to the fact that TenneT does not accurately measure solar generation. This is due to the highly distributed and behind-the-meter nature of solar generation. Even if solar generation data is underreported, if it is done so consistently, it can still be useful for explaining behaviour.

5.3.5. Expected Results

We expected that the day ahead price would be highly correlated with the provided variables; however, the fit might not be as good as possible due to the limitation in the resolution of the data and the missing variables. Regarding the effects of the exogenous variables on the day ahead price, we expect that renewable generation will have a negative effect as it is low on the merit order. Additionally, we expect that the price of gas and CO2 will have an increasing effect on the price of electricity as they will illustrate higher generation costs if demand cannot be met by renewables. Furthermore, we expect that temperature has a negative effect on the price of electricity as it is correlated with higher solar radiation and with summer when less electricity is generally used (Knittel & Roberts, 2005). A high total load value will drive the merit order up, which will likely lead to higher prices in the case of lower renewable generation, especially in winter or on days when there is little wind and solar generation. Additionally, on the effect of the interventions, we expect that both the first and second intervention, adding more players to the European day-ahead market, will have a negative effect on the price of electricity as they will lead to more competition and thus could lead to higher utilization of import and export capacity making the market more efficient and thus leading to lower prices. Lastly, for the invasion of Ukraine, we expect prices to increase due to the gas prices having increased as Russian gas sanctions were put in place. However, it must be noted that these sanctions that came in place later will probably be reflected in the price of gas and, thus, not directly be measurable from the model residuals.

5.3.6. Modelling results

To obtain a SARIMAX model we have followed the Box Jenkins approach. First, the main time-series was tested for stationarity.

KPSS test

We found that using a differencing term of 1 obtained a stationary series, furthermore, by doing seasonal decomposition, we obtained a seasonality of 7 which indicates weekly seasonality. From this, we have trained the model and observed the following results:

Table 5.5: DA Price, Model Parameters and Performance Metrics

Model Parameters	Dependent Variable	Start Date	End Date	RMSE	MAPE
(1,1,2)(2,0,1, 7)	Day Ahead Price (Spot)	07-2020	07-2021	10.75	15.66%

Validating the model against the real-life data gives the following dataset. Which yielded the results as shown in figure 5.3. The blue line describes the training set and the orange line the test set. The green line is the model prediction.

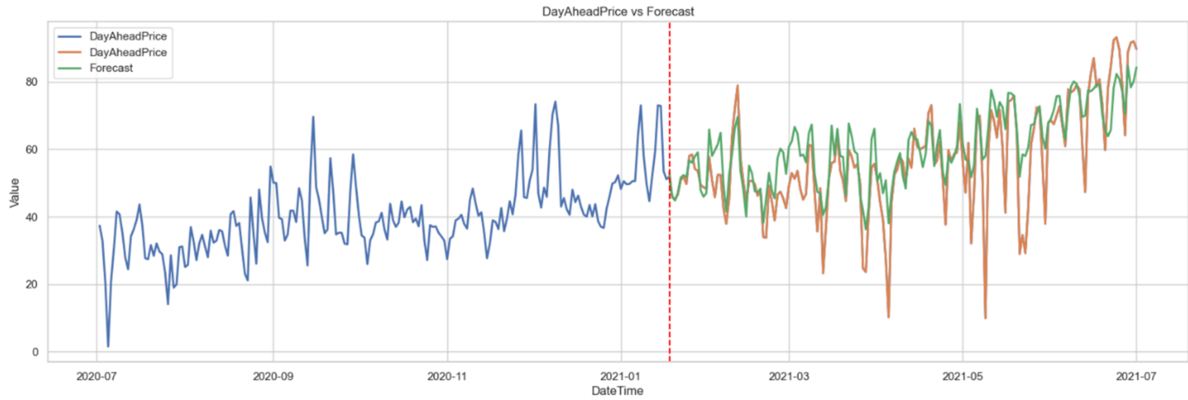

Figure 5.3: DA Price, prediction with real data

Here, we see that the model is able to capture the price for DA products and gives significant results for wind feed-in as well as for the auction price of CO2 and total load. The RMSE and MAPE statistics, together with a visual inspection of the test set, indicate that the model can be used for the intervention analysis.

Table 5.6: DA Price, SARIMAX model coefficients and statistics

Variable	coef	std err	z	P>|z|	[0.025	0.975]
NL_WindOffOutput	-0.0081	0.002	-3.408	0.001**	-0.013	-0.003
NL_WindOnOutput	-0.0063	0.003	-2.144	0.032**	-0.012	-0.001
Holiday	-4.371e-15	1.01e-06	-4.32e-09	1.000	-1.98e-06	1.98e-06
Auction Price euro/tCO2	0.9638	0.470	2.050	0.040**	0.042	1.885
TTF_Price	0.1978	0.751	0.264	0.792	-1.273	1.669
NL_SolarOutput	0.0362	0.053	0.677	0.498	-0.069	0.141
Temperature	0.1543	0.173	0.890	0.374	-0.186	0.494
TotalLoadValue	0.0053	0.001	8.070	0.000**	0.004	0.007
ar.L1	-0.2699	0.240	-1.124	0.261	-0.740	0.201
ma.L1	-0.1770	0.223	-0.794	0.427	-0.614	0.260
ma.L2	-0.4853	0.136	-3.561	0.000	-0.752	-0.218
ar.S.L7	0.2712	1.776	0.153	0.879	-3.210	3.752
ma.S.L7	-0.2185	1.804	-0.121	0.904	-3.754	3.317

5.4. Imbalances in the Netherlands Modelling

5.4.1. Introduction

Having created the experiment for the day-ahead market, we'll now examine imbalances. Here, we'll focus first on predicting imbalance volumes before examining the actual imbalance markets and their capacity activation. As mentioned in the literature review, there are several causes of imbalances. Hirth and Ziegenhagen (2015) have identified five main causes: unplanned interconnect outages, forecasting errors for VRE generation and load, scheduling leaps, and unplanned outages for thermal or hydro plants. These might cause an imbalance, which will need to be restored by using reserves. It is important to mention that these imbalances are not the same as the activation of reserves.

The main assumption for the system's working is shown in the system diagram below in Figure 5.7. This schematic presents where the imbalances are created and what their role is in the system. First and foremost, an imbalance is where the grid is either running faster or slower than required. Thus, looking at the five sources of imbalances presented by Hirth and Ziegenhagen, we see the system balance be presented in the schematic by the "System balance" block. The imbalance comes from the electricity grid and can be solved by activating BSPs. This means that imbalances are a more fundamental process than the activation of balancing capacity, which is why we'll first test the ability to model this less complex system before applying the method to balancing market products.

Table 5.7: Description of Imbalances NL metric

Type	Relevance	Format	Resolution	Source	Period
Imbalances NL 17.1.H	Fundamental for balancing capacity activation and price is the imbalance.	MW	Daily	ENTSO-E	01/07/2020 - 01/07/2021

The imbalance is used by the TSOs to signal to BSPs that capacity activation is needed. Here, we model the imbalance as stated on the ENTSO-E website, which is published by the TSOs themselves. We assume thus that imbalances are largely dependent on the forecasting errors from the BRPs and that they are the signal for the TSO to start balancing.

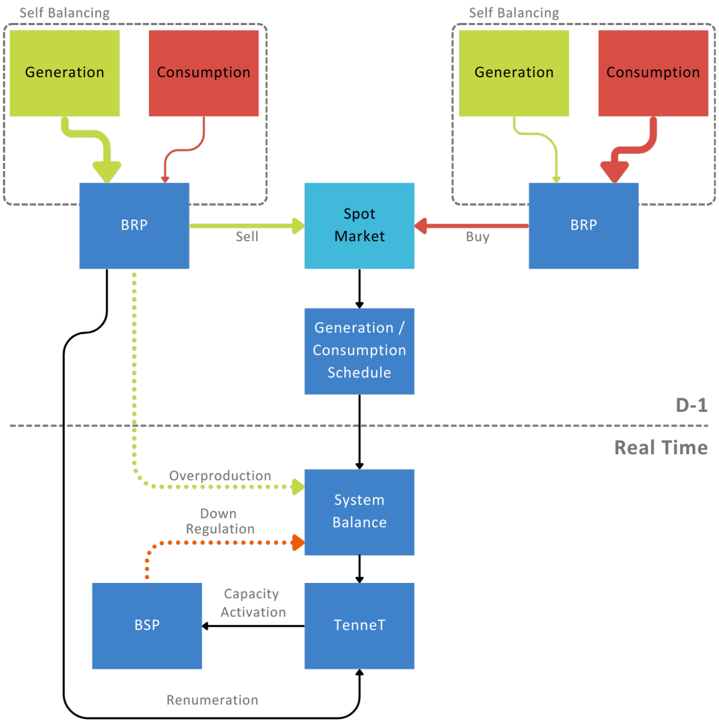

Figure 5.4: Imbalance creation framework

5.4.2. Data selection and descriptive statistics

As can be seen in the table, the main variables taken into account here are those of the forecasting errors. In Chapter 4, the creation of this dataframe is explained in more detail. The table includes the dependent imbalance variable for the Netherlands as well as the explanatory variables such as the forecasting errors for solar off/onshore wind and grid load. Additionally, temperature is included. All data is aggregated to a daily level.

Table 5.8: Statistical summary of variables

Variable	Mean	Min	Max	Std. Dev.	#observations
fcrcoop_imbalance_total_NL	-0.7	-9.7	8.1	2.7	350.0
ForecastErrorTotalWindOnshore	397.3	-3217.9	7049.7	1208.9	365.0
ForecastErrorTotalWindOffshore	87.6	-1157.5	1604.3	433.2	365.0
ForecastErrorTotalSolar	517.3	-987.5	2383.9	585.2	365.0
LoadForecastErrorFCR	-4132.2	-10727.9	6885.8	2629.9	365.0
Temperature	11.2	-2.5	29.0	6.0	365.0

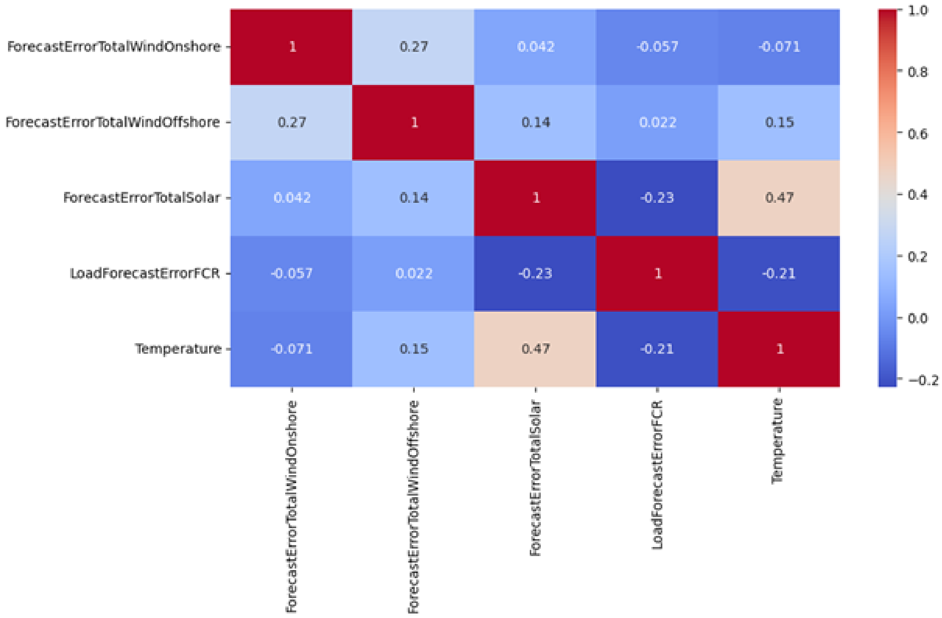

Figure 5.5: Correlation Matrix Imbalances Netherlands

5.4.3. Modelling Limitations

As mentioned before, there are five main causes of imbalance creation. We assume that if all these causes are quantified and included in the data, we will be able to predict the system imbalance. However, there are several limitations to this modelling method, the first is that it assumes a deterministic system that will respond the same to exogeneous inputs. Secondly, as it has already become clear, not all of the exogenous variables are present. We have only included data on forecasting errors and outages, and scheduling leaps have not been included. Additionally, the imbalances get aggregated more than the day-ahead price. The original dataset has a resolution of 15 minutes that is aggregated to a daily level. This means that if the imbalance has both positive and negative peaks during the day, the data will not reflect this, as the peaks will essentially cancel out.

5.4.4. Expected results

Before balancing capacity has to be activated, imbalances need to be registered at the TSO level. The TSO can then activate balancing capacity, which happens on either a European (FCR) or national scale (aFRR, mFRR). As imbalances are the main driver of balancing capacity activation, it is assumed that this process depends more on fundamental drivers of imbalances, such as forecasting errors. Therefore, we expect that a model that incorporates variables that fundamentally drive imbalances, is able to explain these imbalances to a certain degree. It is known that imbalances are caused, and in this model, only forecasting errors are included, impacting final performance and fit.

