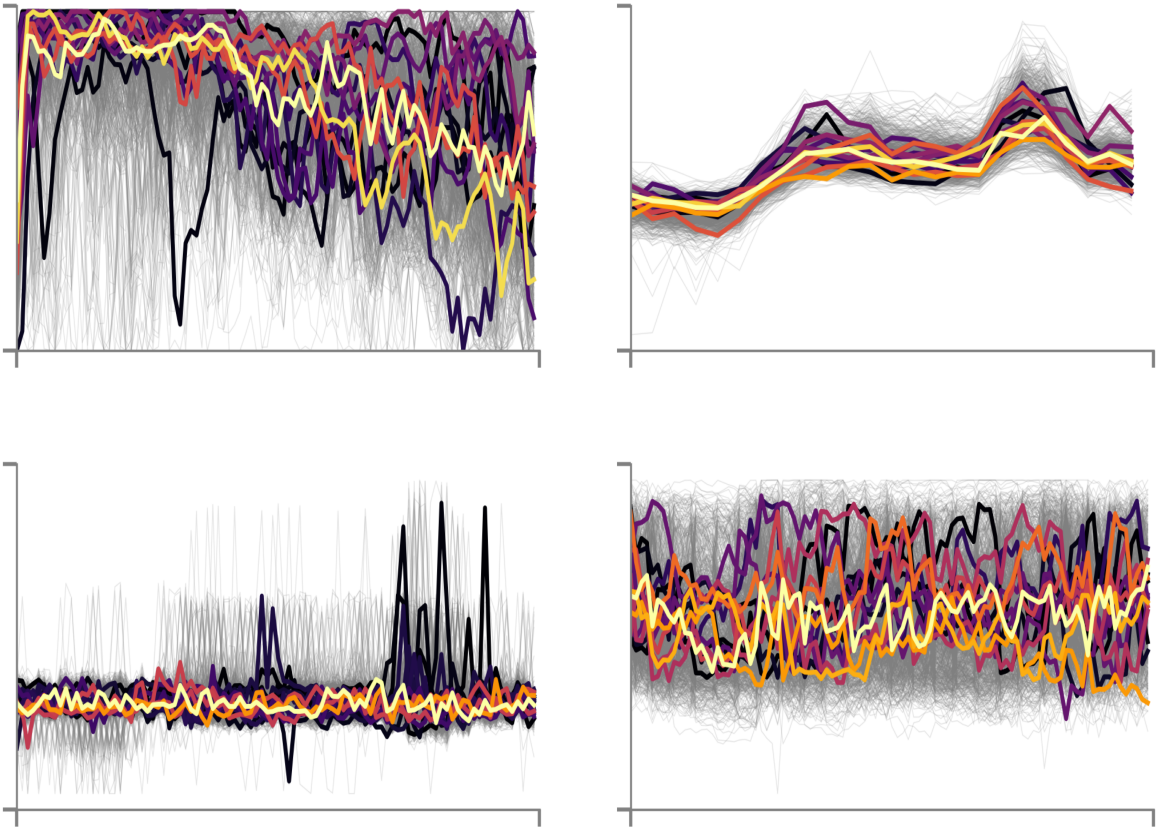
Increasing the Market Value of Wind Power Using Improved Stochastic Process Modeling and Optimization

A Case Study of a Belgian Wind Power Producer

G.J. Deen



TU Delft

Increasing the Market Value of Wind Power Using Improved Stochastic Process Modeling and Optimization

A Case Study of a Belgian Wind Power Producer

by

G.J. Deen

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on March 29, 2019

Student number:	4036409	
Project duration:	March 12, 2018 – March 29, 2019	
Thesis committee:	Dr. S. W. Cunningham,	TU Delft, chair
	Dr. ir. R. Verzijlbergh,	TU Delft, supervisor
	Dr. S. van Langen,	Essent, supervisor
	MSc. T. Kuipers,	Essent, supervisor

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Preface

This thesis is written as part of the MSc programme Engineering and Policy Analysis. It is the end product of a year of research into the possibility to increase the market value of wind power in the Belgian electricity market. It aims to provide both private and institutional decision-makers insight in the potential of advanced bidding strategies, in order to help alleviate some of the short-term negative effects from upscaling wind power. While the short-term effects of scaling up wind power and any other renewable power with an uncertain production are not only negative for their own value, it also has a negative effect on the for the electricity grid, as it leads to strong short term fluctuations and an increased risk of extreme imbalances. As both effects negatively impact the business case for and societal cost of wind power, while the short term scaling up of low carbon power generation seems increasingly urgent, I was and luckily am still thankful and highly motivated to be able to contribute to this subject.

While at the outset the content of the research was loosely defined, the goal was always clear, which was to provide an economic tool to help the business case for wind power generation. The content and scope gradually formed and resulted in a sizable research project, exceeding what was originally intended. However, as the research progressed new insights were obtained, which ultimately led to the choice of expanding the scope and capitalizing on new insights, which ultimately proved valuable. Although the process was very challenging at times, the fact that it led to the research presented in this report leaves me with an overall very positive feeling.

Regarding the knowledge and skillset going in, I am very happy to have chosen the EPA programme, as it provided me not only with a strong analytical toolset, but also with a unique perspective and the confidence to tackle any task. Before starting the EPA programme, I would have never dared to tackle the problems I have during this project. I would especially like to thank Scott Cunningham for introducing concepts such as machine learning and Bayesian statistics in his course on data analytics and visualization, as without it I would have never had the confidence to apply the statistical methods I did during this project. Also, as my second supervisor his insights into probabilistics proved valuable for my research, while his interest with the societal aspects of the research helped shape my final analysis for the better. Also, I would like to thank Remco Verzijlbergh, as his knowledge and insight in the topic at hand, combined with his contagious enthusiasm and the confidence he placed in my ability to solve the problems I faced along the way proved decisive in the direction and success of my research. Especially the literature he suggested has provided me with valuable insights, which ultimately defined my research. Furthermore, the great atmosphere at his company Whiffle provided my with a lot of energy and enthusiasm, which at times inspired me in tackling some of the challenges I faced.

I would also like to thank Stijn van Langen, my first supervisor at my internship at Essent, for his interest, patience and knowledge of the real world considerations of those participating on the electricity market. The freedom I got at Essent to pursue my interests and ideas were valuable in the process of forming the final research project. Furthermore, his down to earth attitude to theoretical concepts proved often right and definitely fruitful in making decisions with regards to my research. Without these suggestions I would not have made the choices I made, which ultimately influenced my research for the better. Also, I would like to thank Taco Kuipers, my second supervisor at Essent, for many fruitful discussions, great table soccer matches, although less fruitful, and for multiple opportunities he set up for me to interact with other departments, to get their take on the matter. This extends to the colleagues of Taco and Stijn in the short-term forecasting team at Essent, who provided me with the opportunity to obtain insight in the operational aspects of a large energy company, a great working environment and who made me feel part of the team.

Finally, I would like to thank my roommates, friends and family for their tremendous support. Specifically I would like to thank my friend Aubin for his seemingly limitless patience in discussing complex statistical concepts, especially those involving Copula models. Furthermore, I am very grateful to my sister Hieke, who greatly helped me in organizing my research during the most challenging final phase of my research and my brother Volkert and his girlfriend Annelieke for their great hospitality and support.

Gerrit Deen
Delft, March 2019

Summary

One of the most significant challenges facing mankind today is that of the Grand Challenge of Climate Change. An important mitigation strategy in recent years has been the upscaling of zero and low carbon electricity generation, with an important role for wind power. A main driver behind the successful upscaling of wind power generation has been cost reduction. However, as wind power producers (WPP) are increasingly exposed to market conditions as support policies are starting to reduce, cost alone is not anymore the only factor driving its business case. Regarding the market value of wind power, a significant decrease is projected, which is mainly due to two effects. First, wind power is highly correlated within a single market, which results in decreased electricity prices when output is high, as it is often collectively high for all WPPs simultaneously. Second, as WPPs are highly correlated, they are often collectively wrong in their predictions, which lead to strong grid imbalances, for which they are increasingly required to pay themselves.

As these factors lead to a decreased value of wind power and this effect is predicted to further decrease as market penetration of wind power increases, this study chooses to focus on what a WPP can do to increase its market value. A review of the literature shows that the best way to do so is to alter the way WPPs offer their electricity on the electricity markets. Hence, the main research question of the research is:

To what extent can WPPs improve the market value of their wind power generation through improved offering strategies?

In order to answer this question a case study is performed on a 147.6MW Belgian offshore wind farm, for which advanced bidding strategies are constructed, aimed at increasing the market value of its generation. Several market strategies are optimized for the day-ahead market and the balancing market. The intraday market is not considered in this study due to its low liquidity. With the conventional market strategy as a reference, multiple market strategies are constructed using two stochastic optimization models and one analytic model. As the reference strategy for a WPP entails offering the wind power forecast on the day-ahead market, three different forecasts are analyzed, both as alternative reference strategies and as alternative inputs to more advanced strategies. The first forecast is the one currently operationally used, the second and third are provided by Whiffle, a company specialized in ultra-high resolution forecasts.

Several inputs are required for the models, which are forecasts for the wind farm, the day-ahead price, the imbalance price and the effect the WPP has on the balancing market (price-maker effect). As only wind power forecasts were available, this study introduces a method is introduced which uses a univariate kernel density model, a robust regression model and a seasonal auto-regressive moving average model to generate forecasts for all three price series. Although this study does not provide a comparison with other methods, when used as inputs to the offering strategies, these forecast are able to reliably increase value.

Regarding the stochastic optimization models, two models are considered. The first is a price-taker model, which assumes the WPP has no influence on the imbalance price. The second is a new formulation of a price-maker model found in literature, which assumes a linear influence on the imbalance price as a stochastic process. Stochastic optimization aims at finding an optimum for an uncertain problem, where all uncertain parameters are represented by a discrete set of scenarios. For this purpose this study introduces a modeling framework. This framework uses applies two modeling steps. First, a conditional multivariate kernel density model is used to quantify the uncertainty surrounding the forecasts of wind power, the day-ahead price, the imbalance price and the price-maker effect. These stochastic processes are then combined in a Gaussian Copula model, which models the dependency between these stochastic processes using the multivariate conditional kernel density models and a multivariate Gaussian distribution. This allows the modeling of the interactions between all of these processes stochastically, which results in a more realistic scenario set. Conventionally, scenario sets for stochastic optimization are constructed using scenario reduction, which assumes all processes are independent from each other. To be able to identify the value of the dependent modeling of the stochastic processes introduced in this framework, the Copula coupled scenario set is benchmarked against a conventionally reduced scenario set.

To be ale to judge the quality and reliability of the inputs of the modeling framework, the modeling itself and the outputs it produces, this study also introduces an evaluation framework. The purpose of this frame-

work is to empower WPPs to make informed decisions with regards to the modeling process, enabling higher quality decision making. The evaluation of the modeling process shows that the model is able to accurately and reliably capture the uncertainty surrounding all stochastic processes, although less so for the price-maker effect. For the analysis of the performance of the different strategies, two different analyses are carried out. The first assumes that a difference in offering strategy does not influence the imbalance price (price-taker assumption), while the second assumes it does (price-maker assumption). Section 8.2 shows that under the price-taker assumption the price-taker optimization model strongly outperforms the other models on expected value, albeit with much higher risk. However, section 2.3 concludes that a change in strategy does in fact influence the imbalance price, for which the mechanism is explained in section 3.4, while section 3.1 shows that the price-taker model results in extreme bidding strategies, which leads to large volumes being traded on the imbalance market. Section 8.3 shows that when the effect on the imbalance price is taken into account, the price-taker optimization model in fact leads to a strongly lowered revenue. This analysis shows that the real-world performance of the price-taker optimization model is strongly negative, both on the expected difference in revenue, as well as on risk. The price-maker models both take the effect of their strategy on the imbalance price into account, which is explained in sections 3.2 and 3.3. However, only the optimization model using the Copula coupled scenario set shows to be able to achieve an increase in revenue with regards to the reference strategy, where a large difference exists between the three forecasts, with one Whiffle forecast clearly outperforming the other two forecasts.

Section 8.4 quantifies the effect the different strategies have on the system imbalance. This analysis shows that all price-taker strategies lead to significant increases in average expected system imbalance, risk for large imbalances and imbalance volatility. All price-maker models positively impact the system imbalance in all three respects. Of the price-maker strategies, again the one using the Copula coupled scenario set outperforms the others. Section 8.5 quantifies the effect of the strategies on the opportunity cost of the system as a whole through trading on the balancing market. This shows that the price-maker strategy using the Copula coupled scenario set is likely to have a positive effect on the system as a whole. This effect is likely to be larger in reality, as a successful strategy results in a dampened imbalance price, which is not captured in this analysis.

In conclusion, this study found a possibility for a 1.17% increase in revenue, from choosing the strategy using the price-maker optimization model using the Copula coupled scenario set with a change in forecast. Based on the results, this study presents several recommendations. First, this study found a large difference compared to other studies found in literature. This is due to the fact that most studies in literature applied the price-taker analysis, whereas this study applied a price-maker analysis for its final conclusions. Although improvements can and should be made to the algorithm this study applies to recreate imbalance prices, not taking the price-maker effect into account significantly skews results, which can result in dramatic negative performance in reality both for the WPP as on a system level. Hence, this study strongly recommends to first find a method to reliably recreate prices before drawing conclusions. Second, this study found that only the price-maker optimization model using the Copula coupled scenario set was able to produce a strategy that outperformed the reference strategy in the price-maker analysis. The other strategies resulted in a varying degree of negative performance. As the processes are dependent, not respecting the dependencies results in an unrealistic scenario set, which results in an unrealistic optimization. Hence, it is strongly recommended to build on the dependency modeling introduced in the modeling framework, when applying the model operationally. The third recommendation regards the formulation of the price-maker optimization model. Section 3.4.1 showed that the sign of the quarter hourly average volume of balancing bids (NRV) activated to counter system imbalances plays a crucial role in determining the imbalance price. This in fact explains the switching behavior of the imbalance price, which is shown in section **??**. Incorporating this effect allows for the modeling of the imbalance price in two separate components as well as the endogenization of the switching behavior of the imbalance price in the optimization model through the NRV process. This increase in resolution of the uncertain processes driving the decision-making process seems a likely source of improvement for the optimization. Furthermore, the positive impact on the total system opportunity cost from the price-maker optimization strategy using the Copula coupled scenario set is explained by the fact that it offers too aggressively, as it causes a switch in the imbalance price when prices are high. This is due to the non-inclusion of the NRV signal in the optimization. Hence, it is strongly recommended to include these effects if the model is to be used operationally.

Contents

Glossary

FSP	Forecast Service Provider
DA	Data Assimilation
BRP	Balance Responsible Party
WPP	Wind Power Producer
PTU	Programme Time Unit
MIB	Marginal Incremental Bid
MIP	Marginal Incremental Price
MDB	Marginal Decremental Bid
MDP	Marginal Decremental Price
NRV	Net Regulation Volume
SI	System Imbalance
α	Additional incentive imbalance price
λ^+	Long imbalance price
λ^-	Short imbalance price
λ^D	Day-ahead price
λ^{Imb}	Imbalance price
λ^Δ	Imbalance difference price
γ^Δ	Price-maker effect
P^D	Power offered on the day-ahead market
P^Δ	Power offered on the balancing market
P	Observed power
$P^{\max}$	Rated power
t	Time of day
k	Forecast horizon
ω	Scenario index

d_t	Timestep
R	Revenue
π_ω	Scenario probability
CVaR_α	Conditional Value at Risk at level α
λ^{Δ^+}	Long opportunity price
λ^{Δ^-}	Short opportunity price
R^{Δ^+}	Long opportunity revenue
R^{Δ^-}	Short opportunity revenue
R^{Δ}	Total opportunity revenue
P^{Δ^+}	Long volume
P^{Δ^-}	Short volume
$\overline{\lambda}^{\Delta^+}$	Average long opportunity price
$\overline{\lambda}^{\Delta^+}$	Average short opportunity price
$\overline{\lambda}^{\Delta}$	Average opportunity price
σ	Volatility
R_{SI^+}	System long opportunity revenue
R_{SI^-}	System short opportunity revenue
R_{SI}	System opportunity revenue
$\epsilon_{t+k\|t}$	Forecast error for forecast issued at time t, for forecast horizon k.
y_{t+k}	Observation of process y
$\hat{y}_{t+k\|t}$	Forecast of process y, issued at time t, for forecast horizon k.
$K_h(\cdot)$	Kernel function with bandwidth h.
$\hat{q}^{(\alpha)}_{t+k\|t}$	Quantile forecast at nominal level α, issued at time t, for forecast horizon k.
$D_k(\cdot)$	Kantorovich Distance
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
NQS	Net Quantile Score

CRPS	Continuous Ranked Probability Score
E_s / P_s	Energy score / Price score
CKD	Conditional Kernel Density
$\Delta\|q\|$	Discrepancy Index
Δq_m	Deviation from perfect calibration of quantile m

1 Introduction

The Fifth Assessment Report of the Intergovernmental Panel on Climate Change states that it is "extremely likely that human influence has been the dominant cause of the observed warming since the mid-20th century (95-100% likely) [55]. One of the main components of this human influence is the emission of anthropogenic greenhouse gasses, which increased by 81% in the period between 1970 and 2010. Due to human influence on climate change, it is projected that the surface temperature will rise and it is deemed "very likely that heat waves will occur more often and last longer, and that extreme precipitation events will become more intense and frequent in many regions (90-100%). The ocean will continue to warm and acidify, and global mean sea level will rise" [55]. These developments are projected to increase risk for "people, assets, economies and ecosystems, including risks from heat stress, storms, extreme precipitation, inland and coastal flooding, landslides, air pollution, drought, water scarcity, sea level rise and storm surges (90-100% likely). Fortunately, mitigation strategies can reduce the likelihood of these effects happening. An important mitigation strategy is the decarbonization of the electricity and heat production sector, as it was responsible for 25% of all anthropogenic greenhouse gas emissions in 2010. This requires the upscaling of low- and zero-carbon electricity generation technologies. There is a long way to go, as the share of low-carbon electricity supply must increase from 30% in 2010 to 80% in 2050 in order to achieve a 50% likelihood of limiting global warming to 2°C in the majority of scenarios drawn up by the IPCC, which is the agreed upon limit which should help reduce the most severe risks [55]. Furthermore, the main driver behind the human influence is *cumulative* CO_2 emissions, which is due to the fact that CO_2 remains in the atmosphere for a very long time. Because of this fact short-term mitigation strategies have a higher weight in mitigation strategies, as delayed mitigation will require much more severe medium to long term mitigation strategies, which would lead to significantly higher mitigation costs in the medium to long term. In recent years one of the main technologies driving the short-term decarbonization of the electricity sector has been wind power. In terms of all electricity generated by renewables in 2017 wind energy was second only to hydro in 2017 [28]. However, as wind power only generated 4.4% of the world's electricity in 2017, while it is one of the main technologies for the short-term upscaling of low carbon electricity generation, there is still a long way to go [6]. Wind power is currently an important energy technology due to its technological maturity, low cost and relatively stable power output compared to solar. Although it has clear benefits as a method for electricity generation, it also has a downside, which is often referred to as 'integration cost' [25]. This consists of increased grid costs, increased costs to balance the grid and reduced utilization of the capital in thermal plants. Although these effects are not considered in this study, considerations with respect to these integration costs are discussed in the reflection.

