Charles University

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

The Impact of Renewable Electricity on the Czech Electricity Balancing Market

Author: **Bc. Amálie Kašparová**

Supervisor: **Mgr. Luboš Hanus**

Academic Year: **2020/2021**

Declaration of Authorship

I hereby proclaim that I wrote my master's thesis on my own under the leadership of my supervisor and that the references include all resources and literature I have used.

I grant a permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 4, 2021

Signature

Acknowledgments

I would like to express my gratitude to my supervisor Mgr. Luboš Hanus for his guidance and valuable comments that improved the final contents of the thesis. Most importantly, I would like to thank Matěj Kouřílek for his endless patience and willingness to help. Lastly, I would like to thank my family for their constant support during the whole studies.

Abstract

As global investments in renewable energy technologies continue to grow, their effects on electricity markets are a challenge for regulators and policymakers. The thesis examines the effects of forecast errors of Czech and German renewable energy sources on the size and volatility of the system imbalance of the Czech balancing market. Using a quantile regression and *ARFIMA-GARCH* models on hourly data, I found that higher solar and wind forecast errors increase the system imbalance in absolute terms and affect the volatility. The results show that the Czech solar and wind forecast errors have significantly higher effect than the German forecast errors on the size and volatility of the system imbalance. The strongest effect on the size and volatility of the system imbalance have the Czech solar forecast errors. Therefore, the Czech government should insist on improving the accuracy and availability of renewable energy forecasts from the transmission system operator ČEPS.

Klasifikace JEL	C14, C50, Q42
Klíčová slova	renewable sources, forecast errors, balancing market, system imbalance

Abstrakt

Vzhledem k tomu, že globální investice do technologií obnovitelné energie nadále rostou, jejich účinky na trhy s elektřinou jsou výzvou pro regulačním orgány i politiky. Práce zkoumá vliv chyb v předpovědi českých a německých obnovitelných zdrojů na velikost a volatilitu systémové odchylky českého vyrovnávacího trhu. S použitím kvantilní regrese a *ARFIMA-GARCH* modelu na hodinových datech, autor nalezl, že vyšší solární a větrné chyby v předpovědích, zvyšují systémovou odchylku v absolutních číslech a ovlivňují její volatilitu. Výsledky ukázaly, že české chyby v předpovědích solární a větrné energie mají signifikantně vyšší vliv na velikost i volatilitu systémové odchylky než německé chyby v předpovědích. Nejsilnější vliv na velikost i volatilitu systémové odchylké má chyba v předpovědi české solární energie. Z těchto důvodů by měla česká vláda prosazovat zlepšení přesnosti a dostupnosti prognóz obnovitelné energie od provozovatele soustavy ČEPS.

Klasifikace JEL C14, C50, Q42
Klíčová slova obnovitelné zdroje, chyba předpovědi, vyrovnávací trh, systémová odchylka

Contents

List of Tables

List of Figures

Acronyms

ACF Autocorrelation function

ADF Augmented Dickey Fuller test

AIC Akaike information criterion

ARFIMA Autoregressive fractional integrated moving average model

BIC Schwarz's Bayesian information criterion

BRP Balance responsible party

BSP Balance service provider

EBGL Electricity balancing guideline

EEX European Energy Exchange

FCR Frequency containment reserve

FFT Fast Fourier Transform

FRCE Frequency restoration control error

FRR Frequency restoration reserves

GARCH Generalized autoregressive conditional heteroskedasticity model

GPH Geweke-Porter-Hudak estimator

H Hurst exponent

KPSS Kwiatkowski-Phillips-Schmidt-Shin Test

LAD Least absolute deviations estimator

LFC Load frequency control

MW Megawatt

OTC Over-The-Counter market

PV photo-voltaic

RES a renewable energy source

RR Replacement reserves

TSO Transmission system operator

Master's Thesis Proposal

Author	Bc. Amálie Kašparová
Supervisor	Mgr. Luboš Hanus
Proposed topic	The Impact of Renewable Electricity on the Czech Electricity Balancing Market

Motivation Over the last decade, a strong emphasis has been put on the promotion of green electricity in the European energy sector. The increasing levels of renewable energy add the uncertainty and variability inherent in electricity grids and impose additional costs for balancing and reserve requirements (Kiviluoma et al., 2012). With greater intra-day uncertainties, the costs from real-time supply-demand imbalances rise for the market participants, creating financial difficulties. It leads to market exits for retailers who fail to hedge effectively, and the operational complexity adds to the costs of transmission system operators (TSOs) (Goodarzi et al., 2019).

Due to the European transition to low carbon energy supply, significant energy surpluses flow into the Czech transmission system, potentially affecting the imbalance volumes and price of electricity in the local market. The source of these surpluses is mainly German wind farms. To what extent German renewable electricity affects Czech imbalance volumes and prices are interesting and important information for government and local electricity providers. The Czech national TSO ČEPS has to buy reserved capacity on balancing market. Thus, the knowing key drives of imbalances is of great importance. Further, with the rise of Internet of Things (IoT) and progress in processing real-time data new business on energy markets has emerged, specifically on balancing market. These new market participants try to make profits and optimise their portfolio on balancing market by forecasting imbalance volumes and prices. Such businesses could profit on the knowledge what is the driver of the imbalance volumes.

Whilst a number of empirical studies were conducted on the topic of renewable electricity and its impact on spot prices, focus on electricity's real-time imbalances has been relatively scant. Bueno-Lorenzo et al., (2013) investigate the relationship between wind energy and imbalance volumes in the Spanish energy market, focusing on developing new pricing scheme to create more efficient electricity market. Other studies primarily focus on thermal power generation in order to mitigate fluctuations in wind energy generations (e.g. Aïd et al, 2016). Next, many researchers such as Kiesel and Paraschiv (2017) and Pape et al., (2016) seek to develop superior forecasting models for electricity imbalance volumes and electricity prices. Both studies use regression methods to forecast imbalance and electricity prices for German intraday market. However, we would like to show the impact of renewable energy on the imbalance volumes. A paper close to our field of interest written by Goodarzi et al. (2019) studies, how renewable energy (wind and solar) forecast errors affect the imbalance volumes.

Hyptheses

1. Hypothesis #1: The Czech renewable power generation increase imbalance volumes on the Czech electricity balancing market.
2. Hypothesis #2: The German wind and solar power generation increase imbalance volumes on the Czech electricity balancing market.
3. Hypthesis #3: Renewable power generation increase volatility of imbalance volumes on the Czech electricity balancing market

Methodology I will use data of day-ahead forecast for wind and solar production in Germany and the Czech Republic and data of actual wind and solar production in Germany and Czech Republic from European Network of Transmission System Operators for Electricity (ENTSO-E). Further, I will employ Load (CZ) from the same platform. These data show instantaneous electricity demand. While data from the Czech market has hourly granularity, German production volumes are given in 15 minutes timestamps. Thus, I will compute and work with hour averages.

To analyse imbalance volumes of electricity, there is a large variety of models that can be used. As in Goodarzi et al. (2019), we can use quantile regression to estimate imbalance volumes, which is an extension of ordinary least squares regression that aims to estimate the median and quantiles of the response variables (Koenker and Bassett, 1978; Koenker and Hallock, 2001). Another popular family of models used in energy market analyses are autoregressive models with external variables (ARIMAX). Such a model takes into account autocorrelation which is usually present in energy markets data while allowing to estimate the effect of other variables on a response variable (Perez-Mora et al., 2015; Cui and Peng, 2015).

Further, to analyse the volatility of imbalance volumes, we can use similar methods as Tashpulatov (2013), Pham & Lemoine (2015), Conejo et. al (2005) who analyse price volatility using autoregressive conditional heteroskedasticity models. Another approach that can also help to estimate the volatility of imbalance volumes is described in Haximusa (2018), where the author estimates cross-border effects of German wind and solar electricity on French spot price volatility. As dependent variable Haximusa (2018) chooses the absolute value of the deviation of the actual hourly French spot price from its daily mean and not standard measure of price variance (e.g. Ketterer, 2014).

Expected Contribution As the levels renewable energy accelerates, it is important to know to what extent is the Czech market, specifically balancing market, affected. The thesis should analyse whether Czech and German electricity produced by renewable resources increases the imbalance volumes on Czech electricity balancing market. Further, the volatility of the Czech imbalance volumes concerning renewable power generation will be analysed. The estimates can be used by both the Czech government and Czech local electricity providers for strategy optimisation.

Outline

1. Motivation: Introduction to the topic and description of the electricity market.
2. Literature Review: Description of estimation techniques used in recent studies of the electricity market.
3. Data: I will describe the data used for analysis.
4. Methodology: A detailed summary of methods I will use in the thesis.
5. Results: I will discuss my findings based on estimated models.
6. Conclusion: Summary of regression results and possible implications for the Czech policy decision-maker.

Bibliography

Kiviluoma, J., Meibom, P., Tuohy, A., Troy, N., Milligan, M., Lange, B., ... & O'Malley, M. (2012). Short-term energy balancing with increasing levels of wind energy. IEEE Transactions on Sustainable Energy, 3(4), 769-776.

Goodarzi, S., Perera, H. N., & Bunn, D. (2019). The impact of renewable energy forecast errors on imbalance volumes and electricity spot prices. Energy Policy, 134, 110827.

Bueno-Lorenzo, M., Moreno, M. Á., & Usaola, J. (2013). Analysis of the imbalance price scheme in the Spanish electricity market: A wind power test case. Energy policy, 62, 1010-1019.

Aïd, R., Gruet, P., & Pham, H. (2016). An optimal trading problem in intraday electricity markets. Mathematics and Financial Economics, 10(1), 49-85.

Kiesel, R., & Paraschiv, F. (2017). Econometric analysis of 15-minute intraday electricity prices. Energy Economics, 64, 77-90.

Pape, C., Hagemann, S., & Weber, C. (2016). Are fundamentals enough? Explaining price variations in the German day-ahead and intraday power market. Energy Economics, 54, 376-387.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. Econometrica: journal of the Econometric Society, 33-50.

Koenker, R., & Hallock, K. F. (2001). Quantile regression. Journal of economic perspectives, 15(4), 143-156.

Perez-Mora, N., Martinez-Moll, V., & Canals, V. (2015). Spanish Renewable Energy Generation Short-Term Forecast. In Proceedings of the ISES Solar World Congress.

Cui, H., & Peng, X. (2015). Short-term city electric load forecasting with considering temperature effects: an improved ARIMAX model. Mathematical Problems in Engineering, 2015.

Conejo, A. J., Plazas, M. A., Espinola, R., & Molina, A. B. (2005). Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. IEEE transactions on power systems, 20(2), 1035-1042.

Tashpulatov, S. N. (2013). Estimating the volatility of electricity prices: The case of the England and Wales wholesale electricity market. Energy Policy, 60, 81-90.

Pham, T., & Lemoine, K. (2015). Impacts of subsidized renewable electricity generation on spot market prices in Germany: evidence from a garch model with panel data.

Ketterer, J. C. (2014). The impact of wind power generation on the electricity price in Germany. Energy Economics, 44, 270-280.

Chapter 1

Introduction

In recent years, the share of renewable energy sources (RESs) production has been increasing rapidly. The energy produced by RESs is not only cheap but also reduce green house gas emissions. Despite significant benefits of these sustainable technologies, there are growing concerns about their direct and indirect adverse effects on electricity markets (Hache & Palle, 2019). In particular, in the context of market efficiency, balancing "real-time" markets have gained a lot of attention (Goodarzi *et al.*, 2019). The increasing amount of RESs add uncertainty and variability inherent in electricity grids and impose additional costs for balancing and reserve requirements (Kiviluoma *et al.*, 2012). Each market participant is financially responsible to keep the system in balance. When real-time production or consumption of market participants differ in comparison with their prior notification to the transmission system operator (TSO), they are exposed to an imbalance settlement charge (Goodarzi *et al.*, 2019). Subsequently, the imbalance settlement charge is reflected in the regulated component of the electricity price for the end customer (Klamka, 2016).

Hence, the knowledge of more accurate electricity generation forecasts is of importance. The UK energy regulator seeks to address this issue and directly rewards or penalizes TSOs for the accuracy of their demand forecasts (Ofgem, 2018). In doing so, it seeks to force TSOs to develop and publish more accurate demand forecasts. This is not a trivial challenge that involves just improving forecasting techniques. Substantial amount of RESs is being installed in the system at end-use level, not as metered generation on the transmission system (Goodarzi *et al.*, 2019). As a result, the production data is harder to capture.

Since the flow of electricity does not respect borders and follows the least resistance path, energy surpluses of cheap green electricity flow from Germany to the Czech transmission system, potentially affecting the imbalance volumes.

The objective of this thesis is to examine to what extent RESs forecast errors contribute to the size of the system imbalance. The productions from the Czech and German RESs are considered. Moreover, the impact of RESs forecast errors on the volatility of system imbalances is the main question of my research. To study effects on the size of the system imbalance I use quantile regression. It allows to explain tails characteristics of the system imbalance distribution. Estimating the tail dependencies of system imbalance distributions are particularly interesting since balancing high fluctuations is expensive. Furthermore, the effects of forecast errors on volatility of the system imbalance is captured using $ARFIMA$-$GARCH$ models. I use $ARFIMA$ process to model the mean equation as the system imbalance has long memory behavior.

Thesis contributes to the existing literature in several ways. There have not been many papers devoted to the effects of RESs on the Czech system imbalance (Klamka, 2016). Especially, there has not been research on effects of RESs forecast errors on the Czech system imbalance. Furthermore, no cross-border dependencies have been considered in the existing literature.

The thesis is structured as follows. Chapter 2 describes the European energy markets with emphasis on the balancing market. Next, I describe the Czech and German energy markets. Chapter 3 provides the revision of existing literature on balancing markets and models used in this thesis. Chapter 4 introduces the process of data collection, data sources, and an overview of data adjustments. Chapter 5 covers the methodology background and introduces the theoretical model. Chapter 6 includes results and discussion of the analysis. Chapter 7 concludes the thesis and provides recommendations for further research.

Chapter 2

The European Energy Markets

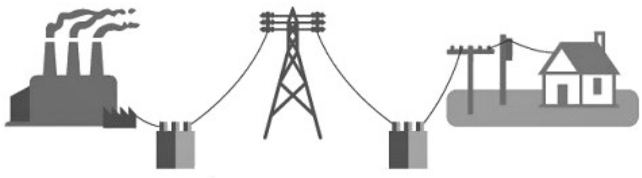
In this section, I describe the European energy market design with a main focus on the European balancing markets and the impact of the increasing penetration of renewable sources. Further, I provide an overview of the German and the Czech energy markets. This Section 2.1 is based mainly on Flášar *et al.* (2016), Erbach (2016) and in the Subsection 2.1.1, I mostly work with papers written by ENTSO-E (2018) and Brijs *et al.* (2015).

2.1 The European Energy Market Design

The EU is in the transition towards a low-carbon economy in order to minimize the emission of greenhouse gasses. The key role in the transition plays the electricity, mainly the use and growing share from renewable sources. The electricity system comprises an organised electricity market and the physical infrastructure for electricity generation, transport and use. The physical interconnection consists of electricity generators, transmission and distribution. Transmission is used for long distance electricity transport, while distribution transports electricity to residential and industrial consumers, to see in Figure 2.1.