In this case, the Dutch imbalance volume is predicted per day from the end of 2019 to the end of 2020. This period has been chosen due to data availability for both the Dutch situation and within the FCR cooperation.

5.4.5. Modelling results

Table 5.9: Model Parameters and Performance Metrics

Model	Dependent Variable	Start Date	End Date	RMSE	MAPE
(1,1,1)(1,0,1, 7)	Total Imbalance (NL)	01-2019	01-2020	2.76	-17.00%

As can be seen from the table above, the model has relatively small prediction errors. From the figure below, it becomes apparent that the model is able to capture the behaviour of the sign and the peaks rather well but that the amplitude seems incorrect.

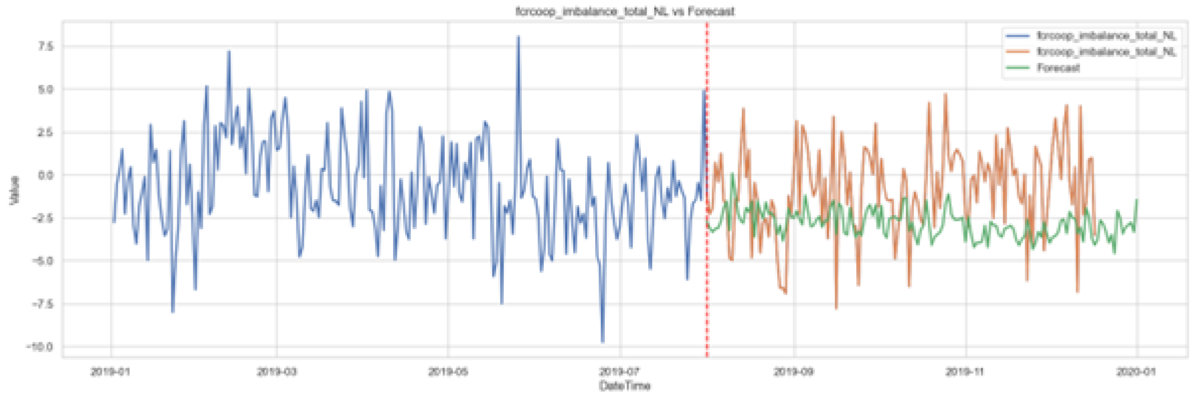

Figure 5.6: Imbalances Netherlands: Model validation results

The imbalances for the Netherlands model were created, and the following coefficient estimates were made.

Table 5.10: Coefficients and Statistical Metrics

Variable	coef	std err	z	P>|z|	[0.025	0.975]
ForecastErrorNLWindOnshore	0.0021	0.001	3.733	0.000**	0.001	0.003
ForecastErrorNLWindOffshore	-0.0026	0.003	-0.969	0.332	-0.008	0.003
ForecastErrorNLSolar	0.0022	0.001	1.869	0.062	-0.000	0.004
LoadForecastErrorNL	0.0001	0.000	0.380	0.704	-0.001	0.001
Temperature	-0.0423	0.105	-0.405	0.686	-0.247	0.163
Holiday	1.6927	1.264	1.339	0.181	-0.785	4.170
ar.L1	0.0789	0.135	0.584	0.559	-0.186	0.344
ma.L1	-0.7143	0.116	-6.177	0.000	-0.941	-0.488
ar.S.L7	0.1062	0.647	0.164	0.870	-1.161	1.374
ma.S.L7	0.0473	0.653	0.072	0.942	-1.232	1.327
sigma2	7.1706	0.787	9.115	0.000	5.629	8.712

From these estimates, it becomes clear that only the forecasting error for wind onshore production has significantly affected the model. Together with the visual inspection of the model, we cannot assume that it is fit for the intervention analysis despite positive error terms.

5.5. Imbalances in FCR Cooperation Modelling

5.5.1. Introduction

Similarly to the modelling of the imbalances in the Netherlands, this model uses the same variables but has forecasting errors calculated for all the countries within the FCR cooperation. Furthermore, the dependent variable also reflects the imbalance in all the FCR cooperation countries.

5.5.2. System description

Similarly to the imbalance variable for the Netherlands, the FCR cooperation imbalance reflects the total imbalance in the FCR cooperation by ENTSO-E 17.1.H. Similarly, the period for the modelling was kept the same for the whole period of 2019, with daily aggregation. As mentioned in the previous chapter, the data for these imbalances is created by merging all the countries in the FCR Cooperation, namely France, Switzerland, Austria, The Czech Republic, Germany, Denmark, Belgium and the Netherlands. Therefore, the system diagram remains the same as figure 5.11.

Table 5.11: Description of Imbalances FCR Cooperation metric

Type	Relevance	Format	Resolution	Period
Imbalances FCR Cooperation	Market of imbalances in the FCR Cooperation	MW	Daily	01/01/2019 - 01/01/2020

5.5.3. Data selection and descriptive statistics

Similarly to the last section, we selected the same variables but now aggregated them on the FCR Cooperation level instead of on the Netherlands. Therefore, we have automatically extended the scope of these variables.

Table 5.12: Statistical summary of variables

Variable	Mean	Min	Max	Std. Dev.	#observations
Full_TotalImbalanceVolume	124.6	-689.1	913.8	272.0	365.0
ForecastErrorTotalWindOnshore	397.3	-3217.9	7049.7	1208.9	365.0
ForecastErrorTotalWindOffshore	87.6	-1157.5	1604.3	433.2	365.0
ForecastErrorTotalSolar	517.3	-987.5	2383.9	585.2	365.0
LoadForecastErrorFCR	-4132.2	-10727.9	6885.8	2629.9	365.0
Temperature	11.2	-2.5	29.0	6.0	365.0

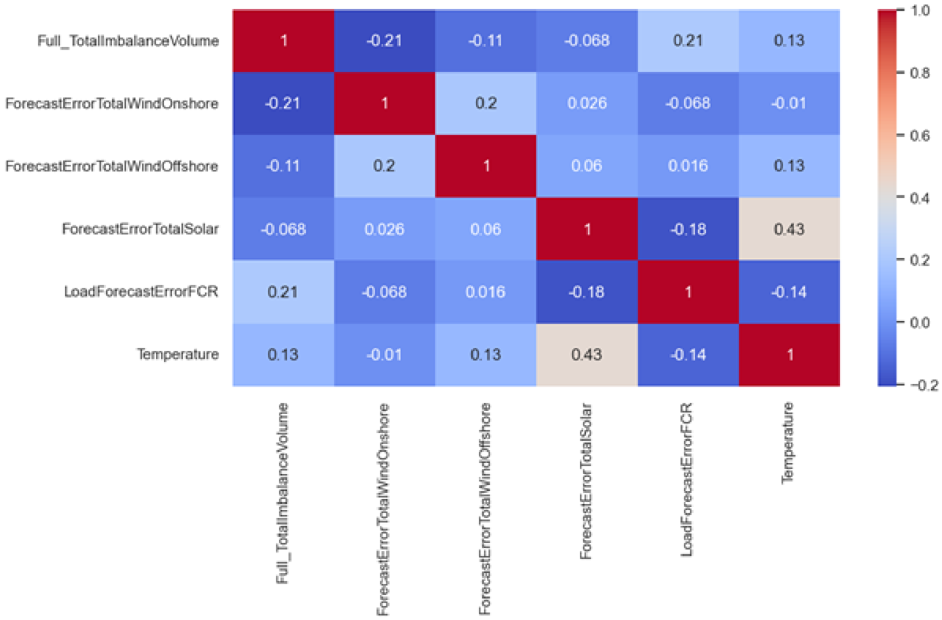

Figure 5.7: Correlation Matrix Imbalances FCR Cooperation

5.5.4. Expected results

Since we have extended the scope of the modelling for the imbalances to all the countries in the FCR cooperation, we assume that the total imbalance will increase. However, since more countries also mean a more diverse and complex dataset, the modelling results may be less predictive than on a smaller scale.

5.5.5. Modelling Results

Table 5.13: Imbalances FCR Cooperation: Model Parameters and Performance Metrics

Model	Dependent Variable	Start Date	End Date	RMSE	MAPE
(1,1,1)(1,0,1, 7)	Total Imbalance	01-2019	01-2020	318.14	-101.54%

From the modelling results, we can see that the results for the imbalance model for the FCR cooperation yield larger error scores than those for the Netherlands alone. Furthermore, from a visual inspection of the model validation graph below, we can see that the model is also not able to accurately predict peaks and does not follow the trend completely.

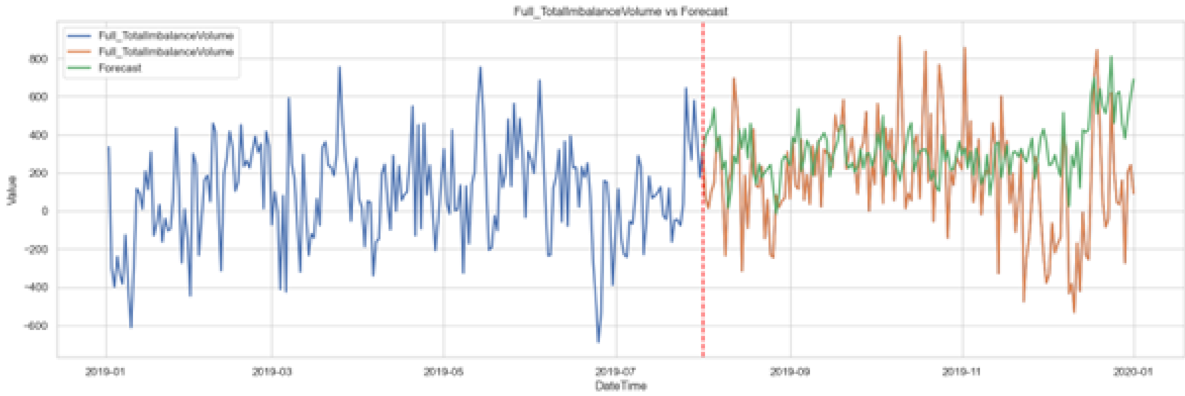

Figure 5.8: Imbalances FCR Cooperation: Model validation results

Table 5.14: Imbalances FCR Cooperation: Coefficients and Statistical Metrics

Variable	coef	std err	z	P>|z|	[0.025	0.975]
ForecastErrorTotalWindOnshore	-0.0420	0.014	-2.968	0.003**	-0.070	-0.014
ForecastErrorTotalWindOffshore	-0.0341	0.039	-0.869	0.385	-0.111	0.043
ForecastErrorTotalSolar	-0.1369	0.038	-3.607	0.000**	-0.211	-0.063
LoadForecastErrorFCR	0.0377	0.010	3.802	0.000**	0.018	0.057
Temperature	-0.0022	7.413	-0.000	1.000	-14.531	14.527
ar.L1	0.2136	0.095	2.244	0.025	0.027	0.400
ma.L1	-0.9297	0.040	-23.191	0.000	-1.008	-0.851
sigma2	4.547e+04	4990.984	9.110	0.000	3.57e+04	5.52e+04

Looking at the coefficients from the model, however, we can see that several variables that were found to be insignificant for the NL scenario are now significant, thus indicating better model performance. Further inspection of residuals (see appendix) indicated that the model has correlated errors and that some variability is not accounted for.

5.6. FCR activation Netherlands Modelling

5.6.1. Introduction

For the analysis of the activated FCR in the Netherlands, we looked at the time period in which dependent variables were available. FCR activation implies that a signal has been obtained from the TSOs, and the balancing capacity has been activated. This is the following step and indicates a more complex system than the system causing imbalances.

Table 5.15: Activated FCR NL: Description of dependent variable

Type	Relevance	Format	Resolution	Source
Activated FCR Capacity	The activated FCR capacity is a measure of activated MWh of capacity	MWh	Daily	ENTSO-E

5.6.2. System description

Similar to the system description from the previous sections, we now focus on the signal that the BSPs obtain from the TSO to activate FCR capacity. This works the same as the other systems but also is dependent on the activation of the other players.

