TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technology
Department of Software Sciences

Andri Busch 185995IAIB

IMPROVING SHORT-TERM POWER MARKETS TRADING STRATEGY FOR WIND POWER PRODUCER

Bachelor Thesis

Supervisor
Marko Kääramees
PhD
Co-Supervisor
Sven Nõmm
PhD

Tallinn 2022

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond
Tarkvarateaduse instituut

Andri Busch 185995IAIB

LÜHIAJALISTE ELEKTRITURGUDE KAUPLEMISSTRATEEGIA ARENDUS TUULEENERGIA TOOTJALE

Bakalaureusetöö

Juhendaja
Marko Kääramees
PhD
Kaasjuhendaja
Sven Nõmm
PhD

Tallinn 2022

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author:	Andri Busch	....
		(signature)
Date:	May 30, 2022	

Annotatsioon

Taastuvenergia toodangu kauplemisega on tootjal kõrgendatud finantsrisk, seoses keerukusega toodangut täpselt prognoosida. Kui tegelik toodang ei vasta prognoosile, arveldatakse ülejääk või puudujääk ebabilansi turul, mille hinnad on seotud kogu elektrituru süsteemi bilansi suunaga. Need hinnad on väga kõikuvad, mis tähendab, et tootjal on risk müüa enda toodangut turuhinnast palju madalamalt.

Käesoleva lõputöö põhieesmärgiks on arendada statistika -ja masinõppe meetoditel põhinev kauplemisstrateegia ning seda realiseeriv tuuleenergia füüsilise kauplemise otsustusmudel. Töö käigus analüüsiti süvitsi äriprobleemi, uuriti erinevaid alternatiivsed lahendused lähtuvalt kättesaadavatest andmetest ja ärilistest nõudmistest ning disainiti tarkvaralahendus otsustussüsteemi näol. Valitud lahenduse oluliseimad tehnilised komponendid on optimeerimisalgoritm ning aegridade ennustusmudel.

Töö rakenduslikuks väljundiks on otsustusmudel, mille abil on võimalik elektrikauplejal operatiivselt informeeritumaid otsuseid teha.

Disainitud lahendused on treenitud kasutades andmeid perioodist 2019-2020 ning testitud reaalsetel turuandmetel ajavahemikus 01/2021 kuni 02/2022, kasutades mudeli väljundeid reaalsete ajalooliste andmetega ning arvutades mudeli mõju majanduslikele näitajatele. Tulemuseks saadi, et mudeli tulemusena on võimalik tõsta tootja kuu keskmist kasumit 12 - 18 % sõltuvalt valitud riskitundlikkuse profiilist. Seejuures juhtimata riskidest tulenevat kulu (bilansikulu) suudeti vähendada kuni 480 %. Kõige täpsemad süsteemi ebabilansi ennustustulemused saavutati rekurrentse närvivõrgu mudeli abil, mille puhul saavutati tulemuseks keskmine absoluutviga 34.93, juuritud keskmine ruutviga 46.23 ning ebabilansi suuna täpsuseks 82.29%.

Töö on kirjutatud inglise keeles ning sisaldab 50 lehekülge teksti, 4 peatükki, 20 graafikut ja 7 tabelit.

Abstract

The sales of electricity from renewable resources involve increased financial risks due to the complexity of forecasting actual power generation. If the actual generation deviates from forecasts, the deviations are financially settled as imbalance. The price of the imbalance is determined from the overall imbalance of the electricity system. This price is very volatile, meaning that the electricity producer has the risk of selling the generation at a much lower cost than the price of the physical power market.

The main goal of the present bachelor's thesis is to develop a short term power market trading strategy, based on statistical and machine learning methods and implement it to an automated decision model. The work explores the previously mentioned business problem and the validity of different alternative solutions, depending on the available data and specific business requirements. The most important components of the technical implementation are the optimization algorithm and a time series forecasting model.

The practical contribution of this work is the automated decision system that helps electricity traders make better informed decisions and reduce computational overhead in fast-paced, short-term trading.

The designed solutions were trained on data from 2019 to 2020 and backtested on real market data between period 01/2021 to 02/2022, analyzing the effect of the model on real economic variables. As a result of this model, the monthly average profit was shown to increase by 12-18 % depending on the chosen risk aversiveness profile. The cost of not managing the risks was reduced by up to 480 %. The most accurate system imbalance forecast results were obtained, using recurrent neural network model, that reached mean absolute error of 34.93, and the accuracy of directional imbalance 82.29%.

The thesis is written in English and contains 50 pages of text, 4 chapters, 20 figures and 7 tables.

List of abbreviations and terms

BRP	Balance Responsible Party
Nord Pool	Main physical electricity exchange
TSO	Transmission System Operator
AIC	Akaike Information Criterion
Day-Ahead Market	Main physical electricity market, bidding done every day at 12:00 CET
Intraday Market	Adjustment market to Day-Ahead
CET	Central Europe Time (+02 UTC)
mFRR	Manual Frequency Restuaration Reserve
ARIMA	Autoregressive Integrated Moving Average
ISP	Imbalance Settlement Period
PDF	Probability Density Function
CDF	Cumulative Density Function
KDE	Kernel Density Estimation
MLE	Maximum Likelihood Estimation
ACF	Autocorrelation Function
PACF	Partial Autocorrelation Function
LSTM	Long Short-Term memory
EVPI	Expected Value of Perfect Information
VSS	Value of Stochastic Solution
SGD	Stochastic Gradient Descent
Adam	Adaptive Moment Estimation
GPU	Graphics Processing Unit
CVaR	Conditional Value at Risk
SES	Simple Exponential Smoothing

Table of Contents

List of Figures

List of Tables

1. Introduction

The energy sector plays an enormous role in the transition towards a green economy, as combustible fuels are the main reason for excess greenhouse gas emissions. Electricity production alone contributed 25% of global greenhouse gas emissions. To combat this, countries and companies are looking for ways to replace fossil fuels with renewable energy generation, and the ones that are developing at the fastest pace are wind and solar power [1]. The increasing volume of renewable electricity generation puts increased pressure on the power system, which has to maintain the balance between supply and demand at all times. To achieve the most effective control of the system, a range of different power markets are put in place. Due to the very volatile nature of wind and solar generation, a power producer must take measures against the risks of the present market.

Since electricity production and consumption must be balanced at all times, it imposes challenges to electricity producers and consumer aggregators, who must also be able to forecast their production or consumption reliably. Unexpected deviations from schedule pose considerable financial risk to producer and are called imbalance costs.

The current work focuses on minimizing the costs associated with participating in the electricity market of a power producer with a large renewable generation resource. The imbalance cost is crucial to every participant in the power market, so it serves as a starting point for optimizing power producer performance. The general goal of this work is to research different options how to make electricity trading operations more efficient, by selecting the portfolio position in between the markets and time horizons. The specific contribution of this work is the development of a decision tool accompanied by a forecasting model, as the environment is highly uncertain. The task of such a decision tool is to find an optimal position of the market participant power portfolio, using the limited amount of information from forecasts on the future state of the market and the production output.

Different methods for attaining the forecast are considered and compared. Forecasting methods are also decoupled from decision models to assess the effects separately.

The research questions of the work are the following.

- How much value can we add by making smarter decisions with the help of decision modeling?
- Which of the forecasting methods (statistical or machine learning) provides the best results in terms of accuracy and in terms of reducing balancing costs when used with the same decision model?
- Is the complexity of more advanced forecasting methods justified in the case of this specific problem?

Section 2 provides the necessary background information on electricity markets, to understand the problem and the solution. A brief overview of the relevant forecasting methods used is also included. In the end, the articles and research on which this work is based is highlighted. Section 3 first formulates the problem and trading strategies analytically. Two first trading strategies are provided as a baseline, to compare with the more advanced strategy, that requires computation and modelling. After the trading strategies, the development of the required imbalance forecasting model is discussed. The section ends with describing the implementation of the decision model to the computer system. Section 4 provides the results of the use of strategies developed in real market settings. The case study is done using the data for the period January 2021 to February 2022. The components of the final results are decomposed and discussed separately. In the final results, the imbalance forecast model is provided.

2. Background

In this section, a brief background is provided on electricity markets and forecasting methods, to understand the essence of the remaining work.

2.1 Electricity markets

Since electricity as a product is uniform, the difference between markets mainly occurs in the traded time horizon. Trading horizons range from long-term contracts (up to 30 years) to real-time buying and selling. Understanding electricity markets is crucial to understand the problem and the respective solution addressed in this work.

Overview of the three different physical electricity markets with different time-horizons and related price levels, where traders can allocate the resources would be following:

- Day-Ahead market – the price of Day-Ahead market acts as a baseline for all other physical market prices
- Intraday (adjustments) market – up to 2 hours before the delivery hour adjustments. Prices are generally less favorable than the Day-Ahead market
- Balancing market – Real-time, most volatile market. Determines the imbalance settlement prices. Prices can be extremely unfavorable or very profitable

This list is not comprehensive but serves the purpose of helping to clarify the background for this work. The development of new system services and markets is a rapidly developing area where the main goal is to control the electrical system in the most economically and environmentally efficient way possible. With the new developments, allocation choices and trading complexity increase as well.

Financial instruments allow for the trade of electricity on the longest horizon. There are a wide variety of instruments, such as options and future contracts, that allow energy investors to hedge their market risks, not participating directly in the physical power market [2]. Financial markets play an important role, but are not the main focus of this work. They are brought out, to help understanding the risk management problem at hand. The two types of markets where a power producer faces an allocation problem are physical power markets and balancing markets, which are reviewed more thoroughly below.

2.1.1 Physical power markets

Physical power market usually refers to a aggregating agent, that pools together suppliers and consumers and serves as a connecting party improving efficiency through market mechanisms. In Europe, NordPool is the biggest physical electricity market, consisting of all Nordic countries (Norway, Sweden, Denmark, Finland), Baltic countries, Poland and the UK.[3] The second largest power exchange is EPEX Spot, that includes West-European countries.[4] The key to understand electricity exchanges is that although these market exchanges connect many countries over long distance, electricity still remains mainly a local product due to physical nature and transmission high losses. Liberalization of the electricity market solves the economical problem of producing electricity at the lowest possible cost when enough transmission capacity is available between areas[5].

The NordPool offers 2 markets: Day-Ahead and Intraday[6]. The main volume of electricity is traded on the Day-Ahead market, and the Intraday market serves the purpose of balancing unexpected changes closer to the delivery hour. Day-Ahead bidding is done every day at 12:00 CET for the next 24 hours. However, the intraday market in the Baltic market area closes 1 hour before the physical delivery hour, allowing accommodating unexpected changes in generation (for example, wind forecast change) or consumption. The time schedule, how different markets are connected is depicted in Figure 1.

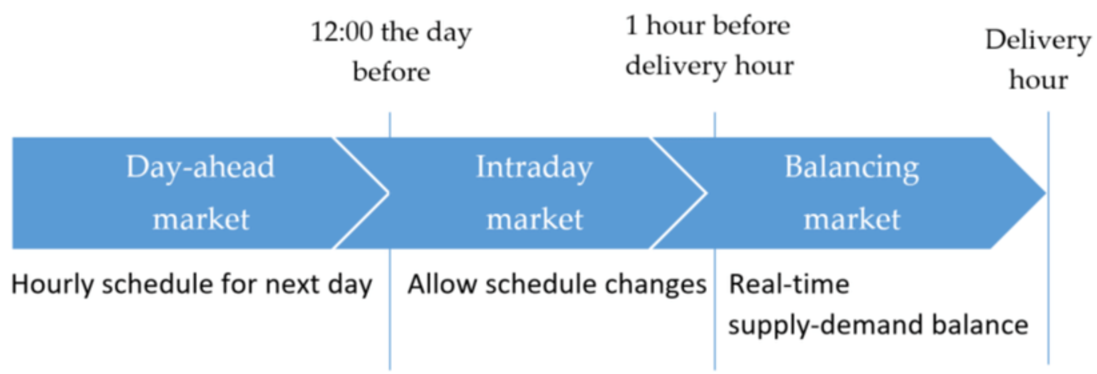

Figure 1. *NordPool market structure. Source: [7]*

2.1.2 Balancing markets

Balancing markets general principle is to keep the system load and generation balanced at all times. That is crucial to keep the alternating current frequency at 50Hz at all times as deviations from the balance may cause sensitive appliances to malfunction. Designing a functioning balancing markets is an hard task concerning both the economical and engineering challenges[8]. These markets are operated by Transmission System Operators (TSO). In our region, three Baltic TSO's collaborate to operate the Baltic Balancing market[9].

Currently, the power system imbalance is determined between 3 Baltic countries, but the system frequency is still managed by the Russian counterparty. If imbalance exists in Baltic balancing area, Russian electricity system keeps the frequency in place and imbalance is settled with financial payments from Baltics. The joining of the continental European synchronization area is a very big project and is expected to be completed by the end of 2025 [10]. Therefore, system imbalance presents a direct outflow of capital from our economy and every measure, that helps to keep imbalances low, will be beneficial. In order to keep frequency independently with the lowest cost possible, many different new markets such as aFRR (Automatic Frequency Restauration) and FCR (Frequency Containment Reserve) need to be implemented and it poses new opportunities for electricity traders.