To drive investment in wind energy, policies have provided various incentives. For instance, subsidies have been awarded to tenders or parties participating with wind power in electricity markets have been exempted from certain penalties, such as paying for grid imbalances [42]. However, recently wind farms have been increasingly participating in the electricity markets. This means they are exposed to certain risks, such as uncertain and volatile market prices and uncertainties in their output. Furthermore, support policies have been steadily reduced, in part motivated by high policy costs, but also due to cost reductions in wind technology, which reduces the need for such support policies [42]. Although this reduction in cost is projected to continue towards the future, a study by Hirth, Ueckerdt and Edenhofer emphasizes a need to not only focus on the cost perspective, but also on the value perspective [25]. This is proposed shift in perspective is not only

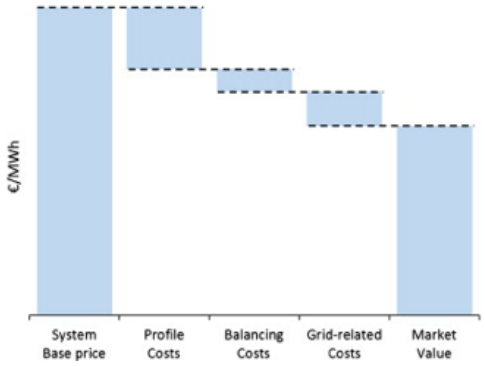


Figure 1.1: The system base price is depicted in the left bar and the market value of wind power is depicted in the right bar. The difference between them can be decomposed into 'Profile', 'Balancing, and 'Grid-related costs [25].

motivated by the increasing reliance on the market value to drive investment, but also due to a corresponding decrease in market value as the installed capacity increases. Hirth et al. present three factors that influence this value most [25], as visualized in figure 1.1: 1) Profile costs, 2) Balancing costs and 3) Grid-related costs. Firstly, profile costs consist of two elements, the 'merit-order effect' and the 'correlation factor'. Both have to do with the average electricity price wind power obtains. The main electricity price in most markets today is the day-ahead price, which for most European countries, including Belgium, consists of a centralized auction process, where all price-quantity offer bids are arranged in a increasing order to form the aggregate supply curve and all price-quantity demand bids are arranged in a decreasing order to form the aggregate demand curve. Where these two curves intersect, the market clearing price is determined. This process is explained more in-depth in section 2.2. Due to the near-zero marginal cost of wind power, its offers enter the supply curve on the left-hand side, which means that these bids (and other near-zero marginal bids) are accepted before higher priced bids, directly shifting the aggregate supply curve to the right and thus directly influencing the price as a result, which is visualized through figure 1.2 [24]. This 'merit order effect' means that when wind energy experiences high output, less expensive generators are activated by the market, leading to a decrease in the electricity price. Unfortunately, wind farms operating within a single market have highly correlated power outputs, which means that this 'renewable shift' in the price is likely to be high when a wind farm experiences high output, thus reducing the price it receives proportionally to its output. The second profile cost factor for wind energy is the 'correlation factor', which stands for the correlation between wind output and demand [24]. When this factor is high, wind output is high in times of high demand, thus realizing a relatively high price for its output. Unfortunately, this correlation is rather low, which further decreases the price for wind energy. This drop in market value due to 'Profile costs' is expected to increase as penetration of wind energy increases, as shown in figure 1.3 [24]. Hirth et al. found this value drop in Germany from 2001 to 2013 to be 17% [24]. Unfortunately, as the output for a wind power producer is uncontrollable and storage is not yet a viable option, 'Profile costs' can currently not be reduced by wind farms.

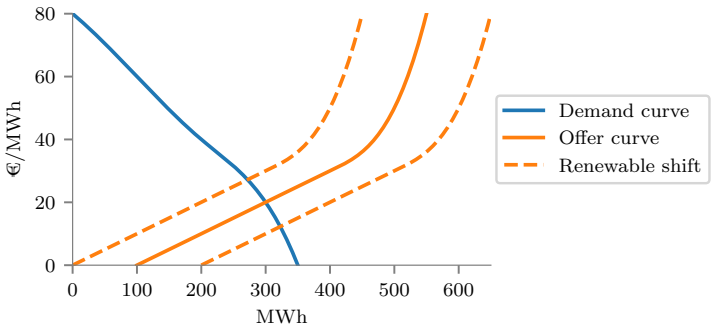


Figure 1.2: The impact of the near-zero marginal cost of renewables on the merit-order. The dotted line indicates the shift caused by renewable offers entering the supply curve on the left-hand side, reducing the price as a result.

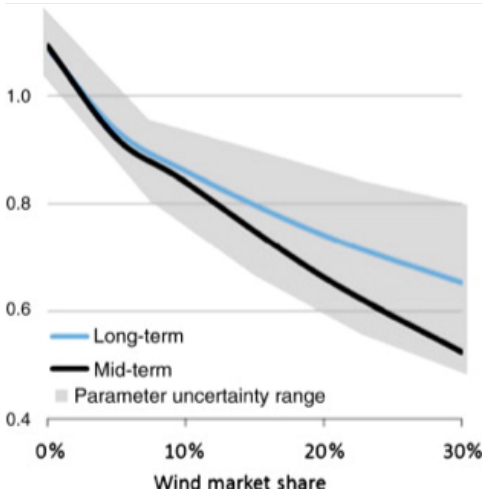


Figure 1.3: Development of the Value Factor as a function of wind market share, where the Value Factor (dimensionless) on the y-axis is defined as the Ratio of average wind revenue in € MWh versus average market revenue in € Mwh. The long-term trend shows the effect of added flexibility in the electricity market, e.g. electricity storage.

The second factor contributing to the decrease in market value for wind energy is 'Balancing costs'. These are costs associated with deviations between what a wind power producer (WPP) offers on the day-ahead market and what it generates in real-time. As mentioned above, due to their near-zero marginal cost WPPs offer their generation at a zero price to the day-ahead market. The party responsible for the wind farms power balance (Balance Responsible Party (BRP)) then has to inform the Transmission System Operator (TSO), which is responsible for balancing the grid, on its scheduled generation. In the Belgian system, the day-ahead market closes at 12:00AM, which means that the wind farm must at a minimum forecast its generation between 12 and 36 hours ahead. The quantity the WPP bids depends on what a farm expects to generate the next day. As wind power output is a direct function of wind speed and wind speed is inherently uncertain, this expected output cannot be known perfectly, which means that the expected output nearly never equals the actual output. The deviation between this forecast and the observed output is called the WPP's 'imbalance'. Depending on whether the overall system imbalance of the electricity grid as a whole was hurt or helped by the WPP's imbalance, the WPP must buy back its shortage or sell its surplus at the 'imbalance price', which can be favorable or unfavorable for the WPP. This price is a result of the TSO activating balancing power bids in order to maintain a balanced grid. This is important as imbalances can lead to component failures, which results in a decreased security of supply, which can have severe economic consequences.

Unfortunately, this price is on average unfavorable, which makes these imbalances constitute as a cost for the wind farm, further decreasing its overall market value. The severity of these costs was analyzed in a study by Mazzi Pinson from 2017 [39]. They studied the impact of uncertainty on market revenue for a 21MW wind farm located in Western Denmark for 2014, which found that when using a typical point forecast the cost due to imbalance was 5% of total revenue. As WPPs currently obtain very slim profit margins if any, this is considerable. This effect is explained more in-depth in section 2.2. The third effect, 'Grid-related costs', relates to transmission constraints, which in some markets lead to locational prices: As wind energy is often generated at locations relatively far from where the demand is located and the grid doesn't always allow certain flows of generated electricity, some markets apply prices that differ between areas. However, as this is not the case in the Belgian market, these effects are not taken into account.

In conclusion, in the Belgian case, wind power's market value is decreased mainly by the first two factors, 'Profile costs' and 'Balancing costs'. As wind farms cannot currently influence 'Profile costs', the main problem in the context of decreasing market value consists of 'Balancing cost'. The next section focuses on ways WPPs can influence this problem.

1.1. Research Problem

The following section provides an overview of research into the problem, after which it concludes by formulating the knowledge gaps and the main research question. The previous section concluded that WPPs can increase the value of wind energy by decreasing balancing costs. It also explained how forecast errors cause these imbalances. A possible solution could be to improve these forecasts and therefore reduce their errors. In a review on the state of the art in wind power forecasting Giebel and Kariniotakis mention that forecasts can also be used in more advanced ways which can enhance value [16]. Hence, two options for increasing value are considered here, firstly through the improvement of forecasts and secondly through the improved utilization of these forecasts. The next two sections discuss the literate on both.

1.1.1. On improving forecasts

This first part discusses the state of the art of wind speed and power forecasting and where possibilities for improvement lie, based on the review by Giebel and Kariniotakis [16]. Forecasts are playing an increasingly important role in today's electricity markets and grids. Especially since in most European countries, the main market is cleared around noon. This means that accurate forecasts crucial for the efficient functioning of the electricity market. Furthermore, forecasts are becoming a central tool to keep the grid balanced, for which the TSO is responsible. To maintain grid balance, the TSO requires market parties to communicate a balanced schedule of their supply and demand typically sometime shortly after closure of the day-ahead market, depending on the specific market. It then bases its balancing strategy on the schedules communicated to it by the different market parties. As these schedules contain both uncertain demand and uncertain supply (e.g. wind or solar power), the actual supply and demand often deviates from these schedules, which means that the grid is often imbalanced. The quality of the forecasts for the uncertain parts of the schedule thus determines the magnitude of these imbalances. The TSO continuously monitors the grid for these imbalances and activates or deactivates generators or flexible demand to maintain balance. Currently, balancing actions are mostly executed by conventional generators, which need to ramp their generation up or down. Unfortunately, such ramping causes the lifetime and efficiency of generators to go down to varying degrees depending on the ramp rate, frequency of ramping and the kind of generator. For the market participants forecasts are also of importance, as they dictate the amount to be bought or sold to ensure a balanced schedule. In the case of imbalances with regards to its schedule, the market participant is forced to participate on the balancing market, on which it on average incurs unfavorable prices. Hence, the quality of its forecast has a direct impact on the performance of its market strategy.

For wind energy, there is a large variety of forecasts being used. The first distinction to be made between them is whether or not a Numerical Weather Prediction (NWP) model is involved. This distinction leads to the first class of forecasts, one that relies solely on time series approaches using statistical models. This type of forecast generally performs well up to a 6 hour horizon (Giebel Kariniotakis, 2017). It relies on historical data, with the main input being Supervisory Control and Data Acquisition (SCADA) data, which consists of measurement data from the generating unit(s), either meteorological or power. As this approach is outperformed strongly after a look-ahead time of 3h, the remainder of this section focuses on forecasts involving NWP's.

The second class is the one that incorporates data from NWPs. The NWP models leverage the equations of motion, known as the Navier-Stokes equations and integrate them forward in time [16]. An important part of this process is the initialization of the equations of the model with the best estimate of the initial atmospheric state. For the determination of these initial conditions Data Assimilation (DA) is applied. This is a process in which observations distributed in time and space are merged together with a dynamical numerical model in order to determine the initial state of the atmosphere as accurately as possible, captured in the initial conditions. NWP's perform well up to a forecast horizon of about 2 weeks, with gradually degrading performance as the forecast horizon increases. An important limiting factor was found by Lorentz: He found that the equations are extremely sensitive to differences in initial conditions, which can cause large deviations in the conditions of the flow at a later time [37]. Part of the loss in forecast performance is hence due to the fact that we cannot perfectly know the conditions at all places at the initial time. This is why currently important improvements in the performance of these models are due to improvements in measurements and the DA process [22]. Due to the highly specialized nature and the scale of the task of data collection and DA, this is not considered within the scope of the WPP.

There is a large variation between the methods employed by the different companies offering forecasts (Forecast Service Providers (FSP)). Generally speaking, operational forecasts provided by FSPs employ a combination of physics and statistics as both are needed for valuable forecasts [16]. Also, there are different chains

in the forecast industry. The first one consists of the global scale NWPs, like the one by the European Centre for Medium-Range Weather Forecasts (ECMWF). A global scale models start by determining their initial conditions from many different types of measures, the Data Assimilation (DA) step. Then the model is run, producing output at a given temporal resolution at certain (low) grid resolution.

The output from these models is then used in several different ways [16]. There are different meso-scale or micro-scale NWPs which use the output from larger scale NWPs to determine the initial and boundary conditions of their own models, where the first refers to the initial state of the atmosphere, while the second refers to the state of the edges of the domain of the smaller-scale model at later times. These NWPs can add value by simulating at a higher grid resolution, at a higher temporal resolution and/or by incorporating their own DA process, with the aim of improving initial conditions. Generally speaking, such a model simulates what happens within the coarser grid of the larger scale NWP. There are many different lower level NWPs, each with different performance in different areas.

Another way forecasts from NWP models can be improved is through post-processing [16]. A first part of this process is the removal of bias. Models have multiple inherent biases and once defined these can be corrected using statistical or machine learning approaches. The standard mean of improvement in this aspect is through model output statistics (MOS). This is generally a statistical regression model aimed at removing biases [17]. Other methods are also used to provide corrections to models, such as Artificial Neural Networks (ANN), autoregressive statistical (AR) models and others. Also used is ensemble MOS, which is used to both correct the individual models and to optimize weights for the blending of forecasts from multiple models. Postprocessing can be applied at any stage of the forecasting chain. While both post-processing and ensemble MOS is something a WPP could do, the purchasing of multiple forecasts can lead to high costs for a WPP, while many FSPs already employ the ensemble method, local conditions and other specialized methods to improve their forecasts, as it is their core business to increase the performance of their forecasts.

Any of these NWPs can be used as input to additional statistical models, where the aim is to find relationships between multiple explanatory variables, for instance between NWPs and (online) Supervisory Control And Data Acquisition (SCADA) data, which consists of measurements of observed meteorological conditions and / or power generation [16]. These models can have the form of explicit grey-box statistical models using advanced autoregressive statistical methods or black-box models such as ANN. The main value of these models lies in the use of historical SCADA data to calibrate themselves, which can be advantageous for the resulting forecast performance. Also, if online SCADA data is available, a self-calibrating recursive model can be advantageous, as it can take advantage of constant updates. This is something that can be within the scope of WPPs, as some are not willing to share SCADA data with FSPs.

Concerning the conversion of wind speed to power, often similar statistical methods are used, which can also take advantage of (online) SCADA power data. Such methods should capture the non-linearities in the relationship between weather data and power data, some due to the shape of the power curve, some due to the characteristics of the terrain or the lay-out of the wind farm itself [16]. Also, it is suggested that NWP models may be able to model power conversion better by making use of a higher temporal resolution, which can impact power value because of the non-linearity of the relationship between wind speed and power output. This can be achieved through statistical or artificial intelligence methods or by incorporating a physical representation of a wind farm directly inside an NWP model.

Regarding the output of the different methods, there are multiple possibilities, which are especially relevant when forecasts are used in more advanced applications. Firstly, the forecasts can be provided as either power forecasts or wind speed forecasts. Secondly, some FSPs provide forecasts with a single value for each time step, referred to as point forecasts, some provide forecasts accompanied by a certain confidence interval, referred to as interval forecasts, while others provide an approximation of the complete probability density at each time step, referred to as density forecasts [42]. Furthermore, all NWP models can produce approximations of the uncertainty of the forecast by generating multiple disturbed simulations using different values for uncertain input parameters / initial conditions, which some do provide in the form of scenarios. However, in practice the distributions of these scenarios currently tend to be relatively narrow, when compared to statistical methods of distribution density estimation, thus they tend to underestimate uncertainty [16].

Concerning the performance of different forecasts several factors need to be taken into account. Firstly, while forecasts can be directly used in their point forecast form, when using more advanced bidding strategies, the value of such forecasts may change a lot. Furthermore, more advanced higher resolution forecasts often perform worse on conventional tests, like the Root Mean Square Error (RMSE) or Mean Average Error (MAE), while they perform better on skill scores when used in a probabilistic context, which are more advanced metrics [42]. It might be that those that show higher skill may perform better when used in more

advanced bidding strategies, but worse when used as point forecasts. Also, some FSPs provide density forecasts directly, which may be interesting, as multiple methods exist for determining forecasts' uncertainty, implying that higher quality density forecasts may be available [16]. However, most forecasts are provided in point form, which means that in order to judge their performance in more advanced use cases, they would first need to be transformed.

Concerning the role of the WPP in improving forecasts, several possibilities for improvement are identified. Firstly, SCADA (online) data can be shared with FSPs, providing them the opportunity to provide a better calibrated forecast. Secondly, WPPs can improve their own statistical postprocessing using SCADA data. Thirdly, the modeling of power conversion can be improved, either by improving the WPPs power conversion models or by providing FSPs the opportunity to directly forecast power by sharing power data. Lastly, when used in more advanced applications, some FSPs provide some information on the uncertainty surrounding the forecast. However, it is currently still common to provide only a point forecast, while the scenarios mentioned before are not (currently) an accurate description of the underlying uncertainty. Although BRPs have multiple possibilities to improve forecasts, Giebel and Kariniotakis (2017) conclude in their review that the main error in short term forecasting models stems from the NWP models. Hence, this research limits the scope of improving forecasts to the higly specialized work of FSPs. As many FSPs exist, each producing different performing forecasts, which can rank differently whether used as point forecasts or as more advanced transformed forecasts such as interval or density forecasts, it seems more relevant for a WPP to best judge the performance of each in their intended use case.

1.1.2. On Improved Forecast Utilization

This section discusses the literate on how WPPs can improve the usage of forecasts with the aim to improve the value of their wind energy portfolio. As mentioned in the previous section, some forecasts may rank differently when judged on conventional metrics, compared to metrics relating to more advanced use cases. Within the scope of this research the best use case of forecasts is the one that leads to the highest market value. The methods with this specific aim are referred to as advanced bidding strategies.