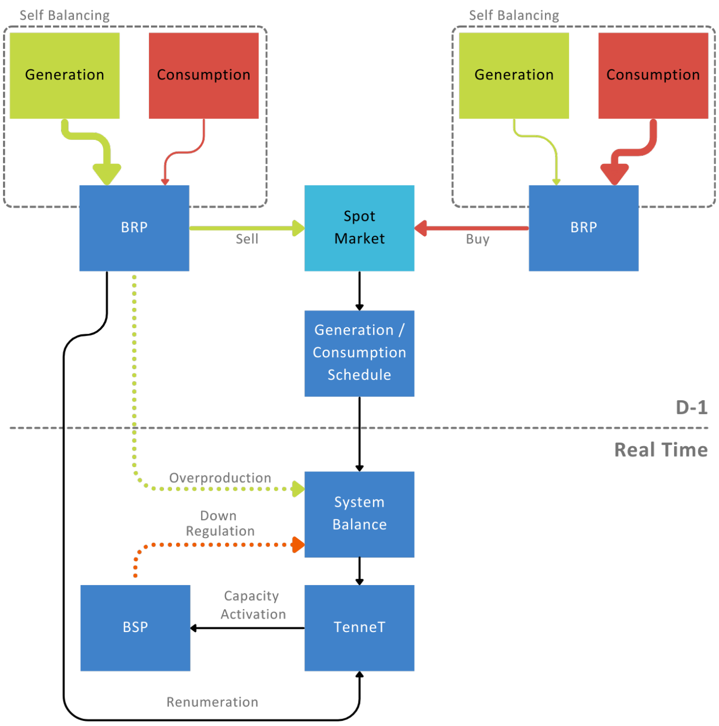

Figure 5.9: Activated FCR NL: System framework

5.6.3. Data selection and descriptive statistics

Similarly to the imbalance created we have selected variables from the final dataframe that could cause the system to be out of balance. For the period of study here, we have the following

Table 5.16: Statistical summary of variables

Variable	Mean	Min	Max	Std. Dev.	#observations
ForecastErrorTotalWindOnshore	297.3	-3612.7	6268.8	1283.4	365.0
ForecastErrorTotalWindOffshore	-189.0	-1939.9	1185.0	557.3	365.0
ForecastErrorTotalSolar	558.7	-1641.5	2833.1	692.9	365.0
LoadForecastErrorFCR	-1792.1	-10552.6	5395.1	2540.9	365.0
Temperature	10.7	-6.0	27.0	6.4	365.0

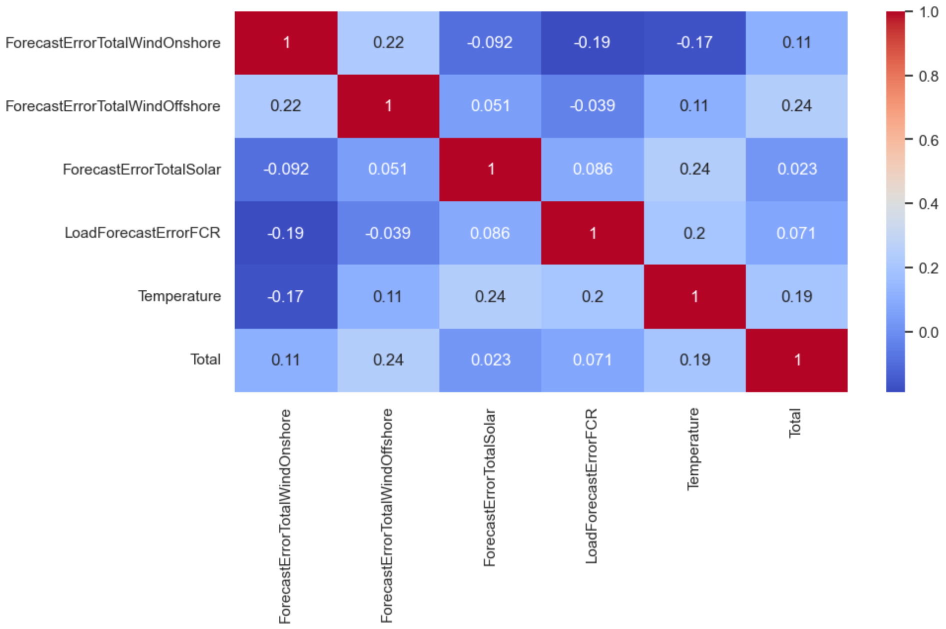

Figure 5.10: Activated FCR NL: Correlation Matrix

5.6.4. Expected results

Given the fact that the FCR activation in the Netherlands is a more complex system than the imbalance creation system we expect that the model performance shows that.

5.6.5. Modelling results

Table 5.17: Activated FCR NL: Model Parameters and Performance Metrics

Model	Dependent Variable	Start Date	End Date	RMSE	MAPE
(6,1,0)(0,0,1,7)	FCR Quantity NL (Activation)	07-2020	07-2021	0.94	7.08%

Interpreting the modelling outcomes in terms of RMSE and MAPE, the scores seem to indicate a model that would perform well in explaining the FCR activation. However, visually inspecting the model performance indicates that there are several behaviours we are not able to capture here as the model remains at a stable level while real life indicates an upward trend

Figure 5.11: Activated FCR NL: Modelling outcome

Table 5.18: Activated FCR NL: Parameter Estimates

Parameter	Coef	Std Err	z	P>|z|	[0.025, 0.975]
ForecastErrorTotalWindOnshore	-6.616e-05	0.000	-0.514	0.607	[-0.000, 0.000]
ForecastErrorTotalWindOffshore	0.0003	0.000	0.747	0.455	[-0.000, 0.001]
ForecastErrorTotalSolar	0.0001	0.000	0.516	0.606	[-0.000, 0.001]
LoadForecastErrorFCR	-5.134e-05	0.000	-0.491	0.623	[-0.000, 0.000]
Temperature	-0.0752	0.074	-1.017	0.309	[-0.220, 0.070]
ar.L1	0.5132	0.165	3.112	0.002	[0.190, 0.836]
ma.L1	-0.8970	0.160	-5.616	0.000	[-1.210, -0.584]
ar.S.L7	0.0826	3.213	0.026	0.979	[-6.214, 6.379]
ma.S.L7	-0.0260	3.225	-0.008	0.994	[-6.346, 6.294]
sigma2	1.8458	0.149	12.414	0.000	[1.554, 2.137]

Additionally, the modelling parameter estimates do not indicate any significant variables that can help explain the FCR behaviour other than the internal variables that are purely based on the model learning from the data. Therefore, we can conclude that this model is also unfit for further intervention analysis.

5.7. Reflection on electricity market modelling

The results in this section seem to indicate that data availability determines model performance. Some important data is not available for balancing markets, but the creation of imbalances seems to be better explainable. Modelling Day Ahead markets is possible with mostly open-source data and yields interesting results. We will use this model for intervention studies. Using the SARIMAX modelling techniques is largely dependent on stable system conditions, and the modelling performance starts to deteriorate when the underlying system changes. Therefore, it is important to acknowledge that long-term modelling in changing complex systems is more difficult with ARIMA models. The imbalance model worked better for the Netherlands than for the FCR cooperation. However, neither model performed sufficiently well in the intervention analysis. It is assumed that the difference in their performance is due to the fact that dataset for the Netherlands is more complete than that for the FCR cooperation. This indicates a data issue more that a modelling issue. It is expected that, when both datasets are at a same level of completeness, the models would perform at a similar level.

6 Intervention Analysis

From the modelling chapter we found the models that are able to do intervention analysis on to be the day ahead market. This chapter will study interventions that might have had an impact on the day-ahead market. First, these interventions will be defined; afterwards, the found interventions will be studied in regard to the models created in the previous chapter. Additionally, we will hypothesize about the effects of the interventions. Lastly, the interventions will be analysed and their sensitivity will be studied. Doing this will answer sub-question 3 and will help in the understanding of the implications for policymakers, forming the basis for sub-question 4.

6.1. Intervention models

From the previous section we have found that the DA model was able to capture the behaviour of the DA market moderately well. Modelling the imbalance and the FCR activation, while interesting from a modelling perspective did not indicate sufficient explainability to do intervention analysis on.

As mentioned in chapter 3, the intervention analysis as performed in this research consists of adding a dummy variable to the model that shifts from a 0 to a 1 with a step function at the time of the intervention. Therefore it is required that the intervention has a discrete point in time.

6.2. Intervention selection

As indicated in Chapter 2, there have been several significant events on the DA market in terms of design and large external events (shown in fig 2.12). Additionally, there were constraints on the time-frame for which data was available for the DA modelling. Therefore, we have selected the following three interventions to study.

Table 6.1: Interventions selected for DA intervention analysis

Intervention	Date
Market coupling of Greece to SDAC	15th of December 2020
Market coupling of Poland, Czech Republic, Romania, Slovakia and Hungary	17th of June 2021
Russian invasion of Ukraine	24th Feb 2022

With these interventions found we can start to make hypotheses about the impact that they might have on the day-ahead price in the Netherlands.

6.2.1. Hypothesis 1

Intervention: Coupling of the Greek market to the SDAC.
Hypothesis: The coupling of the Greek market will lead to a decrease in the Dutch day-ahead electricity price due to increased competition and cross-border electricity trade.
Reasoning: Market coupling is expected to enhance competition and optimize resource allocation, thereby lowering prices.

6.2.2. Hypothesis 2

Intervention: Coupling of Poland, Czech Republic, Romania, Slovakia, and Hungary to the common market.
Hypothesis: The coupling of these markets will further decrease the Dutch day-ahead electricity price due to a larger integrated market and improved supply diversity.
Reasoning: Expanding the integrated market increases supply options and competition, likely reducing prices.

6.2.3. Hypothesis 3

Intervention: Russian invasion of Ukraine.
Hypothesis: The Russian invasion of Ukraine will lead to an increase in the Dutch day-ahead electricity price due to geopolitical instability and disruptions in energy supply.
Reasoning: Geopolitical events can create supply shocks and increase market uncertainty, driving prices up.

The following section will describe the results of adding a dummy variable at the given date of the intervention along with a sensitivity analysis on the intervention date and on the modelling period. This will help to see the robustness of the intervention effect.

6.3. Intervention analysis

As mentioned, we have modelled the DA price similarly to chapter 5 and this section will explain on their effects and on their reliability.

Intervention 1: Coupling of Greece

From the results, it seems like the coupling of the Greek market to the SDAC has had a correlation with a higher price. In reality, it must be said that there are likely several factors that largely impact the validity of this claim. The main one is the validity of the model and the method of modelling the intervention. Firstly, at the time of the intervention, late 2020. Electricity prices were rather stable (low volatility). Our model only computes mean prices and does not have a volatility component; thus, it performs better in times of low volatility. Secondly, due to data limitations, the model does not take into account the effect of import and export on the day-ahead price. This is likely the main method in which the coupling of the Greek market impacts the Dutch day-ahead price.

Table 6.2: Coefficients and P values for specific dates in December 2020

Date	Coef	P value
20-12-2020	-0.99	0.729
19-12-2020	2.25	0.283
18-12-2020	-4.42	0.056
17-12-2020	-6.84	0.000
16-12-2020	4.75	0.023
15-12-2020	3.69	0.018
14-12-2020	1.03	0.645
13-12-2020	3.66	0.134
12-12-2020	-2.65	0.217
11-12-2020	-4.70	0.022
10-12-2020	-5.11	0.002

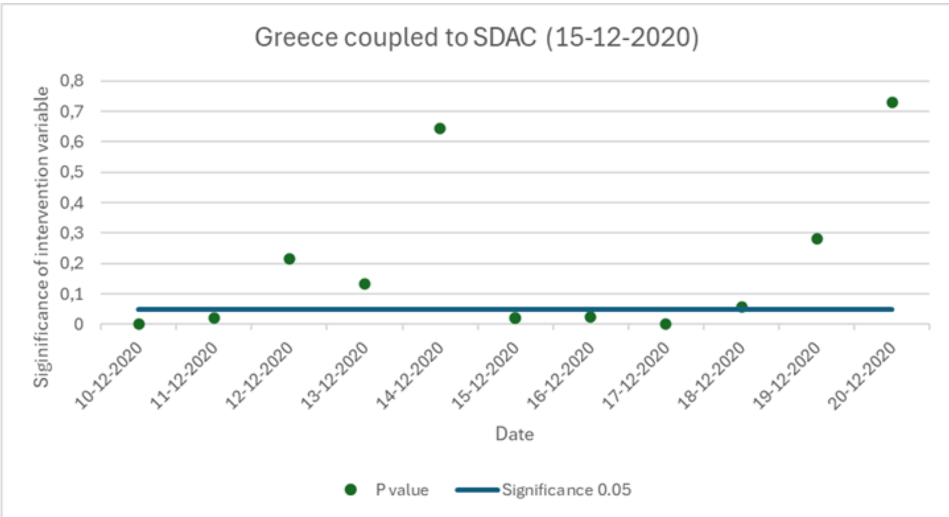

Figure 6.1: Intervention 1: Significance sensitivity analysis

The table indicates that the intervention had a significant effect on the DA price at the time of the intervention and that the price had increased. This opposes the hypothesis that prices would decrease.

Intervention 2: Coupling of Poland, Czech Republic, Romania, Slovakia and Hungary

Results: On June 17th, 2021, several countries joined the single market for day-ahead auctions in Europe. To check if this has had a direct impact on the price of electricity, we modelled the intervention on our existing model and did not see a significant impact.