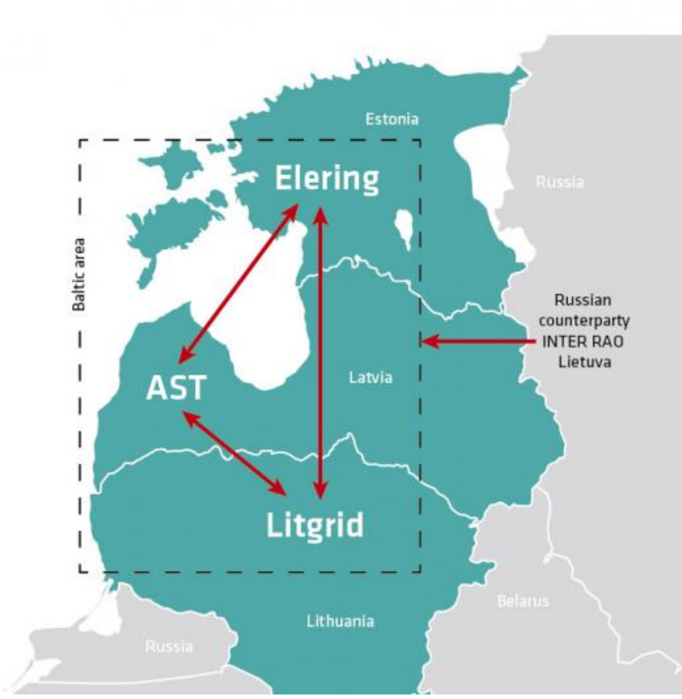

Figure 2. *Baltic synchronous area. Source: [11]*

2.1.3 Trading strategies

The core idea of power trading strategies is to earn the producer the best price for the power generated under market conditions. Bulk of the sold power is sold through Day-Ahead auction and in perfect scenario, power producer would sell all the generation there. But in the real world, the renewable energy producer faces the risk of generating less or more electricity at a certain hour. For example, if they generated less than they sold, they have to buy the difference from TSO, with the current hour imbalance settlement price. That price can be significantly larger than the Day-Ahead or Intraday price, so the producer faces financial risks. Since power generation forecasts cannot be 100% accurate, the producer cannot ignore the risk of being in wrong direction of the imbalance market. So they have mitigate these risks, by placing themselves on a so called safer side of the imbalance market, by having an estimation of what is the system imbalance state and thus the imbalance price at certain hour.

Therefore, the main goal of short-term trading strategies addressed in this work, is to manage financial risks, that come from having uncertain generation quantity.

2.2 Forecasting

To plan how much electricity to sell, the power market trader is highly dependent on renewable energy forecasts. Electricity market forecasting is a very well-developed area of research with a multitude of different forecasting methods. One of the most comprehensive article about different forecasting methods in electricity markets field, is from Weron [12].

The methods considered in this work are statistical and machine learning methods (in 3 referred to as *Computational Intelligence*). The primary difference between these two methods is the ability to handle nonlinearity. Statistical methods are not very well suited to handle nonlinear processes, but they can be more easy to interpret. Machine learning methods on the other hand can handle complex patterns very well and can be very suitable for short-term power forecasting. What unites these methods and separates them from fundamental and agent-based methods, is the underlying principle of relying on mathematical combination of past data, without taking into account the economical and technological cause-effect relationships.

A large body of statistical machine learning models is based on the assumption of non-correlating features. When this assumption is made for the data in this work, important information about the sequential nature of the data is lost. Therefore, they may not be very well suited for short-term forecasting, as the stochastic processes are usually more affected by the dynamics of the process, but can work very well in longer term analysis. These models also have very convenient property of interpretability.

For short-term forecasting, time series forecasting methods are therefore used. One of the most popular approaches in statistical time series forecasting is using Autoregressive Integrated Moving Average (ARIMA) models [13]. The key idea behind these models is that there exists an autocorrelation structure in the process that can be captured into a mathematical model.

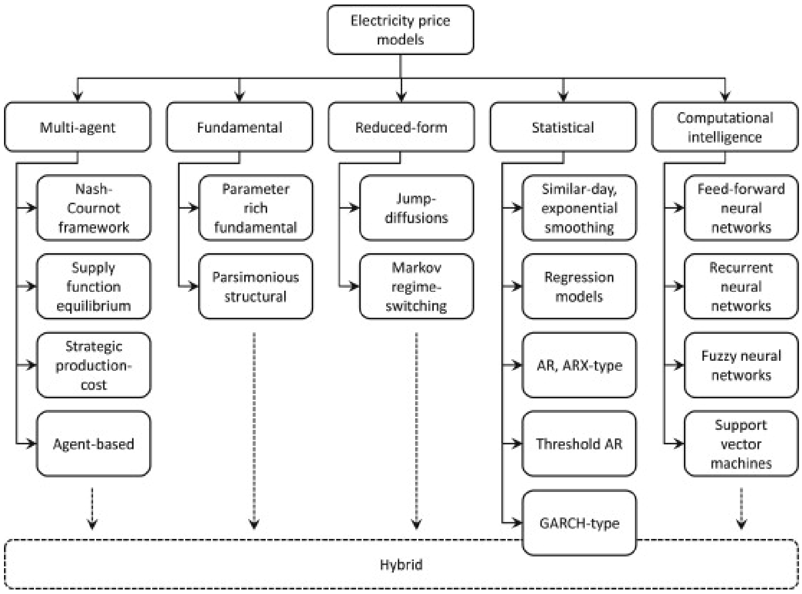

Figure 3. *Electricity market forecasting methods taxonomy. Source: [12]*

A very popular approach in machine learning for time series forecasting is the use of recurrent neural networks to capture the sequential nature into a model. The key feature that differentiates them from the simplest feedforward neural networks is the ability to recall the internal state. One of the main drawbacks of this approach is the exploding or vanishing gradients that appear in the backpropagating training process with longer sequences. The method that solves this problem at least partially is to add gates in each hidden state cell to better control the movement of information in the hidden state 4.

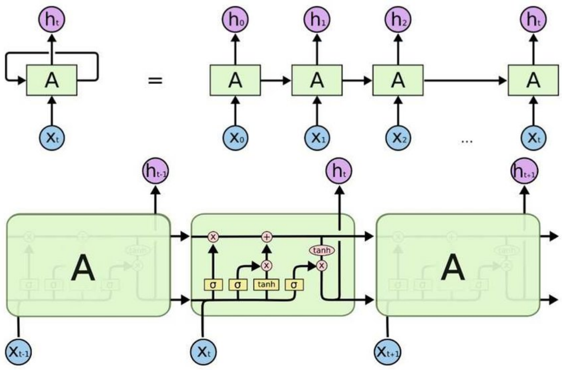

Figure 4. *RNN neural network architecture with LSTM cell. Source: [14]*

Within this work, the quantity of system imbalance is the predicted variable. The results of the forecast are used as an input to other part of the model developed later in chapter 3. Wind power forecasts are used as input and enhanced by turning point forecasts into probabilistic generation forecasts.

2.3 Related works

The theoretical background about many following ideas about trading strategies are mathematically defined in [15]. Many ideas in the practical development of trading strategies follow the work of Pinson, Chevallier, and Kariniotakis in [16], but differ in the market organization framework and conditions, as the recent years have been very turbulent times on electricity markets.

Stochastic programming is widespread in power markets area, since the decisions have to be in situations of imperfect information. A classical work about stochastic programming is by Birge and Louveaux [17], and the models and concepts are also used throughout this work. In terms of specific contributions to electricity trading, stochastic programming is used in many sources. Boomsma, Juul, and Fleten implemented the scenario based stochastic programming model between Day-Ahead and balancing market. The work is focused more towards developing an optimization model and scenario generation is done solely using open data and ARIMA timeseries method [18]. The method of scenario-based uncertainty modeling is also used in [19], [20] and is in contrast to the probabilistic forecast approach, used in [16]. The current work follows the probabilistic forecasting approach, but scenario-based methods can be seen as a further development.

The different offering strategies are analyzed in [21] and the insights of this work are also used to structure the results of this work. The insights about wind production error modeling are derived from [22].

Since risk is a very important aspect to manage in addition to maximizing revenues, many works have added risk component to the objective function and tackled a more difficult problem, finding optimal strategies with the lowest risk [20], [23]. An even further state-of-the-art methods have been proposed to modify the risk attitudes automatically, and this is seen as a further development to the current model [24].

3. Method of trading strategy

In order to further increase the efficiency of trading operations, and more specifically, reduce the balancing costs by 20%, the whole trading process was analyzed for systematic inefficiencies. Different approaches to solving this problem were tested and the most promising one is further described in this section. The proposed method is a smart, multi-step way to place sold electricity positions in the market, depending on our estimate of the future.

At first, an overview is given, of the other methods used by the author, to approach this problem, the reasons, why they were inferior to the selected approach and insights derived from research. Next, an example is made, to illustrate the nature of decisions that a trading strategy must make. After that, the short-term trading strategy is described in general, and basic metrics are mathematically defined. Then, the chapter proceeds with developing the baseline strategies. The bulk of the chapter is dedicated to the development of a more advanced trading strategy against the imbalance market, starting with the formulation of a decision model and developing a wind power generation forecast enhancement method. The system imbalance forecasting methods are discussed later, and the chapter ends with practical implementation to information systems.

3.1 Alternative methods research

The idea for the proposed model developed further in this section is based on insights derived from research done by author at the start of this project. The more general goal of this work was to find a smart way to distribute offers between different markets and time horizons.

The first solution was to choose between the manual frequency restoration market (mFRR) and Day-Ahead market. In this method, the choice between markets was priced as a real option and a positive price indicated a potentially profitable decision. This method turned out to be inferior to the one discussed in this thesis, because of the very noisy nature of balancing market variables. The forecast had to be done 12-36 hours into future, but the research showed, that the best feature for forecasting was the previous values of the forecasted variable. Therefore, by increasing the forecast horizon, the accuracy dropped to the level, where the model did not add significant value. This insight allowed shifting the focus on methods that depend on short-term forecasting and a focus to be more on developing smart methods for decision modeling.

3.2 Example of short-term trading strategies in action

To give the reader an understanding, what decisions are faced and how the proposed strategy helps to improve results, an illustrative example is provided in Figure 5. The figure presents the process, of trading one hour of generation between different markets. The numbers depicted are in the higher end of the actual data, to better illustrate the differences.

The first step shown in the Figure, is the Day-Ahead auction. The amount sold serves as a basis position for all subsequent decisions. The contribution of a short-term strategy starts with the 2 hours before the power generation hour. Then, the wind power trader is faced with a choice, how much electricity to sell or buy on intraday market. Trusting the point forecast corresponds to strategy 2 described below in the dedicated section. This strategy tries to eliminate risk by removing the error between actual generation and nominated position. A more complex algorithm has to take into account the size of the risk and move the nominated position into a safer side of imbalance. That is the essence of strategy 3, which is the main contribution of this work.

3.3 Short-term trading strategy

The electrical system must be in balance at all times to maintain frequency (described in Chapter 2), and this is obtained by centrally managing the system by TSO. All Balance Responsible Parties (BRP) are required to forward their forecasts for both generation and consumption. These forecasts must be confirmed by the TSO and, if confirmed, become nominations. After forwarding their nominations, BRP's have obligation to fulfill production/consumption plans. Everything that exceeds or does not meet the plan is considered an imbalance[25].

Therefore, all BRPs are interested in managing their power portfolio, to reduce the risk of unfavorable imbalance price. The imbalance price is determined after the delivery hour has passed and is based on the highest balancing market regulation (mFRR) bid price. The basic logic of the short-term trading strategy is to have one's portfolio imbalance on the other side of the system imbalance. The logic is shown in figure 6. It is immediately clear, that this logic poses problems of uncertainty. A forecast about the system imbalance in a given hour in the future is needed, as well as a forecast about our own electricity generation.

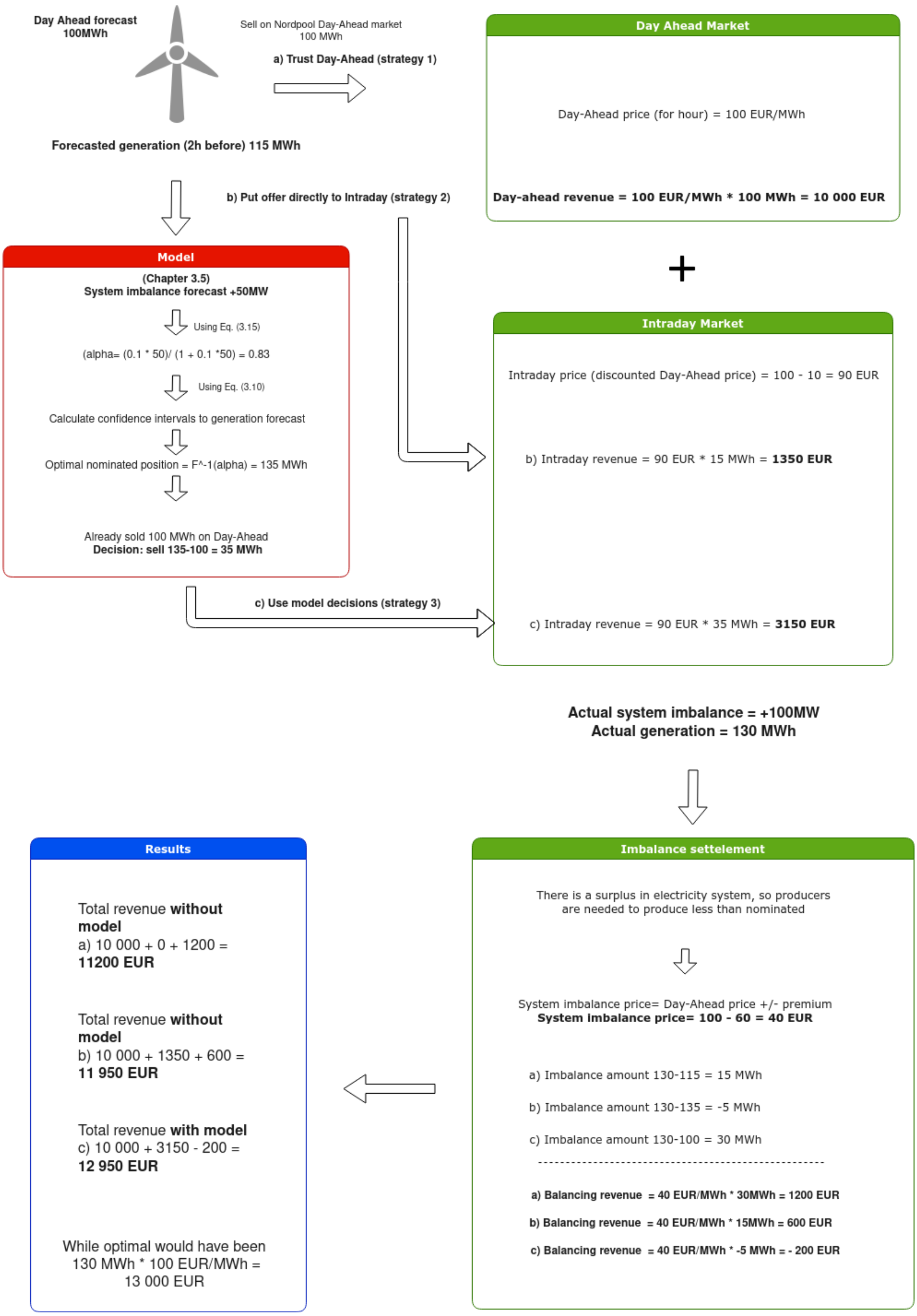

Figure 5. *One hour example of trading decisions in different phases and results with different strategies*