The most straightforward bidding strategy, which is considered the reference strategy, is where forecasts are considered deterministic, thus the forecast power is offered directly on the day-ahead market at a zero price [10]. Fortunately, more advanced strategies exist which show a strong potential in value improvement over the reference strategy. The first study to look at improving bidding for wind power in a more advanced probabilistic framework is that of Matevosyan and Söder from 2006 [38]. In this study the day-ahead market and the imbalance market of the Nordic system are considered to minimize imbalance costs. They constructed scenarios for wind power and imbalance price scenarios as input to a stochastic optimization model. This study modeled the inputs to the optimization using an autoregressive moving average (ARMA) model, where the modeled processes were assumed to be independent from each other. They simulated this strategy for January 2003 for a Danish wind park and they found it led to an increase in total revenue of 5.2%.

A later study that dealt with imbalance costs in a probabilistic framework is that of Pinson, Chevallier and Kariniotakis [47]. This study presents a method to use stochastic information in the form of quantiles from probabilistic forecasts to decrease imbalance costs, where a loss function is specified that includes the cumulative distribution function of the expected power generation and an estimate of imbalance prices in its terms, which were chosen to either be a yearly average (strategy 1) or a quarterly average (strategy 2). Simulation of the method for a year for the Dutch day-ahead and imbalance market showed a decrease in average imbalance cost from €9.13/MWh for the reference strategy, to €6.74/MWh for strategy 1 and €4.04/MWh for strategy 2. The main conclusion of this paper was that stochastic information can be a powerful tool towards combating imbalance costs. In their discussion they mention that the strong performance boost by simply taking quarterly differences of the imbalance price into account should be seen as an indicator that increasing the quality of the inputs strongly improves the outcome of such an optimization.

A study that mostly built on the methods introduced by Matevosyan and Söder is that of Morales, Conejo and Pérez-Ruiz from 2010[43]. Here improvements were made to the underlying mathematical principles of the stochastic optimization model and the options are added to leverage the intra-day market as well as control risk using the Conditional Value-at-Risk at confidence level α (CVaR_α) [43]. Furthermore, using this model in different setups, incorporating a single intra-day market or not, as well as other variations are analyzed in a comparative study for the months February, April, July and November of 2008 for the Spanish electricity market [51]. This study also modeled the inputs to the optimization using an ARMA model, where the modeled processes were assumed to be independent from each other. The day-ahead market in this study closed at 12:00AM and the intra-day market closed at 11:00PM, one hour before the next day. This study firstly

found that all advanced strategies outperformed the reference strategy. Secondly, it found that offering the difference between the forecast used for the day-ahead market and an updated forecast used for the intra-day market provided a 0.7% increase in total revenue. Thirdly, a strategy where an optimized bid for the day-ahead market was applied as a result of the optimization without risk aversion provided a 2.0% increase in revenue. Fourth, incorporating the possbility of intra-day trading in the optimization model without risk aversion led to a 2.7% increase in value. Finally, the paper showed that for all but one month the strategy of optimized bidding without risk aversion on both the day-ahead and the intra-day market yielded the highest result. The month where this was not the case, a strategy that incorporated risk aversion through the CVaR was still superior over the reference strategy. The authors state that this month had much more volatile imbalance prices, which led to the increase in value from risk aversion.

A similar study was carried out by Chaves Avila, Hakvoort and Ramos [9] for the Dutch market, where they simulated risk neutral strategies using the model from Morales et al. (2010) [43] for December 2010 for the Dutch day-ahead market, an hourly intra-day market and the Dutch imbalance market. This study [9] also modeled the inputs to the optimization using an ARMA model, where the modeled processes were assumed to be independent from each other. This study firstly found an increase in revenue of 4.5% when optimizing bids for the day-ahead market without considering the intra-day market. Secondly, when optimizing bids for the day-ahead market and adjusting the bids using updated forecasts for the intra-day market revenue was increased by 15% with regards to the reference strategy. Lastly, it found for the strategy where both the day-ahead and the intra-day market are incorporated in the optimization revenue rose by 21%. Chaves Avila et al. conclude their study by stating that the intra-day market is not very liquid and therefore the revenue increase from participating in it will in practice be lower. Also, this study concludes by stating that improvements in the quality of input to the model is are expected to further increase the value of the strategies, while mentioning that liquidity in the intra-day market is expected to remain a problem for future applications.

Another study that used the model developed by Morales et al. (2010) to study the effects of optimizing bids for yet another market is that by Bertrand and Papavasiliou (2017) [4]. This study also modeled the inputs to the optimization using an ARMA model, where the modeled processes were assumed to be independent from each other. This study researched the effect of optimized bidding when used in the Belgian day-ahead and imbalance market. It does not incorporate the intra-day market in its optimization due to low liquidity. The optimization model by Morales et al. is expanded through the incorporation of a price-maker effect for the imbalance market, for which Bertrand and Papavasiliou [4] claim an analytic solution. Where all previously mentioned studies assume that the wind farm has no effect on the price in both the day-ahead and imbalance market, this study emphasizes the need to model this effect for the imbalance market as larger wind farms can in fact have a strong influence on the imbalance price. It also states that evaluating the value of an optimized bidding strategy without considering the effect the strategy would have had on the market price is invalid, as the price can be strongly influenced by the choice of strategy, especially when the price-maker effect is not incorporated in the optimization. By recreating the imbalance price the study found the difference between a price-taker and price-maker model to be more than five-fold, meaning that a price-taker strategy according to their study would earn less than 20% of the revenue from the price-maker strategy. The price-maker version of their model was simulated for the period between October 2013 and December of 2015 and the revenue was found to increase by 41% compared to the reference strategy.

In conclusion, first multiple studies found significant increases in revenue through the application of advanced bidding strategies. Secondly, several conclude that it is important for the optimization which generates the optimized forecast to use high-quality inputs, which is indicated to be able to improve the performance of the strategy. Thirdly, the incorporation of an intra-day market depends on the liquidity of the specific market in question. Fourth, for some markets it may be important to implement a price-maker version of the optimization model. This was emphasized in the study by Bertrand and Papavasiliou (2017) [4], but as the similarity between the Dutch mixed-pricing imbalance market studied by Chaves Avila et al. (2013) and the Belgian single-pricing imbalance market is strong, the conclusion by [4] implies that results may be different if the imbalance price were also recreated for the analysis in that study. This raises a question whether or not this could also be the case for other studies, as all studies considered here use the observed prices to perform their value analysis, while in fact these prices would in practice possibly be influenced, depending on the size of the wind farm. Finally, all of these studies have in common that the methods used to model the inputs use highly similar models of the ARMA class, which is a relatively simplistic method for modeling stochastic processes. Furthermore, these studies assume independence between the processes, which seems improbable as they are coupled through the electricity market.

1.1.3. Research Gaps

From section 1.1.1 it can be concluded that most of the improvement in forecasts lies with FSPs, as most improvements comes from improvements in the estimation of initial conditions or with the functioning in the NWP model itself. As this is a highly specialized task, this is left out of scope. However, there are some possibilities for improving forecasts that do lie with FSPs. Most notably allowing the FSPs to better calibrate their models using (online) SCADA data, allowing FSPs the opportunity to model power conversion or by WPPs improving calibration using (online) SCADA data or by improving power conversion modeling using (online) SCADA data. However, as any improvements made by an WPP seem to also lie within the scope of FSPs, it is left out of scope in this study.

From section 1.1.2 the main conclusion is that there seems to be significant potential for value improvements through more advanced bidding strategies, which better utilize wind power forecasts. However, choosing the correct model for a specific market seems important as it can lead to significant differences in revenue, as shown in the study by Bertrand and Papavasiliou (2017) [4]. Furthermore, depending on the size of the wind farm in question, the value obtained through a changed strategy cannot directly be determined by applying observed prices, as the change in strategy can impact the price. Hence, not only needs the model be chosen correctly, the evaluation of the optimization results should ideally incorporate the changed prices.

In section 1.1.1 it is mentioned that some forecasts may differ in rank on conventional metrics used to evaluate them in point forecast form, compared to when they are used in a probabilistic context, like in stochastic optimization. How to choose the best one as input to the strategy that provides most value therefore becomes increasingly difficult. For this, the BRP should have insight on how each forecast compares not only on performance in different forms, but also on how each compares in value when applied in different use cases. This is also brought up in 1.1.2 where the 2007 paper by Pinson et al., as well as the 2013 study by Chaves Avila et al. explicitly mention the importance of using high quality inputs to an optimization with the purpose of improving the result. Hence, being able to rank these inputs is an important task for the WPP. Also mentioned in section 1.1.1, some FSPs provide information regarding the uncertainty surrounding their forecasts. However, not all FSPs provide such information, while some FSPs provide information that is not reliable. Hence, even if an FSP can provide such input, being able to judge the reliability and quality of the stochastic information is a task that the WPP needs to be able to perform. Also, as not all FSPs provide information that can be used as input to stochastic optimization, the choice of FSP could be limited, while some forecasts may provide higher quality probabilistic information when transformed to the appropriate form. Hence, it seems like an important skill for WPPs to have. Since the importance of the quality of the inputs is emphasized, transforming these forecasts should be done intelligently with a clear path for improvement, for which clear evaluation guidelines are needed. While the studies mentioned do provide a method to construct the inputs for the advanced offering models, the method used is rather simplistic for modeling stochastic processes, while the studies also assume the processes to be independent, which seems unlikely. Improving the modeling of these processes seems like a likely source of improvement of the result, where a method ideally should accurately and reliably capture the uncertainty surrounding these processes as well as their interdependence.

In conclusion, several concepts seems important for WPPs in order to maximize the value of their wind power generation. First, WPPs need to be able to choose and apply the right model for their market. Second, WPPs need to be able to judge the quality of forecasts from WPPs for their purpose. Third, WPPs need to be able to transform these forecasts so they can be used in more advanced bidding strategies. Last, WPPs need to be able to reliably judge each strategy's value.

The first part of the introduction and the first paragraph of section 1.1.1 mention that the effects of renewable energy on the electricity grid as a whole can be significant. Also, as the Belgian study showed [4], the influence a wind farm has on prices can, depending on the specific strategy and size of the wind farm, be significant. As these prices are changed for the market as a whole, this could mean that while the WPP wins individually, the market as a whole loses. Hence, it seems important to identify what the total imbalance costs will be for the market under strategic bidding. Also, as the WPP bids strategically by indirectly altering its imbalances, the system imbalance is also influenced. As one of the major concerns surrounding renewable energy is the effect it will have on the balance and subsequent reliability of the electricity grid, it is important to not only evaluate what the effects of a value improvement strategy are for for WPPs, but also what the effects will be on the total grid imbalance. Finally, as market design differs between countries, specifics of the market may have important implications with regards to the choice of strategy and its performance. Hence, it is important to provide insight in the relationship between market specifics and strategic offering by WPPs.

1.1.4. Main research question

Section 2.2 concludes in the formulation of four needs. Each of these needs helps address the WPPs need to improve the value of their wind power generation. Hence the main research question is formulated as follows:

To what extent can WPPs improve the market value of their wind power generation through improved offering strategies?

1.2. Research approach and sub research questions

To answer the main research question a case study approach is applied. As explained, the research gaps relate to the value that can be obtained both from different forecasts and bidding strategies that leverage the information from these forecasts. These differences will be evaluated in an analysis on real data of a wind farm currently operating in the Belgian market. This wind farm consists of 24 Senvion 6.15 MW wind turbines, as part of an offshore farm on the Thorntonbank in Belgium (see cover), which has a Power Purchase Agreement (PPA) with Essent N.V. and for which Essent N.V. is the Balance Responsible Party. Also, three different wind power forecasts are provided, one provided by Essent and two provided by Whiffle, a company specialized in ultra-high resolution forecasts.

1.2.1. Sub questions

In order to answer the main research question, the following sub-questions are formulated:

- *What model is best suited for constructing bidding strategies for the Belgian case?* The Belgian market has specific characteristics which require a specific choice of model. The choice of model also dictates the inputs that are required to determine the optimal strategy.
- *What methods can be applied to provide price forecasts?* The models applied in this study require specific forecasts for electricity prices. As these are not provided a modeling method is needed to construct these.
- *What methods can be applied to transform point forecasts to the form required for stochastic optimization?* The optimization model requires specific inputs which need to be provided in specific form. For this purpose a method for transformation is needed.
- *What measures are best able to judge forecasts on performance?* To both guide the choice of FSP for WPPs, as well as guide the choices for choosing methods to transform point forecasts to their probabilistic form, insight is needed in both the performance and the reliability of forecasts in all forms.
- *How can the different strategies be ranked on value?* As pointed out by the study by Bertrand and Pasiviliou (2017) [4], judging strategies by assuming prices do not change seems unreasonable in the Belgian context, especially since the size of the wind farm in this case study is even larger. Hence, a method is needed to both accurately recreate prices as well as on how the resulting revenues can be ranked.
- *What are the effects of strategic bidding on the imbalance costs of the market as a whole and the total system imbalance?* Using the same methods to rank strategies on value, the differences in effect they have on the total market imbalance cost and system imbalance can be ranked. This should provide insight in the societal value of such strategies.
- *What are the effects of market specifics on the composition and performance of bidding strategies?* As the choice of model is dictated by the specifics of the market, insight should be obtained on how these specifics influence the bidding strategy. This should provide insight in design considerations for the markets, as well as insights for market participants on how to formulate their strategy.

1.3. Thesis outline

Before these questions can be answered, several steps need to be carried out. First, as the BRP operates in several markets on which it is dependent for its revenue, insight needs to be obtained on these markets and the system they are a part of, specifically on the mechanisms involved and the actors involved. This is discussed in chapter 2.

Second, insight is needed on the mechanism through which wind power forecasts lead to value, how BRPs use them and how they can be used. This includes the before mentioned advanced bidding strategies, how

these work and which inputs are needed. Also, a method that incorporates the influence these strategies have on imbalance prices need to be taken into account. This is discussed in chapter 3.

Thirdly, a framework is needed for the construction of the inputs for such advanced bidding strategies. This should provide BRPs with a structured methodology, while providing insight on how and where opportunities for future improvements lie. This can be found in chapter 4.

Fourth, a framework is needed to evaluate the construction of said inputs. This includes the evaluation of the point forecasts, which form the basis for the aforementioned transformation. It should also include clear metrics to evaluate the quality and reliability of the transformations involved in constructing the inputs for the more advanced bidding strategies, thus providing a clear structured methodology, which can guide BRP's towards reaching higher quality inputs. This is presented in chapter 5.

Fifth, as price forecasts are needed as a basis for the more advanced inputs to advanced bidding strategies, a methodology is needed to construct these. This is explained in chapter 6.

Chapter 7 presents the results and evaluation of the modeling of the inputs for the advanced bidding strategies. Chapter 8 presents the results of the simulation of advanced bidding strategies, both on the value they create for the WPP as well as on the effects they have on the market and system as a whole. Chapter 9 presents conclusions to the research questions, discusses these and provides recommendations for future research. Finally, chapter 10 presents reflections on the research, including societal aspects and implications for the electricity market, its actors and the way the imbalance market currently functions.

1.4. Research Relevance and Goals

As mentioned in the first part of the introduction, the premise of this research is the need for short term up-scaling of low-carbon electricity generation. As the increase in value could potentially improve the short-term profitability of investment in wind farms and help dampen the medium to long-term value decrease, it aims to be a contribution to the Grand Challenge of Climate Change. Furthermore, the analysis of the effects on the electricity grid and the electricity market as a whole in considering strategies aims to provide insight in the societal value of such strategies, while providing insights for potential market improvements. Also, as the construction of the inputs for the advanced bidding strategy means insight is obtained in the future uncertainty of both the generation of the wind farm as well as the electricity prices involved, decision-makers in the electricity market as well as those responsible for balancing the grid are provided with tools to increase the quality of decision-making. Furthermore, these tools are not limited to the scope of advanced bidding strategies for wind power, as these can be applied to an entire portfolio, including wind, wave, solar, electricity demand and conventional power generation. Furthermore, it both provides private actors the tools to improve the value of their portfolio, which is of societal importance, while providing public actors insight in how their markets might be improved to help achieve societal goals.

Finally, the main goals of this research are to provide WPPs a clear framework they can use to model processes related to those discussed in this research, to provide a clear framework on how to improve any step of the modelling process and to use these to determine advanced bidding strategies. With the aim to empower WPPs to utilize these more advanced techniques. Furthermore, the final stochastic optimization can be operationally used and it can easily be extended to incorporate other generation resources. Lastly, the value analysis should give the BRP confidence that the choices it makes anywhere in the modeling process provides real-world value, while providing the tools to reliably rank and evaluate them.

2

System Overview

2.1. Electricity Market

Formerly EU electricity markets were vertically integrated monopolies, where a single party was responsible generating, transmitting, distributing and supplying electricity. Gradually these vertically integrated companies were unbundled through three pieces of EU legislation in 1996, 2003 and the latest in 2009 [34]. The vertical unbundling led to transmission and distribution becoming regulated natural monopolies, while generators and consumers of electricity compete in a liberalized electricity market. In this market, generators compete to sell electricity to suppliers and large industrial consumers, while suppliers compete in the retail market to sell electricity to consumers. In figure 2.1 an overview is given of the various markets and the actors participating in them.

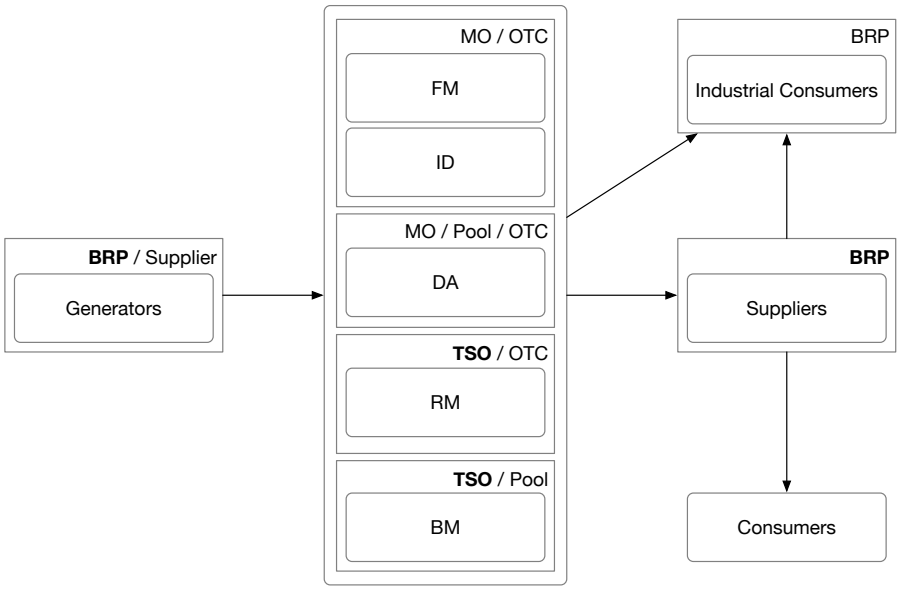


Figure 2.1: The flow of electricity through the electricity market is depicted by the arrows. The upper right corner of each square field indicates what the properties of a market or market party are, where those in bold always apply, while those in regular font *can* apply. BRP: Balance Responsible Party. MO: Market Organizer. TSO: Transmission Systems Operator. OTC: Over The Counter. FM: Futures and forward Market. ID: Intra-Day market. DA: Day-Ahead market. RM: Reserves -and regulation Market. BM: Balancing Market.