Table 6.3: Coefficients and P values for specific dates

Date	Coef	P value
12-6-2021	-7.6379	0.656
13-6-2021	-13.3031	0.388
14-6-2021	20.2781	0.135
15-6-2021	-0.7394	0.977
16-6-2021	5.5631	0.822
17-6-2021	-9.0741	0.639
18-6-2021	2.4897	0.921
19-6-2021	2.6024	0.893
20-6-2021	-7.3208	0.581
21-6-2021	18.5031	0.367
22-6-2021	7.3342	0.727

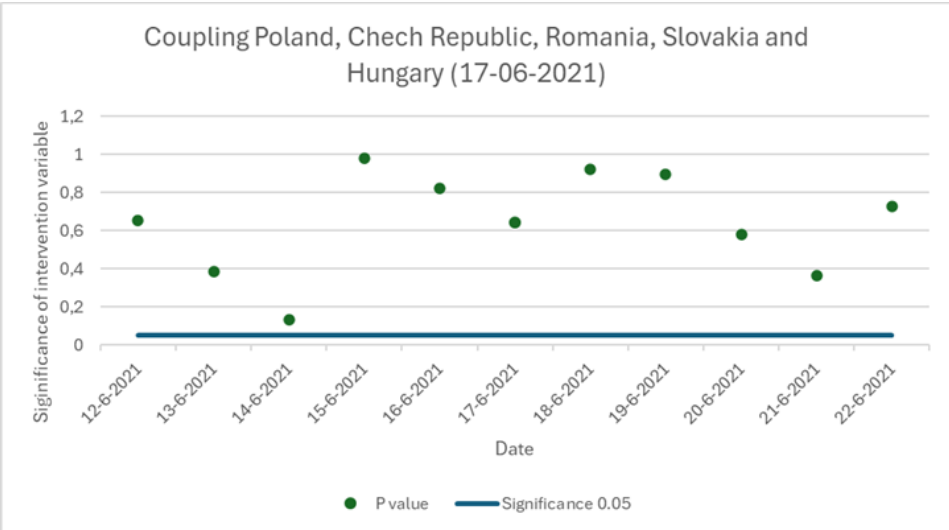

Figure 6.2: Intervention 2: Significance sensitivity analysis

We can see that the significance of this intervention does not reach below P<0.05, and thus, we cannot assume this intervention has had a significant impact on the day ahead price at that moment.

Intervention 3: Russian invasion of Ukraine

On the 24th of February 2022, Russia invaded Ukraine, this event has had a large cascade of effects on energy prices throughout Europe. Here we tested the effect of the intervention on the Day Ahead Price prediction model made earlier by adding a step function to the external regressors. Before the time of the intervention, the variable was set to 0, afterwards to 1.

Table 6.4: Coefficients and P values for the given dates

Date	Coef	P value
19-2-2022	-55.62	0
20-2-2022	16.0377	0.025
21-2-2022	30.2334	0
22-2-2022	6.0299	0.547
23-2-2022	0.51	0.947
24-2-2022	-1.49	0.805
25-2-2022	-3.75	0.453
26-2-2022	50.72	0
27-2-2022	-31.91	0
28-2-2022	13.15	0.06
1-3-2022	0.73	0.101

Figure 6.3: Intervention 3: Significance sensitivity analysis

The table seems to indicate that there is not a significant impact of the invasion of Russia on the electricity price on the 24th of February. However, looking a little further we can see that there is a significant impact which is positive at first and negative after. It can be speculated that this is due to the fact that at the moment of the invasion the day ahead prices already had been determined at which they were fixed.

6.3.1. Intervention limitations and reflections

During the intervention modelling we have seen that the outcome is very sensitive to changing the date of intervention. Additionally, taking this approach assumes a very quick responding system, the moment the intervention takes place, it is reflected in the price. Reality might prove a little slower and thus this can provide a basis for further research seeing if it is possible to spot trend changes after an intervention. Furthermore, model is also very sensitive to timeframe of modelling in general. This is expected due to the fact that there are systematic changes over a long timeframe that are not captured after the model is trained.

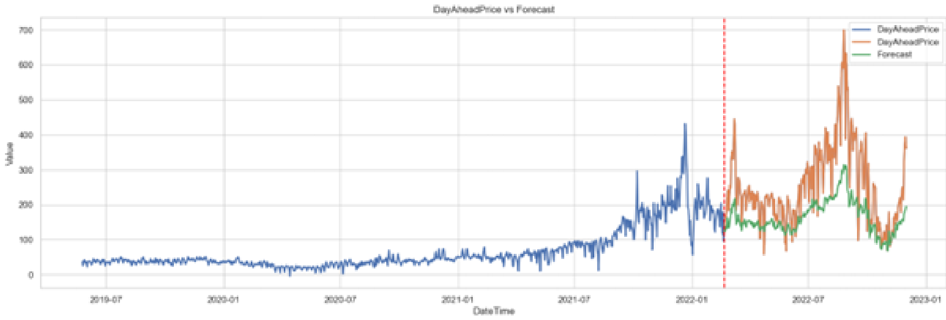

Figure 6.4: Timeframe 1: Training to early 2022

Looking at the fit of the curve for the day-ahead price, we can see that the prediction follows the real value rather closely in Figure 6.4. If, however, we move the training period forward as in Figure 6.5, thus limiting the data that the model includes, we see that the prediction becomes less able to capture the large spikes in the end.

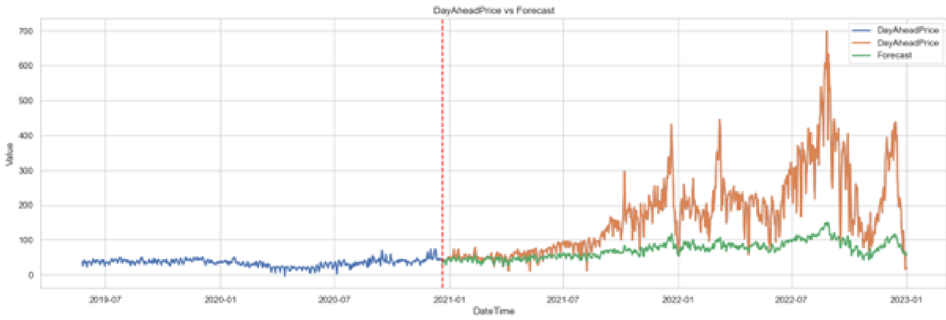

Figure 6.5: Timeframe 2: Training to late 2020

These indicate that the model is very sensitive to the training period which also impacts the interpretation of the intervention outcomes. Further, results will be discussed in the following chapter, after which the thesis will be concluded.

6.4. Implication for policy makers

To answer subquestion 4, the results from this chapter need to be translated into practical results and applicable results for policymakers. First, this chapter shows that not only is it important to reflect on the performance of the model, but it is also important to reflect on the method in which an intervention is modelled. In this chapter, we have modelled interventions as a binary on/off switch, having its full effect from the day of implementation. This approach does not account for transient effects, which is something policymakers should take into account when applying similar methods. An intervention (or policy) might have delayed effects, and the current implementation does not account for this. Furthermore, the results of this research show the importance of data availability and, most of all, uniformity. There have been several policies in place to allow for uniform data collection of the main energy market variables by ENTSO-E, but many of these are not strictly adhered to, which makes it hard to compare different countries and TSOs. One important takeaway is thus that TSOs and regulators should take a close look at the data that they publish and make sure they all adhere to the same standard. Additionally, a large takeaway from this chapter is that models can be very sensitive to their inputs and the method in which they are trained. Models can be a useful tool for the policymaking process as they allow for seemingly clear conclusions to be drawn from large datasets. However, looking at the sensitivity some of these models have to the way the interventions are modelled, for example, they must be perceived with a certain level of caution before policymakers draw conclusions.

7 Discussion

7.1. Introduction of Chapter

This study looked into the analysis of large-scale interventions in electricity markets during system changes. The primary goal was to gain insight into the relationship between electricity market outcomes such as prices and volumes and policy interventions. This was done to see if, with quantitative analysis alone, effects on market outcomes could be attributed to market interventions. To achieve this, several sub-questions needed to be answered, starting with the identification and definition of markets and their changes. Afterwards, a focus was put on modelling the specific markets with the aim of providing a model that could serve as a reference. Finally, with knowledge of market interventions and a reference model, these could be combined to see if significant changes in prices and volumes of these markets could be attributed to the interventions. The main findings are on modelling imbalance markets and on intervention analysis in electricity markets. One of these findings is that modelling imbalances on a Dutch level yielded better results than on an international level. Depending on the particular situation, this was not the expected outcome, as it was assumed that a larger model incorporating more countries could more accurately reflect the interconnected electricity grid. Furthermore, we have observed that several factors, such as forecasting errors, significantly correlate with system imbalances. Additionally, we have found that a simple intervention analysis, as performed here, can give clear results on the impact of the intervention. However, the validity of these results can be investigated further. In this chapter, we'll start by answering the sub-questions. Then, we will reflect on the methods used to analyze electricity markets and explore the limitations of the findings and outcomes of the study. Thereafter, we will look into the contribution and further research.

7.2. Methodological Reflection

We have chosen to perform an interrupted time series analysis due to its ability to capture the impact of interventions in the fields of economics and medicine. ITSA, in combination with an ARIMA model, was chosen as it allows for the attribution of effects and the ability to test hypotheses. The primary objective of the SARIMAX model was to create two scenarios, one with the intervention and one without. Within the model, the intervention was described as a step function, changing from 0 to 1 when the intervention took place. To apply this methodology, significant assumptions had to be made.

Firstly, SARIMAX, and ARIMA modelling in general, assumes a black box, meaning that the relationship between the input and output variables can be modelled one-on-one. In reality, the underlying systems that determine DA prices and imbalances are not linear and change over time. This has had implications on the ability to predict the outcome of the system on a larger timeframe. The way the electricity markets worked was fundamentally different in 2018 than it was in 2023. SARIMAX, however, interprets all the points in time with the same system relations. This has also been observed in the intervention analysis chapter, where we can see the effects that a different modelling timeframe has had on the performance of the DA model. Furthermore, the assumption of a step intervention simplifies the intervention to nothing more than a simple on or off. Assuming an intervention is a simple on or off moment does not capture the whole reality and assumes the system to change significantly overnight. We can assume that market participants anticipate regulatory and other interventions and thus already have an impact before their implementation date. Additionally, the effect of an intervention could also be a delayed effect, in which case our analysis will not attribute the effects.

To summarize, we have found, in accordance with other research, that ARIMA models have difficulties when system relations change. We have tried to account for this systematic change by correcting for exogenous variables and adding intervention effects, which have improved the results.

7.2.1. Method Determination

Electricity markets are complex systems characterized by many interdependencies and non-linear interactions. Therefore, we needed a method of studying intervention effects that could capture this. The ITSA approach was chosen as it greatly simplified the analysis by looking at the change in the dependent variable only after the intervention. There are also other methods to evaluate the impact of an intervention quantitatively. One way is to perform a randomized control trial, with which the attribution of effects is stronger than with interrupted time series analysis. Because we look at historical data and the inability to have a 'non-treatment' group, the evidence from an ITSA is weaker, but the method is much more applicable for large-scale system analysis.

Interrupted Time Series Analysis

Interrupted Time Series Analysis was selected as a primary method due to its use in quantitatively assessing the impact of interventions across various fields such as economics, healthcare, and social sciences. In this study, ITSA was instrumental in identifying shifts in electricity market outcomes, such as price and imbalances following large significant interventions. By establishing a reference of market behaviour prior to the intervention, ITSA allowed for the clear observation of the intervention's impact.

SARIMAX Modelling

The SARIMAX model extends the traditional ARIMA approach by incorporating seasonal effects and exogenous variables, making it useful for the complex environment of electricity markets. SARIMAX modelling was also chosen for its interpretability of the results. A clear regression table gives the coefficients of all exogenous variables and their significance on the dependent outcomes. Other methods, such as artificial neural networks, are not as clear in their attribution of the outcomes to input variables but can perform better for the modelling task. This could be due to the fact that ARIMA models are linear, which makes for clear attribution of outcomes but potentially sacrifices performance.

Initially, when modelling balancing markets only, the SARIMAX model did not perform sufficiently to explain market behaviour. Once testing was done on less complex systems, such as imbalance creation or the day ahead market price, it was concluded that the SARIMAX method is applicable to this research. Traditionally, these models are mostly used to predict the future state of a system, given estimates of external inputs. In this research, however, SARIMAX modelling has been used to attribute systematic change to the intervention variable. This approach has not been explored before in the context of the Dutch electricity market, which is one of the main contributions but also one of the main limitations as there is not a large body of literature against which to compare modelling techniques and performance.

7.3. Data Sources

The quality of the analysis in this research heavily depended on the completeness, reliability, and comprehensiveness of the dataset. While significant time was spent on the selection of data sources to ensure that they provided the necessary temporal coverage and completeness, gathering a dataset that allowed for international comparison proved difficult. Additionally, some data was not included that literature indicates has a role in these systems such as cross border trade and transmission.

7.3.1. Aggregation

Additionally, there were constraints in the number of data points the SARIMAX model could work with. As the goal of this study was to account for changes over a long-term time horizon, data that is available on a 15-minute scale had to be aggregated to be able to be modelled. Choosing this scale was one of the main dilemmas as aggregation sacrifices completeness and also loses details that could be critical, especially for a process such as balancing capacity activation. This also serves as one of the potential explanations of the suitability of the DA prices to be modelled in this way, as this process is more in line with the dataset resolution.

7.3.2. Completeness of Dataset

During initial research, it became clear that modelling electricity markets, especially balancing markets, often depends on variables that are not publicly available. In this case, we were thus limited in explaining the balancing markets' behaviour. While forecasting errors are a strong predictor of imbalances, there are also others that have not been included in this analysis.