The state of our own portfolio is not a very hard problem in case of dispatchable generation such as thermal, nuclear or hydro resources, but becomes a very hard problem in case of wind and solar power resources. The logic in Figure 6 indicates the imbalance position to take in the deterministic case. But in reality, the quantities involved and the uncertainty of the forecasts have to be taken into consideration.

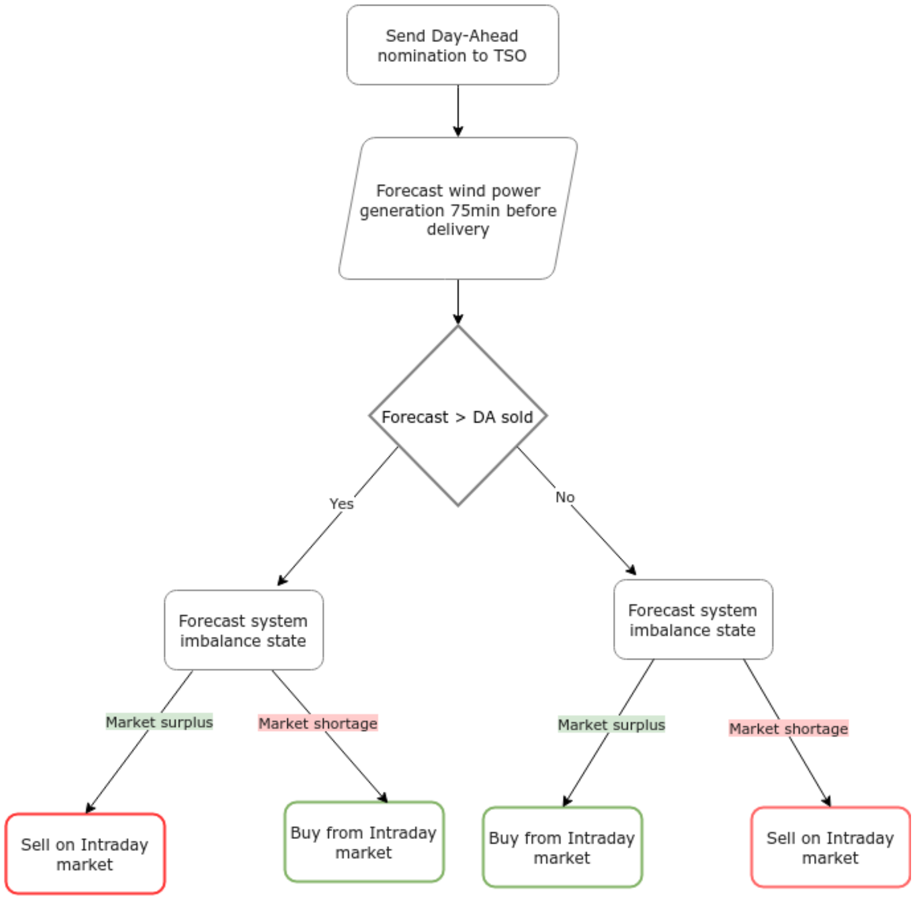

Figure 6. *Short-term trading strategy logic on a conceptual level*

To measure the benefit of proposed strategies, the balancing cost has to be formally defined, as well as the baseline against which we measure the rate of advancement. The objective of the wind power producer is to sell their generation in the Day-Ahead market, and every deviation from that result is considered balancing cost (or profit in the case of positive deviation). The revenue earned at period t is denoted by π_t, the Nord Pool Day-Ahead market price by λ^{DA} and the actual generated quantity in time period t by y_t. In the following, the aggregation period t will be considered 1h, because it is the official aggregation step for the Baltic markets at the time of this work. During 2023-2024, the 15 minute imbalance settlement period (ISP) and 15-minute Day-Ahead and Intraday products will be implemented in the Baltic markets [26]. Most of the logic of this solution is not dependent on the exact aggregation period. Equation 3.1 expresses the optimal revenue for the wind producer and every deviation from it is considered balancing cost. In practice, the actual revenue deviates from the optimal due to reasons outlined above.

$$\pi_t^* = y_t * \lambda^{DA} \tag{3.1}$$

where

π_t^* is optimal profit for period t
λ^{DA} is Day-Ahead market price for price area EE at period t
y_t is Actual generated wind energy in MWh for period t

Balancing cost is defined in Eq. 3.2, as the difference between optimal sales and the revenue of the chosen trading strategy.

$$z_t^i = \pi_t^* - \pi_t^i \tag{3.2}$$

where

z_t^i is the balancing cost of i'th strategy at period t
π_t^i is the revenue of i'th strategy at period t

The imbalance strategy must take into account the imbalance that arises from an error in the generation forecast. That error is purchased or sold to/from the TSO, depending on the sign of imbalance. Equation 3.3 defines the general functional relationships between variables that must be considered in every specific strategy. y_t^{DA} denotes the generation, that is sold on Day-Ahead market, y_t^{ID} is the generation sold on Intraday market.

$$\pi_t^i = (y_t^{DA} * \lambda_t^{DA}) + (y_t^{ID} * \lambda_t^{ID}) + [y_t - (y_t^{DA} + y_t^{ID})] * \lambda_t^{SI} \tag{3.3}$$

where

y_t^{DA} Generation sold in Day-Ahead market at period t

y_t^{DA} Generation sold in Intraday adjustment market at period t

λ_t^{ID} Weighted average price at Intraday adjustment market at period t

y_t Actual generation in MWh at period t

λ_t^{SI} Imbalance settlement price paid to/received from TSO for every MWh at period t

For the wind producer, the nominated quantity to which the generation must match at all times is described by 3.4. It is also possible, that in the future, balancing regulation services will be added to this equation and also count in nominated quantity.

$$y_t = \hat{y}_t^{DA} + y_t^{ID} \tag{3.4}$$

where

$\hat{y}_t^{DA}$ Forecasted generation for Day-Ahead market bidding at period t

y_t^{DA} Adjustments to initial bids sold/bought on intraday market at period t

The imbalance settlement price at hour t is denoted by λ_t^{SI} and depends on the total system balance. The price is based on Day-Ahead market price. If there is not sufficient generation in the system to balance the load, the imbalance price will be high, to provide producers with an incentive to ramp up the production.

In a simplified case, the relations in 3.5 hold for imbalance price. It should be kept in mind that these are true most of the time and exceptions for the rule might occur in market environment. These rules are specific enough in our application to help develop a better imbalance strategy. Official calculation methods for the price of the imbalance are described in [9].

$$
\begin{cases}
\lambda_t^{SI} < \lambda_t^{DA}, x_t > 0 \\
\lambda_t^{SI} > \lambda_t^{DA}, x_t < 0 \\
\lambda_t^{SI} = \lambda_t^{DA}, x_t = 0
\end{cases}
\tag{3.5}
$$

where

x_t System imbalance at period t

3.4 Baseline trading strategies

In this section, a stochastic optimization strategy is devised to improve the wind producer trading results. For comparison, we also include less sophisticated strategies that use given forecasts and automatic intraday trading.

3.4.1 Strategy 1. Trust Day-Ahead forecast

This strategy takes wind generation point forecasts as input and forwards them to the market, without any further actions taken. The strategy is simple, but the wind producer is very exposed to the imbalance price volatility and the rapidly occurring changes in weather situation. The revenue function of this strategy is given by Equation 3.6. The $\hat{y}_t^{DA}$ denotes the given Day-Ahead forecast about wind generation, that is directly forwarded to market.

$$
\pi_t^1 = (\hat{y}_t^{DA} * \lambda_t^{DA}) + (y_t - \hat{y}_t^{DA}) * \lambda_t^{SI} \tag{3.6}
$$

3.4.2 Strategy 2. Intraday market adjustments

Since a closer predictions about the resulting generation can be made closer to the actual generation hour, it is worthwhile to update the nominated production schedule through the Intraday adjustment market. Strategy 2 uses forecasts closer to the delivery time to update the position in a favorable direction. The main difference with the 1'st strategy is the shorter forecast horizon. The strategy still treats forecast outputs as deterministic and does not use any forecasts about the balancing market state.

$$
\pi_t^2 = (\hat{y}_t^{DA} * \lambda_t^{DA}) + (y_t^{ID} * \lambda_t^{ID}) + (y_t - (\hat{y}_t^{DA} + y_t^{ID})) * \lambda_t^{SI} \tag{3.7}
$$

3.5 Strategy 3- Imbalance risk management strategy

Since at the time of making the decision about the quantity to nominate, uncertainty is faced in two variables. First, the actual generation is unknown and second, the imbalance price is not known. The core benefit of stochastic programming, compared to deterministic one, is to find optimal solutions that do not maximize the objective function in a certain set of parameters corresponding to a specific scenario, but find a solution that is optimal, given the range of possible scenarios.

This strategy is the main contribution of this work and is composed of many modules, that work together. The overview of the strategy is brought out in figure 7.

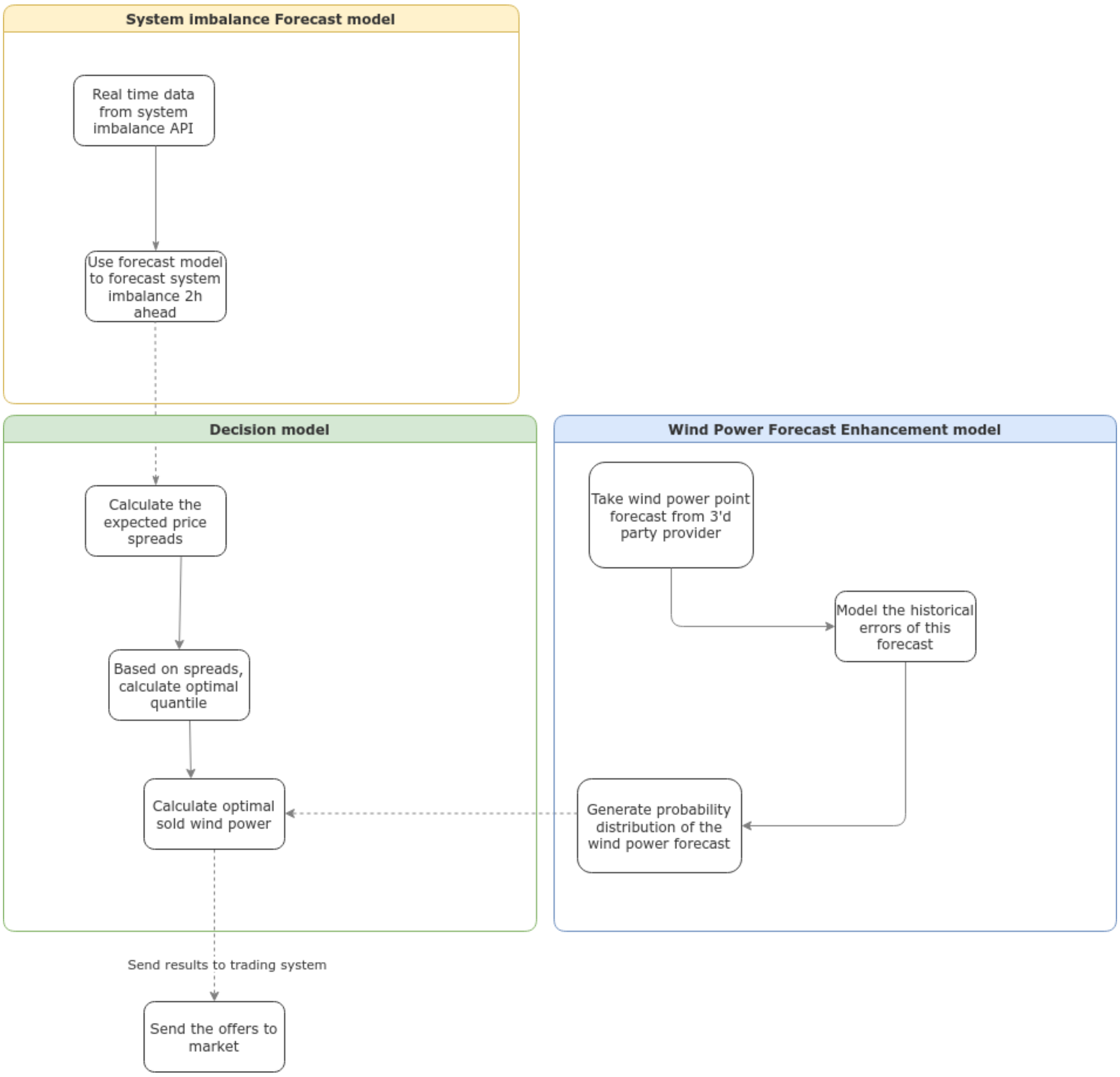

Figure 7. *Overview of different parts in strategy 3*

3.5.1 Decision model

The problem resembles a common stochastic programming problem structure, the News Vendor problem, discussed in [17]. In the classical case, the news vendor has to decide up-front the number of newspapers to buy from the dealer, to sell them with a profit margin. However, the demand is unknown and later, the news vendor can sell back the newspapers with a loss. In our case, the first decision is the amount to sell on the Day-Ahead market and then second decision would be 75-minutes before delivery hour, to adjust the position (what is the final nominated position). In this formulation, it would be a multistage stochastic program, but at first we stick to the two-stage program and assume that we can adjust the nominated value to an optimal level only via selling/buying electricity on the intraday market and get to the optimal imbalance market position. This is formally defined as Assumption 1. The profit margins of selling and the loss of buying back later correspond to the Intraday and Imbalance price spread in our application. In the original problem, the profit margins are known, but in the case currently at hand, they are unknown and depend dynamically on the state of system imbalance.