The electrical grid is the network designed to connect the generators and consumers through the transmission and distribution systems. It has two fundamental properties. First, the electricity supply has to match demand, otherwise, there is a risk of blackouts. Secondly, the electricity flow is uncontrollable as it follows the shortest distance possible, i.e, the least resistant path. Electricity generators comprise power plants with large installed capacity such as nuclear or coal. Generators are divided according to their generation capacity and flexibility with which they can be operated. The firm-capacity generators have the possibility to be turned off and on, such as hydro-electric dams, coal and nuclear power stations. On the other hand, the variable-capacity generators are influenced by hardly predictable factors,

Figure 2.1: Schematic Overview of Electricity Transport



Source: www.eia.gov (2020)

such as wind and sunshine. Regarding the flexibility, I can compare nuclear and hydro power. Generation of nuclear electricity is convenient for longer periods, while the hydroelectric dam can be switched off in a matter of seconds. The flexibility of generation is important when fluctuations in electricity demand occur. Transmission grids connect the generators with local distribution grids. As mentioned above, transmission grids are networked grids transporting electricity over long distance. Hence, high voltages (between 220kV and 1000kV) are used to avoid losses due to the grid's electrical resistance. Transmission networks are run by transmission system operators (TSOs). On the contrary, distribution grids distribute the electricity at low voltages over shorter distances to the end-consumers. Electricity with small scale generation unit, such as smaller renewable sources, can be generally connected to the distribution grid. Distribution networks are run by distribution system operators. End-consumers can be directly connected to the transmission grids if they run a high offtake, such as steelworks. The Figure 2.1 also depicts so-called "step-up" and "step-down" transformers. The step-up transformer is located between the generator and transmission grid and it upscales the voltage to 115kV and more. The step-down transformer lies between the transmission grid and distribution grid and it reduces voltages to 35kV and lower.

European electricity markets are liberalised, i.e., one entity cannot participate in production, transmission, distribution and retail at once. Production, delivery and retail are completely deregulated. On the other hand, transmission and distribution are still regulated by local states due to the monopolistic nature of transmission (Kouřílek, 2019). Each member state regulates its own electricity market, however, the major aim is coupling them into one single EU market (Pavić *et al.*, 2017).

The fundamental market participants are producers and consumers. Further, the trader (with electricity) enters as an intermediary concentrating demand from a number of final consumers and supply from a number of producers. A direct relationship between the producer and the consumer is not usual and it only makes sense for long-term deliveries between a larger producer and a large consumer. Traders resale electricity on the wholesale market in order to maximize their profits. In addition to traders, there is also retail, where the electricity is provided directly to the end-consumer. As with many commodities, there is an exchange where demand meet supply and organized trades take place. An important role in the market has a market operator who registers market participants on energy exchange and arranges imbalance settlement between actual and negotiated deliveries.

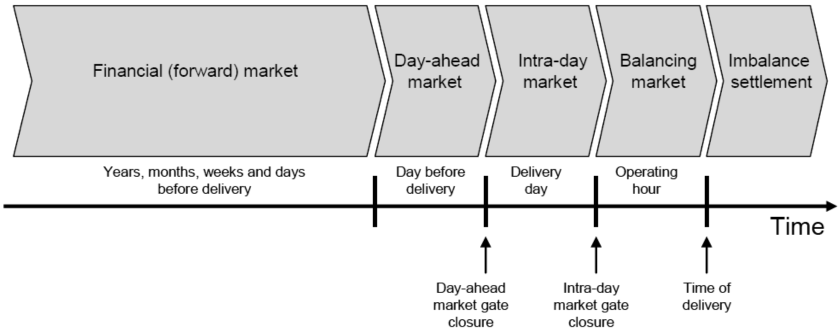
According to the EU Regulation 2019/943 Article 5, all electricity market participants should have balance responsibility, i.e., they are responsible for their imbalances in the system. There are two fundamental regimes of balance responsibility: Market participants are either balance responsible parties (BRPs) or they contractually delegate their responsibility to a BRP according to their own choice. Moreover, each BRP is financially responsible to keep the system in balance.

As already mentioned, electricity has some unique characteristics, such as uncontrollable electricity flow. Furthermore, demand has to match supply at all times. Large volumes of electricity are not possible to store yet. Therefore, the value of electricity differs over time and most electricity transactions involve the delivery at some point in the future. Due to these characteristics of electricity, there are a number of different electricity markets.

Wholesale markets operate at a larger scale and ensure other business transactions between individual business entities in the market (Business to Business, B2B). Each of the wholesale market participants wants to have the maximum number of partners in order to optimize their business portfolio. The prices at the wholesale market are not regulated and are generally lower than retail prices and change in real time. The wholesale and retail market differ not only in the nature of transactions but also in the length of electricity deliveries, the price composition and the settlement of deliveries. The wholesale markets are built on the "take or pay" rule, i.e., both parties are obliged to pay for all deliveries, even if they are not actually realized. On the other hand, the retail market is carried out in order to ensure the consumption of electricity to the end-consumer. Retailers offer electricity contracts and buy electricity from producers. Consumers may choose their retailer. The retail market price differs fundamentally from wholesale market prices and includes both unregulated and state-regulated prices and an environmental tax. State-regulated prices of electricity are adjusted by the national regulatory authorities through price decisions issued annually (Flášar *et al.*, 2016).

The Figure 2.2 shows that the wholesale market can be divided with respect to the time of the product purchase. Future and forward markets are used for long-term contracts, i.e., weeks to years before delivery. This market works more as financial security for the price of electricity in the long term. Thus, there may not always be a physical delivery of electricity in this market (Salavec, 2017). The largest physical

Figure 2.2: Structure of Power Trading



Source: Ruska & Similä (2011)

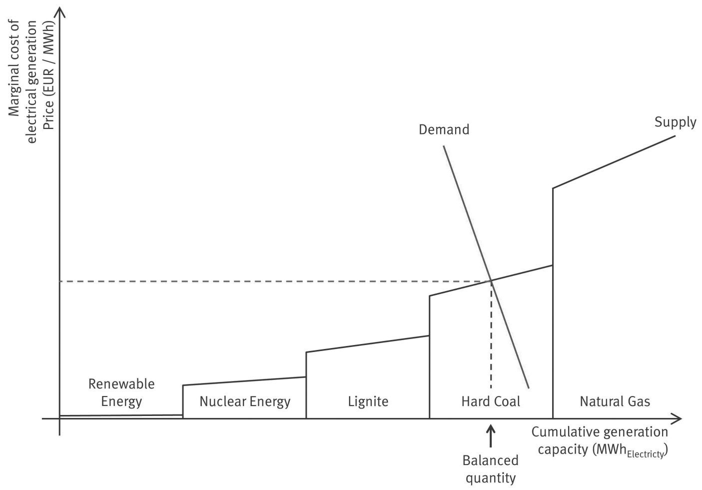
electricity market is the day-ahead market, where trades are organized on the day before the day of the physical delivery. In the day-ahead market, the spot market price for each hour of the delivery day is determined according to the supply and demand bids of market participants (Ruska & Similä, 2011). Based on the supply and demand of each trading hour is also determined the resulting transmission of electricity to or from abroad (Flášar *et al.*, 2016). In the intra-day market, there is a possibility for market participants to adjust their trading position (generation or consumption) at a time very close to the delivery. The exact time differs across countries. Trading in the intra-day market has been growing especially because of the massive investment into renewable energy sources in the EU, i.e., sources whose operation is difficult to predict and depends mainly on current weather fluctuations. The intra-day market opens after the closure of the day-ahead market. The balancing market is organized by the TSO and opens thirty minutes before electricity delivery. The balancing market is used to keep a balance between generation and consumption in real time. A detailed description of the European balancing market and the imbalance settlement is discussed in Subsection 2.1.1.

Electricity can be traded via energy exchange or bilateral trade. Bilateral trade, or the more commonly used term over-the-counter (OTC) market is a classic method of trading, where two parties close a deal directly. The most frequent products on the OTC market are base load, peak load, off peak, high tariff and low tariff (Flášar *et al.*, 2016). Base load is a delivery at constant power during the whole duration of the contract. Peak load is a deliver at constant power from Monday to Friday always in the period from 8.00 to 20.00, while off peak is the exact opposite. High tariff is a delivery of constant power from Monday to Friday always in the period from 6.00 to 20.00. Low tariff is the opposite of high tariff. The first three products base load, peak load, and off peak are standard products throughout Europe, while high and low tariff are mainly recognized in eastern European markets. The disadvantage of OTC market is the credit exposure. When the counterpart goes bankrupt, the trader may face significant even liquidation losses. The second way of trading is the energy exchange. Exchange trading is absolutely anonymous compared to OTC market and the counterparty of each trade is the energy exchange. On the exchange market is allowed to trade with the same standardized products as those on the OTC market. The most significant contribution of energy exchange lies in the price transparency of the market which OTC market does not provide at all. Another advantage of energy exchange lies in the minimal capital requirements necessary for the realization of a large number of transactions. Traders most likely make a large number of transactions and trade in both directions, i.e., they buy and sell the same product several times in a row. The central counterparty clearing house offsets all these transactions (Flášar *et al.*, 2016). European Energy Regulators claim that increased spot and futures trading at energy exchanges lead to more liquid and transparent wholesale markets. Liquid wholesale markets are fundamental to efficient competition and competitive retail markets (Ruska & Similä, 2011).

The main European energy exchange platforms are the Nord Pool, the European Energy Exchange (EEX), and the ICE Energy exchange (ICE ENDEX) (Flášar *et al.*, 2016). Nord Pool is the largest energy exchange in Europe and in terms of financial and physical trade volumes. Nord Pool exchange works in Denmark, Estonia, Finland, Norway and Sweden (Ruska & Similä, 2011). The EEX is a commodity exchange located in Germany. The EEX operates markets for future and spot trading for Germany, Austria, France and Switzerland. The ICE ENDEX organizes energy markets for Benelux countries. The ICE ENDEX is the most liquid European energy exchange platform for gas trading, but in electricity trading, the exchange is the smallest of the "big" European energy exchanges. The spot market organized by energy exchange is a fundamental electricity market and should be a major price-forming place for all market participants. In the spot market, merit order is applied to create prices. It guarantees the efficient physical delivery of electricity (Flášar *et al.*, 2016). Merit order arranges energy sources from the cheapest to the most expensive with respect to the marginal costs. The Figure 2.3 depicts a general merit order showing that RES have the lowest marginal cost of electrical generation followed by nuclear, lignite, coal plants, peaking at gas and oil with highest marginal costs. In 2019, the rising prices for emission allowances met with decreasing spot prices for natural gas and caused structural changes in the average merit order. Thus, the real order of energy sources of Figure 2.3 should be: RES, nuclear, lignite, natural gas and coal.

RESs have almost zero short-run marginal cost given by the nature of sunlight, wind or water (Luňáčková *et al.*, 2017). Nuclear power's marginal costs are also low and the output capacity is usually not adjusted from hour to hour. Fossil fuel

Figure 2.3: Merit Order Effect Mechanism



Source: Conrad & Staacke (2016)

power stations have high marginal costs, which largely depend on fuel and emissions allowance costs (Ruska & Similä, 2011). The Figure 2.3 shows the mechanism of the merit order effect. The demand curve is inelastic since electricity is a necessary product. Because of the low marginal costs of RES, the entire merit order "supply" curve is shifted to the right, i.e., price decreases. The wholesale prices, indeed, have declined in Europe over past years. Drivers such as decline in final electricity consumption, decline in fuel and CO_2 prices, near-collapse of the European emission trading scheme or less expensive coal or natural gas have been blamed for the price slump however, the largest factor contributing to the drop in wholesale prices was the expansion of renewable energy (Hirth, 2018). Further, due to the RES and their characteristics, it is even possible to reach negative wholesale prices. To accomplish these unique characteristics, which cannot be seen in other markets, the electricity demand has to be low and the weather conditions have to be favourable in order to RES cover the substantial part of total demand. As a result, other conventional power stations have to be minimized and the prices have negative values (Fanone *et al.*, 2013).

2.1.1 The European Balancing Energy Market

Electricity markets have to include a convenient mechanism to guarantee a real-time balance between supply and demand of electricity. Thus, thirty minutes before electricity delivery, balancing market opens. This market is organized by the market operator and TSO is the only participant. TSO's role is to ensure that considering the other markets' results, balance between consumption and generation remain close to real time (ENTSO-E, 2018). Well functioning day-ahead and intra-day markets serves as a foundation for the balancing market.

TSO works on both the procurement and the settlement side of the balancing market. TSO calculates the total system imbalance resulting from the imbalances made by BRPs. Further, TSO also compensates this system imbalance by activating reserve capacity provided by Balance Service Providers (BSPs). Subsequently, TSO creates an imbalance settlement with BRPs by applying an imbalance price (Brijs *et al.*, 2015).

BRPs are market participants such as electricity producers, consumers and suppliers. BRPs are financially responsible for the imbalances in their portfolios. Therefore, they should control their individual position to be in balance (ENTSO-E, 2018). The individual position is defined as a sum of the energy volume physically delivered or withdrawn from the system and trades. BSPs are for instance generators or storage operators who are able to help the TSO balance the system by offering balancing services by activation of energy or capacity (ENTSO-E, 2018).

In contrast to the intra-day, day-ahead and forward markets, which are based on forecasts, balancing markets use real time information about the need and value of power adjustments (Brijs *et al.*, 2015; Hiroux & Saguan, 2010; Vandezande *et al.*, 2010). Thus, the activation and reservation of reserve capacity is referred to as the balancing market purchase side, and BSPs are rewarded for their capacity and activation. Cost of capacity reservation are covered by transmission tariffs and costs of activation are passed on to BRPs through an imbalance settlement mechanism (van der Veen *et al.*, 2012; Brijs *et al.*, 2015). Due to the increase in the use of RES in recent years, forecast errors also lead to further demand for balancing actions. TSO purchase reserve capacity mostly from conventional power stations that can be quickly activated in real time to cover imbalances of the system. It is not common to purchase reserve capacity through long-term contracts by TSO because of the requirement to a minimum guaranteed reserve capacity (Brijs *et al.*, 2015). It is possible to procure reserve capacity not only through contracted reserves but also day-ahead.

Due to the occurrences of forecasting errors (e.g., load and renewable generation) and technical failures (e.g., power station outages) in the system, the TSOs get involved in load-frequency control (LFC) processes in order to ensure that the system frequency preserves within permissible limits. Based on different historic developments, European TSOs work with diverse products and processes to keep the frequency and the system in balance. In Europe, balancing energy is organised in accordance with Commission Regulation (EU) 2017/1485 of 2 August 2017 in up to five steps: Frequency containment reserve (FCR), Imbalance netting (IN), Frequency restoration reserves with automatic activation (aFRR), Frequency restoration reserves with manual activation (mFRR), Replacement reserves (RR) (ENTSO-E, 2018).

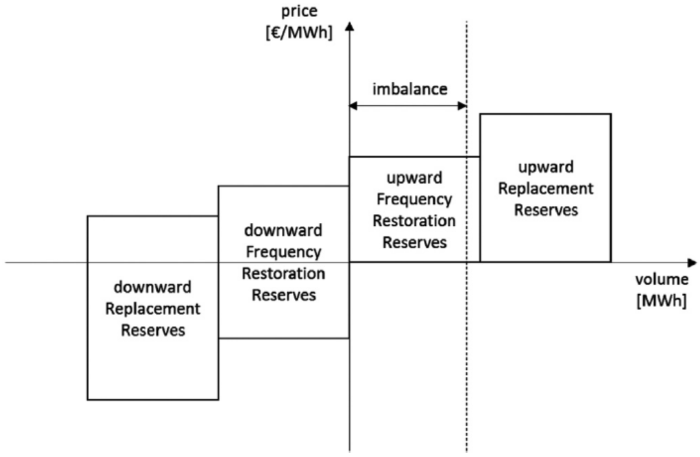
During the first seconds after the imbalance occurs, the FCRs are activated with respect to the measured frequency deviation in order to stabilise the frequency at a steady-state value below 50 Hertz throughout the entire synchronous area. The FCR activation is proportionally executed by control devices, which are implemented in the respective generating or demand units. The returning of frequency to 50 Hertz is performed by aFRR and mFRR. As the power imbalance results in the additional load flows that may exceed the available transmission capacity, the imbalances are regionally compensated by the TSOs in the LFC areas. As a basis for these processes, the TSO shall continuously calculate the deviation between the planned exchange and the measured power exchange of the LFC area that is corrected by the activation of FCR. It results to a value so-called frequency restoration control error (FRCE). The FRCE is used as an input to a frequency restoration controller working with a control cycle of a few seconds. It requests activation of aFRR until the FRCE gets to zero or all available aFRR are fully activated (ENTSO-E, 2020). Additionally, some TSOs use RRs, which are active power reserves that allow to replace or support the required level of FRR. It prepares the activated FRR for possible additional system imbalances, including generation reserves. As before, this process is actuated in the disturbed LFC area. Lastly, IN process is agreed between TSOs and creates a reduction of the amount of simultaneous and counteracting aFRR (ENTSO-E, 2018).