The internal completeness of the dataset was particularly important for the ITSA and SARIMAX models, which rely on continuous time series data. Gaps in the data disrupt the ability to accurately identify trends and seasonal patterns, leading to potential inaccuracies in the results. Therefore, assumptions needed to be made for some of the datasets in the case of missing data. This process involved verifying that all relevant variables were present across the entire dataset, with no systematic gaps or inconsistencies. Consistency checks also involved cross-referencing data from different sources to ensure accuracy. For instance, market price data from different reporting agencies were compared to ensure that they matched and that there were no discrepancies, in particular with solar generation.

7.4. Research Limitations

Given the nature and scope of the research, several limitations must be acknowledged.

7.4.1. Limitations in Knowledge of Electricity Markets

The electricity markets studied in this research are highly complex and constantly evolving. While the research attempted to account for the major factors influencing these markets, the rapid changes in these markets, driven by technological advancements, policy shifts, and different consumer behaviour, means that our understanding is always somewhat incomplete.

This limitation is particularly relevant in the context of modelling market behavior over time. While the models used in this study were based on the best available data and knowledge, there are always unknown factors or future developments that change the observed relationships.

7.4.2. Data Availability and Quality

The analysis was constrained by the availability and quality of data. While significant efforts were made to obtain comprehensive and accurate data, some limitations were inevitable. For instance, certain variables may not have been recorded with the desired frequency or accuracy, and some data points may have been missing or incomplete. These limitations could potentially affect the precision of the model estimates and the robustness of the study's conclusions.

Data quality issues could arise from various sources, such as reporting errors, measurement inaccuracies, or delays in data collection. As part of European integration, TSOs have the responsibility to share their data publicly via the ENTSO-E platform; however, this is not always the case, and a lot of data is still missing or incomplete. Additionally, when doing such a long-term study, there is the risk of the data recording processes changing during the studied timeframe. This became particularly clear when trying to compare balancing capacity contracts with monthly contracts to those with 4-hour contracts. These issues were addressed as much as possible through data cleaning and validation processes, but uncertainty remains. Additionally, the availability of historical data was sometimes limited, particularly for newer markets or collaborations.

7.4.3. Validation of Model and Results

Another challenge was validating the model and its results. While various diagnostic tests were performed to assess the model's performance, the inherent uncertainties in predicting market behavior mean that the results should be interpreted with caution. While some references exist for when a model performs well, determining a clear-cut point for when a model performs sufficiently remains a process of trial and error.

7.5. Contribution to Knowledge Gap

This study contributes to filling a critical gap in the existing literature by providing a quantitative analysis of the impact of large-scale interventions in electricity markets. By applying ITSA and SARIMAX modelling, this research attempts to close the gap between long-term effects and short-term markets. The findings highlight the importance of exogenous factors when assessing the impact of market interventions. Furthermore, this study shows the importance of timeframe selection when using SARIMAX models and the limitations of this approach. This is particularly relevant for policy researchers as they can significantly impact the significance and sign of the intervention effect. This is in line with the findings by Li et al. (2021) that the method chosen to study interventions largely determines the outcome. A similar approach as this research was found in the paper by Papaioannou et al. (2018) on detecting intervention effects in the Greek DA market. This paper, however, did not have a clear sensitivity analysis of the interventions, which led to strong attribution of effects to the particular policy intervention. During this research, we have seen intervention effects change significantly, with the slightest change occurring in initial conditions. Additionally, the research highlights the need for ongoing monitoring and analysis of electricity markets to understand the long-term impacts of policy changes and to make adjustments as necessary.

8 Conclusion

This final chapter of the thesis is meant to provide concluding remarks. It provides a final conclusion with an answer to the main research question. A section on this study's academic and societal relevance is provided. Finally, a personal reflection ends this thesis.

8.1. Conclusion

The main question posed in this thesis was:

> *"What has been the quantitative impact of market interventions on electricity markets in the Netherlands?"*

This question was answered by first identifying a large number of interventions that have taken place during the 2018-2023 time period as well. These showed that several large-scale market design changes have taken place in both the balancing and day-ahead markets which was outlined in Chapter 2. Furthermore, to measure the quantitative impact, SARIMAX models were created in Chapter 5 by using the methodology and data from Chapter 3 and 4. These showed that even though SARIMAX modelling is effective for day-ahead market modelling, currently, the scale of the model and the corresponding dataset make it unable to capture imbalance market behaviour. It was attempted to model imbalance creation but this also showed mixed results. The model that was deemed suitable, that of day-ahead prices for electricity in the Netherlands was analysed by applying several interventions to the model and checking for its significance and checking for sensitivity. This showed that some interventions had a significant impact on the price of electricity. All in all, this thesis has shown that the coupling of Greece to the SDAC has had significant impact on electricity markets in the Netherlands, namely, it raised the price by 3.69. Directly, after saying this however, it must be noted that with all the modelling assumptions, this result's validity can be heavily disputed. The fact that the interventions that were modelled did not account for many external influences or internal system parameters and showed a heavily simplified system.

In conclusion, this study has provided a detailed analysis of the impact of large-scale interventions on electricity markets, focusing on the Dutch day-ahead and balancing markets. Through the use of ITSA and SARIMAX modelling, the research has shown that policy interventions can have significant effects on market outcomes. However, these effects are influenced by various factors, including data quality, model assumptions, and the selection of exogenous variables. Despite its limitations, the study makes a contribution to our understanding of electricity markets and offers insights for policymakers and market participants who want to evaluate interventions. Firstly, the study finds that electricity markets have significantly changed in recent years and has tried to model the effects of these changes. Large-scale market integration brings more market participants but also makes the markets more complex to study, especially in a transitional phase. The study highlights the possibility of modelling electricity markets with limited datasets and shows the routes for improving these datasets based on the literature. It also gives insight into the complexity of studying market interventions, which bridges several domains by looking at price formation more systematically. Overall, this research demonstrates the value of using advanced quantitative methods to analyze electricity markets and provides a solid foundation for future studies. As electricity markets evolve, ongoing research will be essential to ensure that policy interventions are effectively evaluated to promote market stability, efficiency, and sustainability.

8.2. Academic relevance

Referring to the initial research gap in the quantitative evaluation of electricity market policy, this thesis presents a novel perspective on evaluating such markets. The interrupted time series analysis concept had not been applied in the context of electricity markets before and presents academic novelty. However, with the points from the discussion in mind, the application of SARIMAX models for policy evaluation in the context of this research might not directly be a method to be pursued further due to the long timeframe of the policy intervention and the very short timeframe of electricity markets being that divergent that the SARIMAX method is not capable of handling the number of data points that the evaluation would require. With this in mind, the thesis has provided a multidisciplinary view in several academic fields, including economics, engineering, data science, and public policy. It has shown how advanced statistical tools can be leveraged to address economic questions about market design, providing valuable insights to economists, engineers and policymakers engaged in market regulation and energy systems management.

8.3. Societal relevance

The findings of this research carry some societal implications, particularly in the context of market efficiency, policy implications, and the ongoing energy transition. By analyzing the impacts of market interventions on the Dutch electricity markets, this study shows the role that regulatory measures play in market performance and maintaining the grid's reliability. Stable and efficient markets are essential for ensuring that electricity prices remain predictable and affordable, benefiting both consumers and businesses. Furthermore, the insights gained from this research support the broader energy transition by identifying policies that facilitate the integration of renewable energy sources, thus reducing dependency on fossil fuels. For policymakers, this study provides the starting points for evidence-based guidance on how to design interventions that change market dynamics, ultimately contributing to a more resilient, sustainable, and secure energy system.

8.4. Further research

The limitations identified also provide starting points for further research. Further research into modelling the long-term behaviour of balancing markets can be critical if an ITSA is to be performed with other techniques or datasets to better capture the complexities of market behaviour. Furthermore, due to the limited computational ability of the SARIMAX model, research using larger and more refined datasets could offer a clearer picture, particularly for balancing markets. While ITSA has not been very common in electricity market modelling, it has been effectively used in other markets. Applying this framework to other electricity markets could offer insights into how different policies are implemented in different countries.

8.5. Personal reflection

When I applied for the internship position at TNO, I had only a basic understanding of electricity markets. As mentioned in the preface, my interest was sparked by global events, and I was eager to learn more about the future of the electricity market in Europe. Due to my interest in electricity markets, TNO and I decided that the initial focus of the research would be on balancing market development. At the very start, I had not heard of balancing markets before, let alone know how they operate and how to analyse them. Knowing I wanted to take a quantitative approach to analysing electricity markets. Balancing market research was also a niche within TNO; thus, I got a lot of freedom to pick an appropriate approach. After looking into several other quantitative methods, ARIMA modelling was one of the only methods on which previous intervention analyses of electricity markets had been based. However, with the specificity of the question I wanted to ask, little literature was present that took a similar approach. Reflecting on this, taking on a question that had not been studied with this approach before was a great challenge. On the other hand, this also meant a lot of trial and error. If I could start this thesis again, I would have taken a different approach, focusing and scoping the topic further. Additionally, I would focus less on the method's application and more on the research's implications, accepting the inevitable gaps in methodological knowledge that a project of this size brings. Furthermore, this project has shown me the importance of data quality, assumptions while modelling and the limitations of quantitative methods.

During this thesis, I greatly extended my knowledge on the subject and experienced what it is like to research for the first time. During this time, I found out that doing research is about more than just the topic itself.

References

Anchustegui, I. H. (2020). Regulation of electricity markets in Europe in light of the Clean Energy Package: Prosumers and demand response. In *Routledge Handbook of Energy Law*. Num Pages: 17. Routledge.

Autoriteit Consument en Markt (ACM). (2022, June). Goedkeuring latere toetreding TenneT tot aFRR en mFRR platforms | ACM.nl. Besluit. Retrieved August 11, 2023, from https://www.acm.nl/nl/publicaties/goedkeuring-latere-toetreding-tennet-tot-afrr-en-mfrr-platforms

Balansverantwoordelijken (BRPs). (n.d.). Retrieved May 16, 2024, from https://www.tennet.eu/nl/de-elektriciteitsmarkt/nederlandse-klanten/balansverantwoordelijken-brps

Brijs, T., De Vos, K., De Jonghe, C., & Belmans, R. (2015). Statistical analysis of negative prices in European balancing markets. *Renewable Energy*, *80*, 53–60. doi:10.1016/j.renene.2015.01.059

Bunn, D. W., Inekwe, J. N., & MacGeehan, D. (2021). Analysis of the Fundamental Predictability of Prices in the British Balancing Market. *IEEE Transactions on Power Systems*, *36*(2), 1309–1316. Conference Name: IEEE Transactions on Power Systems. doi:10.1109/TPWRS.2020.3015871

CBS. (2023, October). StatLine - Consumentenprijzen; prijsindex 2015=100. Retrieved October 10, 2023, from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83131NED/line?dl=30680&ts=1696941121720

CBS. (2024, March). Nearly half the electricity produced in the Netherlands is now renewable. webpagina. Last Modified: 2024-03-07T15:00:00+01:00. doi:10/nearly-half-the-electricity-produced-in-the-netherlands-is-now-renewable

Do et. al. (2024). Electricity market crisis in Europe and cross border price effects: A quantile return connectedness analysis. *Energy Economics*, *135*, 107633. doi:10.1016/j.eneco.2024.107633

Energy Regulatory Office. (2024, July). Electricity market: Europe is entering the next phase of single market integration. Retrieved June 7, 2024, from https://www.ure.gov.pl/en/communication/news/323,Electricity-market-Europe-is-entering-the-next-phase-of-single-market-integratio.html

ENTSO-E. (n.d.). Frequency containment reserves. Retrieved August 11, 2023, from https://www.entsoe.eu/network_codes/eb/fcr/

ENTSO-E. (2022). Single Day-ahead Coupling (SDAC). Retrieved September 16, 2024, from https://www.entsoe.eu/network_codes/cacm/implementation/sdac/

Eupean Parliament. (2023, March). Internal energy market | fact sheets on the european union | european parliament. Retrieved June 30, 2023, from https://www.europarl.europa.eu/factsheets/en/sheet/45/internal-energy-market

European Central Bank. (2022). The impact of the war in Ukraine on euro area energy markets. Retrieved September 16, 2024, from https://www.ecb.europa.eu/press/economic-bulletin/focus/2022/html/ecb.ebbox202204_01~68ef3c3dc6.en.html

European Commission. (n.d.). Electricity market design. Retrieved October 17, 2023, from https://energy.ec.europa.eu/topics/markets-and-consumers/market-legislation/electricity-market-design_en

European Commission. (2023, March). Reform of the EU electricity market design. European Commission - European Commission. Text. Retrieved August 15, 2023, from https://ec.europa.eu/commission/presscorner/detail/en/IP_23_1591

Gabrielli, P., Wüthrich, M., Blume, S., & Sansavini, G. (2022). Data-driven modeling for long-term electricity price forecasting. *Energy*, *244*, 123107. doi:10.1016/j.energy.2022.123107