Assumption 1-. *Enough electricity can be sold / bought on the intraday market, to reach* π_t^*

To model the problem, a relation between Day-Ahead and Intraday prices needs to be formulated. That is, intraday market prices are lower than Day-Ahead when we want to sell quickly and higher than Day-Ahead, if buying quickly. The discount varies in time, but to make the model more readable, we define it as a constant of 10 EUR, which is based on expert opinion. This is formally defined in the assumption 2. This relation is caused by the constrained time frame, to sell or buy from intraday market. The small time frame for trading (approximately 10 minutes) forces to sell with prices lower than Day-Ahead price and buy with prices higher than Day-Ahead price.

Assumption 2-.

$$\begin{cases} \lambda_t^{ID} = \lambda_t^{DA} + 10, \textit{if buying from Intraday} \\ \lambda_t^{ID} = \lambda_t^{DA} - 10, \textit{if selling to Intraday} \end{cases} \tag{3.8}$$

Optimization of a stochastic program with a recourse function is 3.9. Total revenue is maximized, which is equivalent to reducing balancing costs.

$$\begin{aligned} max\ z = (\lambda^{DA} + \lambda^{ID}) * x\ + \mathscr{D}(x) \\ x >= 0 \end{aligned} \tag{3.9}$$

where

z Revenue for the trading strategy

x Total nominated sold power quantity [MWh] (Day-Ahead + Intraday)

$\mathscr{D}$ is a recourse function that maximizes the expected profit, given the probability distribution of random variable ξ. The random variable is the actual wind generation at hour t. It should be noted, that in order to have further control over the program, the formulation supports adding extra constraints on variables.

$$\mathscr{D}(x) = \mathbb{E}_{\xi} Q(x, \xi) \tag{3.10}$$

where

$Q(x)$ Revenue for the trading strategy, when wind power generation is ξ

$\mathbb{E}_{\xi}$ Expected revenue, when wind generation is ξ

This formulation leads us to a very elegant solution, that can derived by replacing the mathematical expected value with corresponding integral and integrating by parts. The complete derivation is presented in Appendix 5. The objective function is maximized by taking the optimal quantile of the inverse cumulative distribution function of the generation forecast $F^{-1}(\alpha_t^*)$.

$$x_t^* = F^{-1}(\alpha_t^*) \tag{3.11}$$

where

x_t^* Optimal nominated position t

$F^{-1}(X)$ Inverse cumulative distribution of wind generation

α_t^* Optimal quantile for nominated wind production at period t

For every time period aggregate, an optimal quantile α_t^* is found, based on the estimation of price spreads in imbalance market. The optimal quantile is defined in 3.13. S_t^+ is the expected spread between the Day-ahead price for hour t and the DOWN imbalance regulation price, and S_t^- is the spread of the UP imbalance regulation price. Historically, the spreads are equal, making the optimal quantile 0.5, but in a close view, the spread is very volatile and having an estimation about the situation in balancing market, we can estimate the level of spread every hour.

$$\begin{aligned} S_t^+ &= \lambda_t^{DA} - \lambda^{SI}, \text{ if } q_t > 0 \\ S_t^- &= \lambda^{SI} - \lambda_t^{DA}, \text{ if } q_t < 0 \end{aligned} \tag{3.12}$$

where

S_t^+ Imbalance DOWN regulation spread with Day-Ahead price. Only if surplus at period t

S_t^- Imbalance UP regulation spread with Day-Ahead price. Only if deficit in period t

$$\alpha_t^* = \frac{S_t^+}{S_t^+ + S_t^-} \tag{3.13}$$

Price spread asymmetry can be calculated by forecasting the system imbalance a couple of hours in advance. For example, if the system is DOWN regulated (system imbalance > 0), the imbalance price will be more likely below the Day-Ahead price. Therefore, the imbalance spread S^+ will occur more likely and will thus be higher. The optimal quantile α^* will be higher than 0.5. The logic of how the imbalance spread asymmetry function is derived from the forecast system imbalance is shown in 8. The used forecasting methods will be discussed in the section dedicated to Short-Term Imbalance Forecasting.

The value of using this optimization model will largely depend on the accuracy of our forecasts about the generation and the state of the system imbalance that occurs. As mentioned before, the forecasts about generation are given as input, but there is no system imbalance forecast available beforehand. Thus, the forecast model of the system imbalance has to be made.

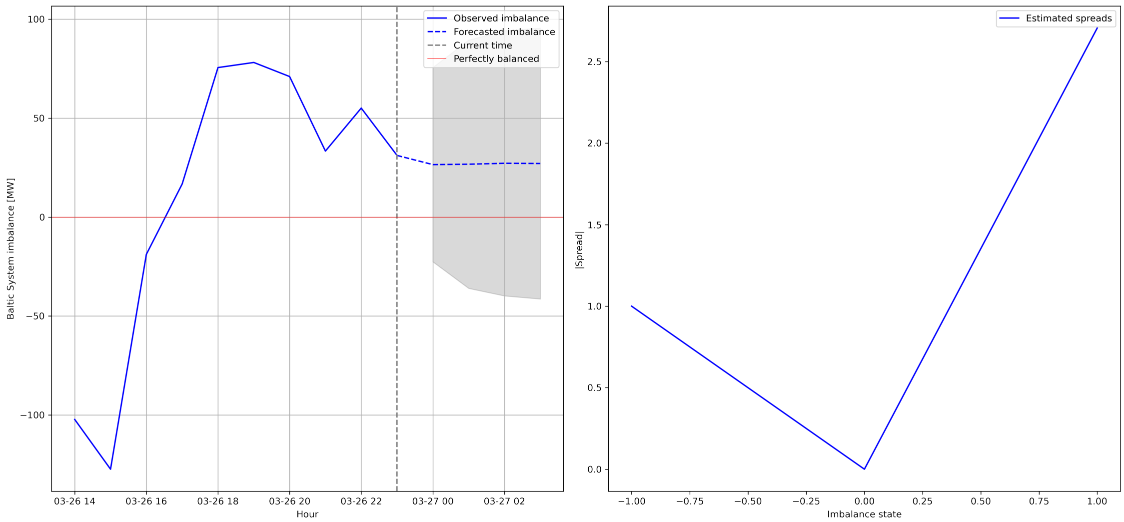

Figure 8. *Calculate spread asymmetry from the system imbalance forecast.* $\alpha^* = 0.7275$ *in the example*

3.5.2 Wind power forecast enhancement for decision model

To use these results, we need to have an accurate quantile function $F^{-1}(X)$, where X is quantile of wind generation in hour t. Wind forecasting is a different topic and outside the scope of this work. A decently accurate point forecasts are available, but the problem arises, that the optimization method under question needs probabilistic forecasts. Thus, we explore the probabilistic features of forecast errors and construct confidence intervals for each separate forecasted value, in order to turn the available point forecast into a probabilistic one. The process corresponds to the module "wind power forecast enhancement" in Figure 7.

To find the empirical probability distribution, we can either use parametric or non-parametric (such as kernel density estimation) methods. Kernel density estimation is used to visually compare the fit of the parametric probability density function. In this work, the parametric distribution is used and to find the parameter vector $\boldsymbol{\theta}$ we use the maximum likelihood estimate 3.14 [27], to fit the parameters for the data. Forecast errors are fitted with the best-fitting probability density function. The data used for fitting is from period 2019-2020, to prevent the information from leaking into the test data set. First, it is tested, if the errors are normally distributed, using Jarque-Bera statistical hypothesis testing [28]. The null hypothesis in this test is that the sample data come from a normally distributed process. The calculated test statistic was 1152 and the corresponding p-value was rounded to 0.0. Therefore, null-hypothesis is rejected on true-alpha level 0.05 meaning that the errors are not normally distributed. Visual confirmation of the fitted Weibull distribution is presented in Figure 9.

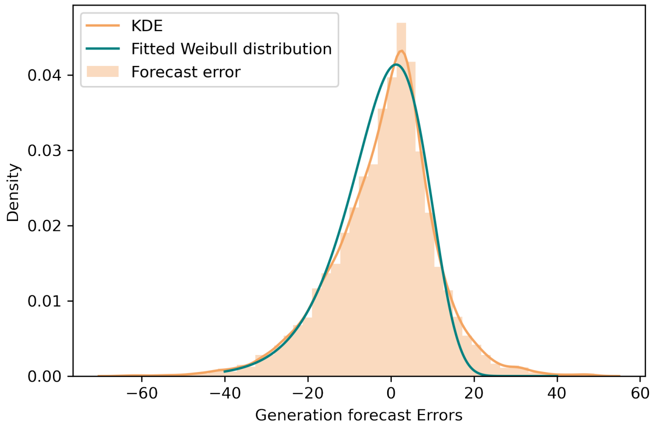

Figure 9. *Fitting the generation forecast errors with Weibull distribution. Error = Forecast-Actual*

Since the error is not normally distributed, a different distribution must be found to model the errors appropriately. Different possible continuous probability distributions are used and fitted with the MLE process. The best fitting distribution was the Weibull distribution, and this is also approved by the literature, which suggests that the error structure comes from the wind speed conversion process to electric energy [22].

$$\hat{\theta} = \underset{\theta \in \Theta}{\arg\max}\, \hat{L}_n(\theta, \boldsymbol{y}) \tag{3.14}$$

where

$\hat{\theta}$ Fitted Weibull distribution parameter vector

$\hat{L}_n$ Likelihood function

$\boldsymbol{y}$ Observed data (in our case, given point forecast errors)

$$\theta = [\kappa, \lambda] \tag{3.15}$$

where

κ Shape parameter in Weibull distribution

λ Scale parameter in Weibull distribution

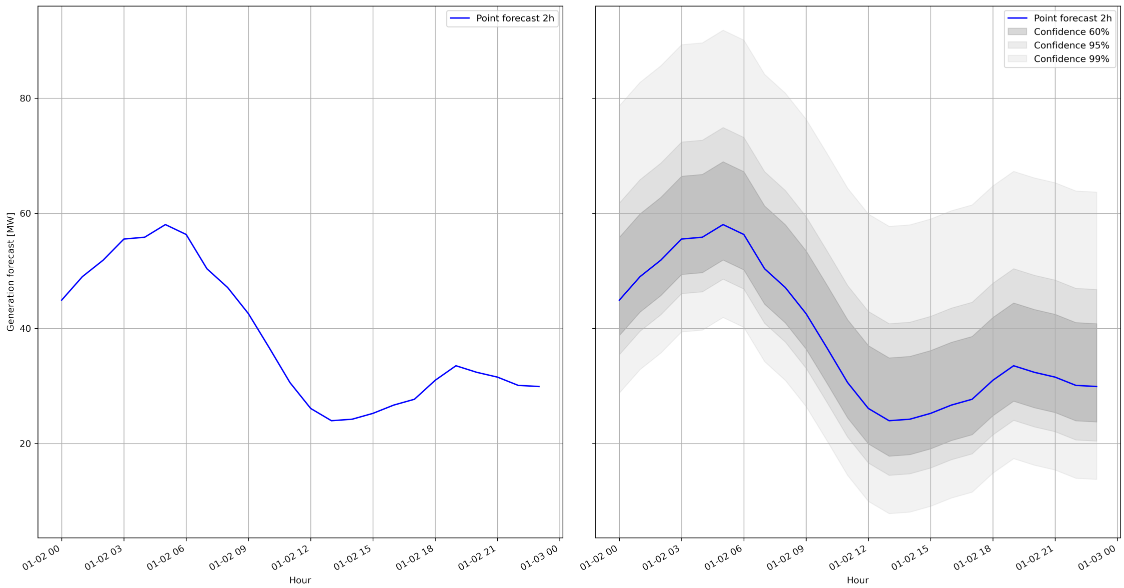

Figure 10. *Wind power generation point forecast turned into a probabilistic one*

3.6 Short-Term imbalance forecasting

To reduce the uncertainty associated with the balance of market prices and our own portfolio imbalance, several forecasting models are implemented to generate a prediction for the stochastic programming model. Both statistical and machine learning forecasting methods are implemented and compared for this task. The random variable has a lot of noise and very accurate prediction cannot arguably, be made. However, in the short-term case, the dynamics of a random process can be used and captured into a model. The training data used for training forecasting models are from 2019-2020. The data from 2021-2022.02 are kept aside for testing.