According to ENTSO-E (2018), there are different steps in the balancing process. Until the balancing energy gate closure time, BSPs can update, e.g., the balancing energy price bids or specific product bids by sending them to their TSO. Next, TSOs send them to the corresponding balancing energy exchange platform together with the available cross-zonal capacities and TSO balancing energy demands. The activation optimisation function accepts the common merit order lists and provides the selected bids to be activated. It aims to balance the system in the most efficient way.

Balancing responsibility is the key element of imbalance settlement. Depending on the imbalance situation, an imbalance cost is imposed on the BRPs that are not in energy balance over that imbalance settlement period (ENTSO-E, 2018). A fundamental characteristic is the calculation of imbalance prices. Two types of imbalance price mechanisms can be identified: a single-pricing scheme and a dual-pricing scheme (Brijs *et al.*, 2015). A single-pricing scheme is the preferable option according to guideline on electricity balancing (EBGL).[1] It is based on marginal activation

[1] The Commission Regulation (EU) 2017/2195 of 23 November 2017, which set detailed rules for the integration of balancing energy markets in Europe (ENTSO-E, 2018)

Figure 2.4: ENTSO-E: Merit Order for Activating Reserve Capacity



Source: Brijs *et al.* (2015)

price of reserves and BRPs who creates positive and negative imbalances in real-time pay a uniform imbalance price. The marginal pricing (or pay-as-cleared pricing) represents the price of the last bid of a standard product activated to cover the energy demand (ENTSO-E, 2018). A dual-pricing scheme is based on different prices when BRPS creates positive and negative imbalances. When the system imbalance sets in, the imbalance price is calculated using the average activation costs of reserve capacity. When dealing with the system imbalance, the imbalance price is usually set up using day-ahead prices (Hiroux & Saguan, 2010; Vandezande *et al.*, 2010). The average price is based on all accepted reserve capacity bids. Imbalance pricing is done during a certain time period, i.e, imbalance settlement period. Presently, countries in the EU apply the imbalance settlement period of fifteen, thirty or sixty minutes. It is required by the EBGL that TSOs have to harmonise the imbalance settlement period by 2025 to fifteen minutes (ENTSO-E, 2018).

The Figure 2.4 depicts the merit order effect for activating reserve capacity. If a negative system imbalance occurs the TSO activates upward reserves. From the right side of the Figure 2.4 it can be seen that this leads to a positive price and the TSO has to pay the BSPs. Recall, the different categories of reserve capacity are FCR, FRR and RR. Since FCR is activated only to stabilize the deviation of frequency after disturbances, it is not discussed any further. Firstly, since FRR is a fast response to the imbalance occurrence, it can be activated comprising of contracted and free bids from power stations. Secondly, the RR is activated as a slow response and it contains not only contracted and free bids from power stations but also an emergency capacity from other TSOs (Brijs *et al.*, 2015).

On the other hand, if a situation with a positive system imbalance occurs the TSO activates downward reserves. The left side of the Figure2.4 shows that the situation creates both negative and positive prices. As a result, the BSP has to pay the TSO or the TSO has to pay the BSP. Downward reserves are provided by the same reserve capacities, i.e. FRR and RR, comprising of contracted and free bids from power stations and in case of RR, possibly emergency capacity from other TSOs. The definition of downward reserves means that generators have to be willing to reduce their output however their energy is already sold on the forward market. Therefore, generators are usually willing to pay the TSO an amount that is equal to their saved operating costs. By this amount, the TSO remunerates the generators (i.e., BRPs) for having an excess supply. Nevertheless, when there is a shortage, providers of downward flexibility may bid positive activation prices, i.e., they are paid for the service provided, which leads to negative imbalance prices. If this happens, BRPs facing a positive imbalance are not remunerated by the TSO for their excess supply but they have to compensate the TSO (Brijs *et al.*, 2015).

2.1.2 The German Energy Market

Germany is a world leader in the use of renewable energy deployment. Their renewable energy policy dates back to 1970. In addition, following the accident at the Fukushima nuclear power plant in 2011, the German government has decided to phase out nuclear energy till 2022, putting Germany at the forefront of the transition to renewable energy, commonly known as Energiewende.

Germany's legislation Renewable Energy Sources Act (EEG), introduced 20 years ago, is responsible for the significant growth of onshore wind, solar and biogas by prioritizing the grid for these energy sources and guaranteeing feed-in tariffs. Feed-in tariffs were set above the market price per kWh for the owners of RES installations that fed power into the grid over a period of 20 years. Subsequent reform replaced the feed-in tariffs by an auction system for new installation of a certain size (Appunn, 2021).

Solar and onshore wind power together with offshore wind and hydro power cover half of the electricity consumption in Germany. According to the Solar Association (BSW), almost one in 10 kilowatt-hours consumed in Germany stemmed from solar energy in 2020 (Meza, 2021). Solar power for the first time accounted for 10 percent of Germany's total electricity generation in 2020 (Meza, 2021). In addition, Germany saw a 25 percent increase in solar rooftop installations in 2020 compared to 2019 and the trend seem to carry on. The increasing installation of photo-voltaic (PV) systems on the roofs of households may contribute to further increases in solar power generation in Germany (Meza, 2020).

After two major reforms of the EEG, the last changes came into force on January

1, 2021. The main goal of the EEG 2021 is to become officially greenhouse gas neutral by the middle of the century. Basically, it means that faster expansion of renewable power sources will be necessary. The installed capacity of PV should increase from 48 GW today to 100 GW, biomass to 8.4 GW, and offshore wind to 20 GW by 2030. For onshore wind energy, 95 GW should be on the grid in 2030 instead of today's 53 GW installed capacity. Renewable energy is supposed to provide 65 percent of total power consumption by 2030 (Appunn, 2021). The overview of the installed capacity of wind and solar in Germany between the years 2016 - 2020 is shown in the Table 2.1.

Table 2.1: Installed Capacity of wind and solar energy sources in Germany

Year	**2016**	**2017**	**2018**	**2019**	**2020**
Solar.GE (MW)	38686	40834	42804	45299	48206
Wind.GE.Onshore (MW)	41168	47042	51633	52792	53184
Wind.GE.Offshore (MW)	3283	4131	5051	6393	7504
Total Capacity (MW)	202803	213816	216198	222381	222968

Source: www.entsoe.eu

2.1.3 The Czech Energy Market

In The National Energy Mix for 2020 solar and wind energy sources represent 9.6 percent and 1.6 percent respectively. The development of installed capacity of RESs between years 2016 and 2020 is shown in Table 2.2. Compared to Germany, the Czech national energy mix is relatively stable.

The Czech Republic is not very suitable for the implementation of wind energy sources (Vágner, 2017; Kouřílek, 2019). For efficient use of wind for electricity production, it is recommended to have an average wind speed higher than 6 m/s. According to Vágner (2017) there is a relatively small and fragmented area with an average wind speed higher than 6 m/s. This does not mean that lower speed locations cannot be used, but they have far worse network conditions to compete with wind sources in higher wind speeds as well as other sources. The fundamental reason for the installation of wind power plants is the revenue resulting from the sale of the electricity produced to the electricity grid. This revenue depends on the wind conditions and the technology used as well as on the price per unit of electricity produced (Hanslian *et al.*, 2012). The current unsubsidized price of electricity delivered to the electricity grid is much lower than the amount needed to pay for the wind turbine investment (Vágner, 2017).

Solar energy sources have a much larger share in the residual mix thanks to Act 180/2005 Sb.. In 2006, due to this act, the setting of subsidized electricity prices from PV power plants came into force and investors were guaranteed CZK 15/kWh. As a result, between years 2008 and 2010, the Czech Republic experienced a solar boom, when due to the guaranteed price of electricity, the annual production of PV power plants increased from zero to approximately 2200 GWh per year (Boček *et al.*, 2019). As in Germany, the number of PV panel installations has increased in the Czech Republic in the last years. In 2018, in the Czech Republic, more than 1500 households asked for reimbursement of part of the installation costs, and the number has been increasing (Boček *et al.*, 2019).

Table 2.2: Installed Capacity of wind and solar energy sources in the Czech Republic

Year	**2016**	**2017**	**2018**	**2019**	**2020**
Solar.CZ (MW)	2067	2027	2040	2049	2061
Wind.CZ (MW)	277	277	308	316	339
Total Capacity (MW)	20627	20188	20845	20820	20576

Source: www.entsoe.eu

Regarding the balancing market, in the Czech Republic, the regulation of system imbalance is managed only by the ČEPS TSO. The Czech balancing market has an imbalance settlement period of one hour and dual-pricing. The size of the electricity reserve capacity is estimated according to the largest power plant unit in the system so that it can be replaced if necessary. In the Czech Republic it is Temelín with generation capacity of 1055 MW.

Chapter 3

Literature Review

Over the years, emphasis on balancing markets and balancing prices, in general, has been relatively scant, while modelling and forecasting of electricity prices has a long history of methodological development. Nevertheless, system imbalances follow similar characteristics as electricity prices thus, I follow some approaches regarding, for example, spot prices.

Electricity prices have leptokurtic distribution with heavy tails and excess kurtosis at the mean (Brooks, 2019). Moreover, they are characterized by high volatility, skewness, volatility clustering, large spikes, mean reversion and seasonality (Hagfors *et al.*, 2016; Fanone *et al.*, 2013). Fanone *et al.* (2013) capture dynamics of hourly EPEX day-ahead spot prices using arithmetic Lévy-based ARFIMA model. Authors use fractional ARIMA model since spot prices in the German electricity market exhibit negative prices and long memory, in addition to the items already listed. Hagfors *et al.* (2016) create quantile regression models for the electricity price in the UK for each trading period showing how sensitivity to various fundamental factors changes (e.g., gas price, coal price, carbon emission prices) across quantiles and time of day. It was discovered that sensitivity to gas is higher, relative to coal at high quantiles and lower at low quantiles. In semi-parametric specifications such as quantile regression, there is no need for assumptions about the distribution of the residuals. A number of models have been invented to detect different price formation processes for normal as well as extreme events, e.g., Karakatsani & Bunn (2008a) use time-varying parameter regression model and regime-switching model.

Various papers have been written on the effects of incorporating RESs on electricity markets, particularly, many researchers have focused on modelling and forecasting electricity prices that are affected by RES implementation. Gelabert *et al.* (2011) analyse the impact of the introduction of RES on the wholesale electricity prices in Spain in the period of $2005 - 2009$. The authors estimate the reduction of 2 EUR/MWh in wholesale price per GWh of electricity produced by the RES. Clò *et al.* (2015) claims, there is a reduction in Italian spot prices by 2.3 EUR/MWh for wind and 4.2 EUR/MWh for solar, on average. Next, Jónsson *et al.* (2010) study day-ahead wind power forecasts and their effect on electricity spot prices in the Western Danish price area of the Nord Pool's Elspot market using a non-parametric regression model. Jónsson *et al.* (2010) show the non-linear impact of wind power forecast and claim that on average the spot price decrease with increased wind power forecast.

Regarding balancing prices and balancing market in general, Nicholson *et al.* (2010) estimate the average marginal effects of wind generation on the balancing-energy market price in the Electric Reliability Council of Texas using ARMAX model. According to Nicholson *et al.* (2010), balancing price data have a persistent daily and weekly pattern. They use method of subtraction of the median hourly weekly pattern from the hourly balancing prices in each year and zone to remove the weekly trend. Nicholson *et al.* (2010) discover that wind generation has a measurable effect on balancing price, with a higher effect during the day (8am - 7pm).

Other studies on balancing markets focus on the minimisation of the cost of RESs (mainly wind energy) integration and the use of ancillary services. Paper written by Vandezande *et al.* (2010) analyse the structure of balancing markets. Authors suggest to not only depend on the support policies but to create an efficient market design capable to deploy wind energy. Moreover, Vandezande *et al.* (2010) require balancing markets to be integrated across borders. Similarly, Bueno-Lorenzo *et al.* (2013) researched the Spanish balancing market and the effectiveness of imbalance prices as market signals. Bueno-Lorenzo *et al.* (2013) came with a new optimal imbalance price strategy since the incomes obtained by the wind power producers were not fully correlated with their contribution to the overall power deviations. Frade *et al.* (2019) dealt with the impact of wind power production on the Portuguese energy market in the period of $2012-2016$, where wind generation accounted for 23% of demand. They discover that wind balancing costs are 2.17 euros per MWh of generated wind energy and suggested setting up incentives for efficient wind forecasting. Further, Swinand & Godel (2012) surveyed the impact of average marginal wind generation on balancing costs of Great Britain and estimated that balancing cost were of about €0.01/MWh, which is significantly less than previously estimated €1–4/MWh (Gross *et al.*, 2006).

With the increasing incorporation of RESs in many countries, the forecast errors that burden the balancing systems have become an important research topic. Aïd *et al.* (2016) studied the problem of optimal trading for a power producer. They further considered the case of errors in the prediction of wind power generation resulting in jumps on the residual demand forecast. Paper closest to my topic of interest is written by Goodarzi *et al.* (2019). They study whether it would be beneficial for Germany TSO to enhance the quality and market availability of RES forecast data.

Using OLS and quantile regression, they investigate the effect of wind and solar energy forecast errors on imbalance volumes and spot electricity prices in one calendar year 2014. Goodarzi *et al.* (2019) use EPEX SPOT intraday quarter-hourly data (7 days a week) and implement two control variables, two-period lagged value of imbalance and two-period lagged value of EPEX Spot price, called adaptive imbalance response and adaptive price, respectively. These variables measure the level of learning of market participants from past events. It is assumed that with the absence of good forecasts, there should be more adaptive behaviour in the price formation (Goodarzi *et al.*, 2019; Karakatsani & Bunn, 2008b). Since the German balancing market closes 30 minutes before the delivery, authors also use two-period lagged realized total load. Goodarzi *et al.* (2019) create dummy variables for thirteen different price levels, per season (winter, summer) and peak/off-peak periods of the day. As a result, Goodarzi *et al.* (2019) find out that with higher wind and solar forecast errors the absolute values of imbalance volumes increase, as well as the effect on real-time prices. The effect is greater at the tails of the distribution. Sirin & Yilmaz (2021) focus on the Turkish balancing market, specifically on the effects of RESs (including wind, solar and run-of-river hydro) on the balancing market prices. They employ quantile regression and the results show that system marginal price declines as RES generation increases. In addition, there is a higher probability of positive imbalance because the positive difference between real-time and forecasted RES variable increases.

Focusing on the seasonality, Tashpulatov (2013) prefer to use smooth sine and cosine functions instead of dummy variables to capture seasonality since it should lead to a more parsimonious model. This is consistent with statements in Koopman *et al.* (2007). The data of my thesis will be processed in the same way.