Hirth et al. (2015). Balancing power and variable renewables: Three links. *Renewable and Sustainable Energy Reviews*, *50*, 1035–1051. doi:10.1016/j.rser.2015.04.180

Hu, J., Harmsen, R., Crijns-Graus, W., Worrell, E., & van den Broek, M. (2018). Identifying barriers to large-scale integration of variable renewable electricity into the electricity market: A literature review of market design. *Renewable and Sustainable Energy Reviews*, *81*, 2181–2195. doi:10.1016/j.rser.2017.06.028

IEA. (2023, October). Electrification - Energy System. Retrieved October 30, 2023, from https://www.iea.org/energy-system/electricity/electrification

Knittel, C. R., & Roberts, M. R. (2005). An empirical examination of restructured electricity prices. *Energy Economics*, *27*(5), 791–817. doi:10.1016/j.eneco.2004.11.005

Kraft, E., Keles, D., & Fichtner, W. (2020). Modeling of frequency containment reserve prices with econometrics and artificial intelligence. *Journal of Forecasting*, *39*(8), 1179–1197. _eprint: https://onlinelibrary.wiley.com/d doi:10.1002/for.2693

Lago, J., De Ridder, F., & De Schutter, B. (2018). Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Applied Energy*, *221*, 386–405. doi:10.1016/j.apenergy.2018.02.069

Lago, J., Marcjasz, G., De Schutter, B., & Weron, R. (2021). Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. *Applied Energy*, *293*, 116983. doi:10.1016/j.apenergy.2021.116983

Li, L., Cuerden, M. S., Liu, B., Shariff, S., Jain, A. K., & Mazumdar, M. (2021). Three Statistical Approaches for Assessment of Intervention Effects: A Primer for Practitioners. *Risk Management and Healthcare Policy*, *14*, 757–770. doi:10.2147/RMHP.S275831

Lucas, A., Pegios, K., Kotsakis, E., & Clarke, D. (2020). Price Forecasting for the Balancing Energy Market Using Machine-Learning Regression. *Energies*, *13*(20), 5420. Number: 20 Publisher: Multidisciplinary Digital Publishing Institute. doi:10.3390/en13205420

Meeus, L. (2020). *The evolution of electricity markets in Europe*. Accepted: 2020-12-15T13:08:01Z. Edward Elgar Publishing. Retrieved August 8, 2024, from https://cadmus.eui.eu/handle/1814/69266

Moitroux, O. (2020). Master's Thesis :€Improvement of decision making for trading in wholesale electricity market.

Mosquera-López, S., & Nursimulu, A. (2019). Drivers of electricity price dynamics: Comparative analysis of spot and futures markets. *Energy Policy*, *126*, 76–87. doi:10.1016/j.enpol.2018.11.020

Newbery, D., Pollitt, M. G., Ritz, R. A., & Strielkowski, W. (2018). Market design for a high-renewables European electricity system. *Renewable and Sustainable Energy Reviews*, *91*, 695–707. doi:10.1016/j.rser.2018.04.025

Newbery, D., Strbac, G., & Viehoff, I. (2016). The benefits of integrating European electricity markets. *Energy Policy*, *94*, 253–263. doi:10.1016/j.enpol.2016.03.047

Ocker, F., & Ehrhart, K.-M. (2017). The "German Paradox" in the balancing power markets. *Renewable and Sustainable Energy Reviews*, *67*, 892–898. doi:10.1016/j.rser.2016.09.040

Papaioannou, G. P., Dikaiakos, C., Dagoumas, A. S., Dramountanis, A., & Papaioannou, P. G. (2018). Detecting the impact of fundamentals and regulatory reforms on the Greek wholesale electricity market using a SARMAX/GARCH model. *Energy*, *142*, 1083–1103. doi:10.1016/j.energy.2017.10.064

Peixeiro, M. (2022). *Time series forecasting in Python*. Shelter Island, NY: Manning Publications Co. Retrieved December 1, 2023, from https://ieeexplore.ieee.org/book/10280601

Pepermans, G. (2019). European energy market liberalization: Experiences and challenges. *International Journal of Economic Policy Studies*, *13*(1), 3–26. doi:10.1007/s42495-018-0009-0

Petrella, A., & Sapio, A. (2012). Assessing the impact of forward trading, retail liberalization, and white certificates on the Italian wholesale electricity prices. *Energy Policy*, *40*, 307–317. doi:10.1016/j.enpol.2011.10.011

Poplavskaya et al. (2021). Making the most of short-term flexibility in the balancing market: Opportunities and challenges of voluntary bids in the new balancing market design. *Energy Policy*, *158*, 112522. doi:10.1016/j.enpol.2021.112522

Regelleistung.net. (n.d.). Www.regelleistung.net > European cooperations > FCR Cooperation. regelleistung.net. Retrieved August 1, 2023, from https://www2.regelleistung.net/en-us/European-cooperations/FCR-Cooperation

Roumkos, C., Biskas, P. N., & Marneris, I. (2021). Manual Frequency Restoration Reserve Activation Clearing Model. *Energies*, *14*(18), 5793. Number: 18 Publisher: Multidisciplinary Digital Publishing Institute. doi:10.3390/en14185793

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Pages: xxi, 623. Boston, MA, US: Houghton, Mifflin and Company.

Tarsitano, A., & Amerise, I. L. (2017). Short-term load forecasting using a two-stage sarimax model. *Energy*, *133*, 108–114. doi:10.1016/j.energy.2017.05.126

TenneT. (n.d.). Dutch ancillary services. TenneT. Retrieved July 18, 2023, from https://www.tennet.eu/markets/dutch-ancillary-services

Turner, S. L., Karahalios, A., Forbes, A. B., Taljaard, M., Grimshaw, J. M., & McKenzie, J. E. (2021). Comparison of six statistical methods for interrupted time series studies: Empirical evaluation of 190 published series. *BMC Medical Research Methodology*, *21*(1), 134. doi:10.1186/s12874-021-01306-w

van der Veen, R. A. C., & Hakvoort, R. A. (2016). The electricity balancing market: Exploring the design challenge. *Utilities Policy*, *43*, 186–194. doi:10.1016/j.jup.2016.10.008

Verbong, G., & van der Vleuten, E. (2004). Under construction: Material integration of the Netherlands 1800–2000. *History and Technology*, *20*(3), 205–226. Publisher: Routledge _eprint: https://doi.org/10.1080/073415 doi:10.1080/0734151042000287970

A

Data Gathering and analysis

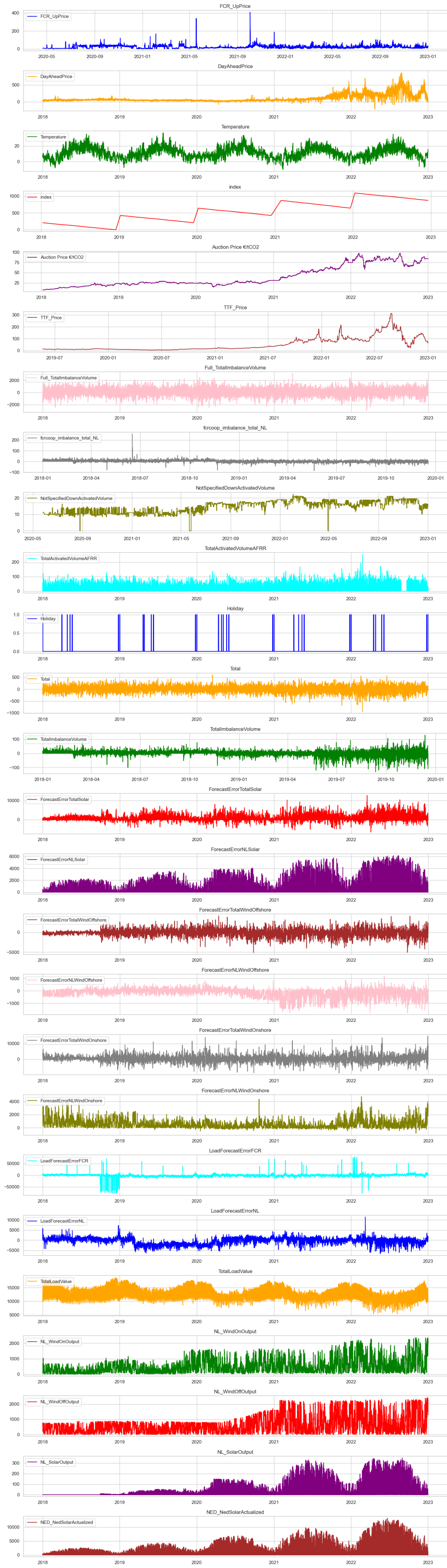

Figure A.1: Final Dataframe graphically

B

Electricity Market Modelling

B.1. Day-ahead price modelling

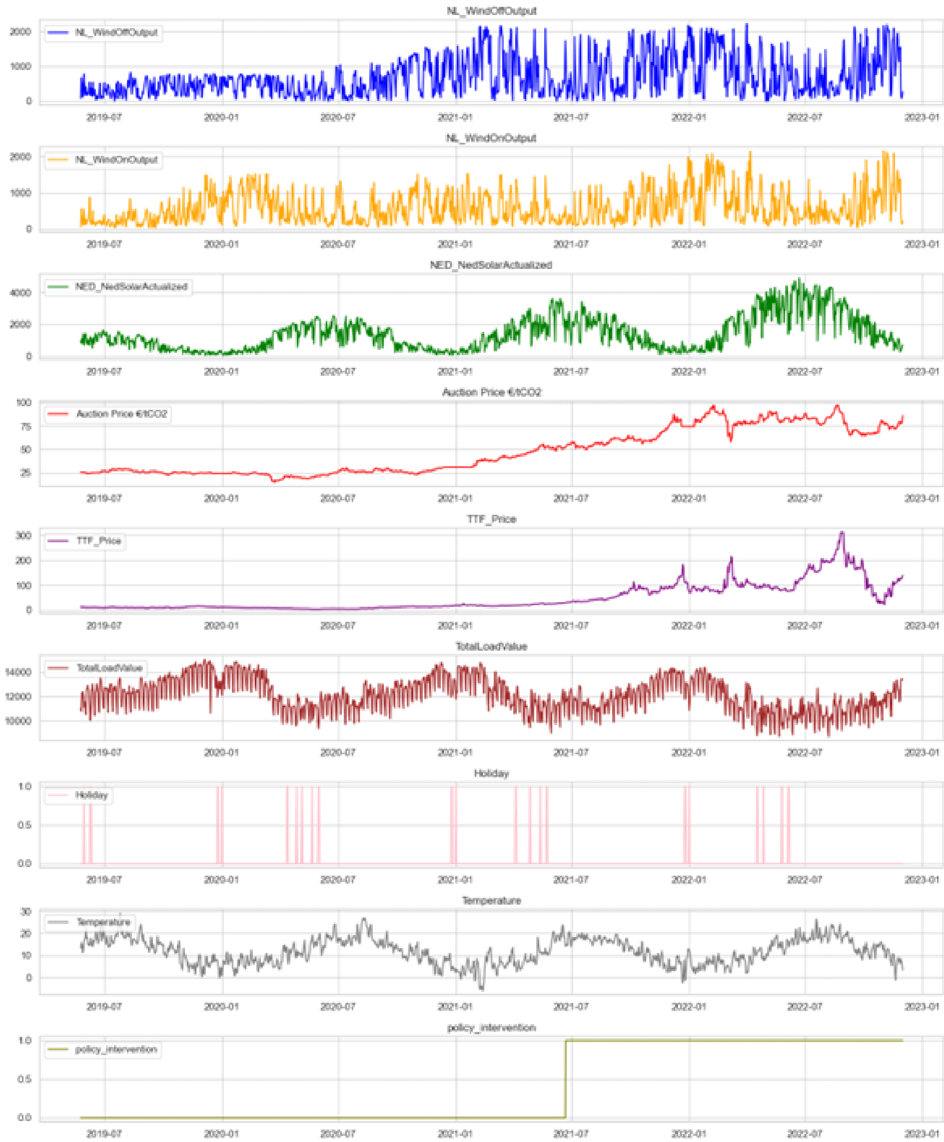

Figure B.1: Time series Day-ahead price

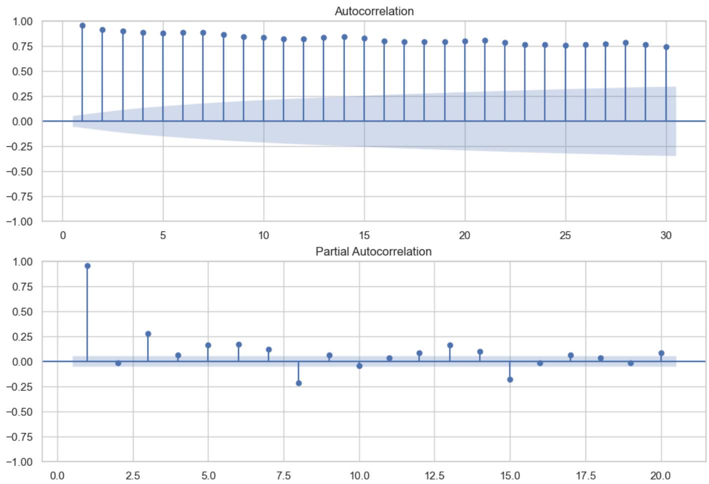

Figure B.2: DA Price, ACF and PACF plots

KPSS test

Table B.1: DA Price, Stationarity Test Results

d	ADF Stats	p-value	ADF Stationary	KPSS Stationary	Stationary
0	4.158837e-01	0.01	False	False	False
1	5.259926e-13	0.10	True	True	True
2	4.564303e-29	0.10	True	True	True