The models are evaluated against naive forecast. We define a naive forecast as extrapolating the last seen value as many steps forward as our forecast horizon. We define naive forecast made at timestep t for forecast horizon h in 3.16.

$$\hat{y}_{t+h} = y_t \tag{3.16}$$

where

y_t Last seen system imbalance quantity at time period t

$\hat{y}_{t+h}$ Forecasted imbalance quantity value, h steps into the future

3.6.1 Statistical forecast models

The author starts the time series forecasting from statistical models. The first model to be examined is the Simple Exponential Smoothing model [29]3.17. The model is a fairly simple model for timeseries that do not have apparent structure (seasonality or trend). It is a slight increase of complexity from naive models.

$$\hat{y}_{t+h} = \alpha y_t + (1 - \alpha) y_{t-1} \tag{3.17}$$

where

α Smoothing factor

$\hat{y}_{t+h}$ Forecasted imbalance quantity value, h steps into the future

$\hat{y}_t$ The current value of the time series up to current point

Second model, that is tested, is the ARIMA (Autoregressive Integrated Moving Average) time series forecasting model. In developing the model, Box-Jenkins method [13] will be followed. This consists of three main steps:

- Check, the timeseries stationarity and seasonal stationarity. Differentiate, if needed to achieve the stationarity
- Define the Autoregressive (AR) and Moving average (MA) components, using visual inspection of autocorrelation function (ACF) and Partial autocorrelation function (PACF). Fit the model using non-linear least squares. Choose model that minimizes AIC (Akaike Information Criterion)
- Test, if the model captures all the autocorrelation patterns in the data, with Box-Ljung Q-test.

The complete process of finding a model structure with the fit is described in Appendix 2 5 and the sources used for development of statistical time series model are [30] and [31]. There are many reasons that make the statistical models very suitable for this forecasting task. First, there is a huge toolbox developed over many decades specifically for similar problems. Secondly, the models are easy to train and maintain, when compared to more complex machine learning models.

3.6.2 Machine Learning forecast models

Although machine learning models are more complex than their statistical counterparts, they usually achieve much better results. One of the main reasons for this is the ability to recognize more complex and longer-term patterns in the data. For example, the ARIMA(3,0,2) model can only take into account a maximum of 3 periods of history, whereas models based on the architecture of recurrent neural networks can easily take into account longer intertemporal dependencies.

Since the model will be deployed and constantly used for decision making, the machine learning forecasting component will need constant administration and preferably provisional measures, to prevent it from failing. Therefore, the method should outperform statistical methods significantly in order to justify the cost of more complicated solution.

For comparison with statistical methods, the LSTM model (*Long-Short Term memory)* [32], will be implemented to forecast system imbalance. After implementing the model, the results of both will be compared with the forecast error metrics and the economical effect of using the machine learning models in conjunction with the proposed decision model. Different hyperparameter settings about the size of hidden layer, input sequence and different optimizers will be tested, to find the best fitting configuration for the forecast problem at hand. The results are provided in the corresponding section in chapter 4.

3.6.3 Forecast evaluation metrics

The forecasts are evaluated using the continuous value forecast evaluation metrics MAE (Mean Absolute Error) 3.18 and RMSE (Root Mean Squared Error) 3.19. In addition, we are interested in the directional accuracy of our model.

$$MAE = \frac{\sum_{i=1}^{N} |(\hat{y}_i - y_i)|}{N} \tag{3.18}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2}{N}} \tag{3.19}$$

where

y_i Actual value

$\hat{y}_i$ Forecasted value

Since the spread is determined by both the sign and volume of the imbalance, we can also interpret the forecasting problem as a binary classification problem. To evaluate the accuracy of the forecast of directional imbalance, we use a confidence matrix. Since in the current problem formulation, there is no significant difference between type I and type II error, we only look at the accuracy measure 3.20.

$$ACC = \frac{TP + TN}{P + N} \tag{3.20}$$

where

TP True positive classifications count (imbalance quantity > 0)

TN True negative classifications count (imbalance quantity < 0)

N All negative class observations (imbalance quantity > 0)

P All positive class observations (imbalance quantity < 0)

3.7 Implementation to information systems

Since the model has to be run every hour and the window, where adjustments could be made, is small, it is worthwhile to automate the solution, so the only thing left to human trader would be to verify the decisions and send the final offers to the market. The timing of model run and operational scheme is shown in 11 by using the delivery hour 12.00-13.00 as an example. This process will be activated in every hour, and runs at all 24h in a day.

The model is implemented as a separate component in a larger trading system. The architecture of the solution is shown in figure 12

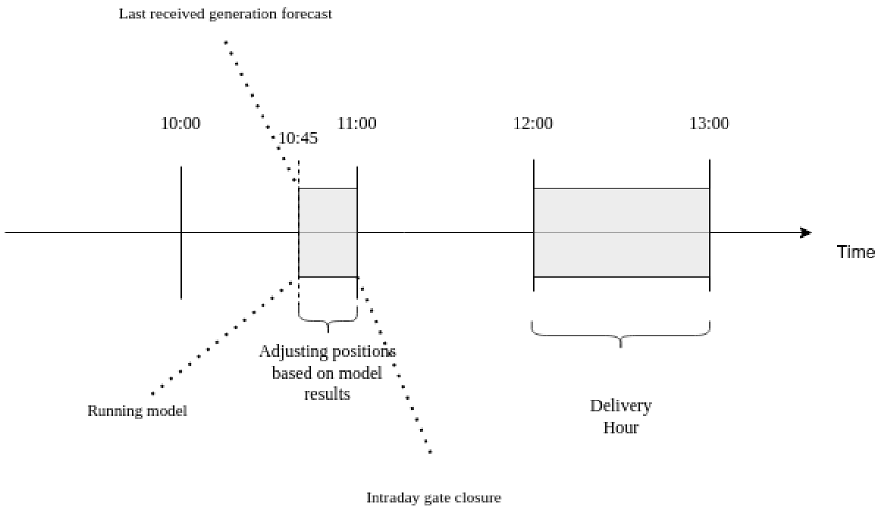

Figure 11. *Model workflow timing description*

3.8 Used technologies

The programming language for implementing the forecast models and optimization framework, was Python 3. Since research and analysis is a major part of that work, Jupyter Notebook interactive notebooks are used as testing grounds for different model development and data analysis. The work largely concerns tabular data, so Pandas library [33] is used throughout the project by having DataFrames as the main data structures. The visualizations are done with Matplotlib graphics library [34]. For statistical models research and implementation, a statsmodels library [35] is used. Statsmodels is a higher-level statistics library that has a similar interface to the R statistics programming language. For computations including probability distributions of random variables, Scipy library [36] was used. For neural network models, the Tensorflow's Keras API [37] and MATLAB Deep Learning toolbox were used.

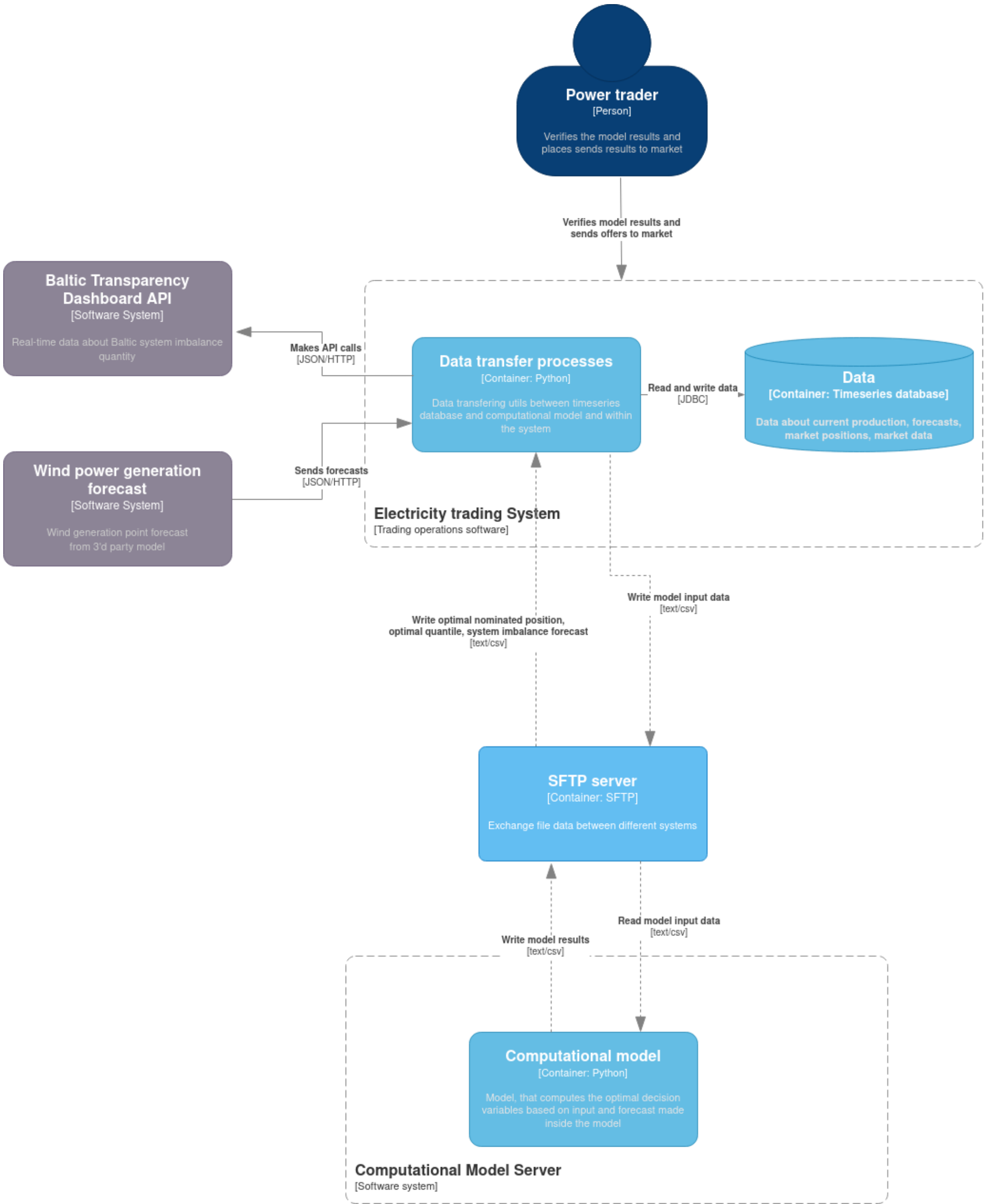

Figure 12. *Component architecture diagram illustrating inputs and outputs of the model*

4. Results of the trading model

In this chapter, the results of applying the previously developed model will be analyzed. First, the model is backtested on historical market data, and the results are interpreted. Then, the common metrics used in stochastic programming problems (Expected Value of Perfect information, Value of Stochastic Solution) are discussed. In the end, the forecasting results are presented as an independent section.

4.1 Short-Term Trading strategies case study

In this section, the results of different trading strategy models are presented and compared with each other. The strategies are tested on historical data from 2021-01-01 to 2022-02-28. In order to fulfil the non-anticipativity constraints, a backtesting framework is developed for the purpose of realistically modelling the results of using the developed model. The goal of non-anticipativity (forecasting models cannot see further in time) is important, to get realistic results.

First, the performances are evaluated on a monthly basis. Figure 13 shows the total performance of the decision model used in conjunction with a naive forecast of imbalances. The naive forecast is used first, in order to bring out the added value of stochastic optimization model without taking into account the added benefit of more accurate imbalance forecasts. Naive forecast is defined in 3.16.

From the figure, it is clear, that the strategy 3 (stochastic optimization strategy) provides constantly better results in terms of revenue in monthly basis. It can also be seen that Strategy 1 (trust the Day-Ahead forecast) performs better than Strategy 2 (intraday adjustments). This is the result of taking smaller risk in Strategy 2.

The total performance is then broken down into components to identify the sources of the results. Figure 14 shows all revenue earned in the Day-Ahead and Intraday markets. The corresponding component in the revenue from the strategy defined in 3.3 is $(y_t^{DA} * \lambda_t^{DA}) + (y_t^{ID} * \lambda_t^{ID})$. It is clear from the figure that strategy 1 and strategy 2 are nearly identical in that term, whereas strategy 3 is constantly achieving higher revenues on a monthly basis. The increase in spot market revenue can only be attributed to the higher amounts sold on

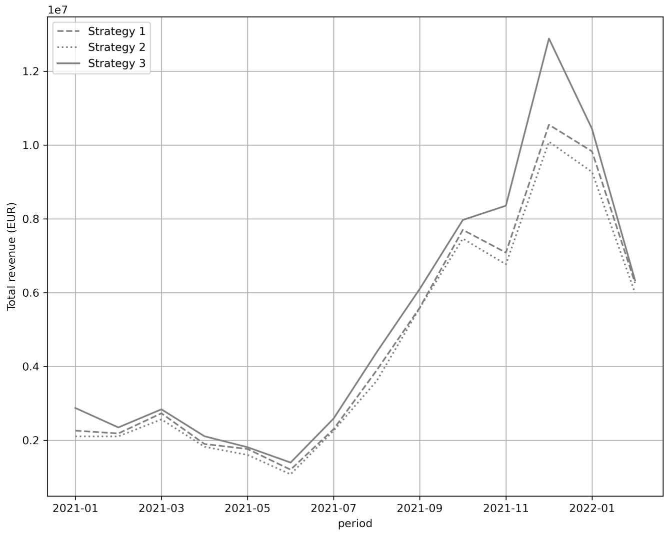

Figure 13. *Different Imbalance strategy monthly aggregated total revenues. Strategy 3 uses naive forecast*

spot markets, since the price is exogenous constant in that case. The outputs of strategy 3 are beneficial for the whole market, as the slight supply increase means lower market prices. Although the effects are small, they are in the right direction.