Furthermore, the effect of RESs on the volatility of system imbalances is part of my study. Ketterer (2014) studies the impact of wind power generation on the spot electricity price level and volatility. Like in Clò *et al.* (2015) using *GARCH* model, he shows that wind energy decreases the price level but increases its volatility. He further claims that price volatility is also smoothed out by the possibility of exports to neighbouring countries. This was used in Haxhimusa (2018), who discuss the volatility of the French spot prices affected by the import of the renewable power produced in Germany and use import, resp. export of electricity from, resp. to Germany as an explanatory variables. Further, Koopman *et al.* (2007) study daily electricity spot prices on four European power markets using periodic seasonal Reg-ARFIMA–GARCH to explain the dynamics in the conditional mean and variance of log prices.

With an increasing number of PV units at the residential level, smart grid components such as smart home and battery energy management systems, renewable energy systems or demand response activities, need to have accurate electricity demand forecasts for the operation of electricity distribution networks to work successfully. Yildiz *et al.* (2017) analyse the impacts of household electricity load consumption profile and PV size on PV self-consumption using clustering model and further analyse the influence of different seasons on the self-consumption forecast. Yildiz *et al.* (2018) implement smart meter data alongside with weather variables to forecast individual household loads using machine learning models.

Chapter 4

Data

In this chapter, I describe the data used in this study and the process of creating the final dataset. Firstly, I depict data sources from which were data collected. Secondly, I discuss why were those data chosen. Thirdly, I describe data transformation followed by descriptive statistics of exogenous variables. Finally, I conduct exploratory data analysis on the dependent variable. The final variables I employ for this study are presented 4.3.

4.1 Data Description

Since European TSOs are obliged to make data available online, it is quite easy to collect all necessary information for my study. Firstly, I use data from the Czech market operator (OTE a.s.) annual reports.[1] Among others, OTE a.s. publishes system imbalance of whole Czech transmission grid for each hour (hereafter referred to as "system imbalance"). System imbalance is in my main point of interest, i.e., my dependent variable. The imbalance values are presented in megawatts (MW). Recall, system imbalance is the difference between instantaneous power production and power consumption. An adequate model has to address both of these elements. To explore the variation in power production I will use both, day-ahead forecast and actual production of Czech and German RESs. Further, I will use physical flows of electricity from Germany to the Czech Republic, i.e., imports. The consumption part of the equation will be addressed by the day-ahead forecast and actual load in the Czech grid. Load represents immediate power consumption in the power grid. All data are available on the ENTSO-E website.[2] For German RESs, I use day-ahead forecast and actual production of wind and solar power sources. German TSOs report these data in quarter-hourly granularity. The value of the measurement is

[1] www.ote-cr.cz
[2] www.entsoe.eu

MW. Likewise, I use the day-ahead forecast and actual production of Czech solar power plants. Further, the actual production of Czech wind power sources will be used. Unfortunately, day-ahead forecasts of Czech wind power production are not available. Thus, I create an arbitrary day-ahead forecast by lagging the actual production of wind. Since Czech wind power production data are time-series with hourly granularity, I lag production data by 24 periods.

Recall, TSO's ultimate goal is to have a balanced grid in every moment. In other words, TSO's job is to keep the system imbalance as close to zero as possible. To do so, production and consumption forecasts are taken into account. However, what TSO cannot predict is an error in the given forecast. Thus, I construct variables called *forecast errors*. For each production type, I simply make a difference between an actual and a forecasted value. The same is done for a consumption, i.e., a load variable. Lastly, I create a variable import which is the amount of imported energy from Germany to the Czech Republic in a given hour. The final dataset consists of 41293 observations from 1 January 2016 to 16 September 2020.

4.1.1 Data pre-processing

Before I created the final dataset with *forecast error* variables, I had to pre-process the data. To have a consistent dataset, I need to resample German data from 15 minutes granularity to hourly granularity. Thus, I take the arithmetic average out of four observation I currently have for each hour. The result is average power (MW) delivered to a transmission grid throughout a given hour. Since data have hourly granularity, it can be said that the resulting value is equal to the amount of energy (MWh) produced during that hour.

Moreover, both Czech Republic and Germany have daylight saving time changes resulting in one day in March with 23 hours and one day in October with 25 hours each year. Therefore, to maintain the consistency of the dataset, I add or delete one observation to the days concerned.

Finally, I have to deal with missing values in the dataset. The dataset contained missing values on both the wind and solar sides. In the case of solar, it is not appropriate to use linear interpolation to fill the missing values, because there is an extreme difference between day and night. Hence, as a replacement for the missing value, I calculate the arithmetic average value of the given hour in a given day and given month across different years. For the wind variables, I use linear interpolation.

4.2 Descriptive Statistics

The general descriptive statistics of the explanatory variables (measured in MW) is summarized in Table 4.1. The times series plots of explanatory variables are shown in the Figure 4.1. The mean of *forecast.error* variables varies, non of them is equal to zero. Variable *Wind.CZ.FE* has the mean closest to zero, however, it may be due to the missing day-ahead forecast, which is replaced by 24 lagged periods of actual production. The highest value of mean is in the variable *WindOn.GE.FE* (103.26 MW), which means that on average, actual production of German onshore wind parks was by 103.26 MW per hour higher than the forecast. Variable *Import* has mean approximately equal to 605 MW, i.e., on average, Germany imports to Czech Republic 605 MW per hour. All variables representing *forecast.error* have positive excess kurtosis, which means they have leptocrutic distributions, i.e, show heavy tails, indicating large outliers. The coefficient of skewness of variables *Solar.FE.CZ*, *Wind.FE.CZ*, *WindOff.GE.FE* and *Load.FE.CZ* is slightly negative (skewed left), i.e., higher number of data points have high values. On the other hand, variables *WindOn.GE.FE* and *Solar.FE.GE* are positively skewed (skewed right), which leave low values more probable. It means that *WindOn.GE.FE* and *Solar.FE.GE* have higher probability to have low forecast errors. In general, RESs *forecast.errors.GE* are characterized by higher extreme values in absolute terms than RESs *forecast.errors.CZ*. This is an obvious information, because many times more wind parks and solar power plants are installed in Germany, i.e., the probability of significant forecast error is higher. Finally, variable *Import* has negative excess kurtosis. In other words, distribution have flat tails indicating small outliers. As assumed, the Anderson-Darling test for normality, reject the null hypothesis of normality for all explanatory variables. Finally, the Augmented Dickey-Fuller test for stationarity was conducted for the explanatory data as well, confirming stationarity.

Table 4.1: Descriptive Statistics

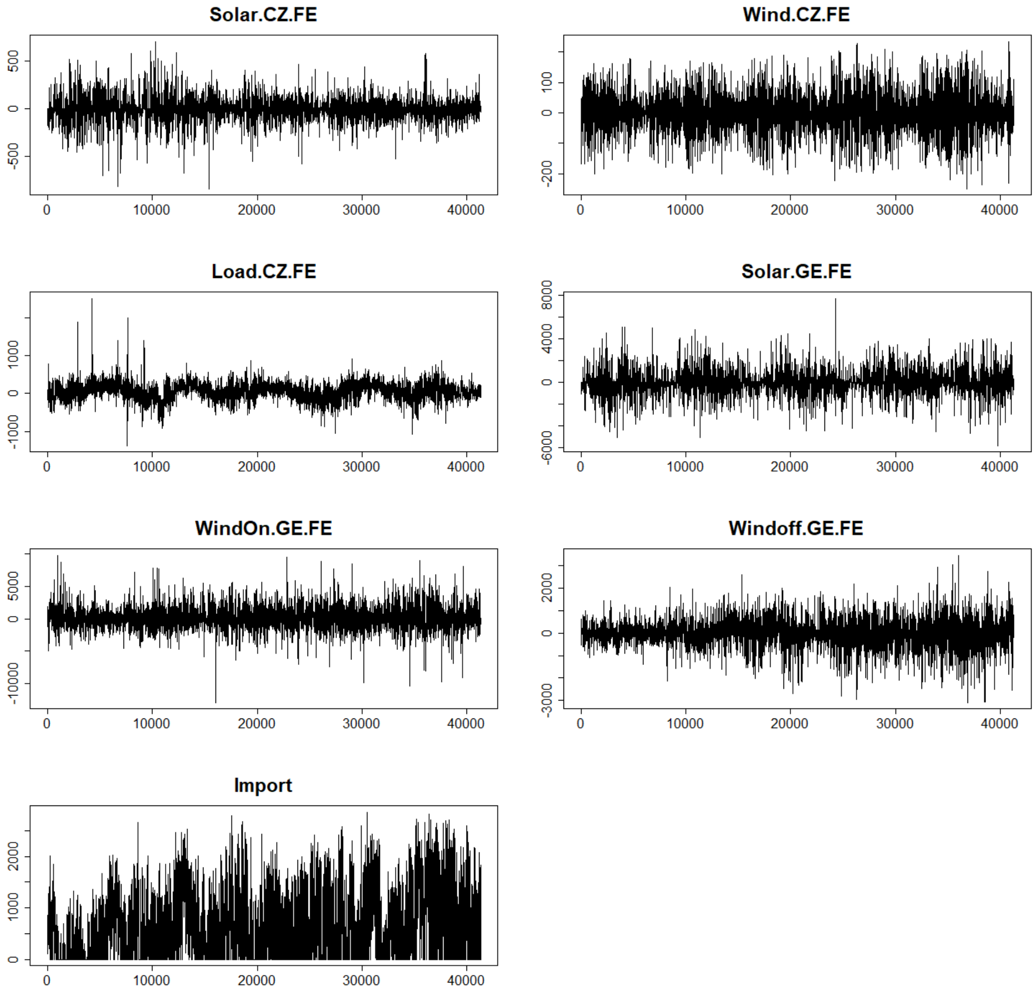
	n	mean	sd	median	min	max	skew	kurtosis
Solar.CZ.FE	41293.00	-2.85	73.85	0.00	-843.00	702.00	-0.41	15.53
Wind.CZ.FE	41293.00	-0.04	57.13	0.00	-250.00	235.00	-0.05	1.07
Solar.GE.FE	41293.00	-0.45	709.78	0.00	-5811.50	7752.25	0.12	9.60
WindOn.GE.FE	41293.00	103.26	1412.95	10.75	-12971.75	9846.25	0.26	4.03
WindOff.GE.FE	41293.00	-25.93	499.86	-16.00	-3096.75	3468.50	-0.49	3.33
Load.CZ.FE	41293.00	56.07	209.41	68	-1383.00	2524.00	-0.17	3.19
Import	41293.00	604.94	635.86	438.00	0.00	2860.00	0.80	-0.41

It is also important to examine the correlation among all variables in the model, correlation matrix is displayed in Table 4.2. There is no significant correlation among the variables, correlation does not exceed value 0.33, in absolute terms.

Table 4.2: Correlation Matrix

	System.Imbalance.CZ	Solar.CZ.FE	Wind.CZ.FE	Load.CZ.FE	Solar.GE.FE	WindOn.GE.FE	WindOff.GE.FE	Import
System.Imbalance.CZ	1.00	0.33	0.03	-0.21	0.12	0.01	-0.02	-0.06
Solar.CZ.FE	0.33	1.00	-0.05	-0.05	0.22	-0.00	0.01	-0.03
Wind.CZ.FE	0.03	-0.05	1.00	0.00	-0.05	0.06	-0.06	0.14
Load.CZ.FE	-0.21	-0.05	0.00	1.00	-0.04	-0.01	-0.00	0.15
Solar.GE.FE	0.12	0.22	-0.05	-0.04	1.00	-0.01	0.01	0.02
WindOn.GE.FE	0.01	-0.00	0.06	-0.01	-0.01	1.00	0.15	0.12
WindOff.GE.FE	-0.02	0.01	-0.06	-0.00	0.01	0.15	1.00	-0.06
Import	-0.06	-0.03	0.14	0.15	0.02	0.12	-0.06	1.00

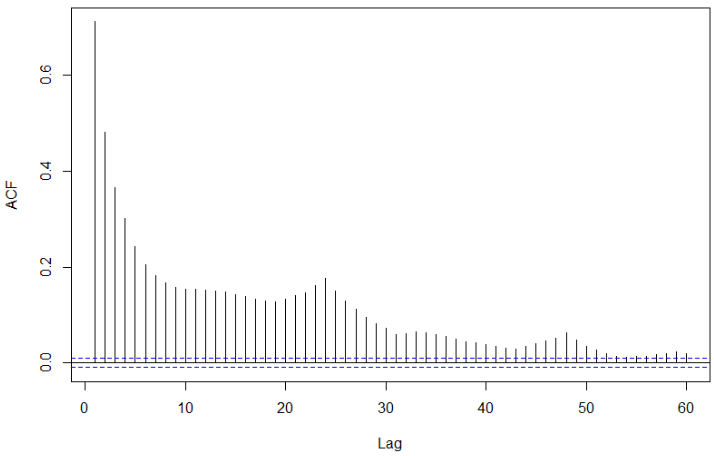
Figure 4.1: Time Series Plot for Explanatory Variables



4.3 Dependent Variable

Firstly, we have to focus on extreme values/outliers, since they have impact on the overall robustness of the whole model. After examining the extreme values in the dependent variable *System.Imbalance.CZ*, I came to the conclusion that the extreme events did not occur together with the extreme values of variables *forecast.errors*,

Figure 4.2: Autocorrelation Function of *System.Imbalance.CZ*



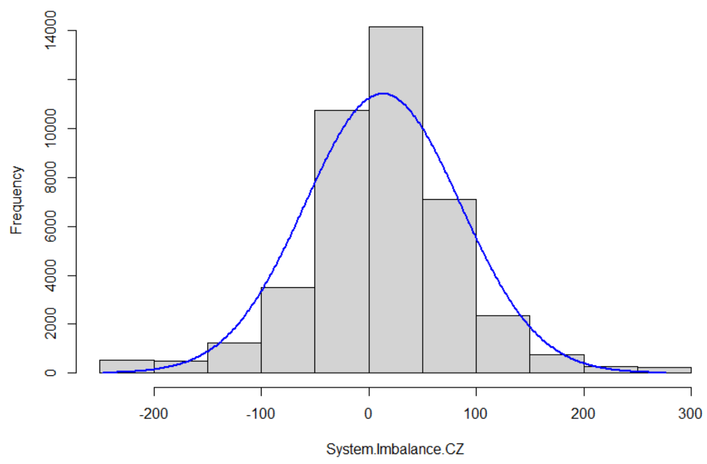
but were caused by outages in the electricity system (for example, 14. September 2018 or 22. January 2020, Šrom (2020)Thus, I decided to use methods used in Haxhimusa (2018), Ketterer (2014) or Kouřílek (2019) and adjust all observations that are three standard deviation from the original *System.Imbalance.CZ* variable. This filter specifies a value equal to three times the standard deviation of the mean for outliers.

Further, according to the Augmented Dickey-Fuller stationarity test, the original *System.Imbalance.CZ* variable is stationary. The autocorrelation function (ACF) of *System.Imbalance.CZ* shows a long-memory dependency, i.e., the decay of the ACF is hyperbolic rather than exponential, to see in Figure 4.2. Moreover, the spikes after 24 hours signal daily seasonality (more detailed in Subsection 4.3.1). Figure 4.3 shows histogram of *System.Imbalance.CZ*. The normal distribution line in Figure 4.3 suggests that *System.Imbalance.CZ* has leptocrutic distribution, which is also confirmed by the Anderson-Darling test for normality rejecting the null hypothesis of normality.