Table B.2: DA Price, additional model statistics

Statistic	Value	Statistic	Value
Ljung-Box (L1) (Q)	0.15	Jarque-Bera (JB)	14.27
Prob(Q)	0.70	Prob(JB)	0.00
Heteroskedasticity (H)	0.99	Skew	0.18
Prob(H) (two-sided)	0.96	Kurtosis	4.42

Table B.3: DA Price, Stationarity Test Results

d	ADF Stats	p-value	ADF Stationary	KPSS Stationary	Stationary
0	4.158837e-01	0.01	False	False	False
1	5.259926e-13	0.10	True	True	True
2	4.564303e-29	0.10	True	True	True

B.2. Imbalances in the Netherlands Modelling

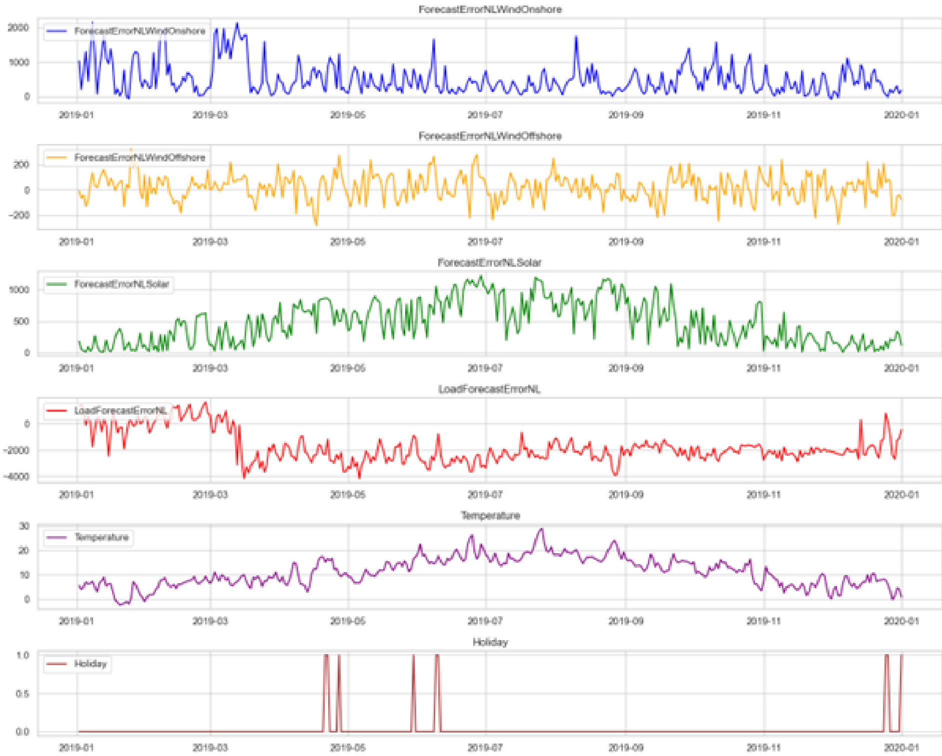

Figure B.3: Time Series Imbalances Netherlands

Table B.4: Model Diagnostic Statistics

Statistic	Value	Statistic	Value
Ljung-Box (L1) (Q)	0.16	Jarque-Bera (JB)	3.40
Prob(Q)	0.69	Prob(JB)	0.18
Heteroskedasticity (H)	1.07	Skew	0.25
Prob(H) (two-sided)	0.79	Kurtosis	3.49

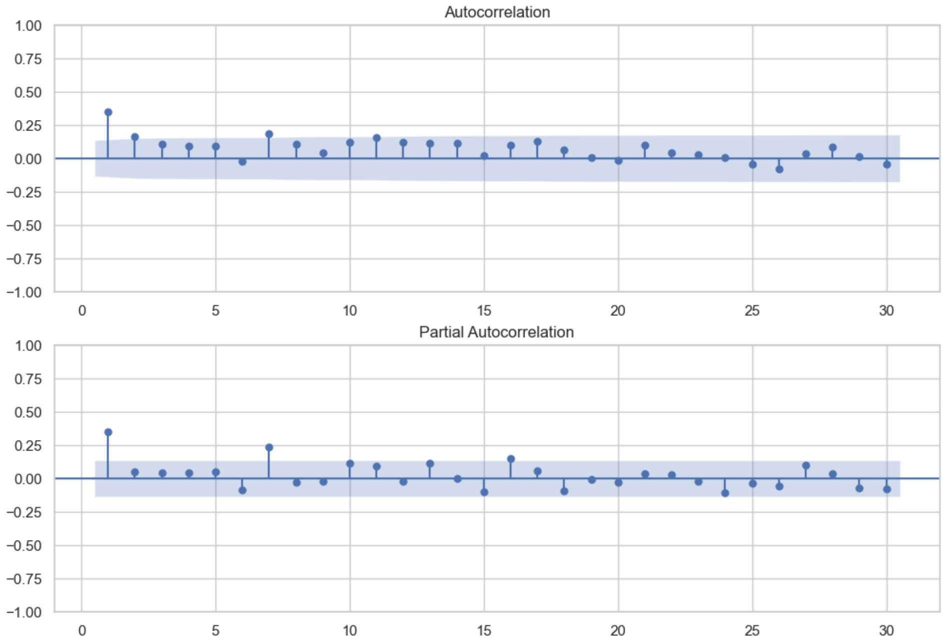

Figure B.4: ACF / PACF Imbalances Netherlands

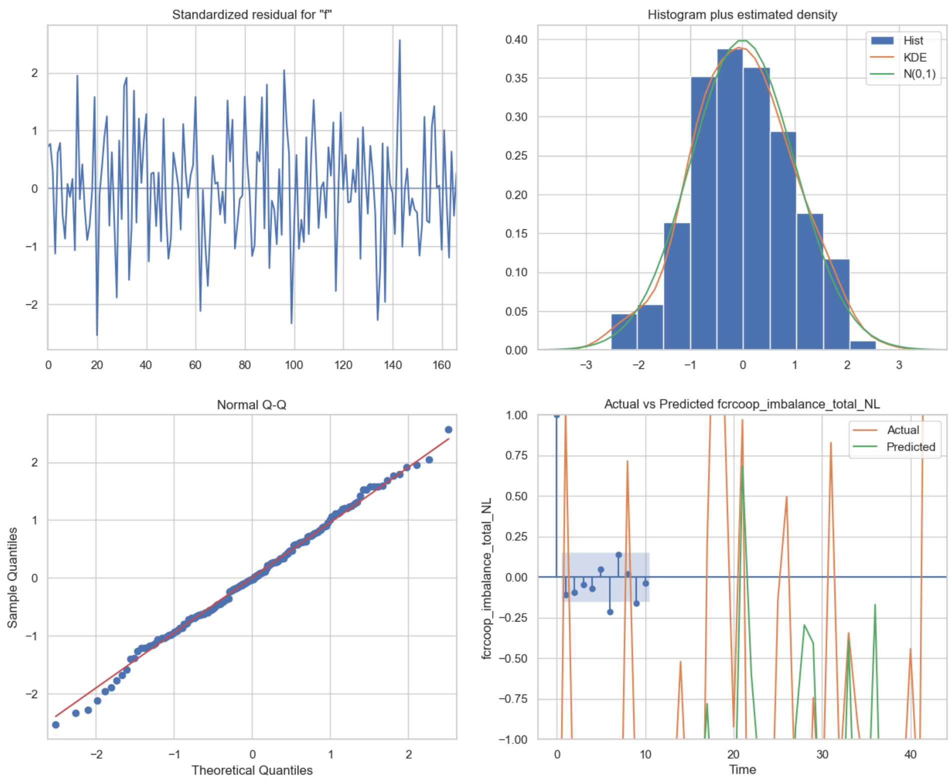

Figure B.5: Residuals Imbalances Netherlands

B.3. Imbalances in FCR Cooperation Modelling

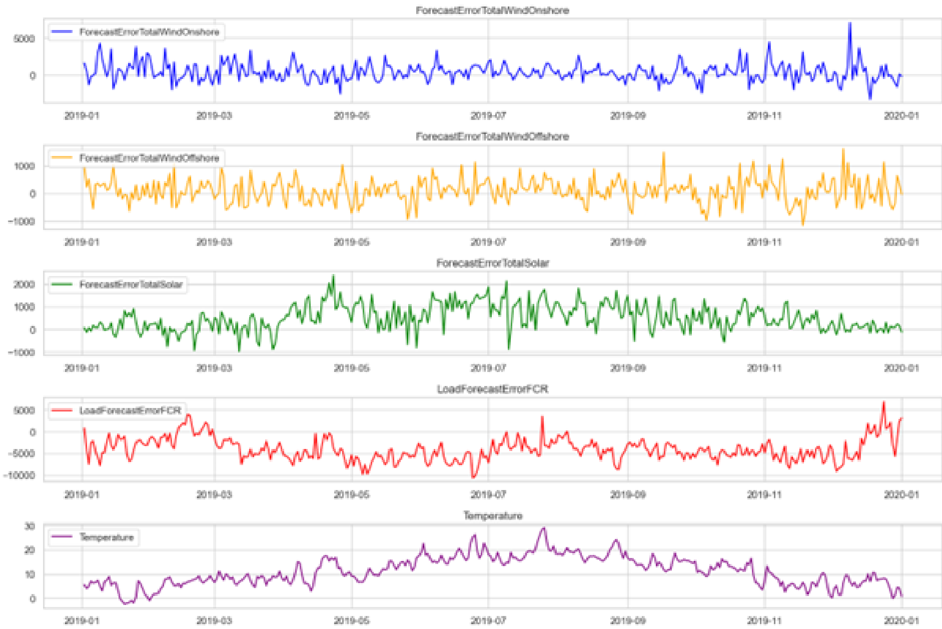

Figure B.6: Time Series Imbalances FCR Cooperation

Table B.5: Imbalances FCR cooperation: Model Diagnostic Statistics

Statistic	Value	Statistic	Value
Ljung-Box (L1) (Q)	0.01	Jarque-Bera (JB)	2.39
Prob(Q)	0.92	Prob(JB)	0.30
Heteroskedasticity (H)	1.06	Skew	0.01
Prob(H) (two-sided)	0.84	Kurtosis	3.59