The next sections discuss the content of figure 2.1. Firstly, the different markets are discussed. Secondly, the different actors in the market are discussed. Although this is done with regards to the Belgian market, the general concepts apply to most EU markets.

2.1.1. Markets

Firstly, a timeline of the electricity market in Belgium is shown in figure 2.2. It shows that some markets start long before delivery starts, while some end after delivery. Below, these markets are discussed in sequential order.

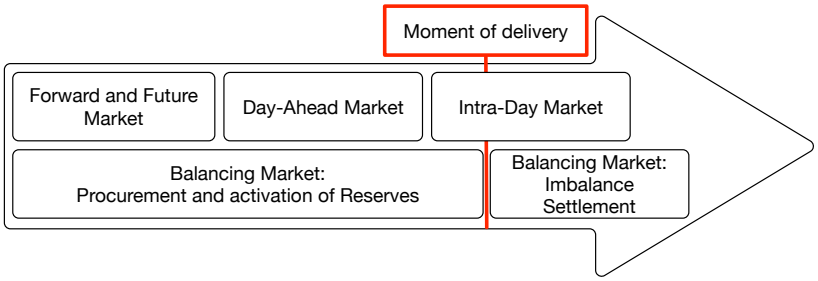


Figure 2.2: The diagram shows when markets start trade relative to the moment of delivery. Note that the Day-Ahead market closes before delivery starts, the Intra-Day market remains open even when delivery has started and closes 2PM after the day of delivery to provide market players the opportunity to close financial transactions. Reserves are procured long before delivery up to the moment of delivery. Activation occurs at the moment of delivery. Imbalances are settled after delivery.

- The first type of market is the futures and forward market (FM). This type of market runs from years before delivery up to the day of delivery. The products traded on these markets consist of contracts to deliver or consume a certain amount of electricity at an agreed upon price at a certain time in the future. Futures are standardized contracts which can be traded on power exchanges, which are managed by market organizers (MO) and trades are made through organized over-the-counter (OTC) trading. The main power exchange for such contracts in the Belgian market are the ICE Endex and the European Energy Exchange (EEX). Forward contracts on the other hand are not standardized and can be freely defined by those involved and these are mainly traded bilaterally OTC without the involvement of a MO. The main motivation for market parties to trade on forward and future markets is to ensure future sales or to reduce their exposure to future price decreases or increases. Such risk-reducing behavior is referred to as hedging. Such markets are not very relevant for WPPs, as the product on this market mainly consist of weekly, monthly, quarterly or yearly base or peak contracts. Base in this context indicates all hours of the time span of the product, while peak indicates peak hours, typically from 8AM to 7PM, also for the time span of the product [10]. As a WPP cannot guarantee its generation day-ahead it certainly cannot for longer time spans, making this market irrelevant for this study.

- The second type of market is the day-ahead market (DA). Market participants interact with this market the day before delivery. This market is of vital importance to the electricity grid, as the market zone has to be balanced after the DA closes at 12:00PM, where the market zone refers to the Belgian market as a whole. This balance requirement means that the sum of scheduled and forecast generation has to equal the forecast demand in the market zone. Electricity on the DA can be traded bilaterally OTC or on the power exchange, which is a pool-based electricity market and is managed by a MO. This means that buyers and sellers submit bids to the market, after which the MO orders all bids to determine a single price for each hour of the next day, which is explained in section 2.2. In Belgium the MO is the Belpex Day-Ahead market (BAM). The prices published by this power exchange serve as an important reference price for all other markets. This market is especially relevant for wind power producers, as it is highly liquid and forecasts for the day-ahead horizon are generally quite reliable.

- The third type of market is the intra-day market (ID). This is a market where electricity is traded continuously after the DA has closed. This allows market participants to correct for changes in their day-ahead schedule, for instance due to updated forecasts for wind power or unexpected outages of power plants. In Belgium the main ID is the Belpex Continuous intra-day Market (CIM). This market is an organized OTC market and is cleared continuously. In theory the ID would be of particular interest for wind power producers, as it would allow them to use updated forecasts to adjust their market position. However, this market for the Belgian context is not very liquid, which results in it not clearing for multiple hours in a day [4, 34].

- The fourth type of market is the regulation market (RM). This market handles the procurement side of the mechanism designed to maintain grid balance. The RM is different from the previous markets as not only energy is traded but also capacity. The MO is Elia, which is the Transmission Systems Operator (TSO) in Belgium, the party responsible for maintaining grid balance. Elia acts as a single buyer in this market, where it contracts market parties to provide balancing services through various mechanisms. The procurement of reserves can vary from a year before delivery up to a day before delivery. Some sellers are remunerated for their reserved capacity, while some are remunerated only for generation. This market is mainly interesting for generators that can vary their generation at will or for large consumers that can adjust their consumption to a certain extent. It is currently not of interest for wind power producers. Although wind power producers could offer asymmetrical reserves, the fact that the capacity of the reserves needs to be communicated at least one day in advance makes it a high-risk procedure both for the TSO and the wind power producer.

- The fifth type of market is the balancing market (BM). This market handles the settlement of imbalances and is organized by the TSO Elia. As market parties' actual generation or consumption deviate from their scheduled, imbalances occur which have an effect on the grid. Whenever total generation deviates from total consumption, the grid frequency starts to deviate, which ultimately can result in a system collapse. The BM is designed to reward those that help reduce system imbalance, either unintentionally through deviations from their schedule (passive balancing) or intentionally as reserves are activated by Elia on an increasing price basis (active balancing), where the most or least expensive activated bid sets the imbalance price, depending on the sign of the average balancing power. Those that cause an increase in system imbalance have to pay an unfavourable price for this restorative action, which is determined through the imbalance price. All participants with an imbalanced portfolio are forced to buy or sell their imbalance on this market, where imbalances are settled ex-post. This ex-post settling of imbalances is due to the fact that on the day itself market parties are allowed to trade, which they can then communicate to the TSO up till 2:00PM the next day, after which all imbalances are settled. The BM is of particular importance as wind power producers cannot know what their actual generation will be when they bid on the DA. This means that they almost always end up participating in the BM. They could provide passive balancing services to this market by applying asymmetrical downward adjustments to their generation, which for instance is offered to wind power producers by a Dutch company named Peeeks. However, this is not considered in this study.

2.1.2. Market actors

- First, there are several institutional actors active in the electricity market. These are acting as market operators (MO), with a special task reserved for the Transmission Systems Operator (TSO) Elia. The MO is typically a nonprofit organization, which manages the marketplace. This task includes the administering of market rules, the clearing of the market and the determination of market prices and the traded quantities of electricity. In the Belgian context there are several MOs active, most notably the ICE Endex, the EEX, the BAM and the CIM. Their main goal is to ensure an effectively functioning market. Elia has a slightly different task, as the markets it operates are geared towards maintaining grid balance, which is Elia's main responsibility in the market context, as Elia's main goal is to ensure a stable secure grid at a low cost.

- Second, there is the private legal entity called the Balance Responsible Party (BRP). The BRP has a responsibility for maintaining a balanced portfolio, which can consist of its own generation, own consumption, but also of electricity traded with other BRPs. It can represent multiple producers, suppliers and industrial consumers. A BRP has to send a program with its so-called nominations for the next day to the TSO. This program contains the planned generation or consumption for every unit that falls under the legal responsibility of the BRP. This legal entity is obligated to send this balanced portfolio each day to Elia at 2:00PM, after which it can only alter its portfolio on the ID, after which it is forced to settle its final imbalance on the BM.

- Third, there are several parties actively trading on the electricity market. First, there are producers. These are entities that own generation units and are in charge of their functioning. A producer can own multiple generation units or be owned by a supplier. It can sell electricity to the market or it can be contracted bilaterally by a supplier. The main goal of a producer is to maximize its profits from the sale of electricity through any interaction with markets, consumers and / or suppliers. It can participate in

the FM, ID, DA, RM and BM. An important type of producer in the context of this research is one that has non-dispatchable generation (e.g. wind or solar). If a producer has non-dispatchable generation, it must cope with the uncertainty of its generation. A producer with such generation needs to participate in short-term markets (DA, ID and BM) to deal with this uncertainty, as it can take advantage of the increased accuracy of shorter term forecasts of its generation. A producer can be a BRP, but it can also have another party acts as BRP. Nonetheless, it must always be part of a BRP's portfolio as its generation directly impacts grid balance. Second, there are industrial consumers. These end-users can participate directly on the electricity market, or they can buy their electricity from suppliers. If they are active on the electricity market they can participate in the FM, DA, ID, BM and RM, where its balancing services consist of demand-response services. If it participates on the RM its consumption is controlled to help balance the grid. If it participates in the electricity market directly, it must be a BRP or have another party act as BRP for its portfolio. Its goal is to maximize utility and minimize cost from its electricity consumption. Third, there are suppliers. These provide electricity to consumers, including industrial consumers. They can buy their electricity by contracting generators bilaterally, by owning generators themselves or by buying electricity from the market. It can also sell electricity from specific generators to specific markets to optimize the profit from its portfolio, as it can occur that electricity can be bought cheaper than it can generate it using its generation units. Its main goal is to maximize the profit it obtains from selling electricity to its customers, while minimizing its cost of procurement. Generally, its profit margin is narrow, which means it is important for this party to buy or generate its electricity at the lowest possible price. This is important, as in a liberalized electricity market customers may change supplier when a lower price can be obtained by switching. Last, there are consumers, which can also include large industrial consumers. These have contracts with suppliers to obtain their electricity. A consumer aims to minimize their cost from electricity and to maximize its utility from its usage.

2.2. Short Term Electricity Markets

This section discusses the main mechanisms involved in the markets most relevant to WPPs in Belgium and it aims to provide insight in how the main electricity prices are determined. As mentioned in the previous sector, producers with non-dispatchable generators are highly dependent on short-term markets, as they cannot guarantee accurate planning of their generation in the long-term. In practice this leaves WPPs dependent on three markets, which are the DA, ID and BM. As mentioned in section 1.1.2, the ID unfortunately still suffer from limited liquidity and is not always cleared. Because of this reason the next section limits itself to the DA and the BM. First, the main economic principles behind the electricity market are discussed, with a focus on European markets. Second, specifics for the DA in Belgium are discussed. Third, specific for the BM in Belgium are discussed. Lastly, differences with other pricing mechanisms for the BM are discussed.

2.2.1. Market design

This section discussed the fundamental principles behind the design of the electricity market, based on the book 'Economics of Electricity Markets' by Biggar and Hesamzadeh (2014) [5]. The objective of an electricity market is to maximize economic welfare. Although different formulations of economic welfare are possible, it generally refers to the sum of economic benefits and costs. Any arrangement that maximizes this sum is said to be efficient. This economic efficiency can be broken down into short-run and long-run efficiency, where short-run refers to maximizing welfare using the existing stock of assets and long-run efficiency refers to changing the stock of assets such that welfare is maximized over time. Several means exist with the aim of achieving the end of economic efficiency, such as competitive markets, but also other institutional arrangements.

The problem of dealing with economic efficiency for electricity markets is referred to as the problem of optimal dispatch. There are many tasks that need to be performed efficiently if an electricity industry is to achieve an overall efficient outcome. Said tasks include the efficient usage of and investment in production, consumption and network resources. Different markets have differences in how this is arranged in terms of responsibility, incentives, information and governance. The main differences in how the market processes are arranged are between the extent to which network constraints are taken into account, the time period of operation of the wholesale market and the realization of real-time balancing.

Concerning long-run efficiency, a well predictable wholesale market price, when the result of an efficiently functioning market, provides a clear and efficient incentive for the need for investments. Although some issues may occur concerning long-run efficiency in electricity markets, the focus of this section lies

with the short-term efficiency, as it to a large extent determines long-run efficiency, while this research focuses on short term strategy decisions rather than investment decisions.

Concerning short-run efficiency, in the case of electric energy, where the good is homogenous and the market participant is assumed to be a price-taker, consumers will consume where their marginal utility equals the market price and producers will produce where their marginal cost equals the market price. This market price then needs to be at a level where welfare is maximized. This short-run welfare maximum tends towards an equilibrium where the aggregate marginal utility and the aggregate marginal cost are the same. In a competitive market the price that achieves this equilibrium is set through an adjustment process, through adjustments in the market price. If this price is too high, the rate of production exceeds the rate of consumption, where some producers will not be able to sell the amount they want at that price. This leads to some producers cutting their price, which results in a decrease of the market price. If the price is too low, the rate of consumption exceeds the rate of production, where some consumers will not be able to buy the amount they want at that price. This leads to some consumers increasing the price of their bids, resulting in an increased market price. This process means that a decentralized market process will lead to an economically efficient outcome, provided participants are price-takers, where the price signal can be regarded as the 'invisible hand' coordinating the market towards economic efficiency.

In the case where physical constraints are incorporated in the market process, such a decentralized market process is not able to achieve economic efficiency. To be able to achieve economic efficiency in this case, a smart market is required involving a central market operator. This market operator collects the marginal cost and marginal utility curves of the participants and performs a constrained optimization to maximize economic welfare, which results in prices that determine the rates of production and consumption for each participant. Given the constraints, the prices can differ, depending on whether constraints are binding or not, which is referred to as locational pricing. However, in Europe there are no wholesale markets that take network constraints into account, as each market typically has a price that holds for the market as a whole, where each country is a market zone. This is typically referred to as zonal pricing. As no physical constraints are incorporated in the market process within a zone, there is not an explicit need for a central market operator. However, as transaction costs are low when the market process is centralized and automated, there is typically high liquidity in the centrally organized wholesale market. As the market price is approximated well by the wholesale price, bilateral trades tend to bid strategically with respect to their expectation of the wholesale price, which means that bilateral trades tend towards an equilibrium. Hence, when pool trading is not mandatory, which is when network constraints are not taken into account, bilateral trades and pool trades tend to achieve the same result. Hence, the main differentiation in European market lies with shorter term markets.

The wholesale market referred to is the day-ahead market, which in Europe typically closes around noon the day before delivery. The reason it does so is because the electricity market has historically been dominated by conventional controllable generation, which have material costs associated with transitioning from shut-down to operational and vice-versa. Due to this reason the optimal dispatch task is actually an intertemporal dispatch task, which is why forecasts are so important for the current electricity market. Hence, for an efficient electricity industry, efficiently using a given set of generation and consumption assets, information is required on ramp rates, minimal capacity and start-up costs and on projected generation and demand. These non-convexities make the task of optimal dispatch difficult, as there may be no feasible solution. Although it is believed that organizing the process of optimal dispatch centrally can result in economic efficiency, achieving efficient incentives in large vertically integrated organizations is difficult in practice. Hence, the market reforms mentioned earlier were directed at achieving a competitive market, aiming to provide proper incentives to increase economic efficiency. This allows for market participants to incorporate the non-convexities in their bid, which theoretically should lead to market efficiency, given adequate levels of competition. Unfortunately, information exchange between market parties is limited, while the electricity grid requires a constant balance between supply and demand. Although a centrally limited market process could theoretically clear the market at a high resolution continuously, limits to information exchange mean this currently results in too high transaction costs. Hence, the wholesale market is cleared at a limited resolution, typically between one hour and 5 minutes, which means a supplementary mechanism is required to ensure real-time balance for the grid.

As the main market is cleared day-ahead, power systems deal with imbalances due to forecast errors and unplanned interruptions in supply and demand. The real-time balancing of the grid could in theory also be achieved by the optimal dispatch process, but high frequency application of this process is currently not feasible. Because of this, TSOs are tasked to deal with these contingencies on a continuous basis. Although ideally,

shorter-term markets could help alleviate some of these imbalances, currently their liquidity is still relatively low, as supply and especially generation are currently relatively well predictable and the uncertain component within aggregate generation is still relatively small. Hence, most of the deviation between day-ahead schedules and real-time observations is dealt with by the TSO. The TSO currently activates balancing services by responding to changes in the frequency and observations deviations through constant grid metering. For the provision of these balancing services the TSO often contracts balancing service providers some time before delivery. Unfortunately, because these contracts are mostly determined some time before delivery, the activation prices of balancing services do not reflect the actual real time price of electric energy, which introduces an inefficiency. Although procurement of balancing services is an important aspect of the balancing market, for the WPP interest lies with the allocation of costs of the balancing process. This is where in European market the main differentiation lies between markets and which has the most significant implications for a WPPs offering strategy. Hence, this main difference is further discussed in section 2.2.4.

2.2.2. Day-Ahead Market

Although bilateral trading is strictly also part of the day-ahead market, this section limits itself to discussing the BAM, as the price resulting from trades on the BAM serves as a reference for other markets. The Belgian day-ahead price λ^D results from a trading process where electricity is traded one day before actual delivery. Participants to the BAM can submit their orders in hourly blocks each day until 12PM the day before delivery. After this market is settled, each BRP has to submit a balanced portfolio to the TSO (Elia) by 2:00PM the day before delivery, containing all planned generation or consumption on a plant level at a quarter-hourly resolution.

The day-ahead price is a result of the following market mechanism: All participants submit their generation and demand bids to the market stating quantities on offer for specific prices, after which these price-quantity pairs are aggregated and sorted, building offer and demand curves. This is done for each programme time unit (PTU) of the market, which in the case of the BAM means for each hour of the following day. The market-price is then the point where these two curves intersect, thus the price which clears the market. This determines the quantities to be generated and consumed along with their corresponding price. This is a single price market, meaning that a single price applies to all trades within a single time period. Below an example of demand and offer curves for the BAM is provided. The point where the two curves intersect determines the price, which in this case was €67.65.