4.3.1 Seasonality

Next, I need to identify long-term variations and seasonality in the data. The autocorrelation function (ACF) can be used to show some of the seasonal pattern, to see in Figure 4.2. It is generally assumed that spot prices are characterized by high (i.e., hourly) and low (i.e. annual, half-annual, seasonal) periodic components, which also corresponds to the behavior of the system imbalances (Fanone *et al.*, 2013). Any time series data can be understood as a combination of cosine and sine waves with

Figure 4.3: Histogram of *System.Imbalance.CZ*



varying periods, which is useful to determine cyclic behaviour. To find out structure of the cyclical behavior, and identify dominant frequencies in time series a spectral (Fourier) analysis is used. Employing Fourier analysis, it is possible to model time series with seasonal components such as:

$$A\cos(2\pi\omega t + \phi) \tag{4.1}$$

where A denotes the amplitude, i.e, define the maximum absolute height of the wave, ω is the frequency and ϕ is the phase shift, i.e., the starting point, in angle degrees, for the cosine wave. The Equation 4.1 can be rewritten using trigonometric identity this way:

$$A\cos(2\pi\omega t + \phi) = \beta_1\cos(2\pi\omega t) + \beta_2\sin(2\pi\omega t), \tag{4.2}$$

where $\beta_1 = A\cos(\phi)$ and $\beta_2 = -A\sin(\phi)$

Frequency can be rewritten as $\omega = 1/T$, i.e., fraction of the finished cycle completed in a single time period. The Equation 4.2 suggests to use variables $\cos(2\pi\omega t)$ and $\sin(2\pi\omega t)$ with coefficients $\beta_1 = A\cos(\phi)$ and $\beta_2 = -A\sin(\phi)$ as explanatory variables to model the seasonality. Fourier transformation of time series $y(t)$ on domain $[0, T]$ is defined as:

$$y(\omega) = \int_0^T y(t)e^{-i\omega t}dt$$

where i is the imaginary unit ($i^2 = -1$) (Tashpulatov, 2013). I use the Fast Fourier Transform (FFT) analysis, which is a way to reduce the complexity of the

Fourier transform. The frequency is set as $\omega = 2\pi/T$ and FFT determines the most significant values of T as: $8297, 20736, 12, 24$ and 168. The value 8297 approximately corresponds to a year seasonality, while $12, 24$ and 168 precisely match with half-daily, daily and weekly seasonality. The value 20736 does not correspond close enough to 2 years seasonality, therefore I do not use it.

Table 4.3: Variables

Variables	Description
System.Imbalance.CZ	The electricity supply-demand imbalance. It is negative when the electricity demand exceeds the supply and the TSO needs to activate extra reserve, and positive otherwise.
Solar.CZ.FE	Czech actual solar production minus its day-ahead forecast
Wind.CZ.FE	Czech actual wind production minus its day-ahead forecast
Solar.GE.FE	German actual solar production minus its day-ahead forecast
WindOn.GE.FE	German actual wind production minus its day-ahead forecast from wind parks situated on land
WindOff.GE.FE	German actual wind production minus its day-ahead forecast from wind parks situated in North and Baltic Sea.
Load.CZ.FE	Czech actual consumption minus its day-ahead forecast
Import	Energy imported from Germany to Czech Republic
cosDay	Controls for daily seasonality determined as $cos(2\pi t/24)$
sinDay	Controls for daily seasonality determined as $sin(2\pi t/24)$
cosHalfDay	Controls for half-daily seasonality determined as $sin(2\pi t/12)$
sinHalfDay	Controls for half-daily seasonality determined as $sin(2\pi t/12)$
cosWeek	Controls for weekly seasonality determined as $cos(2\pi t/168)$
sinWeek	Controls for weekly seasonality determined as $sin(2\pi t/168)$
cosYear	Controls for yearly seasonality determined as $cos(2\pi t/8294)$
sinYear	Controls for yearly seasonality determined as $sin(2\pi t/8294)$

Chapter 5

Methodology

This chapter describes theoretical framework that I use for estimating to what extend the RESs forecast errors affect the Czech system imbalance. Firstly, I use quantile regression based on Goodarzi *et al.* (2019). As well as Goodarzi *et al.* (2019) or Karakatsani & Bunn (2008b), I introduce control variable, adaptive system imbalance, in this thesis. Adaptive imbalance is the imbalance two-periods before delivery and measures how much market participant learn from the past events. My study differs from Goodarzi *et al.* (2019) in seasonality approach, since I use smooth sine and cosine functions to control for seasonality (Tashpulatov, 2013). Further, I focus not only on Czech RESs but I also consider cross-border transmission with Germany, i.e., RESs of Germany. Secondly, I employ *ARFIMA-GARCH* model to capture effects of variables RESs *forecast.errors* on the volatility of Czech imbalance volumes. This is based on methodologies developed by Fanone *et al.* (2013) and Koopman *et al.* (2007). I also describe several formal tests used in the analysis.

5.1 Quantile Regression

The Czech system imbalances are leptocrutic and have time varying volatility. Quantile regression is a semi-parametric model, where we do not have to make such a restrictive assumptions, e.g., on distributions of data. In addition, there is no need for specific distributional form for the error term (Sirin & Yilmaz, 2021). Thus, this model is beneficial for my research. Due to quantile regression, we can examine the relationship between dependent and explanatory variables throughout the entire distribution, creating a more complete picture of how fundamental factors affect the imbalance volumes in different volume ranges (Hagfors *et al.*, 2016). Estimating the tail dependencies of imbalance distributions accurately is crucial from risk perspective for the electricity market participants.

Koenker & Bassett (1978) introduced quantile regression as a natural extension of a linear regression, which attempts to compute a set of regression functions, each of which corresponds to a different quantile of the conditional distribution of the dependent variable, in my research, system imbalance. Since least squares give large weight to large deviations, least absolute deviations (LAD) estimator was introduced to solve at least part of the problem (Greene, 2002). The LAD estimator optimize:

$$\min_{\boldsymbol{\beta}\in\mathbb{R}^p} \sum_{i=1}^{n} \left| y_i - \mathbf{x}_i'\boldsymbol{\beta} \right|$$

where $\{y_1,\ldots,y_n\}$ is a random sample of the dependent variable $Y, x_i \in \mathbb{R}^p$ is the covariate vector corresponding to the ith observation y_i (Chen & Wei, 2005). LAD estimator minimize absolute deviations, hence it is more resistent to changes in extreme values. The LAD estimator estimates the median regression (measure of central tendency), which is only a special case of quantile regression. To get estimates of the different conditional quantile functions, we change absolute values to $\rho_\tau(.)$ and solve (Koenker & Geling, 2001):

$$\hat{\beta}\left(\tau\right) = \min_{\boldsymbol{\beta}\in\mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau \left(y_i - \mathbf{x}_i'\boldsymbol{\beta} \right)$$

where $\rho_\tau(u) = u(\tau - I(u < 0))$ is the linear "check function" and it gives asymmetric weights to the error according to the quantile and the sign of the error (Koenker & Geling, 2001; Koenker & Bassett, 1978). Then the general quantile regression model is as follows:

$$Q\left[y \mid \mathbf{x}, \tau\right] = \mathbf{x}'\boldsymbol{\beta}(\boldsymbol{\tau}) \tag{5.1}$$

such that $\text{Prob}\left[y \leq \mathbf{x}'\boldsymbol{\beta}(\boldsymbol{\tau}) \mid \mathbf{x}\right] = \tau$ for $0 < \tau < 1$.

One possible problem regarding this approach is that the estimated conditional quantile function may have a "quantile crossing" problem. In other words, $Q[y \mid \mathbf{x}, \tau]$ is not monotonically increasing in the quantiles. He (1997) propose restricted version of regression quantiles (RRQ) that prevents the occurence of crossing while maintaining sufficient modelling flexibility. I will follow this approach.

Quantile regression model I use to analyze to what extend RES forecast errors affect the Czech system imbalance is described by following equation:

$$\begin{aligned} System.Imbalance.CZ = {} & \beta_1 + \beta_2 System.Imbalance.CZ_{t-2} \\ & + \beta_3 Load.CZ.FE + \beta_4 Solar.CZ.FE \\ & + \beta_5 Wind.CZ.FE + \beta_6 Solar.GE.FE \\ & + \beta_7 WindOn.GE.FE + \beta_8 WindOff.GE.FE \\ & + \beta_9 Import + \beta_{10} cosDay + \beta_{11} sinDay \\ & + \beta_{12} cosHalfDay + \beta_{13} sinHalfDay + \beta_{14} cosWeek \\ & + \beta_{15} sinWeek + \beta_{16} cosYear + \beta_{17} sinYear + \varepsilon \end{aligned}$$

Testing the significance of quantile regression is still well under investigation and no single approach has yet received general support (Hagfors *et al.*, 2016; Volgushev *et al.*, 2013).

5.2 Formal Tests

In this section, I describe some formal tests used in the analysis of *ARFIMA-GARCH* models. First, I present the stationarity tests. Then I focus on the detection of heteroskedasticity of residuals and lastly, I introduce the goodness of fit tests.

5.2.1 Augmented-Dickey Fuller Test

Augmented Dickey-Fuller test (ADF) is a unit root test for stationarity. It is an extension of Dickey-Fuller test by Dickey & Fuller (1979) since it allows for higher-order autoregressive processes (Brooks, 2019). The following regression is employed:

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + u_t \tag{5.2}$$

where $H_0 : \psi = 0$, i.e, series contains a unit root, and $H_1 : \psi < 0$, series is stationary. Variable u_t is white noise.

5.2.2 Kwiatkowski-Phillips-Schmidt-Shin Test

Another test for stationarity is the KPSS test introduced by Kwiatkowski *et al.* (1992). While the ADF test has as the null hypothesis a presence of unit root (time series is integrated of order 1), the KPSS test has the presence of unit root as the alternative hypothesis, while null hypothesis is the stationarity of time series.

Specifically, the null hypothesis is that deviations of the series are short memory (Lee & Schmidt, 1996).

5.2.3 ARCH Engle's Test

Engle (1982) introduced $ARCH - LM$ test as a standard approach to detect autoregressive conditional heteroscedasticity. In uncorrelated time series there is still possibility of serial dependence due to a dynamic conditional variance process. If residuals are heteroscedastic, the squared residuals are autocorrelated and the series have autoregressive conditional heteroscedastic ($ARCH$) effects. $ARCH - LM$ test is computed from an auxiliary test regression, and the null hypothesis is that a series of residuals r_t exhibits no conditional heteroscedasticity ($ARCH$ effects), against the alternative that an $ARCH(q)$ model describes the series (Sjölander, 2011). The null hypothesis of no $ARCH(q)$ is examined by running the following regression:

$$r_t^2 = \hat{\delta}_0 + \sum_{s=1}^{q} \hat{\delta}_s r_{t-s}^2 + e_t \tag{5.3}$$

where there is at least one $\hat{\delta}_s \neq 0, s = 1, ..., q$. The test statistic is the Lagrange multiplier statistic TR^2 and it is asymptotically distributed as a chi-square with q degrees of freedom. T is the sample size and R^2 is the coefficient of determination from fitting the $ARCH(q)$ model for a number of lags q via regression.

5.2.4 Akaike and Bayesian Information Criterion

Information criteria measure of the goodness of fit, i.e, they are an alternative to adjusted R^2. They have the same role as the adjusted R^2 in the standard scoring problems, so that the models are penalized for model complexity, i.e., having too many independent variables. Leading examples are the Akaike information criterion (AIC) and the Bayes information criterion (BIC). For k being the number of parameters, N the number of observations and L the maximized values of the likelihood function of the estimated model, AIC and BIC are defined as:

$$AIC = -2\ln(L) + 2k$$

$$BIC = -2\ln(L) + k\ln(N)$$

The model with the lowest AIC or BIC is considered the best. BIC penalizes more for inclusion of additional parameters with increasing number of observations. BIC is the most suitable for predictions, while AIC is the best on theoretical basis.

5.3 ARFIMA-GARCH models

This section is devoted to the hypothesis that RESs forecast errors increase the volatility of the Czech system imbalance. For this purpose, I use generalized autoregressive conditionally heteroscedastic ($GARCH$) model. As already noted, the system imbalance is a long memory process (to see in the Figure 4.2), hence I use autoregressive fractionally integrated moving average ($ARFIMA$) model to capture the mean equation.

5.3.1 Long-memory Process

The long memory process was first emphasized in hydrological data based paper by (Hurst, 1951). Long memory process, is characterized by the Hurst parameter or self-similarity parameter, meaning that an object consists of sub-units on multiple levels that resemble the structure of the whole object (Stadnytska *et al.*, 2010). The probability, there is the dependence or coupling between events can be indicated or measured by the estimated Hurst parameter, or the Hurst exponent (H) (Liu *et al.*, 2017). Self-similarity implies that H is different from 0.5. H lies in the interval $\langle 0, 1\rangle$ and $H = 0.5$ stands for no dependence between observations of the increment process $X_t = Y_t - Y_{t-1}$, i.e, correlations at all non-zero lags are zero. If $H < 0.5$, time series is anti-persistent process, i.e, negative correlations between increments. Conversely, $H > 0.5$ implies persistence, i.e, positive correlation between increments. For each value of $0 < H < 1$, exists exactly one Gaussian process X_t that is the stationary increment of a self-similar process Y_t (Beran & Terrin, 1994). X_t is known as fractional Gaussian noise (*fGn*) and Y_t is called fractional Brownian motion (*fBm*) (Mandelbrot & Van Ness, 1968)[1]. In addition to Hurst parameter, other basic properties of long memory processes are a hyperbolically decaying ACF, to see in Figure 4.2 and diverging spectral density as the frequency ω tends to zero such that:

$$f(\omega) \sim c_f|\omega|^{-\alpha}, \omega \to 0 \tag{5.4}$$

where $f(.)$ is a spectral density of stationary X_t, $\alpha \in (0, 1)$ and $c_f > 0$. Conversely, the short memory processes have quick and exponential decay of the ACF and their spectral density converges to finite constant.

Using $ARIMA$ methodology (Box & Jenkins, 1970), Hosking (1981) showed that self-similar hyperbolically decaying autocorrelations of time series can be parsimoniously modeled using the differencing parameter d (Stadnytska *et al.*, 2010). The

[1] Gaussian process is stationary process, i.e., constant mean and constant variance, while Brownian motion is a non-stationary process with stationary increments (Stadnytska *et al.*, 2010)

process of fractionally integrated moving average *ARFIMA(p,d,q)* is stationary and invertible for $d \in (-0.5, 0.5)$. Parameters p and q model the short memory process and parameter d captures the long-memory dependencies between values. Values from $-0.5 < d < 0$ indicate antipersistance of the process, while process with $d > 0$ demonstrates persistence. Hence, finite long memory can be modeled using d from interval $(0, 0.5)$. For $0.5 \leq d < 1$, the process is non-stationary but mean reverting.

For *fGn* and increments of *fBm*, the following relationship holds:

$$d = H - \frac{1}{2} \tag{5.5}$$

Thus, there are two main ways to model the long-memory dependence: using H within the scope of fractal analysis or using the fractional differencing parameter d from *ARFIMA* (Stadnytska *et al.*, 2010). In this study, I model the long-memory dependence by means of the second approach, i.e., d within the *ARFIMA* framework.

5.3.2 Fractional ARIMA

ARFIMA(p,d,q) is an natural extension of *ARIMA(p,d,q)*. Let's first describe processes introduced by Box & Jenkins (1970) on which *ARFIMA* is based.

ARIMA

ARIMA is an extension of *ARMA* models. The *ARIMA* models can only capture short run memory property, since d is limited in the range of integer order (for$d \geq 0$) (Liu *et al.*, 2017). Specifically, series are considered as *ARIMA(p,1,q)*, when the process $y_t - y_{t-1} = (1-L)y_t$ is stationary and invertible *ARMA(p,q)* with parameter $d = 1$, i.e, *ARMA(p,q)* series are differenced once. *ARIMA* can be described as follows:

$$\Phi(L)(1-L)^d X_t = \Theta(L)\varepsilon_t$$

where the polynomials $\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$ and $\Theta(L) = 1 - \theta_1 L - \cdots - \theta_q L^q$. Coefficients $\phi_1, \cdots, \phi_p$ are autoregressive parameters of the model, $\theta_1, \cdots, \theta_q$ are moving average parameters.