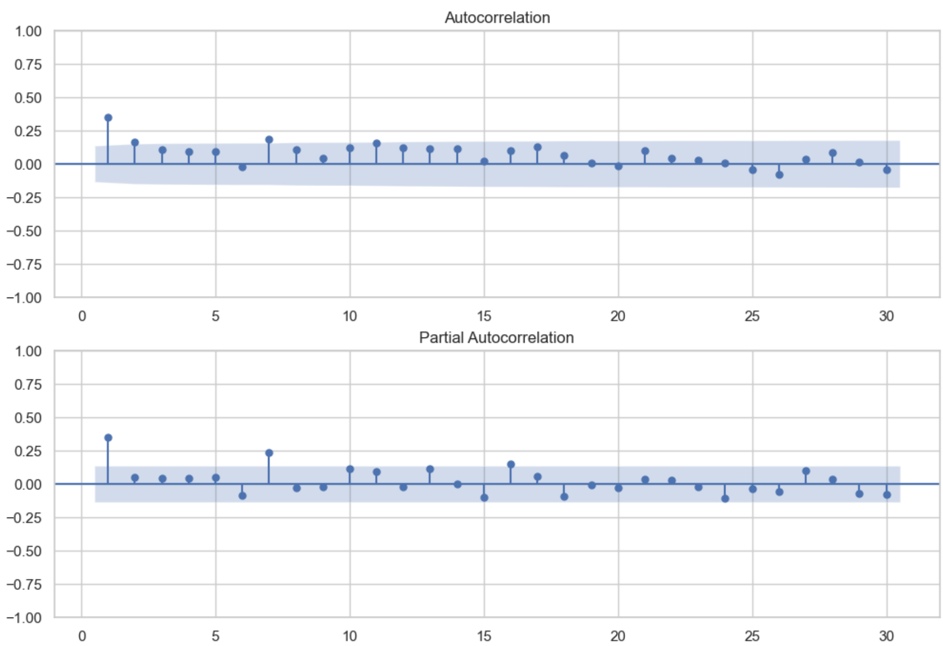

Figure B.7: Imbalances FCR Cooperation: ACF / PACF

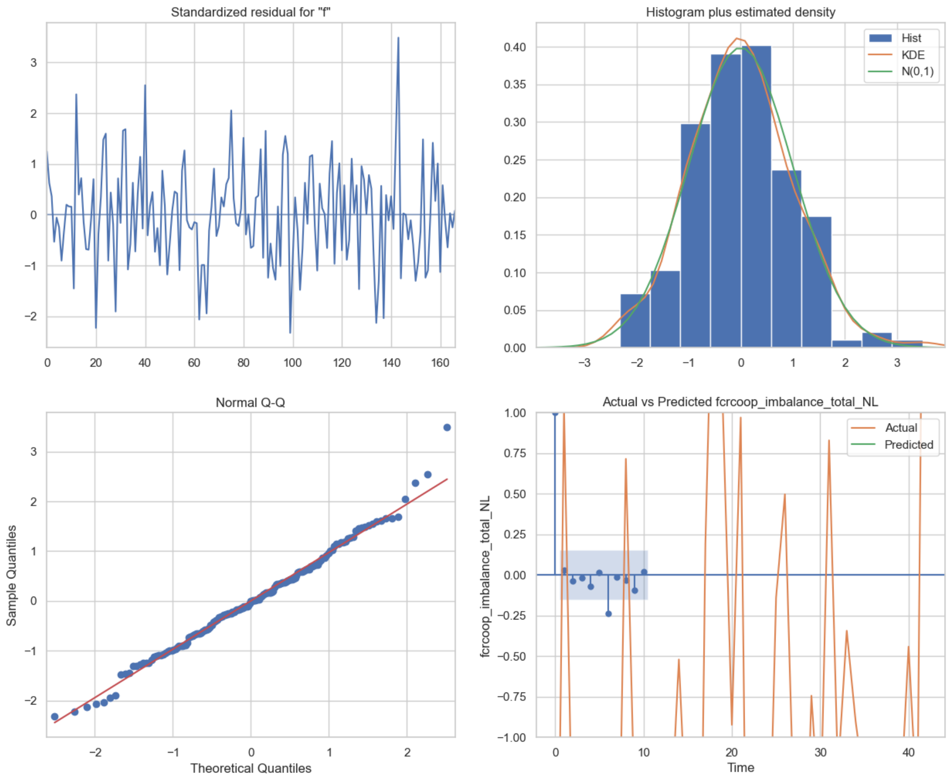

Figure B.8: Imbalances FCR Cooperation: Residuals

B.4. FCR Activation Netherlands Modelling

Table B.6: Activated FCR NL: Diagnostic Statistics

Statistic	**Value**
Ljung-Box (L1) (Q):	3.96
Prob(Q):	0.05
Jarque-Bera (JB):	2379.97
Prob(JB):	0.00
Heteroskedasticity (H):	0.97
Prob(H) (two-sided):	0.90
Skew:	-1.77
Kurtosis:	21.74

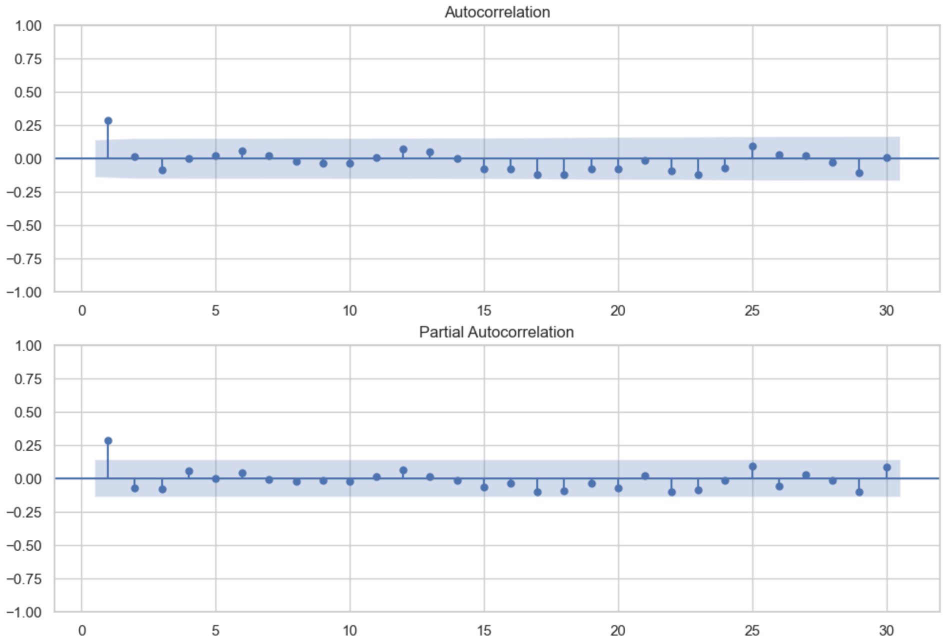

Figure B.9: Activated FCR NL: ACF / PACF

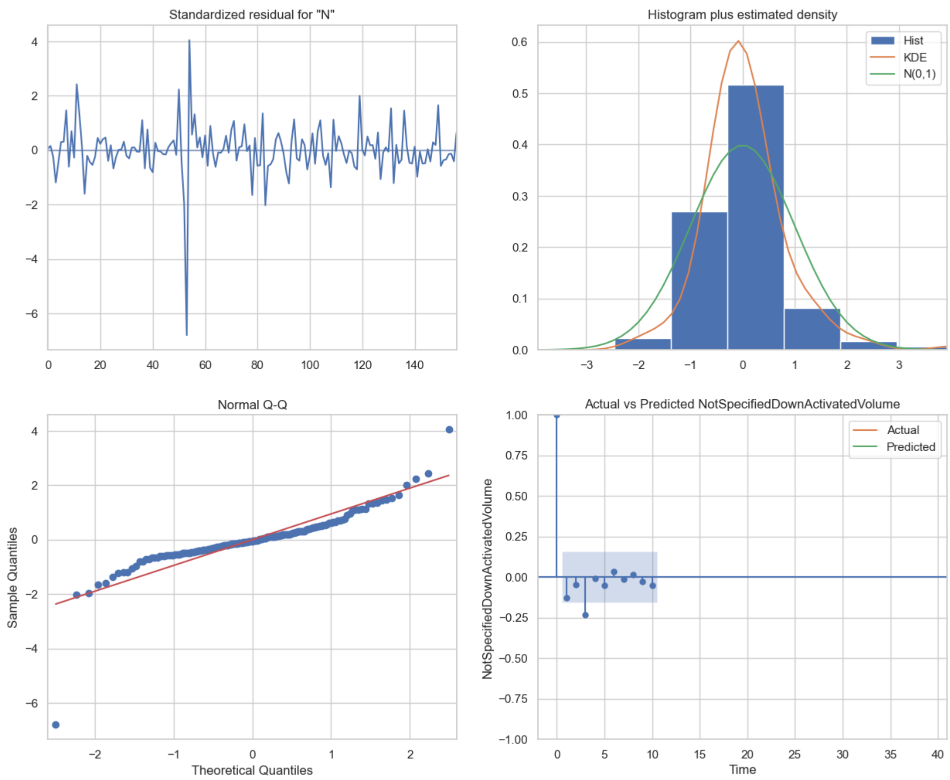

Figure B.10: Activated FCR NL: Residuals

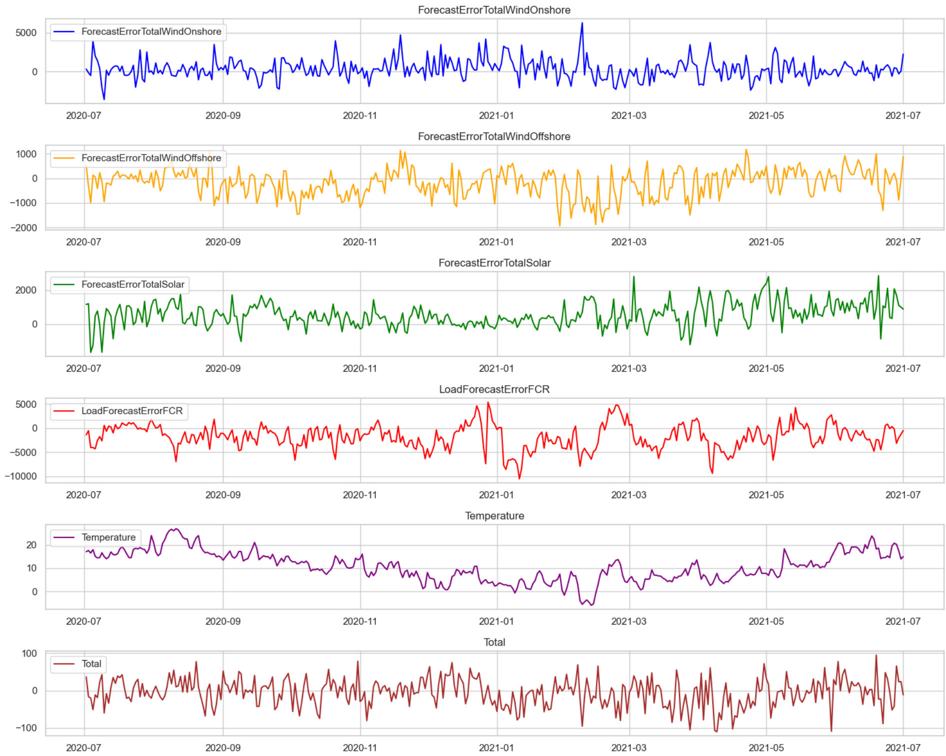

Figure B.11: Activated FCR NL: Time series

C

Market Design Parameters

C.1. FCR Market Design Parameters

Table C.1: FCR Market Design Parameters

Frequency Control Reserve (FCR)	
Formal access requirements	
Explicit restrictions for certain types of service providers	No restrictions
vRES access to the balancing market	Yes
Capacity provision	Mainly voluntary
Specific products for DER	No
Pooling conditions	
Pooling	Allowed
Approach to prequalification	Unit-based for FCR
Technical prequalification criteria	
Activation speed and duration	100% activation in 30 seconds **7**

Table C.2: FCR Market bidding Parameters

Bid Related Requirements	
Minimum bid size	1 MW (1-MW increments)
Bid symmetry	Symmetrical
Volume (Netherlands, 2023)	111 MW
Timing-related characteristics of bidding	
Gate Opening Time	D-7 at 11:00 CET
Gate Closing Time	D-1 at 8:00 CET
Product resolution	4 hours
Renumeration	
Pricing rule (remuneration of awarded bids: pay-as-bid (PaB) vs. marginal pricing (MP))	Marginal Pricing

C.2. aFRR Market Design Parameters

Table C.3: aFRR Market Design Parameters

Automatic Frequency Restoration Reserve (aFRR)	
Formal access requirements	
Explicit restrictions for certain types of service providers	No restrictions
vRES access to the balancing market	Yes
Capacity provision	Mainly voluntary
Specific products for DER	No
Minimum MW (Netherlands, Q3/Q4 2023)	324 MW Up, 362 MW Down
Pooling conditions	
Pooling	Allowed
Approach to prequalification	Pool-based for aFRR and mFRR
Technical prequalification criteria	
aFRR	Response in 30 seconds Full activation within maximum 15 minutes
Ramp rate	Minimum 20% of volume of the bid per minute and maximal 100%

Table C.4: aFRR Market Bidding Configuration

Market Requirements	
Bid symmetry	Asymmetric
Procurement of capacity & energy	Split
Non-precontracted energy bids	Yes
Product resolution	Every 15 minutes
Capacity Bid Related Requirements	
Frequency of bidding – capacity	Yearly and monthly
Frequency of market clearing – capacity	Once a year & once a quarter
Bidding requirement	D-1 at 14:45 CET
Bids transferable to other BSP	D-1 at 23:00 CET
Energy Bid Related Requirements	
Frequency of bidding – energy	Every 15 minutes
Auction	D-1 at 9:00 CET
Frequency of market clearing & activation – energy	Every 15 minutes
Bid Size	min. 1 MW – max. 999 MW (1-MW increments)
Renumeration	
Pricing rule (remuneration of awarded bids: pay-as-bid (PaB) vs. marginal pricing (MP))	PaB for capacity MP for energy

C.3. mFRR Market Design Parameters

Table C.5: mFRR Market Configuration

Manual Frequency Restoration Reserve (mFRR)	
Formal access requirements	
Explicit restrictions for certain types of service providers	No restrictions
vRES access to the balancing market	Yes
Capacity provision	Mainly voluntary
mFRRda volume (Netherlands, Q3/Q4 2023)	980 MW Up, 965 MW Down
Pooling conditions	
Pooling	Allowed
Approach to prequalification	Pool-based
Additional agreements	BRP's notification and agreement required; currently no independent aggregators for balancing products
Technical prequalification criteria	
Activation speed and duration	Activation within 15 minutes
Ramp rate	Minimum 7% of volume of the bid per minute for aFRR and MRR*
Bid-related requirements	
Minimum bid size	1 MW (1-MW increments)
Bid symmetry	Asymmetric
Procurement of capacity & energy	Split
Energy bid adjustment	No
Auction	D-1 at 9:00 CET
Contract period	1 day (0:00 – 0:00)
Renumeration	
Pricing rule (remuneration of awarded bids: pay-as-bid (PaB) vs. marginal pricing (MP))	PaB for capacity, MP for energy