Figure 15 shows all revenues earned due to imbalances in the nominated position and actual generation. The corresponding component in the revenue of the strategy defined in 3.3 is $y_t - (y_t^{DA} + y_t^{ID})) * \lambda_t^{SI}$. First, it can be seen that strategy 2 is superior to strategy 1, by increasing the balance income (decreasing the cost), since the monthly balance income is constantly higher. The result is expected, as is intuitively clear, that closer to the delivery hour, the generation forecasts are more accurate, and the error between the nominated and actual quantities shrinks. What is more interesting is the constant higher revenues in case of stochastic optimization strategy (Strategy 3).The increase in revenue cannot be interpreted just by selling more electricity. The increase in revenue comes from smarter placement of positions. The choices of strategy can be explained with figure 6.

Table 1 shows the average monthly results for each strategy. For strategy 3, the decision model with the LSTM forecast model with asymmetry multiplier 0.2 is selected. The imbalance cost is defined in 3.2. In Table 1, the "+" sign indicates profit from associated uncertainty. The table shows, that the initial goal of decreasing balancing cost by 20% is achived by large, when compared the result to the baseline, which is strategy 2. The balance cost was transformed into a balance profit and was decreased by over 400% compared to the scenario in which no trading model is used. What this table also shows, is that there is

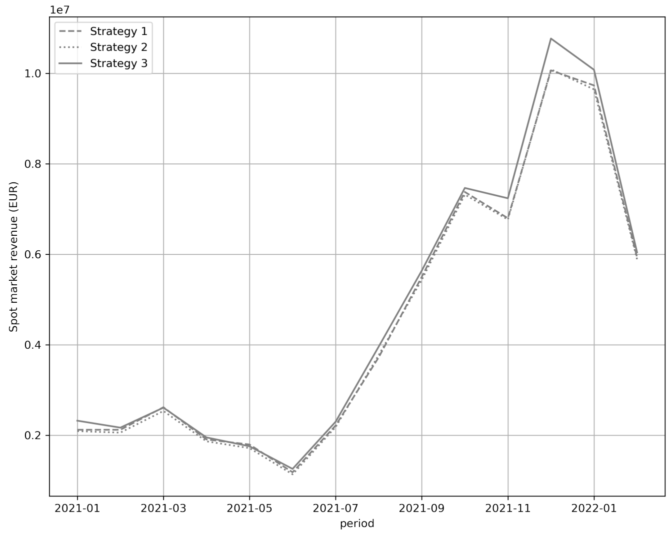

Figure 14. *Different Imbalance strategy monthly aggregated spot market revenues*

a cost to managing risks. Trying to get 0 imbalance costs, Strategy 2 trades on the intraday market at a discounted price. The result is that the strategy is more stable, but in general loses money.

Table 1. *Monthly average results of different strategies. (Units in thousands)*

Metric	**Strategy 1**	**Strategy 2**	**Strategy 3**
Total monthly revenue (10^3 EUR)	4663	4453	5257
Perfect generation forecast revenue (10^3 EUR)	4618	4618	4618
Balancing profit (10^3 EUR)	45	-165	639

Table 2 illustrates the risk profiles of all different strategies. It can be seen that Strategy 3 outperforms all other strategies in that regard as well, while making larger income at the same time. The 95th percentile of the imbalance loss shows the probability that the loss incurred in a certain hour is smaller than the given value in the table. The table shows, that Strategy 3 decreased the 95th percentile of losses and increased the profits by twofold. Although strategy 2 also reduces extreme losses, it makes less profit by doing that. Therefore, it can be concluded, that Strategy 3 not only earns higher profits, but also does that with smaller risk of losing money.

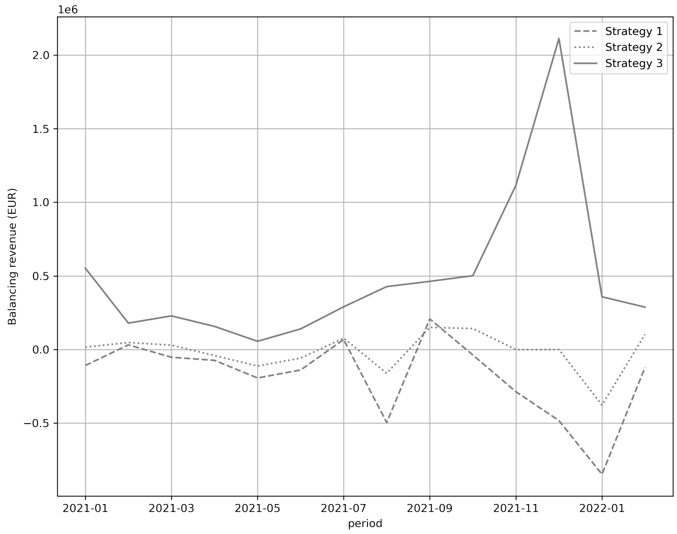

Figure 15. *Different strategies monthly aggregated imbalance settlement revenues*

Table 2. *Risk assessment on hourly losses and profits*

Metric	**Strategy 1**	**Strategy 2**	**Strategy 3**
Max. hourly imbalance cost (10^3 EUR)	-26.98	-20.86	-14.78
Max. hourly imbalance profit (10^3 EUR)	67.25	45.67	94.38
95'th percentile of loss (10^3 EUR)	-3.70	-2.22	-1.64
95'th percentile of profits (10^3 EUR)	2.68	2.28	5.837

Table 3 provides an overview of the different variations in Strategy 3. The models differ in the asymmetry multiplier and the imbalance forecasting model. The asymmetry multiplier can be thought of as *how much do we trust our forecast*, leveraging the results of the decision made by the forecast. The 2 choices of multiplier is presented, as the higher multiplier would render the model results invalid, due to market influencing. The imbalance market is not large, and large values start to affect the prices.

By selecting the best variation, the risk-reward is considered, by taking into account the profit and also the minimizing the possibility of having a large loss. The best performing model is LSTM with an asymmetry multiplier of 0.2. The naive model, which forecasts the last seen value into the future, also performs quite well, taking into account the simplicity of model compared to more complex models. The best-performing configurations are highlighted in the table. SES in the table stands for Simple Exponential Smoothing forecast model.

Table 3. *Strategy 3 results with different forecast models and parameters*

Used forecast model	**Naive**		**ARIMA**		**SES**		**LSTM**	
Asymmetry multiplier	0.1	0.2	0.1	0.2	0.1	0.2	0.1	0.2
Spot monthly revenue (10^3 EUR)	4685	4735	4625	4688	4648	4688	4772	4870
Balancing monthly revenue (10^3 EUR)	490.5	585.3	365.8	441.2	449.7	532.2	473.8	386.8
Total monthly revenue (10^3 EUR)	5175	5320	4991	5129	5234	5129	5158	5257
95'th percentile of imbalance loss (10^3 EUR)	-1.99	-2.32	-1.82	-2.06	-2.12	-2.44	-1.49	-1.64
Day-Ahead (10^3 MWh)	732.87	732.87	732.87	732.87	732.87	732.87	732.87	732.87
Intraday (10^3 MWh)	146	172	123	142	144	171	124	143
Balancing profit (10^3 MWh)	136	159	117	134	136	160	116	132
Balancing profit	557.3	702.2	373.3	511.8	616.1	511.1	540.5	639
p-value (Balancing profit)	0.27	0.18	0.43	0.30	0.28	0.29	0.25	0.21

A statistical hypothesis testing is conducted, to determine, whether the balancing profit results are statistically significant. For the test, Student's T-test [38] was chosen to test if the sample mean is equal to the given expected value. The T-test was chosen due to the very small sample size (n=14, each data point is monthly mean). Null hypothesis means that there is no significant difference between population mean and sample mean. As seen in Table 3, the null hypothesis is not rejected for any values at significance level $\alpha = 0.05$, therefore, there is enough evidence, to say that the results represent expected values of these strategies. Furthermore, the p-value also shows how strong is this statement, and it can be seen that the argument is the strongest in the case of ARIMA model with asymmetry multiplier 0.1.

The reason why Naive model with 0.2 asymmetry multiplier was not selected, although it has the highest mean monthly balancing profit, is that the risk is also the highest. The final selection of the model used is choice of management, and this model provides the input to make that decision.

The energy metrics are the volume of certain market participation (in absolute value). Therefore, they do not have to sum up to a certain value, but show where certain forecast models tended to place the offers in larger quantity. The pattern is clear that, in general, all models placed higher quantities on the imbalance market in case of higher asymmetry

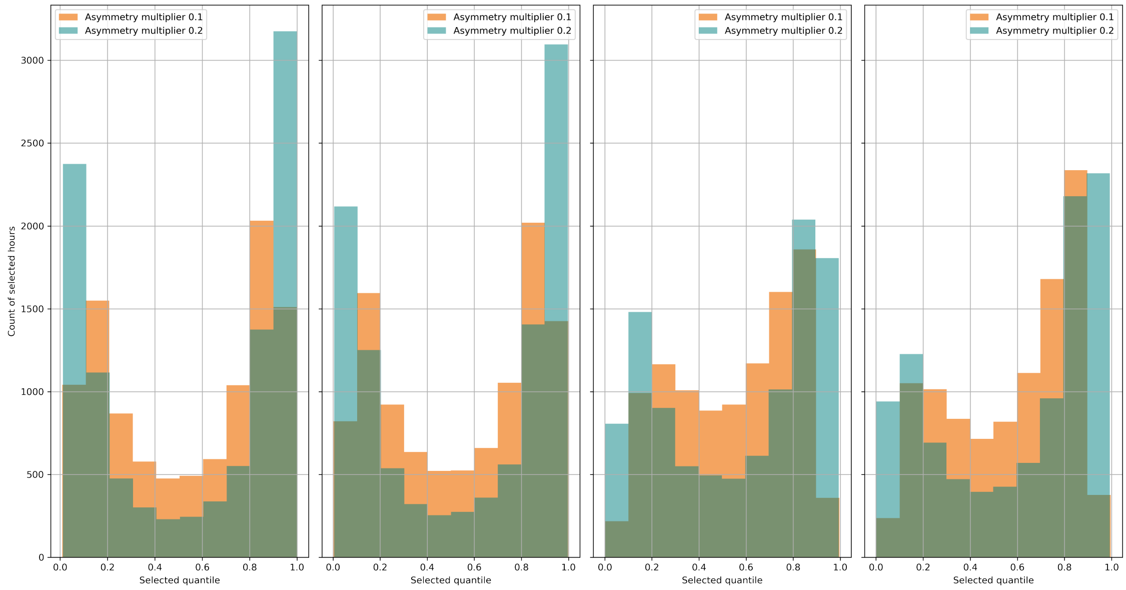

Figure 16. *Decisions about optimal quantile from different forecasts. a) Naive, b) Exponential smoothing, c) ARIMA, d) LSTM*

multiplier. That behavior is expected, but can also increase the risk of the model, which is also seen as a higher 95th percentile of imbalance loss, in case of higher asymmetry multiplier. This loss can also be interpreted as a Value of Risk (VaR) measure at a certain confidence level (for example 95%). The interpretation of this metric would be that the hourly loss stays below the presented value, 95% of the time.

The model decisions are illustrated in Figure 16. It is seen, that the more simple forecast models place much more positions on end of the probability distribution. This is expected, as the naive model, for example, forecasts the last seen value, and does not have any averaging mechanism. More complex forecast models generally rather underestimate and do not forecast extreme values. This kind of behavior is wanted in this case, as the main goal is to reduce the cost by placing the imbalance position on the safe side of the market. It is also seen from the figure that the more two more advanced strategies and the two simpler strategies have very similar decision behavior, which can be attributed to the forecast accuracy.

To determine, whether it is reasonable to invest into better system imbalance forecasting method, it is reasonable to calculate how much value the better system imbalance forecast would add, in order to make better decisions about increasing the complexity of forecast model. For this, we use the expected value of perfect information (*EVPI*), that is commonly used in stochastic programming problems. The metric is defined in [17].