The stationarity conditions are those of an $AR(p)$:

$$z^p \Phi\left(z^{-1}\right) = 0 \equiv z^p - \phi_1 z^{p-1} - \cdots - \phi_p = 0 \Leftrightarrow |z_i| < 1$$

for $i = 1, \cdots, p$. The invertibility conditions are those of an $MA(q)$ process:

$$z^q \Theta\left(z^{-1}\right) = 0 \equiv z^q + \theta_1 z^{q-1} + \cdots + \theta_q = 0 \Leftrightarrow |\tilde{z}_i| < 1$$

for $i = 1, \cdots, q$.

ARFIMA

ARFIMA is a generalization of *ARIMA*, where d is allowed to be any real value as follows:

$$\Phi(L)(1-L)^d X_t = \Theta(L)\varepsilon_t$$

where X_t is stationary and $d \in (-0.5, 0.5)$. Then X_t is called a *fractional ARIMA(p,d,q)* process. As already mention, the parameter d captures the long-term behavior, while parameters p and q allow for modelling short-term behaviour.

The operator $(1-L)^d$ can be defined in a natural way by using binomial expansion for any real number d with Euler's Gamma function $\Gamma(.)$ (Liu *et al.*, 2017):

$$(1-L)^d = \textstyle\sum_{k=0}^{\infty} \begin{pmatrix} d \\ k \end{pmatrix} (-1)^k L^k$$
$$\begin{pmatrix} d \\ k \end{pmatrix} = \tfrac{d!}{k!(d-k)!} = \tfrac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}$$

Thus, after application of the fractional differencing operator d, we get:

$$(1-L)^d X_t = \frac{\Theta(L)}{\Phi(L)}\varepsilon_t$$

where $(1-L)^d X_t$ is the ordinary *ARMA(p,q)* process.

For $-\pi \leq \omega \leq \pi$, X_t has the spectral density $h(\omega)$:

$$\begin{aligned} h_{ARFIMA}(\omega) &= \frac{\sigma_x^2}{2\pi} \frac{\left|\theta\left(e^{-i\omega}\right)\right|^2}{\left|\phi\left(e^{-i\omega}\right)\right|^2} \left|1-e^{-i\omega}\right|^{-2d} \\ \text{where,} & \\ h_{ARMA}(\omega) &= \frac{\sigma_x^2}{2\pi} \frac{\left|\theta\left(e^{-i\omega}\right)\right|^2}{\left|\phi\left(e^{-i\omega}\right)\right|^2} \end{aligned} \tag{5.6}$$

in comparison to the ordinary *ARMA*, the *ARFIMA* process adds one more filtering.

5.3.3 Estimation of d

Over the years, a number of methods have been developed to estimate the long memory parameters H and d (Beran & Terrin, 1994). Estimators of d can be divided into two groups, semiparametric and parametric. I implement both of them. I use the procedure in line with approach of Stadnytska *et al.* (2010).

First, I use Geweke-Porter-Hudak (GPH) and Sperio estimators both belonging to the group of semiparametric estimators. The GPH estimates d by means of the linear regression of the log periodogram on a deterministic regressor (Geweke

& Porter-Hudak, 1983). The GPH is the OLS estimator of the slope parameter in this regression, taking into account only the lowest frequency ordinates of the log periodogram. Reisen (1994) came up with the Sperio estimator, which estimates the memory parameter d employing the regression equation that uses the smoothed periodogram function for modeling the spectral density. Both GPH and Sperio estimator can be applied to non-stationary data, i.e. the range of d does not have to lie in interval $(0, 0.5)$.

Secondly, I use parametric maximum likelihood method to estimate d optimizing the fit of the assumed *(p,d,q)* model to the autocovariance function of the data (Stadnytska *et al.*, 2010). Specifically, I use approximate maximum likelihood method of Haslett & Raftery (1989). In addition, contrary to the GPH and Sperio estimators, this method provides estimates of the short-memory dependencies p and q. Moreover, approximate ML method can be applied only on persistent stationary data. i.e., d is limited to interval $(0, 0.5)$.

5.3.4 GARCH

In *ARIMA* processes, the error term ε_t is considered as white noise. Nevertheless, variance of the error terms is not always constant. Hence, it makes sense to consider a model describing heteroscedastic errors and how the variance of the errors evolves. Models suitable for this purpose are e.g., *ARCH* and *GARCH* models (Brooks, 2019).

The *GARCH* model was developed by Bollerslev (1986) and Taylor (1986). It is a generalization of *ARCH* model, where autocorrelation in volatility is modelled so that it allows the conditional variance of the error term to depend on the previous value of the squared error (Brooks, 2019). In addition, the *GARCH* model allows the conditional variance to depend on its previous lags. *GARCH(1,1)* is the most robust and most commonly used modeling procedure and it has the following form:

$$\begin{aligned} a_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{aligned} \tag{5.7}$$

where $(\alpha_1 + \beta_1) < 1$. Due to GARCH model, volatility clustering might be well detected. Large past shocks of return and volatility in $t-1$ gives large shocks to volatility in time t.

5.3.5 ARFIMA-GARCH Model Estimation

The aim of my model is to find out whether variables *forecast.errors* have effect on the volatility of the dependent variable *System.Imbalance.CZ*. The *ARFIMA(p,d,q)*-

GARCH(1,1) model is defined as follows:

$$\begin{aligned}\Phi(L)(1-L)^d X_t &= \boldsymbol{\beta} Z_t + \Theta(L) a_t \\ a_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \boldsymbol{\gamma} Y_t\end{aligned} \tag{5.8}$$

where Z_t demonstrates periodic *sin* and *cos* functions capturing seasonality and Y_t are RESs *forecast.errors* variables, to see in Table 4.3. RESs variables *forecast.errors* consist of positive and negative values with mean close to zero. I follow Samiev (2013) and use these variables in absolute terms to accurately capture the volatility of the system imbalance.

Following the approach of Stadnytska *et al.* (2010), the process of estimation of the fractional differencing parameter d and the *ARFIMA(p,d,q)* model is as follows. Firstly, I need to visually check the ACF function that needs to show extremely slow decay indicating a finite long memory. Further, I need to employ the unit root and stationarity tests. Specifically, I use ADF test with the null hypothesis that series has a unit root ($H_0 : d = 1, H_1 : d < 1$), and KPSS test with null hypothesis that series is stationary ($H_0 : d = 0, H_1 : d > 0$). If the ADF test rejects null hypothesis and KPSS tests does not reject null hypothesis, it implies that $0 \leq d < 0.5$. On the other hand, if both the ADF and KPSS tests reject their null hypothesis, it implies that $0 < d < 1$. Lastly, if ADF test is not significant, while KPSS test is significant, it implies that $d \geq 1$. Hence, the series is non-stationary and I use the classical differencing and check the ACF again.

If the first scenario happens, i.e, series is stationary, I can use approach introduced by Wagenmakers *et al.* (2004). Wagenmakers *et al.* (2004) suggest determining the maximum likelihood of a time series using the *ARMA* and *ARFIMA* models, and then selecting the best model using the *AIC*. According to Torre *et al.* (2007), *BIC* provides better results than *AIC* especially for low autoregressive or moving average coefficients. If $AIC(ARFIMA) > AIC(ARMA)$, there is no long memory in the series, i.e, $d = 0$. On the other hand, if $AIC(ARMA) > AIC(ARFIMA)$, there is a finite long memory in the data and I can use approximate maximum likelihood method of Haslett & Raftery (1989) controlling for the short-memory parameters p and q.

If the second scenario occurs, i.e., $0 < d < 1$, the pre-estimates of d need to be obtained using, e.g., Sperio or GPH estimator, where the d has to lie in the interval $(0, 0.5)$. Moreover, I need to check the confidence intervals of these estimators. If parameter d might lie in the non-stationary area ($d > 0.5$), further tests are necessary to solve the problem of stationarity. Stadnytska *et al.* (2010) use the Fractional Dickey-Fuller test of Dolado *et al.* (2002). This test is based on significance testing of autoregressive coefficient as follows:

$$\Delta^{d_0} x_t = \phi \Delta^{d_1} x_{t-1} + u_t \tag{5.9}$$

where u_t is a stationary process and the hypothesis are: $H_0 : d = 0.5, H_1 : d < 0.5$. If I confirm stationarity, I can again use the approximate maximum likelihood method of Haslett & Raftery (1989) where I can control for short-memory behaviour with parameters p and q. I choose the *ARFIMA(p,d,q)* model based on corresponding value of AIC. Finally, I fit the *ARFIMA(p,d,q)-GARCH(1,1)* described in Equation 5.8 via maximum likelihood estimator. Residuals are not normally distributed thus, instead of Gaussian innovations, I use *GARCH* with t-distributed innovations.

Chapter 6

Results and Discussion

In this chapter, I describe the results of the quantile regression and the *ARFIMA-GARCH* model introduced in the previous Chapter 5.

6.1 Quantile Regression

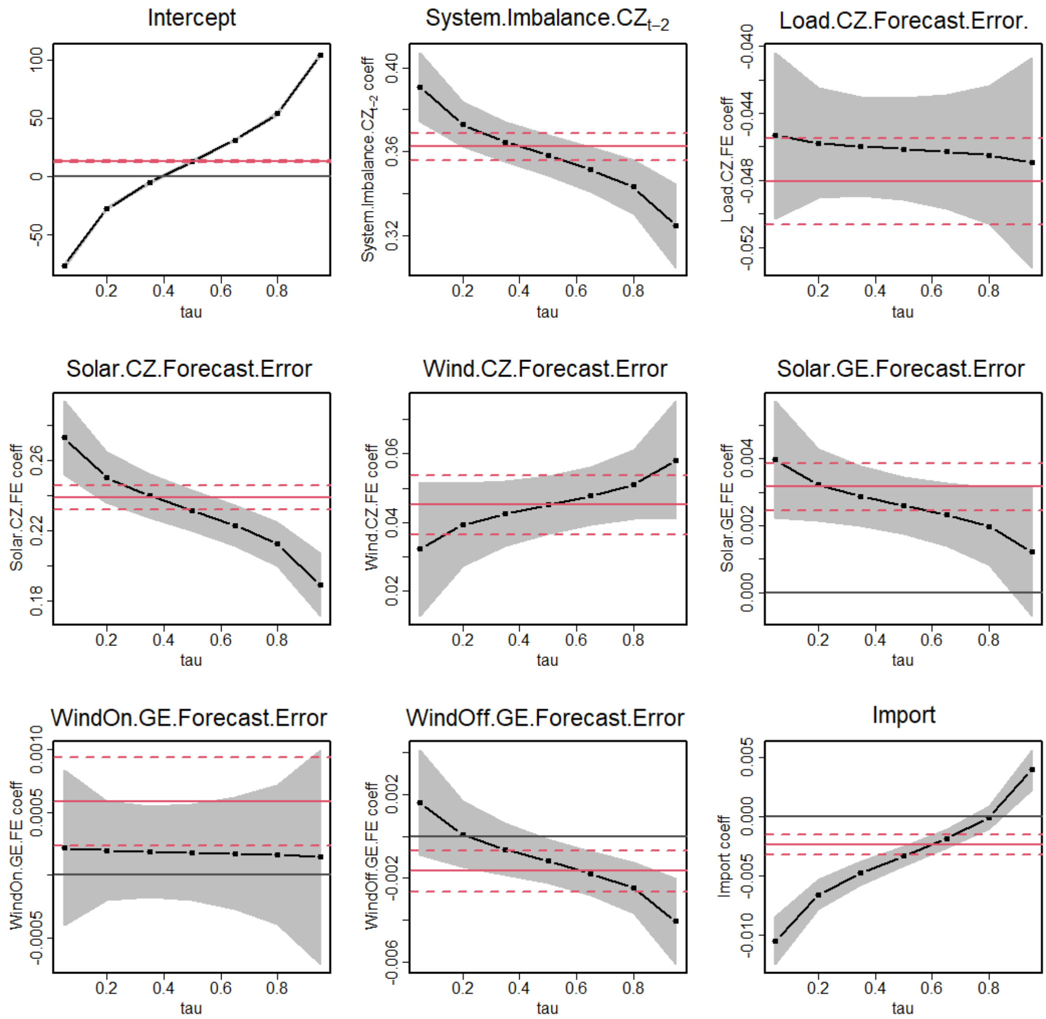
The results show that in absolute terms, RESs variables *forecast.error* increase *System.Imbalance.CZ*, to see in Table 6.1. All in all, solar forecast errors have a larger impact than wind forecast errors on the Czech system imbalance.

Recall, all *forecast.error* variables have leptokurtic distribution with the mean close to zero, containing both negative and positive values. The same applies to the dependent variable *System.Imbalance.CZ*, since negative and positive system imbalance, emerge when negative and positive forecast errors occur, respectively.

In Figure 6.1, for the 9 most important coefficients (including the intercept), I plot 7 different quantile regressions estimates for τ from 0.05, to 0.95 (solid line with dots). The point estimates for each covariate might be explained as the impact of a one-unit change of the covariate on the *System.Imbalance.CZ* holding other covariates fixed. Each plot in Figure 6.1 has a horizontal quantile (τ) scale and the vertical scale in MW indicates the covariate effect. The dashed line in each of the plots presents the ordinary least squares estimate of the conditional mean effect. The 95% confidence intervals for the least-squares estimate are the two dotted lines. The grey area shows the 95% pointwise confidence interval for the quantile regression estimates (Koenker & Geling, 2001).

In all of the panels of Figure 6.1, most of the quantile regression estimates are out of the confidence intervals for the OLS regression, at least at some point and specifically at the tails of the distribution. It suggests that the effects of the covariates might not be constant across the conditional distribution of the explanatory variable (Koenker & Geling, 2001; Koenker & Machado, 1999). The only exceptions are the

Figure 6.1: OLS and Quantile Regression Estimates for *System.Imbalance.CZ* Model



variable *Load.CZ.FE* and *Wind.CZ.FE*, where the quantile regression results seem to be consistent with the OLS results. Nevertheless, to confirm my visual inspection I use ANOVA test that the slopes are the same at all quantiles. As I expected, I do not reject the null hypothesis for variables *Load.CZ.FE*, *Wind.CZ.FE*, meaning the coefficients of quantile regression do not differ from the OLS estimates. For the remaining six variables I reject the null hypothesis that coefficients are the same across quantiles at least at 5% significance level. In addition, we can see that variable *WindOn.GE.FE* is statistically insignificant.

Table 6.1 suggests that both Czech solar and Czech wind forecast errors have positive significant effects on the system imbalance. Czech solar forecast errors have a stronger effect on the system imbalance across all quantiles.