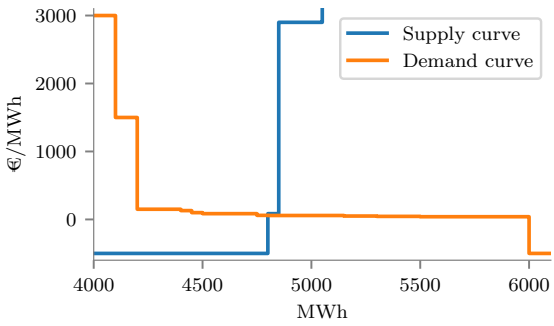


Figure 2.3: Day-ahead supply and demand curves for 2018-12-11 hour 16. Where the two curves intersect the market-clearing price is determined, where λ^D = £67.65

2.2.3. Balancing Market

This price is a result of the process which aims to keep the grid balanced, which it hardly ever is without intervention by the TSO through this market. This is due to the fact that many BRPs have sources of uncertainty in their portfolios, as they for instance contain uncertain demand or uncertain (renewable) generation. To keep the grid balanced these imbalances need to be compensated for by balancing services. In order to compensate market parties for their services a market is managed by Elia, the Belgian TSO. Although this market consists of different types of reserves (primary, secondary and tertiary), the main basis behind its functioning is as follows. As the system becomes unbalanced regulation volume is activated by the TSO to offset the imbalance. Although most balancing service is procured through various contracts and market mechanisms,

these can be aggregated to form offering curves, as Elia activates reserve bids on an increasing price basis. An example of such a curve is shown in figure 2.4. Demand consists of power demand (not energy) to the extent that Elia needs to continuously compensate any imbalance. This is distinctly different from the day-ahead market in the following respect: The marginal price setting bid for the BAM is set by the last activated generator within a certain PTU. Although BRPs' imbalances are determined on a quarterhourly resolution, the actual balancing action occurs at a higher resolution. The marginal price setting bid is determined by the most expensive generator needed for balancing action within a PTU. This can result in a price setting bid which differs substantially from the average power being delivered by the balancing services. This is visualized in figure 2.4, which is explained below.

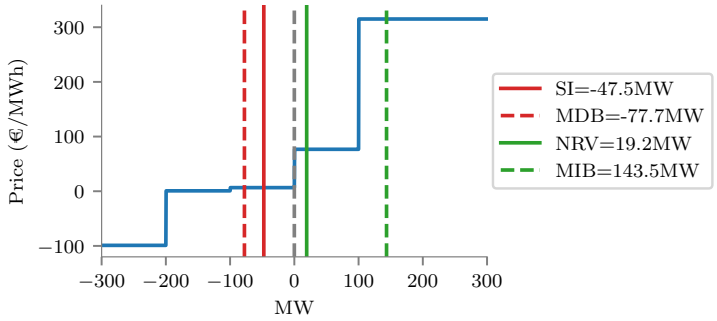


Figure 2.4: Imbalance market offer and Demand curves for 2018-01-05 from 07:30-07:45. MIB = Marginal Incremental Bid. MIP = Marginal Incremental Price = €86.17. MDB = Marginal Decremental Bid. MDP = Marginal Decremental Price = €6.59. SI = System Imbalance average. NRV = Net Regulation Volume average.

Figure 2.4 shows the supply curve for a specific quarterhour period in 2018, for which the data is published by Elia daily at 5:00PM the day before delivery. The NRV line shows the average Net Regulation Volume, which is the volume generated by reserves needed to compensate the system imbalance for this specific PTU. This is determined for each PTU. A positive NRV refers to an overall increase in grid injections or decrease in grid-offtakes, while a negative NRV refers to an overall decrease in grid injections or increase in grid-offtakes. The prices for each PTU are determined according to two prices: 1) The Marginal Incremental Price (MIP), which is the highest price paid by Elia for upward activations for a PTU, referred to as the Marginal Incremental Bid (MIB). 2) The Marginal Decremental Price (MDP), which is the lowest price received by Elia for downward activations for a given PTU, referred to as the Marginal Decremental Bid (MDB). The imbalance price is then set as the MIP if the NRV for the PTU is positive, meaning on average power was injected to the grid to help restore balance. It is set as the MDP if the NRV is negative for the PTU, meaning on average power was taken off from the grid to help restore balance. The demand curve in figure 2.4 consists thus of both the MDB and the MIB, which are always vertical lines, as demand is inelastic and determined by the minimum or maximum of the NRV signal within a PTU.

As balancing services operate at a high resolution, often within a PTU both positive and negative activations occur. In figure 2.4 the marginal bids are shown as the dotted lines, which show that for this specific PTU this was indeed the case, where both a MDB and a MIB existed. It also shows that although the average system imbalance was -77.7MW, the average NRV was only 47.5MW. Furthermore, even though the NRV was 47.5MW, the MIB was much higher at 143.5MW. This means that the higher resolution signal of the balancing power being activated within a PTU that actually determines the price can be very different from the average volume of activated reserves, the NRV. Furthermore, it can occur that the MDP and MIP are very different within a PTU, even when the NRV is quite small. In such a case, the influence of a single generator can determine whether the MDP or the MIP actually sets the price, resulting in a very difference imbalance price for that PTU.

Looking at the curve itself, several observations can be made. As with the day-ahead supply curve, it is a step function, which explains why the imbalance price is quite stable as long as the sign of the NRV does not change, even though the price setting bids do change relatively strongly between PTUs. Note that the resolution of the supply curve shown is 100MW, while in reality it consists of more steps. Unfortunately, a higher resolution supply curve is not made available, which means that the one shown should be considered a rough estimate of the curve.

Finally, in Belgium there are formally two prices, the short price λ^-, which applies to all negative imbalances of BRPs, and the long price λ^+, which applies to all positive imbalances of BRPs. These prices only differ when the absolute system imbalance exceeds a threshold of 140MW, when Elia can apply an additional incentive α, which means that the Belgian balancing market applies a single pricing system, whereas the Dutch market applies a mixed pricing system and the Spanish and Danish markets apply a dual pricing system, which are discussed in the next section. As the absolute SI exceeded 140MW for the situation shown in figure 2.4, α was set by Elia, which was €2.61/MWh in this case, resulting in λ^+ at €310.00/MWh and λ^- at €312.61/MWh. Table 2.1 shows what BRPs pay or receive in different system situations. Afterwards, each BRPs' imbalances for a given quarter-hour are aggregated after which it pays according to the volume of its total imbalance and a price determined by the mechanism explained above.

BRP situation	Net downward regulation	Net upward regulation
Long	MDP-α	MIP
Short	MDP	MIP+α

Table 2.1: Imbalance price definitions

2.2.4. Balancing Market Pricing Mechanisms

The main designs for imbalance market pricing mechanisms are the dual pricing, mixed pricing and single pricing systems. This section discusses their specifics, with a specific focus on effects on passive balancing by the WPP. Their functioning is illustrated through figure 2.5.

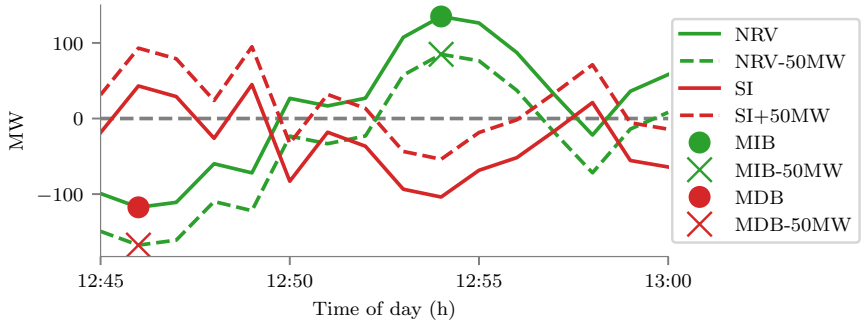


Figure 2.5: Illustration of the NRV and SI signals for 2018-01-03. The change in imbalance of 50MW causes a shift in average NRV from 10.65MW to -39.34MW. Figure shows the Net Regulation Volume (NRV), System Imbalance (SI), Marginal Incremental Bid (MIB) and Marginal Decremental Bid (MDB).

In a dual pricing system, two prices are determined by the sign of either the system imbalance or NRV. For clarity, the NRV is assumed. When a WPP has a positive imbalance in the non-shifted situation of figure 2.5, the WPP sells its imbalance at the day-ahead price. When it has a negative imbalance, it buys it back at the long price. This means the WPP can optimize to minimize its loss from trading on the balancing market. In the case of the 50MW shift in the NRV signal, the NRV changes sign which means the result for the WPP shifts. This mechanism results in a revenue for the TSO, as passive balancing is not rewarded while it does create value.

In a mixed pricing system, there are also two prices, which are determined by the existence the MDP and MIP. If there is an upward regulation bid and a downward regulation bid, two prices exist. This means that whatever imbalance a WPP has, it always incurs a loss if two prices exist, where one price is likely to be less favorable than the other, as the supply curve to the balancing market is asymmetric. If there is only an upward regulation or a downward regulation bid, both the long and short price assume the price resulting from that single bid. This means that in this case the mixed pricing system has the same price as the single pricing ceteris paribus. However, because the existence of an opposite regulating bid determines the existence of a negative opportunity price versus a positive opportunity price, the WPP must be much more careful with its strategy. This means the WPP both aims to minimize its losses and maximize its gains. In the non-shifted case in figure 2.5, there would be two prices in the mixed pricing system, where the extent to which the WPP

is penalized for its imbalance depends on the specific imbalance price with respect to the day-ahead price the WPP incurred for that PTU. If the WPP were to be have a positive imbalance in this situation, the WPP would be hurt by the 50MW shift, as the MDB increases in amplitude, which influence the price in an unfavorable direction.

In a single pricing system, there is one price, although sometimes a penalty can be applied. In the single pricing system, the sign of the NRV determines whether the MDP or the MIP is the price setting bid. This means that it can choose a strategy that provides as much balancing power but does not lead to a sign change in the NRV, thus maximizing its gains. The mechanism of the single-pricing system means it functions similarly to the dual-pricing system, except for the fact that it does not incur the day-ahead price when it successfully provides passive balancing, but incurs the imbalance price, which will be favorable in the case of successful passive balancing.

Summarizing, for the dual and single pricing systems, the sign of the NRV determines whether or not the WPP incurs a penalty for its imbalance bid, where in the single pricing system it is rewarded if it helps the system and in the dual pricing system it is not. This makes it ceteris paribus more attractive for a WPP to operate in a single pricing system over a dual pricing system. As in an optimization passive balancing is given a higher weight in the single pricing case versus the dual pricing case, it is expected that passive balancing will be more significant in a single pricing system. With regards to the mixed pricing system, all depends on how likely it is for two prices to occur. If there are two prices, the aim of the WPP will be to minimize its losses, but it will always lose. If there is a single price, it aims to obtain an opportunity gain, but is less likely to do so due to the higher likelihood of a opposite bid to occur versus the sign of the NRV to switch.

2.3. Revenue Realization for Wind Power

Concerning the value of wind power generation, the most straightforward tool used to judge its performance is by applying the reference strategy, as introduced in section 1.1.2. This means the forecast is offered on the day-ahead market at a zero price, after which the total revenue from both the day-ahead market and the balancing market is calculated. The realization of this process can be observed in figure 2.6, which is discussed below.

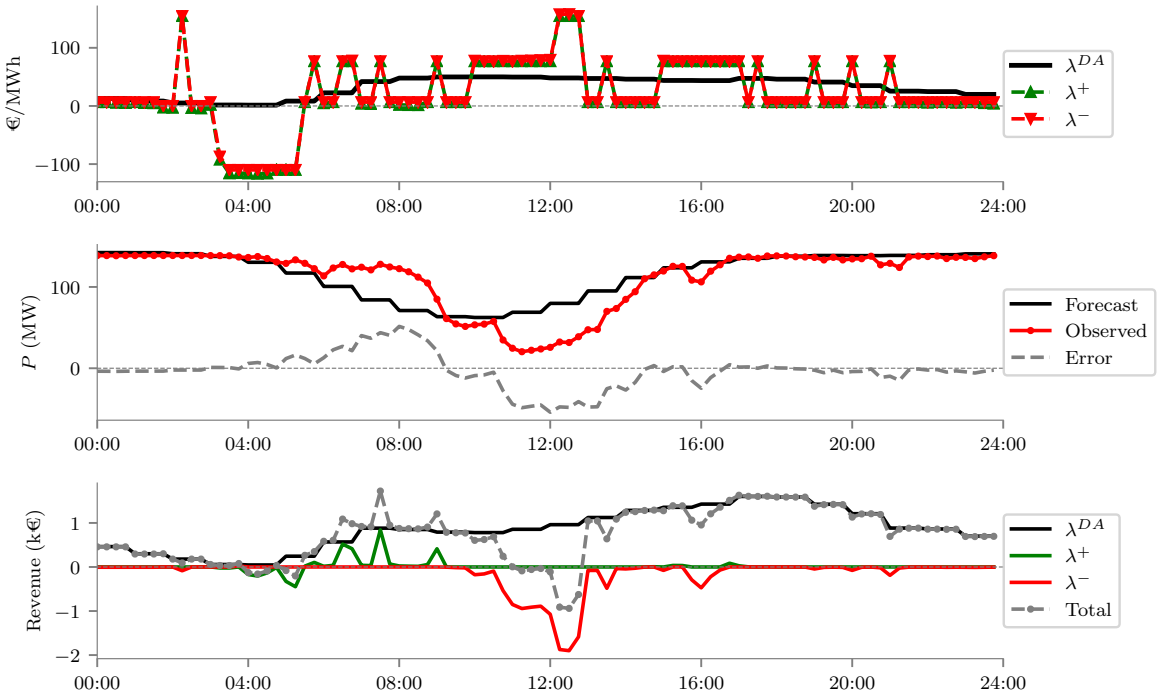


Figure 2.6: a: Belgian Electricity market prices for 2018-01-02.
b: Power values for the Essent forecast.
c: Market revenue for the wind farm

First, figure 2.6a shows the three main electricity prices that most affect the revenue of a wind farm. In summary, these effect the WPP as follows:

- The day-ahead price λ^D. This price is currently the most important electricity price, as it serves as a reference for many market transactions. The noon closure of the market is the reason why day-ahead forecasts are issued in the morning before the closure of the day-ahead market.
- The long price λ^+. This price is the real-time electricity price, which is used to price grid imbalances. Each day before noon each BRP is required to inform the grid operator how much it is going to produce or consume for each generating unit, which for WPPs is the amount sold on the day-ahead market. If the market participant deviates from this plan, real-time imbalances occur, for which the market participant has to pay. These deviations are determined in blocks of 15 minutes. If the market participant produces more or consumes less than it had communicated, it is forced to sell this excess at λ^+.
- The short price λ^-. This price is also the real-time electricity price, but applies when a market participant produces less or consumes more than it had informed the grid operator. It is then forced to buy the shortage at λ^-. As can be observed in figure 2.6a, this price is mostly equal to or very similar to λ^+. This is also true for longer periods, as it hardly ever deviates (strongly) from λ^+ in the Belgian system, as the SI not often exceeds |150| MW.

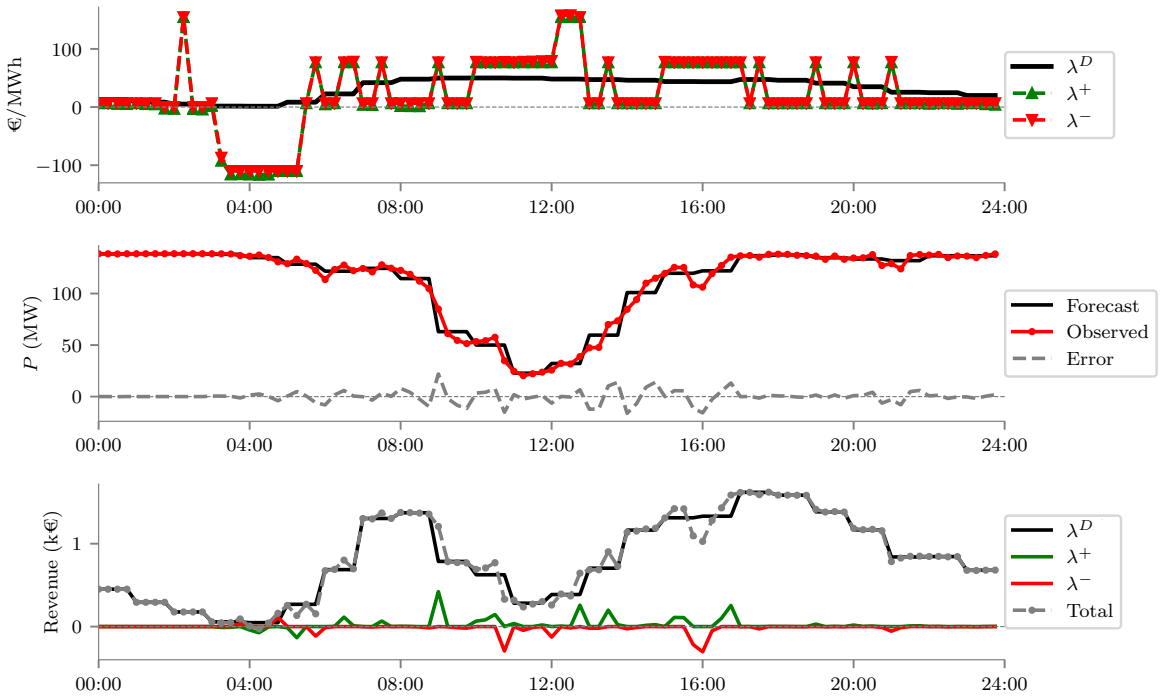


Figure 2.7: a: Belgian Electricity market prices for 2018-01-02.
b: Power values for a perfect forecast.
c: Market revenue for the wind farm

Figure 2.6b shows the amount the wind farm submitted to the DA, which for wind farms is currently always offered at a zero price, to guarantee that all electricity is sold. Alternatively a wind farm could offer no electricity to the day-ahead market and sell everything on the real-time balancing market, but in practice this option is much less profitable. The convention for determining the quantity to offer is to offer the best forecast, in order to minimize imbalances. From the figure, a step function can be observed, which is due to the fact that the bids to the day-ahead market are required to be in hourly blocks. The red line shows the actual generation by the wind farm, which is much less smooth and deviates from the forecast. This deviation is the amount a wind farm is forced to sell or buy on the balancing market. This means the following power variables are involved in determining the revenue:

- P^D: The amount offered on the day-ahead market. In the reference strategy this is the forecast.
- P: The amount actually generated.
- $P - P^D = P^\Delta$: The deviation between what was offered and was was generated. In the reference strategy this is the forecast error.

Figure 2.6c shows the market revenues as a result of each price and power volume. The first thing to note is that it is possible for a wind farm to profit from its own imbalances, for instance if in a specific period the wind farm generated more than it sold on the day-ahead market (the producer is *long*) and λ^+ for that period is higher than λ^D, meaning it sold its excess generation at a higher price than it would have gotten on the day-ahead market. This situation can be observed around the 8:00 mark. It is also possible to lose heavily due to imbalances, which can be observed around the 12:00 mark. Here the forecast error is large and negative (the producer is *short*), while λ^- is higher than the day-ahead price, meaning it has to buy back its shortage at a price that is higher than it would have cost to sell less on the day-ahead market. This results in a significant net loss for that period, even though the wind farm did in fact contribute electricity to the grid.