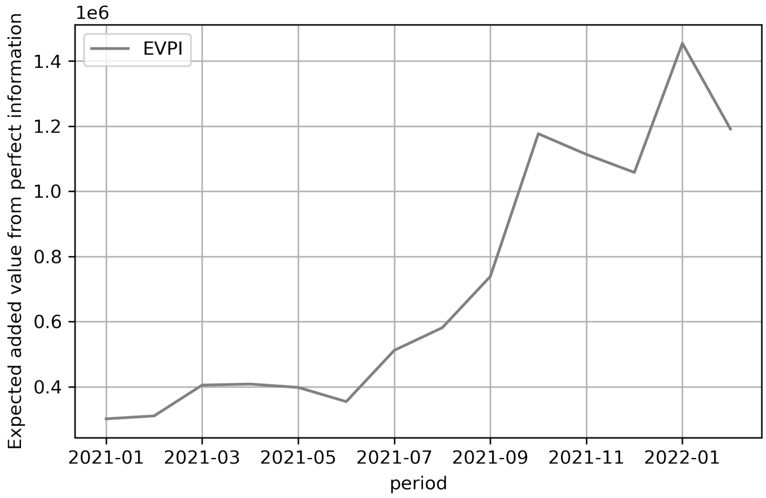

Figure 17. *Monthly expected added value from perfectly accurate imbalance quantity forecast*

$$EVPI = WS - SP \tag{4.1}$$

where

WS Wait-and see solution value, perfect information about system imbalance at all times

SP Stochastic program obtained value with naive forecast

The monthly mean EVPI is presented in Table 4 and Figure 17, which shows the maximum attainable value of a perfectly accurate forecast, when used with the current decision model. From the table, it can be interpreted, that if it is possible, to obtain a perfect forecast about Baltic system imbalance quantity, it would improve the value of this decision model from our current best model by 15.73% and is worth over 800 thousand EUR per month in average, which is a great improvement. It should be emphasized that the values correspond to a specific generation portfolio and the value does not account for the changes in the market conditions that appear from the changed position. For this reason, the value is expected to be lower than that in practice.

Table 4. *Expected value of perfect information*

Metric	**Expected value of perfect information**
Mean EVPI (EUR)	837,118
Mean EVPI (%)	15.73 %

In order to measure the impact of the proposed stochastic model, we can calculate the value added by using a random variable in the objective function instead using the expected value of that random variable. That is the value of the stochastic solution (*VSS*) [17], which tells how much performance can be improved just by using the stochastic programming solution instead of the deterministic optimization problem.

It can be seen from Table 5, that just by using the proposed optimal quantile selection mechanism, performance can be improved by 11.5 % on average when compared to the model, where wind power generation is treated as a random variable. The median values are brought out separately, because of the highly different November-December period, which has a large effect on mean values. Figure 18 shows the monthly variability of value added by stochastic solution.

$$VSS = SP - EEV \tag{4.2}$$

where

EEV Mean Earned revenue with expected value solution, $\alpha = 0.5$, with naive forecast

SP Mean Revenue earned by stochastic programming solution with naive forecast (strategy 3)

Table 5. *Value of stochastic solution*

	Value of stochastic solution
Mean VSS (EUR)	572,441
Mean VSS (%)	12.05 %
Median VSS (EUR)	279,306
Median VSS (%)	5.88%

4.2 Imbalance forecast modeling results

First, the results of timeseries models, where imbalance is treated as a regreression problem, are brought out. Forecasts are made 2 hours ahead, due to the aforementioned time schedule figure 11. From Table 6, it can be seen that the more complex forecasting methods brought more accurate results, with the most complex model bringing the most accurate results.

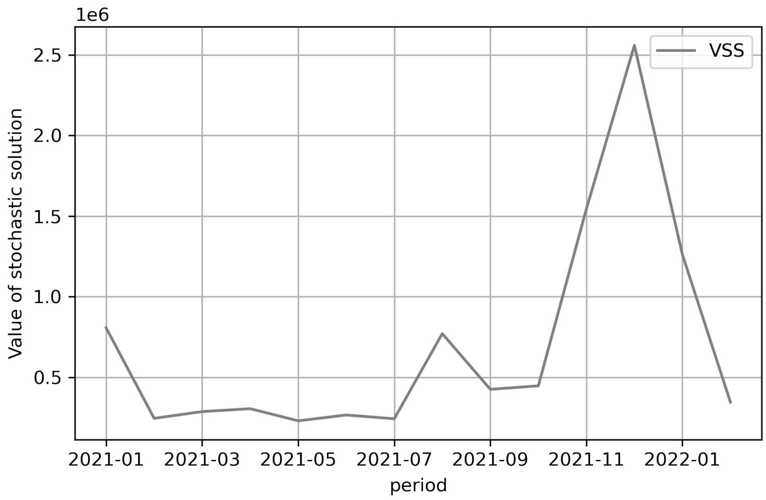

Figure 18. *Monthly added value from using stochastic solution*

Table 6. *Errors of timeseries forecasting methods in regression task*

Model and forecast horizon	**MAE**	**RMSE**	**Directional accuracy (%)**
Naive, 2h	56.03	73.85	67.26%
Exponential Smoothing, 2h	55.17	72.56	66.76%
ARIMA(3,0,2), 2h	49.56	64.63	67.59%
LSTM, 2h	**34.93**	**46.23**	**82.29%**

Previously, the imbalance was modeled as a real value and modeled as a regression problem, but it can also be treated as a binary classification problem by labeling all points according to their sign. The imbalance quantity below 0 is class Short and above 0 is class Long.

Table 7. *Hyperparameter configurations explored in LSTM model*

Parameter	**Value**
Learning rate	{0.001, 0.01, 0.1}
Epochs	100, 150, 300, 500
Optimizer	{Adam, SGD}
Loss function	MSE
Number of hidden units	{64, 128, 512, 1024}

The selected models were trained on a NVIDIA GeForce GTX 1050 Ti GPU, which has 4GB memory. The best model was obtained, using the highest number of hidden units that fit into machine memory (hidden vector length 1024), therefore remembering the most information between time steps. The optimal number of trained epochs was chosen by using a validation set next to training set, in order to determine, where the validation accuracy starts decreasing and identify possible overfitting. The number of epochs with given best configuration, was found to be about 150 and the time elapsed to train this model, with previously mentioned hardware, was around 15 minutes. The best optimizer for this model was found to be Adam, the best learning rate was 0.1.

5. Summary

In the work, a decision tool was developed for the wind power producer to help reduce the costs of portfolio imbalance. In order to reduce these costs, a model was developed, that consists of a decision model and a system imbalance forecasting model. It was shown in the case study, that the developed model is able to increase the total revenue from 12.5 % to 18 % which corresponds to monetary increase of up to 600,000 EUR in month with the best model, compared to the baseline strategy.

The balancing cost was turned into a balancing profit, by using the uncertainty present in the volatile balancing market to the strategy's favor. The balance cost was thus improved by up to 480% on a monthly average with the best model, which exceeds the goal.

The answers to the research questions formulated at the start of this research got answered as follows:

- It is shown that by using only naive forecast, but modeling the generation forecast error and using a smart strategy with optimal quantile selection, monetary value can be added around 500,000 euros in month 3, while also decreasing the risk.
- The machine learning forecast provided better results in terms of accuracy and also in terms of reduced risk and monetary performance. As a suprise, naive model provided good results in conjunction with the decision model, but the worst results in terms of accuracy.
- In the author's opinion, to justify the complexity of using a deep learning model in live environment, further experiments have to be carried out and the cost of maintaining the model has to be seriously evaluated. The simpler model showed only slight less better results, and have better interpretability. Therefore, these are implemented first

There are many possible ways to further improve the decision model. Further developments can be separated into two categories: 1) improvements to the decision model and 2) improvements to forecasting methods.

The improvements to the decision model would be as follows.

- To use scenarios and scenario reduction, to better capture the auto-correlation structure in wind generation forecast
- Add CVaR risk multiplier to objective function, to better model risk-averse strategies and add further control over the possible outcomes

The improvements to forecasting methods would be:

- Use transformers neural network architecture to forecast imbalance prices. There is more possible value added from a more accurate forecast as the perfect information value was shown to be over 800,000 EUR in month in addition to the naive model 4.1.
- Combine multiple time series into the forecast to capture relations between processes that influence system imbalance

As a result of this work, a decision model was developed, to help reduce the financial risk of wind producer, by trading more effectively in short-term power markets. This is beneficial for the wind producer, helping to develop the industry at a faster pace to increase the amount of electricity produced from renewable sources.

Bibliography

[1] International Energy Agency. *Global Energy Review 2021*. 2021. URL: https://iea.blob.core.windows.net/assets/d0031107-401d-4a2f-a48b-9eed19457335/GlobalEnergyReview2021.pdf (visited on 03/26/2022).

[2] S.J. Deng and S.S. Oren. "Electricity derivatives and risk management". In: *Energy* 31.6 (2006). Electricity Market Reform and Deregulation, pp. 940–953. ISSN: 0360-5442. DOI: https://doi.org/10.1016/j.energy.2005.02.015. URL: https://www.sciencedirect.com/science/article/pii/S0360544205000496.

[3] NordPool. *Bidding Areas*. 2022. URL: https://www.nordpoolgroup.com/the-power-market/Bidding-areas/ (visited on 03/26/2022).

[4] EPEX SPOT. *About EPEX SPOT*. 2022. URL: https://www.epexspot.com/en/about (visited on 03/26/2022).

[5] Tooraj Jamasb and Michael Pollitt. "Electricity Market Reform in the European Union: Review of Progress toward Liberalization and Integration". In: *The Energy Journal* 26.6 (2005). Electricity Market Reform and Deregulation, pp. 11–41. ISSN: 1944-9089. DOI: 10.5547/ISSN0195-6574-EJ-Vol26-NoSI-2. URL: https://www.jstor.org/stable/23297005.

[6] NordPool. *Product Specifications Nordic and Baltic market areas*. 2022. URL: https://www.nordpoolgroup.com/494b1d/globalassets/download-center/rules-and-regulations/%5C-product-specifications--nordic-and-baltic-market-areas-valid-from-16.03.22.pdf (visited on 03/27/2022).

[7] Jurate Jaraite-Kazukauske et al. *Market structure of the Nordic power market*. 2022. URL: https://www.researchgate.net/figure/Market-structure-of-the-Nordic-power-market-Source-Authors-own-illustration-based-on_fig6_333263046 (visited on 03/27/2022).

[8] Nicoló Mazzi and Pierre Pinson. “10 - Wind power in electricity markets and the value of forecasting”. In: *Renewable Energy Forecasting*. Ed. by George Kariniotakis. Woodhead Publishing Series in Energy. Woodhead Publishing, 2017, pp. 259–278. ISBN: 978-0-08-100504-0. DOI: https://doi.org/10.1016/B978-0-08-100504-0.00010-X. URL: https://www.sciencedirect.com/science/article/pii/B978008100504000010X.

[9] LITGRID AB Elering AS AS "Augstsprieguma tīkls”. *Baltic Balancing Market Rules*. 2017. URL: https://elering.ee/sites/default/files/public/Teenused/Bilansiteenus/Annex%5C%2010%5C%20Baltic%5C%20balancing%5C%20market%5C%20rules_30112017.pdf (visited on 03/27/2022).

[10] Elering AS. *Synchronisation one-pager*. 2021. URL: https://elering.ee/sites/default/files/2021-06/Sync_onepager_ENG_210x297%5C%2B5mm_print.pdf (visited on 03/27/2022).

[11] *Piiriülene Ebabilanss*. 2021. URL: https://www.elering.ee/piiriulene-ebabilanss (visited on 04/12/2022).

[12] Rafał Weron. “Electricity price forecasting: A review of the state-of-the-art with a look into the future”. In: *International Journal of Forecasting* 30.4 (2014), pp. 1030–1081. ISSN: 0169-2070. DOI: https://doi.org/10.1016/j.ijforecast.2014.08.008. URL: https://www.sciencedirect.com/science/article/pii/S0169207014001083.

[13] George EP Box. “GM Jenkins Time Series Analysis: Forecasting and Control”. In: *San Francisco, Holdan-Day* (1970).

[14] Jian Fu et al. “Condition Monitoring of Wind Turbine Gearbox Bearing Based on Deep Learning Model”. In: *IEEE Access* PP (Apr. 2019), pp. 1–1. DOI: 10.1109/ACCESS.2019.2912621.

[15] Eilyan Y. Bitar et al. “Bringing Wind Energy to Market”. In: *IEEE Transactions on Power Systems* 27.3 (2012), pp. 1225–1235. DOI: 10.1109/TPWRS.2012.2183395.

[16] Pierre Pinson, Christophe Chevallier, and George N. Kariniotakis. “Trading Wind Generation From Short-Term Probabilistic Forecasts of Wind Power”. In: *IEEE Transactions on Power Systems* 22.3 (2007), pp. 1148–1156. DOI: 10.1109/TPWRS.2007.901117.

[17] John R Birge and Francois Louveaux. *Introduction to stochastic programming*. Springer Science & Business Media, 2011.

[18] Trine Krogh Boomsma, Nina Juul, and Stein-Erik Fleten. "Bidding in sequential electricity markets: The Nordic case". In: *European Journal of Operational Research* 238.3 (2014), pp. 797–809. ISSN: 0377-2217. DOI: https://doi.org/10.1016/j.ejor.2014.04.027. URL: https://www.sciencedirect.com/science/article/pii/S0377221714003695.

[19] Mansour Hosseini Firouz. "Optimal offering strategy considering the risk management for wind power producers in electricity market". In: *International Journal of Electrical Power and Energy Systems* 49 (July 2013), pp. 359–368. DOI: 10.1016/j.ijepes.2013.01.015.

[20] Juan M. Morales, Antonio J. Conejo, and Juan PÉrez-Ruiz. "Short-Term Trading for a Wind Power Producer". In: *IEEE Transactions on Power Systems* 25.1 (2010), pp. 554–564. DOI: 10.1109/TPWRS.2009.2036810.

[21] Morteza Rahimiyan, Juan M. Morales, and Antonio J. Conejo. "Evaluating alternative offering strategies for wind producers in a pool". In: *Applied Energy* 88.12 (2011), pp. 4918–4926. ISSN: 0306-2619. DOI: https://doi.org/10.1016/j.apenergy.2011.06.038. URL: https://www.sciencedirect.com/science/article/pii/S0306261911004272.