Table 6.1: Quantile Regression Results

	Dependent variable:		
	System.Imbalance.CZ		
	0.05	0.5	0.95
	(1)	(2)	(3)
$System.Imbalance.CZ_{t-2}$	0.390***	0.358***	0.325***
	(0.011)	(0.005)	(0.012)
Load.CZ.FE	−0.0453***	−0.046***	−0.047***
	(0.004)	(0.002)	(0.004)
Solar.CZ.FE	0.273 ***	0.231***	0.189***
	(0.014)	(0.007)	(0.014)
Wind.CZ.FE	0.032 ***	0.045***	0.058***
	(0.009)	(0.005)	(0.012)
Solar.GE.FE	0.004***	0.003***	0.001
	(0.001)	(0.0005)	(0.001)
WindOn.GE.FE	0.0002	0.0002	0.0001
	(0.0005)	(0.0002)	(0.0005)
WindOff.GE.FE	0.002	−0.001	−0.004**
	(0.001)	(0.0006)	(0.001)
Import	−0.010***	−0.003***	0.004***
	(0.001)	(0.0005)	(0.001)
cosDay	23.590***	0.553	−23.037***
	(0.905)	(0.550)	(1.039)
sinDay	7.804 ***	1.959***	−4.026***
	(0.905)	(0.447)	(0.901)
cosHalfDay	−3.490*	3.509***	10.677***
	(1.023)	(0.441)	(0.964)
sinHalfDay	5.700 ***	3.865***	1.986**
	(1.046)	(0.406)	(0.919)
cosWeek	2.260***	−0.918	−4.173**
	(0.800)	(0.425)	(1.063)
sinWeek	3.006***	3.114***	3.225***
	(0.830)	(0.413)	(0.983)
cosYear	−5.863***	−6.374***	−6.897***
	(0.706)	(0.490)	(1.016)
sinYear	−4.180***	−0.923	2.412***
	(0.819)	(0.454)	(0.806)
Constant	−76.198***	13.288***	104.920***
	(1.183)	(0.404)	(1.046)
Observations	41,293	41,293	41,293
Note:			*p<0.1; **p<0.05; ***p<0.01

Czech solar forecast errors have the largest impact at the lowest 0.05 quantile, i.e., the effect is more pronounced on negative system imbalance. Specifically, one MW increase in *Solar.FE.CZ* leads to 0.273 MW increase in the system imbalance. At 0.95 quantile, one MW increase in *Solar.CZ.FE* leads to 0.189 MW increase in the system imbalance. Recall, the 0.95 quantile is associated with positive forecast errors and positive system imbalance, while 0.05 quantile with negative forecast errors and negative system imbalance. The higher effect at the 0.05 quantile shows that shortage of solar energy, compared to predicted values, is more challenging to compensate than the positive excess.

Regarding *Wind.CZ.FE*, its estimates do not differ from the OLS estimates. Thus, *Wind.CZ.FE* does not have a different effect at the tails of the distribution. Nevertheless, there is missing official day-ahead forecast data for this variable. I replaced them with 24 periods lagged actual wind production data. Thus, the real effect might be different.

Regarding German forecast error variables, the only statistically significant variable at 0.05 quantile is the *Solar.GE.FE*. The variable *WindOn.GE.FE* is statistically insignificant and *WindOff.GE.FE* is statistically significant at 0.95 quantile. As we can see in Figure 6.1, variable *Solar.GE.FE* has a statistically significant effect till 0.85 quantile. The strongest effect of *Solar.GE.FE* is at the 0.05 quantile, specifically, one MW increase in *Solar.GE.FE* leads to 0.004 MW increase in the system imbalance. Variables *WindOn.GE.FE* and *WindOff.GE.FE* have a questionable impact on the system imbalance. The variable *WindOff.GE.FE* indicates that if wind parks situated in the North and Baltic sea produce more energy than the forecasted value, it reduces positive system imbalance in the Czech Republic. This result is misleading and might result from the extreme distance between Czech Republic and North and Baltic sea. Moreover, if German wind onshore forecast errors do not have impact on the Czech system imbalance, it is not probable that German wind offshore forecast errors would have any impact.

In line with my expectation, *Load.Error.CZ* has a negative effect on the system imbalance. Since load is an instantaneous offtake of power the higher it is, the lower is imbalance. Load forecast error has quite a uniform effect over the whole range of the distribution.

Next, I examine *Import* variable. According to Figure 6.1, coefficients of import increase as the quantile level increases. This variable worsens the system imbalance at both tail quantiles. At the 0.95 quantile, electricity inflow from Germany increases the positive system imbalance. At the 0.05 quantile the effect of variable *Import* goes against my original beliefs. It says that an electricity inflow from Germany worsens the negative system imbalance, i.e., there is a lack of electricity. This result might be explained by the fact that the actual amount of electricity imported from Germany was lower than predicted.

Lastly, I introduce the adaptive system imbalance $System.Imbalance.CZ_{t-2}$. Adaptive system imbalance positively affects the *System.Imbalance.CZ*. It means that negative values are followed by negative values, and positive values are followed by positive. In other words, high system imbalance persists in the balancing market. $System.Imbalance.CZ_{t-2}$ has a stronger effect at the 0.05 quantile. It suggests that it is more difficult to deal with negative system imbalances than with positive system imbalances.

Most of the variables controlling for the seasonality are significant at least at one quantile. Since my main interest are the tails, I can conclude that all periodic variables significant at least at 5% significance level. All coefficients of seasonal periodic variables at 0.05 quantile have positive values. On the other hand, *cosDay*, *sinDay*, *cosWeek* and *cosYear* have negative values at the 0.95 quantile.

6.2 ARFIMA-GARCH model

I have captured the effects on the mean equation using the quantile regression in the previous Section 6.1. Using *ARFIMA* model to interpret the mean equation is complicated, since applying fractional differencing changes the units of the problem. In this section, the goal is to identify if any of RESs *forecast.error* variables affect the variance *System.Imbalance.CZ*.

6.2.1 Mean Equation

As already mentioned in Section 4.3, there is an extremely slow decay in the ACF indicating a finite long memory, to see in Figure 4.2. For both ADF test and KPSS test I reject their null hypothesis at 1% significance level, ($t_{ADF} = -25.112, t_{KPSS} = 1.3486$). These results imply that $0 < d < 1$. Pre-estimates of d performed by GPH ($\hat{d} = 0.141, SE_{\hat{d}} = 0.0473$) and Sperio ($\hat{d} = 0.119, SE_{\hat{d}} = 0.0159$) estimators confirm the result of visual inspection of ACF. The 95% confidence intervals, $(0.0487, 0.234)$ from GPH estimator and $(0.088, 0.15)$ from Sperio procedure do not include values of d from non-stationary area. Thus, there is no need for further testing such as

Fractional Dickey-Fuller test and I can confirm finite long memory. Hence, I can use approximate maximum likelihood method method of Haslett & Raftery (1989) to estimate *ARFIMA(p,d,q)* models.

I estimate different specification of *ARFIMA(p,d,q)* model and choose the one with the lowest AIC, to see in Table A.1 in Appendix. For each model, I use exogenous variables *sinDay*, *cosHalfDay*, *sinHalfDay* and *sinWeek* since other periodic seasonality variables turned out to be statistically insignificant. The AIC neither BIC of the models significantly differ. To achieve a parsimonious model, I choose the $ARFIMA(1, d, 1)$, where all coefficients are significant, the values of autoregressive and moving average coefficients are not close to one and it exhibits little autocorrelation in the residuals, to see in Figure A.1 in Appendix A.[1]

6.2.2 Variance Equation

Next, I examine squared residuals obtained from the $ARFIMA(1, \hat{d}, 1)$. According to ACF in the Figure 6.2, the squared residuals exhibit significant correlations, indicating time-varying variance. To confirm the visual impression, I perform ARCH-LM test with the null hypothesis of no heteroskedasticity across the residuals. The results reject the null hypothesis, suggesting conditional heteroskedasticity is present in the residuals. The ACF of squared residuals is used to determine the order of ARCH effect. I use $GARCH(1, 1)$ since it exhibits the best fit and higher-order GARCH, e.g., $GARCH(2, 1)$ was not significant.

Results of the final $ARFIMA(1, d, 1) - GARCH(1, 1)$ model is presented in the Table 6.2. Fractional differencing parameter d is equal to 0.225. *ARMA* coefficients of the mean equation are equal to 0.323 and 0.138, respectively. Each of the coefficients is significantly less than one suggesting a stationary and invertible process. The seasonal periodic variables are all positive and significant. Regarding the conditional variance equation, the $GARCH(1, 1)$ coefficients are equal to 0.200 and 0.413, respectively. Both coefficients are statistically significant at 1% significance level. Their sum is 0.614 conforming stationarity, but denoting movements in the conditional variance persistent, implying long-lasting periods of high volatility. The results of conditional variance clearly state that variables *Solar.CZ.FE*, *Wind.CZ.FE* and *Solar.GE.FE* affect the volatility of the system imbalance. Variables *WindOn.GE.FE* and *WindOff.GE.FE* were excluded from the variance equation since they were insignificant. As in the mean equation, the highest effect has *Solar.CZ.FE* with coefficient equal to 7.874 and the lowest effect has *Solar.GE.FE* with coefficient equal

[1] I observe that the ACF of the residuals still exhibits a daily periodic component. This means that it might be useful to use a more sophisticated method for modelling the seasonal component, however, this is not the goal of my research. These results are similar to Fanone *et al.* (2013), who studies hourly EPEX spot prices data with similar behaviour.

to 0.189. ARCH-LM test confirms adequately fitted ARCH process till 7th lag. The variable shape is equal to 6.01 indicating that residuals follow t-distribution.

Figure 6.2: Squared residuals of ARFIMA(1,$\hat{d}$,1)

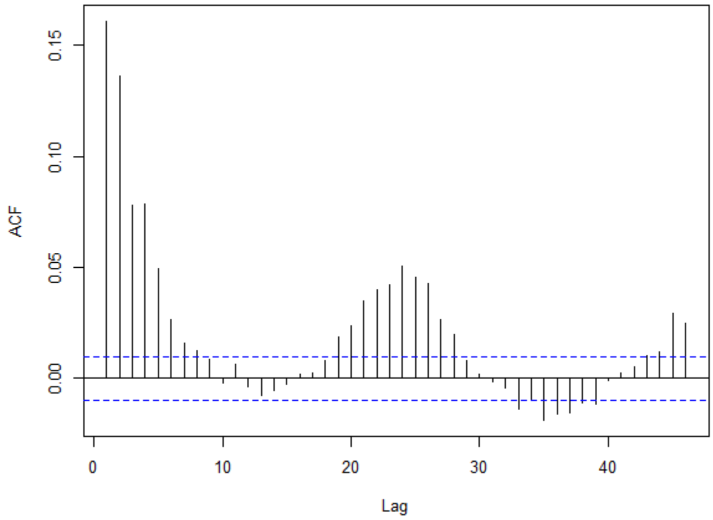


Table 6.2: Results $ARFIMA(1, d, 1)GARCH(1, 1)$ model

Mean Equation				
Variable	Coeff	Std. Error	t-value	p-value
μ	18.806	3.460	5.436	0.00001
$AR(1)$	0.323	0.018	17.630	0.00001
$MA(1)$	0.138	0.013	10.709	0.00001
d	0.225	0.009	24.133	0.00001
sinDay	4.486	0.660	6.799	0.00001
cosHalfDay	1.081	0.511	2.117	0.034
sinHalfDay	5.203	0.519	10.018	0.00001
sinWeek	5.256	1.015	5.180	0.00001
Conditional Variance Equation				
α_1	0.200	0.008	25.099	0.00001
β_1	0.413	0.016	25.134	0.00001
Solar.CZ.FE	7.874	0.524	15.015	0.00001
Wind.CZ.FE	1.778	0.280	6.361	0.00001
Solar.GE.FE	0.189	0.035	5.432	0.00001
shape	6.010	0.173	34.750	0.00001
AIC	10.542			
Log-likelihood value	-217647.1			

6.3 Discussion

The final value of the system imbalance in a given hour determines the amount of regulating energy activated by the Czech TSO ČEPS. Since each BRP is financially responsible for its caused imbalances, the subsequent settlement imbalance price is reflected in the regulated component of the electricity price for the end customer (Klamka, 2016). Reducing the absolute level of imbalances will reduce costs for all, TSO, market participants and end consumers.

Based on the results above, I conclude that larger Czech wind and solar production forecast errors increase the Czech system imbalance. The findings show that this relationship is more influential for Czech solar forecast errors, having the strongest effect at the lowest quantile. It might be due to the fact that the share of solar energy is significantly higher than the share of wind energy. In National Residual Mix, the share of solar energy is 9.6%, while the share of wind energy only 1.6%.

Not only Czech but also German solar forecast errors have statistically significant effects on the size of the system imbalance. In general, German forecast errors have significantly lower effect on the Czech system imbalance. Germany is the leader in the transition to a green energy sector with 45% of installed capacity in RESs. As already mentioned, electricity flows the least resistant path. Since the Czech Republic is neighbouring Germany, there were excess power spillovers from the German grid destabilizing the Czech electricity system. In the beginning of 2017, phase shifters were installed to regulate disruptive electricity flows. This result might suggest that phase shifters work accurately. Another explanation of the significantly lower effect of German forecast errors on the Czech system imbalance might be due to the fact that the German electricity market has better day-ahead predictions. The German electricity market works with 15 minutes data. This higher frequency allows to further decrease uncertainty in weather conditions, which affect the production of RES and subsequently improve their forecastability.

As already written, there is an increasing number of RESs embedded in the transmission system at end-use level, creating difficulties in obtaining production data (Goodarzi *et al.*, 2019). While there are technical and land restrictions for wind farms installation especially at the residential level, the opposite is true for PV power plants. According to Wehrmann (2020), solar arrays with a capacity rating below 10 kWp accounted for over two-thirds of the total 1.6 million German installations in 2018. Recall, Germany saw a 25% increase in solar rooftop installations in 2020 compared to 2019 and the trend continues. In the Czech Republic, the trend seems to be similar but on a much smaller scale. Thus, the difference in significance of solar and wind forecast errors might be driven by uncontrolled residential PV generation. Yildiz *et al.* (2018) deal with this problem and propose models, which could improve forecastability of households production based on data from smart metering.

Regarding adaptive system imbalance lagged by two hours, the results show that for ČEPS it is more difficult to compensate activation of upwards energy reserves than the activation of downward reserves. In other words, it is more complicated to increase production of power plants than to reduce their power. Moreover, ČEPS has to pay the BSPs when activating the upwards reserves, while activating downward reserves can result in both scenarios: ČEPS pays the BSP or the BSP pays the ČEPS.

Lastly, RESs forecast errors affect the volatility of the system imbalance. Variables Czech solar forecast errors, Czech wind forecast errors and German solar forecast errors influence volatility clustering of the system imbalance. It means that high system imbalance tends to persist due to these variables, implying persisting costs for ČEPS and market participants. Volatility clustering complicates control over the balancing market. The Czech solar forecast errors have the highest impact on the volatility as in the mean equation. This might be again due to six time higher installed capacity of PV plants compared to wind parks in the Czech Republic. The low effect of German solar forecast errors on the volatility might be due to well-captured spillovers of electricity by the phase shifters at the Czech-German borders. The lower significance of wind forecast errors might be again due to the uncontrolled residential PV production, as mentioned above.

The overall results are in line with Goodarzi *et al.* (2019) and Sirin & Yilmaz (2021). Nevertheless, Goodarzi *et al.* (2019) have stronger effects of the wind forecast errors especially at high quantiles, where real-time production exceeds real-time consumption. It is reasonable since even in 2015 the wind onshore and offshore production had a total share of about 13% in German electricity production. Moreover, wind production is spread throughout whole day, unlike the solar production.

Chapter 7

Conclusion

The renewable energy sector has developed rapidly since the beginning of this century. In 2014, the European Union adopted its Energy and Climate Framework, where each member should reduce the greenhouse gas emissions by at least 40% by the year 2030 in comparison with 1990 levels. Environmental benefits and low operation costs exceed the high capital cost required to install these technologies. On the other hand, increased inputs of renewable energy into the energy market create complexity. According to Fanone *et al.* (2013), RESs installation substantially impacts the dynamics of intra-day electricity prices by increasing the likelihood of negative prices.

Electricity has to be consumed at the time of its production because electricity travels almost instantaneously and follows the least resistant path. These basic characteristics are of importance because of the intermittency of electricity generation from RESs caused by the weather conditions. Thus, accurate forecasts of wind and solar energy production are essential for well functioning balancing market and energy markets in general. Moreover, better forecasts would reduce balancing cost for market participants and operational costs for TSOs. As electricity does not respect borders, it is important for governments and local retailers to know to which extent is the balancing market influenced by foreign green energy production.