When it comes to comparing forecasts, an interesting comparison is that between a specific forecast and the theoretical perfect forecast, as this allows one to identify the theoretical value of improving them. For this purpose a perfect forecast is constructed, which can be observed in figure 2.7b.

The first thing to note is that the forecast error in figure 2.7b is often nonzero, which is due to the market rules, as bids to the day-ahead market are submitted in hourly blocks, whereas imbalances are determined at a quarter hourly resolution. The second thing to note is that the total revenue is very close to the day-ahead revenue in figure 2.7c. Clearly, both negative and positive peaks in revenue from the imbalance market are mostly avoided by this perfect forecast. Overall, for this specific day the total market revenue for the forecast in figure 2.6 was 11% lower compared to the forecast in figure 2.7. This makes it possible to compute revenue loss due to forecast errors. For instance, for 2017 this specific loss in revenue due to forecast errors for the Essent forecast was 6.5%.

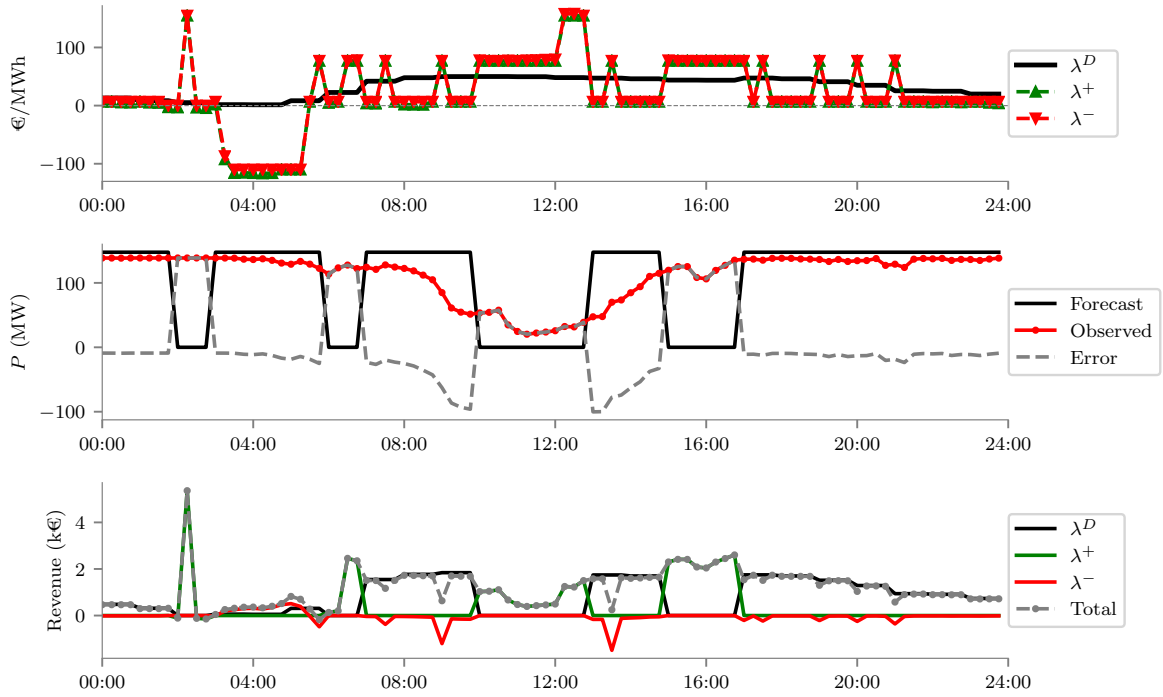


Figure 2.8: a: Belgian Electricity market prices for 2018-01-02.
b: Power values for a perfect bidding strategy.
c: Market revenue for the wind farm

The basic working principle behind the advanced bidding strategies, which were briefly discussed in section 1.1.2, can best be explained visually through figure 2.8, which shows the theoretical optimal bidding

strategy for the same day in 2017. Here the hourly bid to the day-ahead market was either the maximum capacity of the farm $P^{\max}$, which is the maximum bid the market operator allows, or zero. The rationality behind the optimal bid is that of arbitrage opportunity. This is the simultaneous purchase and sale of a product on two different markets to take advantage of a price difference. So a wind power producer can profit from buying power from the day-ahead market and selling it on the balancing market when λ^D is higher than λ^+. On the other hand it can profit from buying from the balancing market and selling on the day-ahead market when λ^- is lower than λ^D. Of course, the purchase and sale of electricity on these two markets is not simultaneous, which makes it a temporal arbitrage opportunity. This bidding strategy does take into account that bids are on an hourly resolution, so a power producer would only buy from the day-ahead market and sell on the balancing market and vice versa when it is profitable to do so for the entire hour. Compared to the reference strategy, the revenue from this theoretical optimal bid was 55% higher for the day depicted in figure 2.8. When considering the whole of 2017, the revenue increase from using this optimal bidding strategy would have been 98% compared to the reference strategy.

Unfortunately, not only does a market participant not know these prices in advance, it also influences the prices, especially since the wind farm in this study is quite large. As was shown in section 2.2.3 on the mechanism determining the imbalance price, it is entirely possible if a wind power producer anticipates a certain arbitrage opportunity, that it changes the sign of the NRV, which means that an opportunity becomes a threat: When the NRV is positive, the imbalance price is determined by the MIP. However, if a WPP strategically provides much balancing power it is possible to shift the sign of the NRV to negative, which means the imbalance price is determined by the MDP. The situation in figure 2.4 is not uncommon, as often within a PTU both incremental and decremental bids are activated, where the difference between the MDP and MIP is large. While the WPP would successfully provide balancing power in the case of an initially positive NRV, if the sign of the NRV changes, it will be paid a much lower price than it would have, had the sign of the NRV not changed. This means it ends up hurting itself through its effect on the imbalance price. As the revenues reported in this section do not take the effect a producer has on the price into account, even under perfect information these revenues are not realistic. However, the quantities presented here do provide some indication that value can be obtained. As value improvement from forecast improvement is limited and forecast improvement is marginal and the process of improvement is mostly exogenous to the WPP, it seems valid to try and obtain some of the large potential value increase from arbitrage opportunity. Hence, the next chapter focuses on how this can be done through advanced bidding strategies.

3

Stochastic Optimization Model

As discussed in section 1.1.2, one of the first studies on this subject is that of Matevosyan and Söder from 2006 [38]. A study that made important innovations on the methods introduced by Matevosyan and Söder is that of Morales, Conejo and Pérez-Ruiz from 2010. This study introduced improvements to the underlying mathematical principles and included the option to leverage the ID as well as control risk using the Conditional Value-at-Risk at confidence level α (CVaR$_\alpha$) [43]. These models are designed for dual-pricing systems under a price-taker assumption. So far there have been two case studies that implemented a form of this model in markets similar to the one studied here. Those are the one by Chaves Avila et al. [8], which implemented a price-taker model in the Dutch market, a mixed pricing imbalance market, and the one by Bertrand and Papavasiliou (2017), where both a price-taker and a price-maker model were implemented for the Belgian single pricing market [4]. This study both applies the price-taker version of the model by Morales et al.(2010) [43] and a reformulated price-maker version of the model. In their study, Bertrand and Papavasiliou also claim a closed form solution to the price-maker version of the model. For completeness sake this study will also test their claim and study this version of the model. In the next section the price-taker of the model is explained. Secondly, a price-taker version of the model is discussed. Lastly, the closed form solution to the price-maker model is discussed. The tools used for implementing these models are the Pyomo[21], Numpy and Scipy libraries in Python. The optimization models are solved using the CPLEX solver from IBM, which interfaces with the model implemented using Pyomo.

3.1. Price-Taker Optimization Model

The model from Morales et al. (2010)[43] assumes two prices exist in the balancing market. As this strictly speaking is the case for the Belgian balancing market, Bertrand and Papavasiliou did not change the original model. However, for the first half of 2018, λ^+ and λ^- only diverged 0.31% of the time. Furthermore, the average value of the penalty α where α is nonzero was €2.73, which is relatively small as the mean absolute value of the imbalance price without α was €47.90. Hence, given the limited divergence of both prices and the limited impact of α on the imbalance price, this study does not take α into account. A further change this study introduces with regards to the formulation by Morales et al.(2010) [43] is the resolution of the model, which is increased to a quarterhourly resolution from an hourly resolution. As forecast errors and imbalance prices can be volatile, when peaks of both line up the hourly mean of both can give a skewed result of their interaction. The model is formulated as follows:

$$\text{Maximize}_{P_t^D,\forall t;P_{t\omega}^\Delta,\forall t,\forall \omega}$$

$$\xi\{R\} = \sum_{\omega=1}^{N_\Omega}\sum_{t=1}^{N_T} \pi_\omega d_t (P_t^D \lambda_{t\omega}^D + P_{t\omega}^\Delta \lambda_{t\omega}^{\text{Imb}}) \tag{3.1}$$

$$s.t. 0 \leq P_t^D \leq P^{\max}, \forall t \tag{3.2}$$

$$P_{t\omega}^\Delta = P_{t\omega} - P_t^D, \forall t, \forall \omega \tag{3.3}$$

$$P_t^D = P_{t'}^D, \forall t, \forall t' : t' = t-1 \forall \frac{t}{4} \in [1,2,...,\frac{N_T}{4}] \tag{3.4}$$

Where P_t^D is the amount that is bid on the day-ahead market for time t, π_ω is the probability of scenario ω and $P_{t\omega}^{\Delta}$ the imbalance at time t for scenario ω. Equation 3.4 provides the added constraint necessary to deal with the difference in resolution between the variables and parameters for the day-ahead market part of the model and the variables and parameters for the balancing market part of the model.

In the optimization model, P_t^D only has one index, which means that the optimization determines the value for P_t^D that optimizes the overall profit, given all possible realizations of the stochastic processes indexed by both t and ω, for the entire scenario set of size N_Ω. This method of optimization, where uncertain parameters are represent by a discrete scenario set of size N_ω is called stochastic optimization. This method is shown to outperform similarly defined robust optimization counterparts in terms of robustness of its outcomes, especially when risk is controlled inside the stochastic optimization model [42]. Furthermore, it provides a guaranteed single optimum contrary to a robust optimization model, given that the stochastic processes are modeled correctly and are accurately represented by the scenario set [42]. Finally, the model can be extended to include risk measures, but also other processes of interest, thus providing a flexible framework on which WPPs can build.

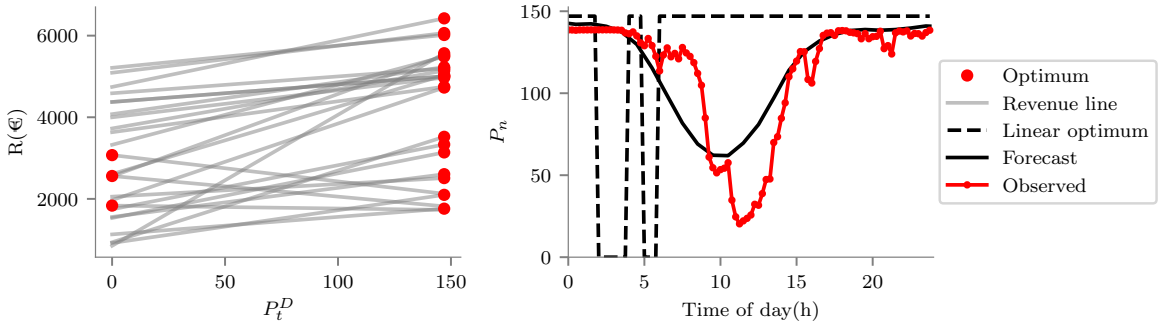


Figure 3.1: a: Linear optimization lines for P_t^D
b: Linear optimal bid for 2018-01-02

Compared to the original version of the model, which was designed by Morales et al. (2010) [43] to be used in a dual pricing system, the functioning of this model is different. In a dual pricing system, passive balancing is not rewarded, as the balancing price for successful passive balancing is set to be the corresponding day-ahead price for that PTU, while the penalty for unsuccessful passive balancing is set at the imbalance price. Hence, the optimum of the linear model lies there where as much of the cost from imbalances is avoided, which means it is never rewarding to go beyond the point where most unfavorable imbalance revenue is avoided. This means that the optimum for the original model can be seen as a cost minimization model, which is also how it is reformulated in chapter 7 of 'Integrating Renewables in Electricity Markets' [42]. However, as passive balancing is rewarded in single and mixed pricing systems, the model becomes a revenue maximization model, which has important implications for the linear version of the model. This means that the optimum always lies at either 0 or $P^{\max}$, which is in accordance with the arbitrage logic as presented in section 2.3. This is illustrated in figure 3.1, where for 2018-01-02 the optimization is approximated by varying P_t^D for each timestep in steps of size 1 MW.

As the capacity of the wind farm in this study is 147.6MW and the mean absolute system imbalance in Belgium for the first half of 2018 was 119MW, it seems irrational to adopt this price-taker assumption. Especially as it seems likely that the sign of the NRV can change given the magnitude of the strategic bids, thus converting opportunities into threats, which was also one of the main conclusions of the paper by Bertrand and Papavasiliou (2017) [4]. Hence the next section focuses on the price-maker formulation of the model.

3.2. Price-Maker Optimization Model

Bertrand and Papavasiliou [4] extend the model by Morales et al.(2010) [43] by incorporating a price-maker term in the model. They model this term as a static parameter by evaluating the relationship presented in the scatter in figure 3.2. This scatter shows a ordinary least square linear regression line with intercept, from which a linear coefficient can be obtained. For the data for the first half of 2018 this coefficient was -0.2189 $\frac{€}{MWh^2}$. This coefficient can be interpreted as follows: If a producer increases the system imbalance, the price will decrease, as more downward regulation needs to be activated. These downward regulation bids are

willing to pay for not generating electricity, albeit at a lower price than they had sold their electricity on the day-ahead market. As bids are activated in a price increasing order, bids are first activated that are willing to pay a high price for not having to produce. Simultaneously, some consumers may be willing to increase their consumption at a favorable price. However, as more bids are activated, at some point base load generators will be required to reduce their output, which for them may represent a cost as it is difficult to ramp back up, their efficiency reduces or variable output may hurt the generation unit. For upward regulation bids the logic is straightforward, as when more bids are activated for upward regulation, more expensive generators need to become active, while consumers will ask for a higher price to adjust their consumption downward, as this will reduce their utility.

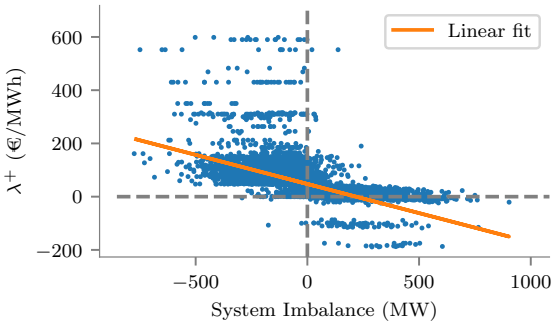


Figure 3.2: Scatter plot of System Imbalance versus λ^+ for observed data from 2018-01-01-2018 - 2018-06-30.

The model formulated by Bertrand and Papavasiliou (2017) [4] incorporates the price-maker term as the slope of the linear influence of the system imbalance on λ^+. The model from that study still contains two prices and is at an hourly resolution. However, as with the linear version of the model, this study assumes a single price and increases the resolution of the model. This leads to the following formulation of the model:

$$\text{Maximize}_{P_t^D,\forall t;P_{t\omega}^{\Delta},\forall t,\forall \omega}$$

$$\xi\{R\} = \sum_{\omega=1}^{N_\Omega}\sum_{t=1}^{N_T} \pi_\omega d_t (P_t^D \lambda_{t\omega}^D + P_{t\omega}^{\Delta}(\lambda_{t\omega}^{\text{Imb}} + P_{t\omega}^{\Delta}\gamma_{t\omega}^{\Delta})) \tag{3.5}$$

$$s.t. 0 \leq P_t^D \leq P^{\max}, \forall t \tag{3.6}$$

$$P_{t\omega}^{\Delta} = P_{t\omega} - P_t^D, \forall t, \forall \omega \tag{3.7}$$

$$P_t^D = P_{t'}^D, \forall t, t' : t' = t-1 \forall \frac{t}{4} \in [1,2,...,\frac{N_T}{4}] \tag{3.8}$$

Where $\gamma_{t\omega}^{\Delta}$ is the price-maker term with unit $\frac{€}{\text{MWh}^2}$, which represents linearly what the influence of the park imbalance is on the imbalance price.

This study presents another change from the model by Bertrand and Papavasiliou (2017) [4]. Whereas they represent the price-maker term $\gamma_{t,\omega}^{\Delta}$ as a constant, this study proposes to model this term as a stochastic process. This is a common method for modeling price-maker effects in stochastic optimization, as is explained in a conventional power producer example in chapter 5 of 'Decision Making Under Uncertainty in Electricity Markets' by Conejo, Carrión and Morales(2010) [10]. Conejo et al. (2010) model this term as the slope of the inverse ordered supply curve to the balancing market. As an approximation to the supply curve to the balancing market is published daily at 5:00PM day-ahead by the TSO Elia, this method can also be applied here. This seems like a likely improvement to the model by Bertrand and Papavasiliou, as the 2007 study by Pinson, Chevallier and Kariniotakis [47] concluded that an increased resolution of uncertain parameters involved in the optimization is an important source for increased value from the optimization. Furthermore, both studies by Chaves Avila, Hakvoort and Ramos (2013) and by Rahimiyan, Morales, and Conejo (2010) indicate that important improvements to the solution are to be expected as the quality of the modeling of the stochastic processes involved increases [9, 51]. As the price-maker effect is published at a quarter hourly resolution, while Bertrand and Papavasiliou model it as a constant, this seems like a likely source of improvement to the

optimization. Hence, this study models this term as the slope of a least squares linear representation of the supply curve to the balancing market, as visualized in figure 3.3.

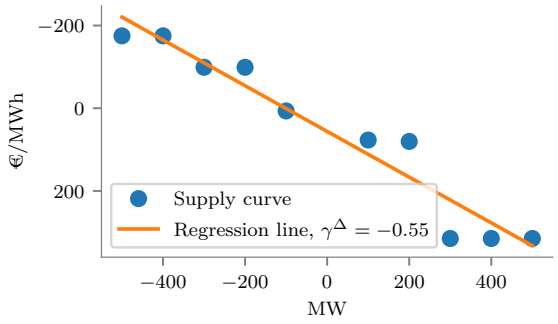


Figure 3.3: Inverse ordered supply curve to the balancing market for 2018-01-01 00:15:00

This formulation with $\gamma^{\Delta}_{t\omega}$ as a stochastic process is a combination of the price-maker formulation, as introduced in equation 3.1 with a price-maker model for conventional generators, as presented in chapter 5 of 'Decision Making Under Uncertainty in Electricity Markets' [10]. Below the new quadratic price-maker optimization is illustrated, where the optimization is approximated by steps of 1MW for 2018-01-02.