[22] J.V. Seguro and T.W. Lambert. "Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis". In: *Journal of Wind Engineering and Industrial Aerodynamics* 85.1 (2000), pp. 75–84. ISSN: 0167-6105. DOI: https://doi.org/10.1016/S0167-6105(99)00122-1. URL: https://www.sciencedirect.com/science/article/pii/S0167610599001221.

[23] Jethro Browell. "Risk Constrained Trading Strategies for Stochastic Generation with a Single-Price Balancing Market". In: *Energies* 11 (June 2017). DOI: 10.3390/en11061345.

[24] Jeremie Bottieau et al. "Automatic Risk Adjustment for Profit Maximization in Renewable Dominated Short-Term Electricity Markets". In: *International Transactions on Electrical Energy Systems* 31 (Sept. 2021). DOI: 10.1002/2050-7038.13152.

[25] Elering AS. *Elektrituru käsiraamat*. 2017. URL: https://elering.ee/sites/default/files/elektrituru-kasiraamat.pdf (visited on 03/27/2022).

[26] Elering AS. *Baltic Balancing Roadmap*. 2021. URL: https://elering.ee/sites/default/files/2021-10/Baltic%5C%20balancing%5C%20roadmap_v6_0.pdf (visited on 04/16/2022).

[27] In Jae Myung. “Tutorial on maximum likelihood estimation”. In: *Journal of Mathematical Psychology* 47.1 (2003), pp. 90–100. ISSN: 0022-2496. DOI: https://doi.org/10.1016/S0022-2496(02)00028-7. URL: https://www.sciencedirect.com/science/article/pii/S0022249602000287.

[28] Carlos M Jarque and Anil K Bera. “Efficient tests for normality, homoscedasticity and serial independence of regression residuals”. In: *Economics letters* 6.3 (1980), pp. 255–259.

[29] Robert G Brown and Richard F Meyer. “The fundamental theorem of exponential smoothing”. In: *Operations Research* 9.5 (1961), pp. 673–685.

[30] Rob J Hyndman and George Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2018. URL: https://otexts.com/fpp2/.

[31] *Unit Root Tests*. 2021. URL: https://faculty.washington.edu/ezivot/econ584/notes/unitroot.pdf (visited on 04/21/2022).

[32] Sepp Hochreiter and Jürgen Schmidhuber. “Long Short-term Memory”. In: *Neural computation* 9 (Dec. 1997), pp. 1735–80. DOI: 10.1162/neco.1997.9.8.1735.

[33] Wes McKinney. “Data Structures for Statistical Computing in Python”. In: *Proceedings of the 9th Python in Science Conference*. Ed. by Stéfan van der Walt and Jarrod Millman. 2010, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a.

[34] J. D. Hunter. “Matplotlib: A 2D graphics environment”. In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/MCSE.2007.55.

[35] Skipper Seabold and Josef Perktold. “statsmodels: Econometric and statistical modeling with python”. In: *9th Python in Science Conference*. 2010.

[36] Pauli Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[37] François Chollet et al. *Keras*. https://keras.io. 2015.

[38] Student. “The probable error of a mean”. In: *Biometrika* (1908), pp. 1–25.

[39] Yosiyuki Sakamoto, Makio Ishiguro, and Genshiro Kitagawa. “Akaike information criterion statistics”. In: *Dordrecht, The Netherlands: D. Reidel* 81.10.5555 (1986), p. 26853.

Appendices

Appendix 1 - Derivation of the Solution to the General News Vendor Problem

This appendix presents the derivation steps needed, to find the used solution. The derivation for the general news vendor problem is presented in [17], but in order to apply it to the problem of finding optimal imbalance position, we have to make some modifications, because in the original problem, the spreads are known constants. In the current case, the spreads are also random variables, and determined from system imbalance forecast.

The general newsvendor problem is written as:

$$
\begin{aligned}
min\ z = c^T \boldsymbol{x}\ + \mathscr{D}(\boldsymbol{x}) \\
s.t\ Ax = b \\
x >= 0
\end{aligned}
\tag{1}
$$

$$\mathscr{D}(x) = E_{\xi}[-qmin(\xi, x) - rmax(x - \xi, 0)]$$

where

$\mathscr{D}$ Expected value of profit on certain realization of ξ

q Spread when imbalance surplus (over 0) (S^- in model)

r Spread when imbalance deficit (below 0) (S^+ in model)

$$\begin{cases} x = 0 \text{ if } \mathscr{D}'(0) > 0 \\ \text{a solution of } c + \mathscr{D}'(x) = 0, \text{ otherwise} \end{cases}$$

The expected value can be constructed as an integral according to definition.

$$\mathscr{D}(x) = \int_{-\infty}^{X} (-q\xi - r(x-\xi))dF(\xi) + \int_{X}^{\infty} -qx\, dF(\xi) =$$

$$= -(q-r)\int_{-\infty}^{X} \xi\, dF(\xi) - \ rxF(x) - qx(1-F(x))$$

$$\int_{-\infty}^{X} \xi\, dF(\xi) = xF(x) - \int_{-\infty}^{X} F(\xi)d\xi$$

$$\mathscr{D}(x) = -qx + (q-r)\int_{-\infty}^{X} F(\xi)d\xi$$

Now, the derivative can be taken, to obtain the optimal solution for the problem.

$$\mathscr{D}'(x) = -q + (q-r)F(x)$$

$$x^* = F^{-1}(\frac{q-c}{q-r})$$

where

F^{-1} Inverse cumulative distribution function of generation forecast random variable

and q and r are the spread values, that we can write as one constant $\alpha^* = \frac{q-c}{q-r}$.

Appendix 2 - ARIMA model selection

In order to model the random process, that generated the observed timeseries, first, the timeseries have to be stationary (weak stationarity). That means the moments of random variable do not change over time. To check that, we can check, whether the process has unit root ($|\phi| < 1$). Augmented Dickey-Fuller test null-hypothesis is that the process does not have a unit root.

Augmented Dickey-Fuller test statistic

$$y_t = \beta \boldsymbol{D}_t + \phi y_{t-1} + \sum_{j=1}^{p} \psi_j \Delta y_{t-j} + \epsilon_t \tag{2}$$

Alternative test for testing for unit-root is Phillips-Perron test, but since it ignores all the serial correlation in test-regression, it is more prone to heteroscedacity, than ADF test. In order to test stationarity only, KPSS (Kwiatokowski-Phillips-Schmidt-Shin) test is frequently used and I will use it with ADF, to see if the results agree.

KPSS test statistic

$$\begin{aligned} y_t &= \beta \boldsymbol{D}_t + \mu_t + u_t \\ \mu_t &= \mu_{t-1} + \epsilon_t, \\ \epsilon_t &\sim WN(0, \sigma_\epsilon^2) \end{aligned} \tag{3}$$

The data about Baltic system imbalance is from 2021-01-01 - 2022-02-28 and consists of 10175 observations. From 19, it is clear, that this is AR process and the AR rank is at least AR(2).

To objectively verify, whether the selected ranks are best fitting, we use AIC criterion [39], given in 4. Intuitively, if the AIC value increases, adding new term to the model is justified.

Akaike Information Criterion

$$\text{AIC} = 2k \ln(L) + 2(p + q + k + 1) \tag{4}$$

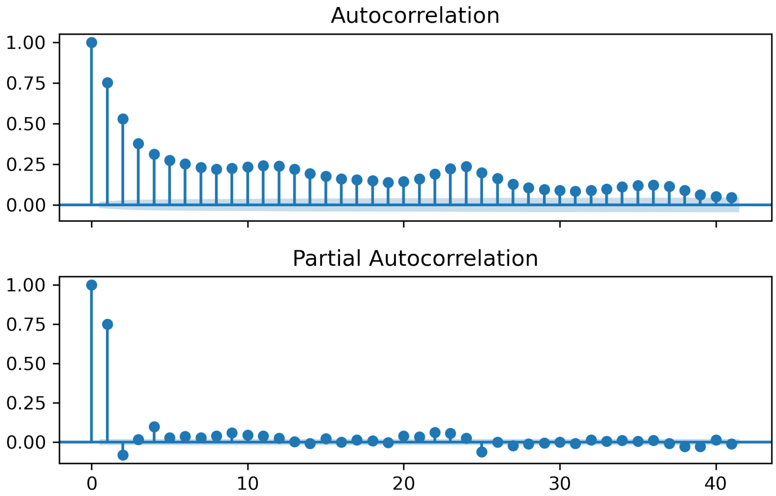

Figure 19. *ACF and PACF*

where

k AR terms in model

q MA terms in model

L Likelihood of the model

The best fitting model regarding this criterion is ARMA(3,2).

Model 1: ARMA, using observations 01.01.2019–31.12.2020
Dependent variable: System imbalance quantity
Standard errors based on Hessian

	Coefficient	Std. Error	z	p-value
const	12.9940	2.22019	5.853	0.0000
ϕ_1	0.663240	0.00263490	251.7	0.0000
ϕ_2	0.986984	0.00156464	630.8	0.0000
ϕ_3	−0.669410	0.00230938	−289.9	0.0000
θ_1	0.0533484	0.00499381	10.68	0.0000
θ_2	−0.936027	0.00454599	−205.9	0.0000

Mean dependent var	12.98023	S.D. dependent var	72.27561
Mean of innovations	−0.021430	S.D. of innovations	48.11183
R^2	0.556856	Adjusted R^2	0.556755
Log-likelihood	−92729.85	Akaike criterion	185473.7
Schwarz criterion	185528.1	Hannan–Quinn	185491.6

			Real	Imaginary	Modulus	Frequency
AR						
	Root	1	1.4442	0.0000	1.4442	0.0000
	Root	2	-1.0021	0.0000	1.0021	0.5000
	Root	3	1.0322	0.0000	1.0322	0.0000
MA						
	Root	1	-1.0055	0.0000	1.0055	0.5000
	Root	2	1.0625	0.0000	1.0625	0.0000

The goal of the ARMA model is to remove any autoregressive patterns from the residuals. To test that, we test the residuals with correlogram and Q-statistic.

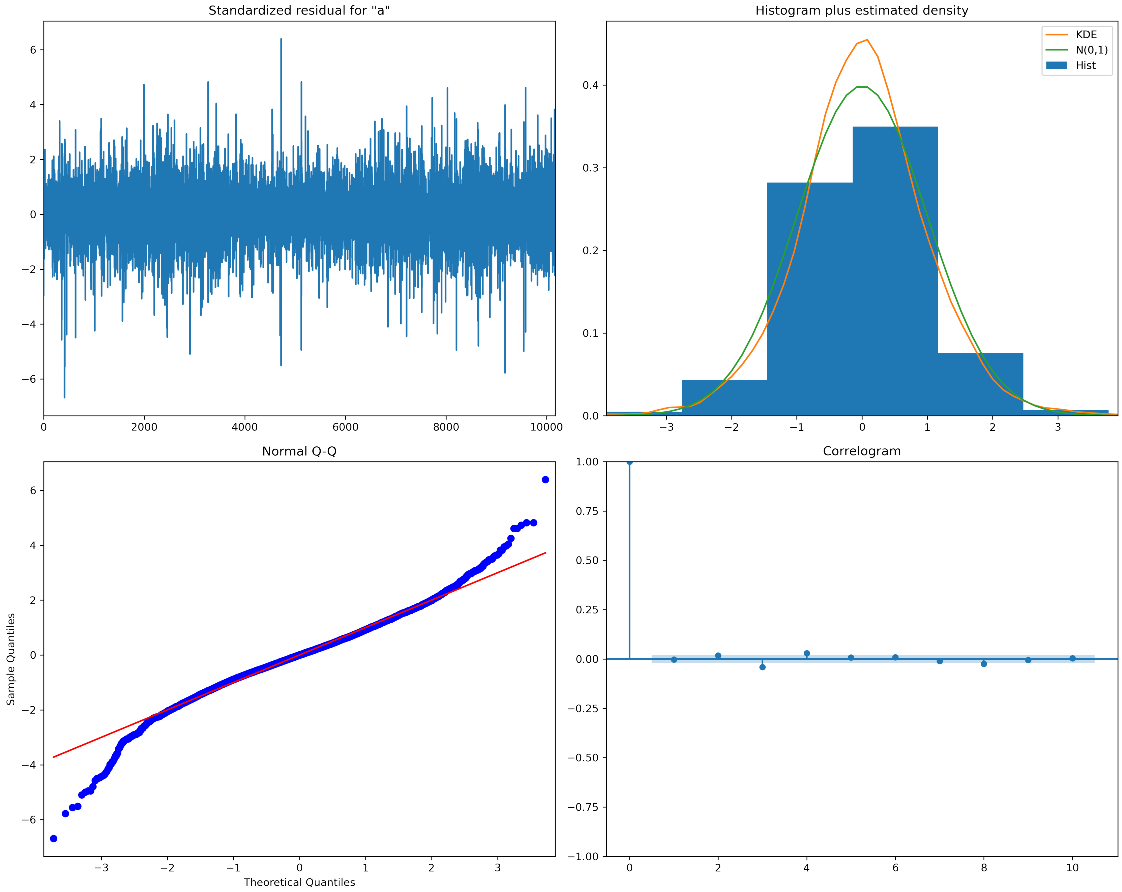

Figure 20. *Residual diagnostics of fitted ARMA(3,2) process*

Appendix 3 - Non-exclusive licence for reproduction and publication of a graduation thesis

I, Andri Busch

1. Grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis "Improving Short-Term Power Markets Trading Strategy for Wind Power Producer", supervised by Marko Kääramees and Sven Nõmm

   1.1. to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;

   1.2. to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.

2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.

3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

May 30, 2022