In the thesis, I confirm that larger Czech wind and solar forecast errors increase the Czech system imbalance in absolute terms. The findings show that this relationship is the most influential for Czech solar energy forecast errors. The effect of Czech solar forecast error is the strongest across all quantiles. I confirm that larger German solar forecast errors increase the Czech system imbalance. Nevertheless, German solar forecast errors have significantly lower effect on the Czech system imbalance. I build on approach developed by Goodarzi *et al.* (2019), using the quantile regression.

Further, the results show that Czech solar and wind forecast errors and German solar forecast errors increase the volatility of the Czech system imbalance. To achieve this result I used *ARFIMA-GARCH* model, since the Czech system imbalance follow long-memory behavior.

To reduce power imbalances in the Czech electricity grid, market settlement procedures should be changed to reflect the temporal characteristics of RESs. Currently, the Czech Republic uses hourly settlement. Nevertheless, this does not reflect inter-hourly variability and intermittency associated with RESs. Shorter settlement periods would therefore better reflect the true value of RESs and mitigate the imbalances in the balancing market. As already mentioned in Chapter 2, The EBGL requires all TSOs to harmonise the imbalance settlement period by 2025, to a period of 15 minutes. Further research, after the implementation of this shorter settlement period, would be an interesting extension of this thesis.

In addition, TSOs should be motivated to provide more accurate wind and solar forecasts of the market. As already mentioned in Chapter 1, the UK energy regulator has incorporated an accuracy target of demand forecasts into its regulatory regime (Ofgem, 2018). In 2015, the EU adopted regulation forcing all national TSOs with more than one percent feed-in of wind or solar power generation per year to publish day-ahead generation forecast for these energy sources. These two principles should be combined and extended to the balancing market in order to obtain more accurate forecasts. This is not a trivial challenge that involves just improving forecasting techniques. Substantial amount of RES is being installed in the system at the end-use level, not as metered generation on the transmission system (Goodarzi *et al.*, 2019). As a result, the production data is harder to capture. Regulatory intervention should also include issues of greater data transparency of end-user behaviour (Goodarzi *et al.*, 2019).

Lastly, further research may contain complex effects of RESs on the energy market as a whole. In this thesis, interactions between day-ahead, intra-day and balancing market are not discussed.

Bibliography

Aïd, R., P. Gruet, & H. Pham (2016): "An optimal trading problem in intraday electricity markets." *Mathematics and Financial Economics* **10(1)**: pp. 49–85.

Appunn, K. (2021): "What's new in germany's renewable energy act 2021." `https://www.cleanenergywire.org/factsheets/whats-new-germanys-renewable-energy-act-2021`.

Beran, J. & N. Terrin (1994): "Estimation of the long-memory parameter, based on a multivariate central limit theorem." *Journal of Time Series Analysis* **15(3)**: pp. 269–278.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* **31(3)**: pp. 307–327.

Box, G. E. & G. J. Jenkins (1970): "Time series analysis."

Boček, J., J. Cibulka, & Štěpán Sedláček (2019): "Dobrodružství fotovoltaiky. proč se ze solární naděje stala zatracovaná cesta české energetiky?" .

Brijs, T., K. De Vos, C. De Jonghe, & R. Belmans (2015): "Statistical analysis of negative prices in european balancing markets." *Renewable Energy* **80**: pp. 53–60.

Brooks, C. (2019): *Introductory econometrics for finance*. Cambridge university press.

Bueno-Lorenzo, M., M. Á. Moreno, & J. Usaola (2013): "Analysis of the imbalance price scheme in the spanish electricity market: A wind power test case." *Energy policy* **62**: pp. 1010–1019.

Chen, C. & Y. Wei (2005): "Computational issues for quantile regression." *Sankhyā: The Indian Journal of Statistics (2003-2007)* **67(2)**: pp. 399–417.

Clò, S., A. Cataldi, & P. Zoppoli (2015): "The merit-order effect in the italian power market: The impact of solar and wind generation on national wholesale electricity prices." *Energy Policy* **77**: pp. 79–88.

CONRAD, H. & G. STAACKE (2016): "Current issues and challenges in the geography of energy in germany." *Bulletin de l'association de géographes français. Géographies* **93(93-1)**: pp. 26–45.

DICKEY, D. A. & W. A. FULLER (1979): "Distribution of the estimators for autoregressive time series with a unit root." *Journal of the American statistical association* **74(366a)**: pp. 427–431.

DOLADO, J. J., J. GONZALO, & L. MAYORAL (2002): "A fractional dickey–fuller test for unit roots." *Econometrica* **70(5)**: pp. 1963–2006.

ENGLE, R. F. (1982): "A general approach to lagrange multiplier model diagnostics." *Journal of Econometrics* **20(1)**: pp. 83–104.

ENTSO-E (2018): "Electricity balancing in europe."

ENTSO-E (2020): "European network of transmission system operators for electricity: Balancing report."

ERBACH, G. (2016): "Understanding electricity markets in the EU." *European Parliamentary Research Service* .

FANONE, E., A. GAMBA, & M. PROKOPCZUK (2013): "The case of negative day-ahead electricity prices." *Energy Economics* **35**: pp. 22–34.

FLÁŠAR, P. *et al.* (2016): "Trh s elektřinou: Úvod do liberalizované energetiky." *Praha: Asociace energetických manažerů* .

FRADE, P. M., J. P. PEREIRA, J. SANTANA, & J. CATALÃO (2019): "Wind balancing costs in a power system with high wind penetration–evidence from portugal." *Energy Policy* **132**: pp. 702–713.

GELABERT, L., X. LABANDEIRA, & P. LINARES (2011): "An ex-post analysis of the effect of renewables and cogeneration on spanish electricity prices." *Energy economics* **33**: pp. S59–S65.

GEWEKE, J. & S. PORTER-HUDAK (1983): "The estimation and application of long memory time series models." *Journal of time series analysis* **4(4)**: pp. 221–238.

GOODARZI, S., H. N. PERERA, & D. BUNN (2019): "The impact of renewable energy forecast errors on imbalance volumes and electricity spot prices." *Energy Policy* **134**: p. 110827.

GREENE, W. (2002): "Econometric analysis."

Gross, R., P. Heptonstall, D. Anderson, T. Green, M. Leach, & J. Skea (2006): "Bthe costs and impacts of intermittency: An assessment of the evidence on the costs and impacts of intermittent generation on the british electricity network,[uk energy research centre, london." *P UK Energy Research Council, London* .

Hache, E. & A. Palle (2019): "Renewable energy source integration into power networks, research trends and policy implications: A bibliometric and research actors survey analysis." *Energy Policy* **124**: pp. 23–35.

Hagfors, L. I., D. Bunn, E. Kristoffersen, T. T. Staver, & S. Westgaard (2016): "Modeling the uk electricity price distributions using quantile regression." *Energy* **102**: pp. 231–243.

Hanslian, D., J. Hošek, & J. Štekl (2012): "Aktualizovaný odhad realizovatelného potenciálu větrné energie na území čr." *Praha, Ústav fyziky atmosféry AV ČR* .

Haslett, J. & A. E. Raftery (1989): "Space-time modelling with long-memory dependence: Assessing ireland's wind power resource." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **38(1)**: pp. 1–21.

Haxhimusa, A. (2018): "The effects of german wind and solar electricity on french spot price volatility: An empirical investigation." .

He, X. (1997): "Quantile curves without crossing." *The American Statistician* **51(2)**: pp. 186–192.

Hiroux, C. & M. Saguan (2010): "Large-scale wind power in european electricity markets: Time for revisiting support schemes and market designs?" *Energy Policy* **38(7)**: pp. 3135–3145.

Hirth, L. (2018): "What caused the drop in european electricity prices? a factor decomposition analysis." *The Energy Journal* **39(1)**.

Hosking, J. R. M. (1981): "Fractional differencing." *Biometrika* **68(1)**: pp. 165–176.

Hurst, H. E. (1951): "Long-term storage capacity of reservoirs." *Transactions of the American society of civil engineers* **116(1)**: pp. 770–799.

Jónsson, T., P. Pinson, & H. Madsen (2010): "On the market impact of wind energy forecasts." *Energy Economics* **32(2)**: pp. 313–320.

Karakatsani, N. V. & D. W. Bunn (2008a): "Forecasting electricity prices: The impact of fundamentals and time-varying coefficients." *International Journal of Forecasting* **24(4)**: pp. 764–785. Energy Forecasting.

Karakatsani, N. V. & D. W. Bunn (2008b): "Intra-day and regime-switching dynamics in electricity price formation." *Energy Economics* **30(4)**: pp. 1776–1797.

Ketterer, J. C. (2014): "The impact of wind power generation on the electricity price in germany." *Energy economics* **44**: pp. 270–280.

Kiviluoma, J., P. Meibom, A. Tuohy, N. Troy, M. Milligan, B. Lange, M. Gibescu, & M. O'Malley (2012): "Short-term energy balancing with increasing levels of wind energy." *IEEE Transactions on Sustainable Energy* **3(4)**: pp. 769–776.

Klamka, D. (2016): "Vliv fotovoltaických zdrojů na systémovou odchylku v České republice."

Koenker, R. & G. Bassett (1978): "Regression quantiles." *Econometrica* **46(1)**: pp. 33–50.

Koenker, R. & O. Geling (2001): "Reappraising medfly longevity." *Journal of the American Statistical Association* **96(454)**: pp. 458–468.

Koenker, R. & J. A. Machado (1999): "Goodness of fit and related inference processes for quantile regression." *Journal of the american statistical association* **94(448)**: pp. 1296–1310.

Koopman, S. J., M. Ooms, & M. A. Carnero (2007): "Periodic seasonal reg-arfima–garch models for daily electricity spot prices." *Journal of the American Statistical Association* **102(477)**: pp. 16–27.

Kouřílek, M. (2019): "The impact of german renewable electricity on czech electricity spot prices." .

Kwiatkowski, D., P. C. Phillips, P. Schmidt, & Y. Shin (1992): "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?" *Journal of econometrics* **54(1-3)**: pp. 159–178.

Lee, D. & P. Schmidt (1996): "On the power of the kpss test of stationarity against fractionally-integrated alternatives." *Journal of econometrics* **73(1)**: pp. 285–302.

Liu, K., Y. Chen, & X. Zhang (2017): "An evaluation of arfima (autoregressive fractional integral moving average) programs." *Axioms* **6(2)**: p. 16.

Luňáčková, P., J. Prša, & K. Janda (2017): "The merit order effect of czech photovoltaic plants." *Energy Policy* **106**: pp. 138–147.

Mandelbrot, B. B. & J. W. Van Ness (1968): "Fractional brownian motions, fractional noises and applications." *SIAM Review* **10(4)**: pp. 422–437.

Meza, E. (2020): "German 2020 solar power generation tops 2019 record already by october." `https://www.cleanenergywire.org/news/german-2020-solar-power-generation-tops-2019-record-already-october`.

Meza, E. (2021): "Germany's solar boom set to continue in 2021." .

Nicholson, E., J. Rogers, & K. Porter (2010): "The relationship between wind generation and balancing-energy market prices in ercot: 2007–2009. golden, colorado: National renewable energy laboratory; 2010."

Ofgem (2018): "Policy decision on the electricity system operator regulatory and incentives framework from april 2018." .

Pavić, I., M. Beus, H. Pandžić, T. Capuder, & I. Štritof (2017): "Electricity markets overview—market participation possibilities for renewable and distributed energy resources." In "2017 14th International Conference on the European Energy Market (EEM)," pp. 1–5. IEEE.

Reisen, V. A. (1994): "Estimation of the fractional difference parameter in the arima (p, d, q) model using the smoothed periodogram." *Journal of Time Series Analysis* **15(3)**: pp. 335–350.

Ruska, M. & L. Similä (2011): "Electricity markets in europe: business environment for smart grids." *VTT Technical Research Centre of Finland* .

Salavec, J. (2017): "Trh s elektřinou - specifika, účastnící trhu a rozdělení." `https://oenergetice.cz/elektrina/trh-s-elektrinou/trh-s-elektrinou`.

Samiev, S. (2013): "Garch (1, 1) with exogenous covariate for eur/sek exchange rate volatility: On the effects of global volatility shock on volatility."

Sirin, S. M. & B. N. Yilmaz (2021): "The impact of variable renewable energy technologies on electricity markets: An analysis of the turkish balancing market." *Energy Policy* **151**: p. 112093.

Sjölander, P. (2011): "A stationary unbiased finite sample arch-lm test procedure." *Applied Economics* **43(8)**: pp. 1019–1033.

Stadnytska, T., S. Braun, & J. Werner (2010): "Analyzing fractal dynamics employing r." *Nonlinear dynamics, psychology, and life sciences* **14(2)**: p. 117.

Swinand, G. P. & M. Godel (2012): "Estimating the impact of wind generation on balancing costs in the gb electricity markets." In "2012 9th International Conference on the European Energy Market," pp. 1–8. IEEE.

Tashpulatov, S. N. (2013): "Estimating the volatility of electricity prices: The case of the england and wales wholesale electricity market." *Energy Policy* **60**: pp. 81–90.

Taylor, S. J. (1986): "Forecasting the volatility of currency exchange rates." *International Journal of Forecasting* **3(1)**: pp. 159–170.

Torre, K., D. Delignières, & L. Lemoine (2007): "Detection of long-range dependence and estimation of fractal exponents through arfima modelling." *British Journal of Mathematical and Statistical Psychology* **60(1)**: pp. 85–106.

Vandezande, L., L. Meeus, R. Belmans, M. Saguan, & J.-M. Glachant (2010): "Well-functioning balancing markets: A prerequisite for wind power integration." *Energy policy* **38(7)**: pp. 3146–3154.

van der Veen, R. A., A. Abbasy, & R. A. Hakvoort (2012): "Agent-based analysis of the impact of the imbalance pricing mechanism on market behavior in electricity balancing markets." *Energy Economics* **34(4)**: pp. 874–881.

Volgushev, S., M. Birke, H. Dette, & N. Neumeyer (2013): "Significance testing in quantile regression." *Electronic Journal of Statistics* **7(none)**: pp. 105 – 145.

Vágner, V. (2017): "Větrné elektrárny včera, dnes a zítra (díl 2.)."

Wagenmakers, E.-J., S. Farrell, & R. Ratcliff (2004): "Estimation and interpretation of 1/f α noise in human cognition." *Psychonomic bulletin & review* **11(4)**: pp. 579–615.

Wehrmann, B. (2020): "Solar power in germany – output, business & perspectives." `https://www.cleanenergywire.org/factsheets/solar-power-germany-output-business-perspectives`.

www.eia.gov (2020): "Electricity explained." `https://www.eia.gov/energyexplained/electricity/delivery-to-consumers.php`.

Yildiz, B., J. Bilbao, A. Sproul, & J. Dore (2017): "A method for classifying households to help forecasting their photovoltaic electricity self-consumption patterns." In "Asia Pacific Solar Research Conference," .

Yildiz, B., J. I. Bilbao, J. Dore, & A. B. Sproul (2018): "Short-term forecasting of individual household electricity loads with investigating impact of data resolution and forecast horizon." *Renewable Energy and Environmental Sustainability* **3**: p. 3.

Šrom, J. (2020): "Vnitrodenní trh s elektřinou a nestandardní stavy v elektrizační soustavě."

Appendix A

Appendix

Table A.1: *AIC* & *BIC* of *ARFIMA*$(p,\hat{d},q)$

ARFIMA(p,d,q)			
Fitted Model	$\hat{d}$	AIC	BIC
(1, d, 0)	0.2233	10.600	10.602
(1, d, 1)	0.2547	10.596	10.598
(1, d, 2)	0.2204	10.596	10.598
(2, d, 1)	0.2455	10.596	10.598
(3, d, 1)	0.2013	10.595	10.597
(4, d, 1)	0.2029	10.595	10.597

Figure A.1: Residuals of ARFIMA(1,$\hat{d}$,1)