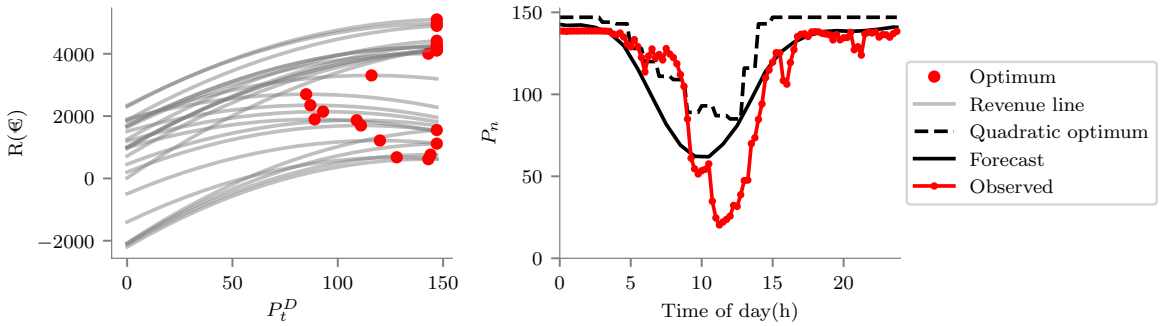


Figure 3.4: a: Quadratic optimization lines for P^D_t
b: Quadratic optimal bid for 2018-01-02

The figure shows that the optimum bid stays closer to the original forecast. The expectation is that this will result in more realistic optimal offers, meaning better real-world performance. Even though the linear approximation to the supply curve is an abstraction, the inclusion of the price-maker effect as a stochastic process is expected to add quality to the optimization in several ways. The first is on the resolution of the term, as was shown by Pinson, Chevallier and Kariniotakis [47]. The price-maker effect as specified here should increase as less regulation capacity is available, which would force the optimization to bid more conservatively. Vice-versa, the term should decrease when more regulation capacity is available, which should result in more opportunist bids. The second is on the inclusion of it as a stochastic process, which is analogous to moving from a deterministic optimization to a stochastic optimization, where the second is shown to be superior when stochastic processes are involved [42].

3.3. Price-Maker Analytic Model

For completeness, this section discusses the analytic version of model 3.5, which according to Bertrand and Papavasiliou (2017) [4] solves the problem under the assumptions that the imbalance price can be represent by one price, that the stochastic processes can be considered independent and that the probability weighted sum of the scenarios result in their respective forecast. The result of this, for which the proof can be found in their paper, is as follows:

$$P_t^D = \hat{P}_t + \frac{\hat{\lambda}_t^{\text{Imb}} - \hat{\lambda}_t^D}{2\hat{\gamma}_t^{\Delta}} \tag{3.9}$$

Where $\hat{P}_t$ is the forecast for wind power, $\hat{\lambda}_t^{\text{Imb}}$ is the forecast for the imbalance price, $\hat{\lambda}_t^D$ is the forecast for the day-ahead price and $\hat{\gamma}_t^{\Delta}$ is the forecast for the price-maker term.

Although the version of the model by Bertrand and Papavasiliou (2017) again presents $\hat{\gamma}_t^{\Delta}$ as a static term, here it is assumed to be dynamic at the same resolution as $\hat{\lambda}_t^{Imb}$. However, as the bid is made at an hourly resolution, P_t^D is resampled as its hourly mean. The closed form reformulation can be interpreted as follows. The numerator is the expected difference between the imbalance price and the day-ahead price, which represents the expected reward when the WPP decides to alter its bid away from the forecast. The denominator represents the influence the WPP has on the imbalance price. The larger this term becomes, the closer the altered bid will be to the forecast. This formulation allows the bid to exceed the minimum and maximum capacity of the wind farm, which is why when this occurs, the optimal solution is projected on the interval within its minimum and maximum capacity, thus respecting constraint 3.6.

3.4. Evaluating Strategy Performance

Concerning the evaluation of the performance of each strategy, the studies by Rahimiyan et al. (2011) and Chaves Avila et al. (2013) applied a price-taker recreation of revenues [9, 51], after which they compared the total revenues. However, this method firstly does not provide insight in the statistical differences between the revenue of one strategy over the other. Secondly, these assume a difference in strategy has no influence on the price. Thirdly, the performance of the strategy is not explained by the sum of the revenue. Fourth, no insight is obtained on the system effects or societal value of the strategy. This section addresses these issues by first presenting a method to reconstruct the price to obtain more realistic revenues, secondly by presenting a method to compare different sets of revenues statistically, thirdly by introducing a method to explain the performance of a strategy and finally by presenting a method to determine system effects and societal value.

3.4.1. Imbalance Price Reconstruction

The method of computing revenues as done in the studies by Rahimiyan et al. (2011) and Chaves Avila et al. (2013) is explained in 2.3 [9, 51]. However, as in mixed pricing and especially in single pricing balancing markets the WPP is incentivized to provide passive balancing, its optimal bid can result in much larger deviations from its forecast compared to dual pricing balancing markets. Therefore, the assumption that prices are not influenced seems unreasonable. This is especially relevant in the case of the linear version of the model, where either 0 or the maximum power of the wind farm is bid. Furthermore, the extent to which a wind farm successfully provides passive balancing determines the extent to which the price is influenced unfavourably for the WPP. This is due to the fact that when passive balancing is provided, fewer reserves need to be activated, thus the price is dampened. To be able to provide a fair comparison between more aggressive and more conservative approaches to the optimization, the imbalance price needs to be reconstructed. Because of this Bertrand and Papavasiliou (2017) presented a method to reconstruct the price for wind farms of different sizes in order to compute a corrected revenue [4]. The supply curve shown in figure 3.3 was used to construct a piece-wise linear function, thus enabling the usage of the published supply curve for each PTU of the balancing market to recreate the imbalance price. However, they did not take into account that a shift in NRV changes the price setting bid. Furthermore, the supply curve to the balancing market is not a piece-wise linear function, but a step function, as the price can remain constant while the price setting bid fluctuates strongly within the bounds of a single step, which explains the relative stability of the imbalance price between switches. Hence, here a revised version of their algorithm is presented.

As explained in section 2.2.3, the price is determined by the sign of the Net Regulation Volume (NRV) and the price setting incremental (MIB) or decremental bid (MDB), through the price determined by the supply curve to the balancing market together with an optional value for α. The final price can be increased by

α by the TSO when the absolute system imbalance exceeds 140MW. Using this knowledge an algorithm is introduced to model the influence a difference in strategy would have had on the balancing price. The main principle behind the mechanism presented here is that the change in strategy directly alters the magnitude of the price setting MDB and MIB. An approximation to these bids can be obtained as the TSO publishes the NRV signal at resolution of one minute, thus allowing the computation of the MDB and the MIB as the minimum and maximum of the signal within a PTU. This principle is visualized in figure 3.5a, which shows how a change in strategy of a single WPP would have altered the SI, NRV, MIB and MDB for a single PTU of the balancing market. As an altered bid would not have changed the electrical signal of the WPP and thus not that of the total system imbalance at a higher resolution than that of the PTU, but only alter the imbalance by a constant for a PTU, the SI for that PTU would be altered by that constant. As the NRV and thus the MDB and MIB are a direct response to the SI, the assumption behind the price reconstruction model is that they are altered by the same amount opposite to the change in SI. This altered SI, NRV, MDB and MIB then allow the reconstruction of the price. A limitation in this algorithm is that the supply curve published by Elia is an approximation of the actual supply curve to the balancing market, with step sizes of 100MW, which means that the price-maker effect is quite crude. Furthermore, Elia mentions on its site that some constraints may change the shape of the actual curve compared to the published curve, mainly grid constraints and the resolution of the approximation, which make the approximated supply curve not a reliable function for simulation of prices. However, it may allow for additional insights in the real-world performance of a strategy relative to the price-taker revenue evaluation.

As for each PTU the NRV signal is published at a minute resolution by Elia, together with the approximated supply curve, the price can be recreated by first obtaining the MDB and MIB by determining the minimum and maximum of the NRV signal. This bid can then be altered by the inverse amount the strategy is altered, after which the supply curve can be applied to obtain the recreated price. This is visualized in figure **??**b, which shows what the effect would have been on the MDB, MDP, MIB and MIP if a bid to the day-ahead market had been 50MW lower, resulting in a 50MW increase in system imbalance and a 50MW decrease in NRV, MDB and MIB. The figure shows that the differences in steps of the supply curve lead to different impacts on the MDP and MIP. Taking the possible sign change of the NRV into account, this may allow for a more realistic recreation of the balancing price. As Elia also publishes the value for α for each PTU, the algorithm also checks whether or not α should still be applied, depending on whether or not the absolute value of the system imbalance exceeds 140MW.

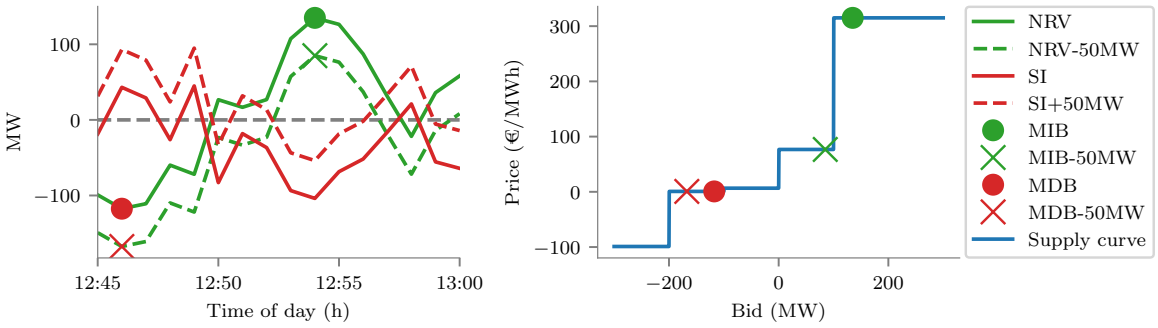


Figure 3.5: Illustration of price recreation of the imbalance price for 2018-01-03. The change in imbalance of 50MW causes a shift in average NRV from 10.65MW to -39.34MW.
a: Net Regulation Volume (NRV), System Imbalance (SI), Marginal Incremental Bid (MIB) and Marginal Decremental Bid (MDB).
b: Supply curve to the balancing market with the MDB and MIB, where the MDB remains the same at €0.74/MWh, the MIP goes from €315.00/MWh to €76.71/MWh, and the price goes from €315.00/MWh to €0.74/MWh.

3.4.2. Revenue Comparison

The studies by by Rahimiyan et al. (2011), Chaves Avila et al. (2013), Bertrand and Papavasiliou (2017) all present their conclusions based on the total revenue of a certain strategy for a certain time period [4, 9, 51]. While this is an important indicator for the existence of a difference between strategies, there are more robust measures to indicate statistical differences between strategies. As the total revenues are the sums of all

revenues for a time period, these can be seen as part of a statistical distribution, focus is shifted to the distributions of those revenues. In econometrics it is common to compare different sets of revenues rather than their sums. This provides insight whether the difference in sum is due to a structural difference between two sets of distributions. This can be determined through the concept of *stochastic dominance*. This concept is introduced by Hadar Russell (1969), where first order stochastic dominance determines whether distribution $F(x)$ should be preferred over distribution $G(y)$ by evaluating:

$$F(x) \leq G(y) \forall x, y \in [a, b] \tag{3.10}$$

Where a is min(x, y) and b is max(x, y). This means that the cumulative distribution function F(X) must lie under cumulative distribution function G(Y) for all values of x and y. This case is shown in figure 3.6b for distribution $f_1(x)$. This figure shows that the cumulative density functions can be used to determine whether or not a distributions of revenues is shifted towards the right, which means it would consistently result in higher revenues.

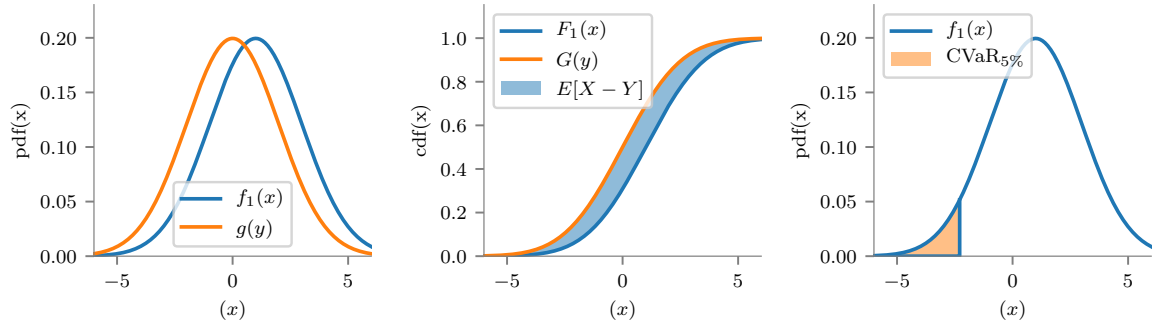


Figure 3.6: Visualization of determining first -and second-order stochastic dominance, where $f_1(x)$=N(1,2) and $g(y)$=N(0,2).
a: Probability Density Functions.
b: Cumulative Density Functions and expected difference of the mean.
c: Conditional value at risk.

A problem with the first-order stochastic dominance test in equation 3.10 is that often cumulative distribution functions intersect, which means that this test often does not hold. However, it may still be the case that one distribution is shifted to the right of the other, which means that one would result in a higher revenue than the other. This case is shown in figure 3.7 for distribution $f_2(x)$. Distribution $f_2(x)$ is clearly to the right of distribution $g(y)$ and will result in a higher expected revenue, but does not strictly first-order dominate distribution $g(y)$.

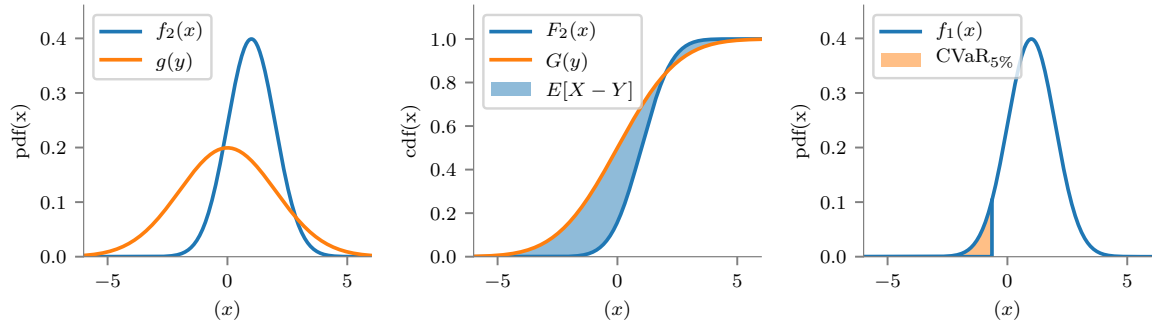


Figure 3.7: Visualization of determining first -and second-order stochastic dominance, where $f_2(x)$=N(1,1) and $g(y)$=N(0,2).
a: Probability Density Functions.
b: Cumulative Density Functions and expected difference of the mean.
c: Conditional value at risk.

In the case of the strategies in this study, especially when comparing those from similar wind power forecasts, gains in revenue are marginal and revenue distributions can be very close, which can lead to intersecting cumulative density functions, which means that there is no strict first-order dominance but there is more revenue to be expected from one over the other. To still get a measure on the quantity of the shift in revenue, this magnitude of this shift can be quantified by the expected difference of the mean:

$$E[X-Y] = \int_a^b (G(y) - F(x))dx \tag{3.11}$$

Where $E[X-Y]$ is the expected difference of the mean, a is $\min(x,y)$ and b is $\max(x,y)$. This score is the surface area of the difference between G(x) and F(x) and is regarded as the main ranking criterion of the strategies. Conveniently, $E[X-Y]$ is the expected mean difference between the two strategies' revenues, which means that this integral reduces to the difference of the average of the two revenue samples:

$$E[X-Y] = \frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{n}\sum_{i=1}^{n} y_i \tag{3.12}$$

Although this provides insight in what magnitude of difference in revenue can be expected between the strategies, the value of one strategy over the other can also lie in its risk. For instance, if two strategies have the same expected mean, one may have a smaller spread. Such a smaller spread in revenue, without a reduced mean would be interesting for market participants, as it would reduce risk without sacrificing revenue. An example of such a case is shown in figure 3.8 for distribution $f_3(x)$. Especially a distribution which does not only have a shift in mean but also a reduced spread would be preferred, as it would result in a reduced risk, combined with an increased expected revenue, which is the case for distribution $f_2(x)$ as shown in figure 3.7. A concept for testing this effect is the Conditional Value at Risk (CVaR), which is expressed as [42]:

$$\text{CVaR}_\alpha = \int_a^{F^{-1}(\alpha)} x f(x) dx \tag{3.13}$$

Where $f(x)$ is the probability density function of the revenue distribution, α is the confidence level, where $0 \leq \alpha \leq 1$, $F^{-1}(x)$ is the inverse cumulative distribution function of $f(x)$ and a is the minimum value of sample X. This means that the conditional value at risk is the expected value of the α% lowest revenues. This means it provides a suitable metric for risk-averse producers, as it quantifies the expectation for low revenues. For the samples in this study, this reduces to the mean of the lowest α% of the revenues:

$$\text{CVaR}_\alpha = \frac{1}{n_\alpha}\sum_{i=1}^{n_\alpha} x_i \forall x_i \in \{x_i, x_{i+1}, ..., x_{n_\alpha}\} : x_{i+1} \geq x_i \tag{3.14}$$

Where n_α is the number of samples in set X of size n, which equates to αn. This means that the the CVaR_α for a discrete sample reduces to the mean of the 5% lowest values.

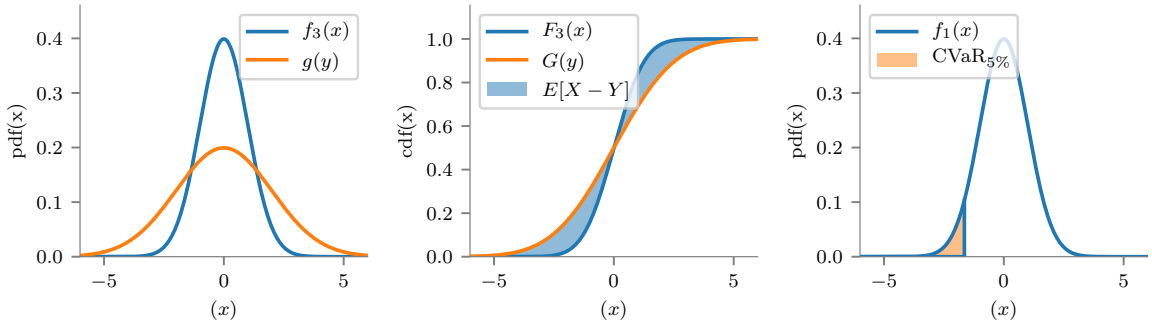


Figure 3.8: Visualization of determining first -and second-order stochastic dominance, where $f_3(x)$=N(0,1) and g(x)=N(0,2).
a: Probability Density Functions.
b: Cumulative Density Functions and expected difference of the mean.
c: Conditional value at risk.

In conclusion, the shift of one distribution of revenues can be quantified by the expected difference of the mean, while risk can be quantified by the Conditional Value at Risk. For the three distributions $f_1(x)$, $f_2(x)$ and $f_3(x)$ table 3.1 shows in the differences in expected mean and CVaR_α compared to distribution $g(y)$, where α is chosen as 5%. Both distributions $f_1(x)$ and $f_2(x)$ show the same shift in mean, but since $f_2(x)$ has a smaller spread it results in a higher increase in CVaR_α, making it the preferred strategy. Distribution $f_3(x)$ shows no increase in expected mean, but does show a significant increase in CVaR_α, greater than that for distribution $f_1(x)$. This means that a risk-averse producer may want to consider the strategy resulting in revenue distribution $f_3(x)$.

Distribution	$E[X-Y]$	$\Delta\ \text{CVaR}_\alpha$
$f_1(x)$	1	0.05
$f_2(x)$	1	0.15
$f_3(x)$	0	0.10

Table 3.1: Differences between the three distributions in figures 3.6, 3.7 and 3.8 and reference distribution $g(x)$.

3.4.3. Strategy and market quantities

Although the revenue comparison provides insight in which strategy to choose, interest also lies with the composition of each strategy. To analyze the composition of each strategy an analysis is introduced, which aims to provide insight in the interaction between the strategy and the market, thus explaining its performance. First, imbalance prices are redefined as opportunity prices relative to the day-ahead price as follows [42]:

$$\begin{aligned} \lambda^{\Delta^+} &= \lambda^+ - \lambda^D \\ \lambda^{\Delta^-} &= \lambda^- \lambda^D \end{aligned} \tag{3.15}$$

The first thing to note is that this redefinition is based on the characteristics of the imbalance price, which is typically higher than the day-ahead price when the system is in up-regulation and lower than the day-ahead price when the system is in down-regulation. This means that the balancing prices are strongly related to the day-ahead price and system imbalances. Thus, when looking at relationships between forecasts and the imbalance price, these should be redefined in order to help identify these relationships. The second thing to note about the redefinition is that the second definition is positive when λ^- is higher than the day-ahead price, meaning that when this term is positive it hurts to be short as the park is forced to buy back its imbalance at a higher price than it would have had to pay on the day-ahead market. By defining it in this way, it becomes easier to identify the valence of the relationship a specific forecast has with market prices. Secondly, a negative imbalance indicates that the producer has to buy on the balancing market. When the producer sells a negative amount and it pays a negative λ^{Δ^-}, the buy results in a positive opportunity.

The redefined opportunity prices help explain how a strategy can help the revenue of a WPP. Firstly, a WPP experiences opportunities and threats. It seizes an opportunity when it can buy back what it is short at a lower price than it has sold it at on the day-ahead market or when it can sell what it is long at a price that is higher than what it would have sold it at on the day-ahead market. It experiences threats when it buys back at a higher price or sell at a lower price relative to the day-ahead price. The strategy can help the revenue for a WPP in four ways. First, it can reduce its threats when short by being short less and less often when $\lambda^{\Delta^-} \geq 0$. Second, it can seize opportunities by being short more and more often at when $\lambda^- \leq 0$. Third, it can reduce its threats when long by being long less and less often when $\lambda^{\Delta^+} \leq 0$. Fourth, it can increase its opportunities by being long more and more often when $\lambda^{\Delta^+} \geq 0$. The extent to which it reduces threats or increases opportunities can be quantified by computing the opportunity revenues it incurs for being short, long and in total:

$$\begin{aligned} \mathrm{R}^{\Delta^+} &= \sum_{t=1}^{N_t} P_t^{\Delta^+} \lambda_t^{\Delta^+} \\ \mathrm{R}^{\Delta^-} &= \sum_{t=1}^{N_t} P_t^{\Delta^-} \lambda_t^{\Delta^-} \\ \mathrm{R}^{\Delta} &= \mathrm{R}^{+} + \mathrm{R}^{-} \end{aligned} \tag{3.16}$$

Where $P_t^{\Delta^+}$ are imbalance where $P_t^{\Delta} > 0$, $P_t^{\Delta^-}{}_t$ are imbalance where $P_t^{\Delta} < 0$. This can then be split in the volume traded through being short or being long by computing the total volume:

$$\begin{aligned} P^{\Delta^+} &= \sum_{t=1}^{N_t} P_t^{\Delta^+} \\ P^{\Delta^-} &= \sum_{t=1}^{N_t} P_t^{\Delta^-} \\ P^{\Delta} &= \sum_{t=1}^{N_t} P_t^{\Delta} \end{aligned} \tag{3.17}$$

This can then be used to compute the average opportunity price the producer incurs for being long, short and on average, which provides insight in the extent to which it manages to reduce threats and increase opportunities for being short and long seperately:

$$\begin{aligned} \overline{\lambda}^{\Delta^+} &= \frac{\mathrm{R}^{+}}{P^{\Delta^+}} \\ \overline{\lambda}^{\Delta^-} &= \frac{\mathrm{R}^{-}}{|P^{\Delta^-}|} \\ \overline{\lambda}^{\Delta} &= \frac{\mathrm{R}}{P^{\Delta^+} + |P^{\Delta^-}|} \end{aligned} \tag{3.18}$$

If a producer is better able to seize opportunities more and experience threats less, the relationship between its imbalance and its experienced opportunity is positive. When it less able the relationship between its imbalance and its experienced opportunity is negative.

Another relationship that is analyzed is that between the strategy and the system imbalance. The logic behind this is as follows: when the SI is positive, i.e. the system is long, the net regulation volume (NRV) is typically negative relative to this imbalance and vice versa. This means that NRV represents the volume increased or decreased to the grid to offset system imbalances. Coming back to the earlier statement that balancing prices tend to increase relative to the day-ahead price as the NRV becomes increasingly positive and that they tend to decrease relative to the day-ahead price as the NRV becomes increasingly negative, it is likely that there is a clear relationship between system imbalances and balancing prices. Hence, as a forecast's errors strongly correlate with the SI, there is a strong correlation with less favorable previously defined prices. Hence, correlations between forecast errors and the system imbalance are also analyzed, where a lower correlation implies a more beneficial relationship.

The main method for this analysis consists of the evaluation of correlations, which are defined as:

$$\rho_{x,y} = \mathrm{corr}(x, y) = \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y} \tag{3.19}$$

Correlation is a standardized measure, which makes it easy to interpret.

3.4.4. System Effects and System Costs

The last sub question in section 1.2 states the need for insight on the impacts of a strategy on the system as a whole. To this effect, firstly, analyses are carried with respect to the system imbalance. The system imbalance can be influenced in several ways. First, the system imbalance can be dampened by a strategy, which means that on average it helps reduce system imbalance. As the system imbalance is centered around 0MW, this

is analyzed for the absolute system imbalance. As the system imbalance is known, as well as the imbalance caused by the Essent forecast, the system imbalance for each strategy can be directly computed. Then the expected difference of the mean can be computed, using equation 3.12, which allows for the comparison between strategies. Second, the absolute system imbalance, like the revenue, has tails to its distribution. Unlike the revenue distributions, the least favourable values lie in its right tail, as these contain the extreme imbalances, which lead to an increased risk of blackouts. To this end the CVaR_α is computed for the highest α% of system imbalances as follows:

$$\text{CVaR}_\alpha = \frac{1}{n_\alpha} \sum_{i=1}^{n_\alpha} x_i \forall x_i \in \{x_i, x_{i+1}, ..., x_{n_\alpha}\} : x_{i+1} \leq x_i \tag{3.20}$$

Finally, an important issue with system imbalances consists of its volatility, as due to ramp constraints of balancing generators this places a lot of strain on the system. To this end, volatility of the system imbalance at the quarterhourly resolution is an important metric. Volatility in time series is measured by the standard deviation:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \overline{x})^2}{N-1}} \tag{3.21}$$

Concerning system costs, these can be computed as the total opportunity loss of the system as a whole. For this purpose the redefined opportunity prices from section 3.4.1 can be used. For the system as a whole, much of the opportunity cost is offset by positive opportunities, as many parties profit and many parties lose out within a single PTU, the net opportunity loss or gain is determined by the excess system imbalance, which is the sum of all negative and positive imbalances within a PTU. This means that the opportunity for the system as a whole can be computed as:

$$\begin{aligned} \text{R}_{\text{SI}^+} &= \text{SI}^+ \lambda^{\Delta^+} \\ \text{R}_{\text{SI}^-} &= \text{SI}^- \lambda^{\Delta^-} \\ \text{R}_{\text{SI}} &= \text{R}_{\text{SI}^+} + \text{R}_{\text{SI}^-} \end{aligned} \tag{3.22}$$

This means that the opportunity for the system as a whole are computed identically to the opportunities for the WPP. This opportunity is in practice a negative value, as imbalances represent a cost for the system. This results in a sample of opportunity costs for the system as a whole, which are influenced by the strategies through λ^{Imb} and through the system imbalance itself, which means that the opportunity costs can be used to infer the influence of the strategy on the societal costs of system imbalances. For this purpose the expected difference of the mean and CVaR_α are used.

4

Framework for Stochastic Process Modeling

To be able to obtain the maximum value from of a forecast, it first needs to be transformed to a scenario forecast, as this enables the user to apply stochastic optimization [42]. The stochastic optimization models from chapter 3 require a structured scenario set as input, which forms a representation of all stochastic processes involved in the optimization. This chapter aims to provide a framework on the modelling process leading to this scenario set. First, in section 4.1 an oversight is given on the modelling steps involved. Second, in section 4.2.1 the first step is discussed, which is the construction of density forecasts. Third, the second step is discussed in section 4.3, which is the generation of scenarios. Last, the final step is discussed in section 4.4, which is the construction of a scenario tree. The tools used for implementing the models in this chapter are the Statsmodels, Numpy and Scipy libraries in Python.

4.1. High Level Overview

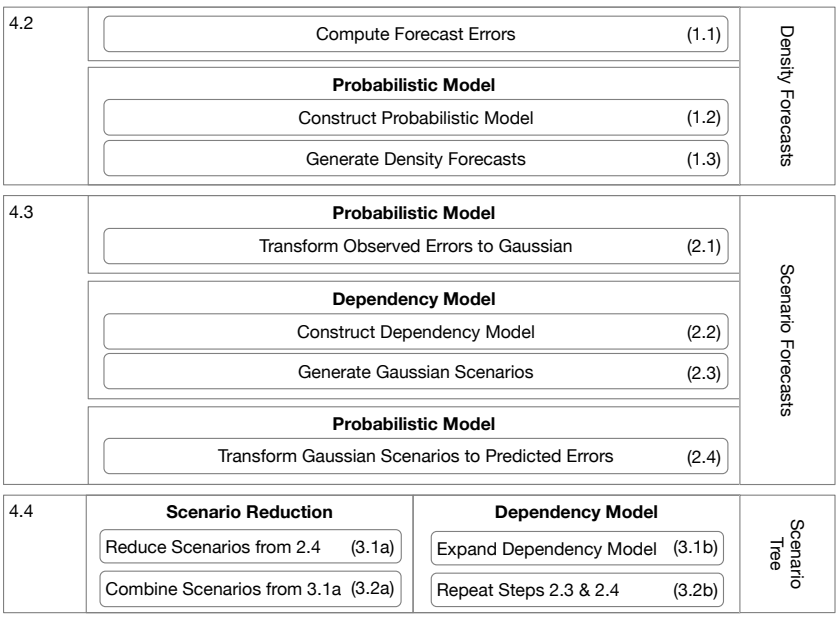


Figure 4.1: Framework for Stochastic Process Modeling. Numbers on the left show which section of chapter 4 discusses the steps. Each step is shown as a rounded rectangle. Where a step is inside a larger rectangle, the model or method involved is shown in bold. To the right of each step, the output of the step is indicated.

Most forecasts are produced as single value forecasts, aiming to provide the best estimate of future values of a stochastic process [15]. However, stochastic optimization in the case of wind power bidding has shown that significant gains can be made when information on the uncertainty of the stochastic processes is accurately captured and utilized, as discussed in section 1.1.2. The method of stochastic optimization requires that these stochastic processes are represented by a set of discrete scenarios, where these scenarios capture how uncertainty evolves in amplitude and between timesteps. In order to move from a point-forecast to such a scenario set, several models need to be applied in several steps, which are shown in diagram 4.1. The first model needed is a probabilistic model, which describes the probability distribution around the forecast, providing the uncertainty in amplitude. The second model needed is a model that captures the interdependence of the uncertainty between timesteps. The third model needed is a model that combines stochastic processes in a scenario tree, which can be directly used in stochastic optimization.

4.2. Construction of Density Forecasts

Regarding the construction of probabilistic forecasts from point forecasts, much research has been carried out in recent years. Methods applied in these studies can be divided in parametric and non-parametric methods. A non-parametric methods makes no assumption on the shape of the underlying distribution, whereas a parametric methods fits a distribution defined by only a few parameters to the data. A recent review on probabilistic wind power forecasting concluded that although parametric approaches can be computationally advantageous, the assumption that the underlying distribution depends only on a few parameters does not seem reasonable in the case of wind power [63]. Furthermore, Pinson and Kariniotakis (2010) concluded in a study on prediction intervals for wind power forecasts that the non linearity of the power curve makes the underlying distribution not subject to any known parameter type distribution, which makes the performance of the non-parametric method in practice superior [49]. Finally, as this study focuses on a single wind farm, computational intensity is less of an issue as a limited number of predictive models is needed. Furthermore, as this framework also aims to construct models for electricity price processes, the added flexibility from non-parametric methods seems appropriate. Hence, focus is turned to non-parametric methods.

Two methods with proven performance in wind power density forecasting are quantile regression and kernel density estimation. The first has been successfully applied to the wind power case by Møller, Aalborg Nielsen and Madsen (2008) [] and to day-ahead electricity prices by Jónsson, Pinson, Madsen and Aalborg Nielsen (2014) [30]. This method consists of approximating a probability density function by constructing a discrete set of quantiles, after which through interpolation a density function can be approximated. Although quantile regression is shown to be successful at density forecasting in these cases, there are some reasons why this method may not be particularly suitable for a generalized framework. First, an arbitrary number of splines can be applied in the regression framework, with varying risk of crossing quantiles, which makes the construction of reliable predictive densities difficult. Second, the choice of the number of discrete quantiles used to approximate the density function determines the resolution of the approximation. When more complex density shapes are approximated, preventing quantile crossing can become difficult as more quantiles are needed to accurately capture the underlying distribution [33]. This may prove troublesome in the case of the imbalance price, which has a more complex underlying distribution. Although the method of quantile regression has shown to be successful in applications in wind power density forecasting and day-ahead electricity price density forecasting, the difficulty of using an arbitrary number of splines and an arbitrary number of discrete quantiles to accurately capture more complex shapes, while preventing quantile crossing, makes it difficultly generalizable.

The second method, kernel density estimation, has been successfully applied to wind power density forecasting by Taylor and Jeon in two studies [29, 56]. In the 2015 study by Taylor and Jeon [56], their kernel density method is tested against the quantile regression method. The main conclusion from this study with regards to their difference is that in terms of performance the two methods are very similar. Also they conclude that although the quantile regression approach is competitive in terms of performance, the issue of crossing quantiles makes it difficult to implement and in need of significant more validation before it can be applied, compared to the kernel density estimation method. On the other hand, the kernel density method can be easily applied to other more complex shapes, as it provides a continuous approximation of the underlying density function, rather than a discrete approximation. Especially when dealing with more complex shapes it can provide well performing density forecasts with little effort. This is important for the current framework, as it should not only capture wind power forecast and day-ahead forecast probability densities, but also those of the imbalance price and the price-maker effect. However, a downside to this method is

the need for many data points, especially when estimating densities in higher dimensions, as data sparsity increases. A downside to the method of kernel density estimation is that the choice of the specific kernel function and the bandwidth of this kernel used to estimate probabilities can lead to different results. However, the more flexible and more easily generalizable method of kernel density estimation seems like a more suitable choice for this framework. Hence, the choice is made to apply the method of kernel density estimation, which is explained in the next section.

4.2.1. Constructing a probabilistic model

The specific method applied here is conditional kernel density estimation (CKD), which was successfully applied by Jeon and Taylor to the case of wind power [29, 56]. This method allows the modelling of a specific dependent variable, conditional on one or more independent variables. The dependent variable here is the forecast error, rather than the observation. This is because the forecast error is an approximately zero mean process and has reduced variance, making it easier to model compared to the observation. This leads to the first transformation, which determines the forecast errors (step 1.1):

$$\epsilon_{t+k|t} = y_{t+k} - \hat{y}_{t+k|t} \tag{4.1}$$

Where index t stands for the time at which the forecast was issued, index k stands for the forecast horizon, y_{t+k} is the observation of process y, $\hat{y}_{t+k|t}$ stands for the forecast of process y issued at time t and $\epsilon_{t+k|t}$ stands for the observed forecast error. The notation $t+k|t$ stands for the fact each forecast value is conditional on the knowledge of the stochastic process up to time t, where the | is the conditional symbol from probability theory.

Although the forecast error has better statistical properties than the target variable, the error data can show varying distributions, which for instance in the case of wind power means that for high or low forecast power the distribution is relatively narrow and skewed, whereas for intermediate forecast power the error distributions are relatively wide and centered. Also, the skewness does not vary smoothly for varying values of $\hat{y}_{t+k|t}$. This is where the flexibility of CKD is beneficial, as a quantile regression model would require an arbitrary number of splines to obtain satisfactory sharpness, with a varying risk of crossing quantiles, making it harder to accurately capture the underlying distribution and to capture it to sufficient detail. As CKD does not have such limitations, the expectation is that with relatively little effort a satisfactory sharpness can be obtained, resulting in a user friendly yet competitive framework.

Regarding the factors that can be used to predict the error, a review on uncertainty analysis of wind power forecasting by Yan et al. (2015) found that wind speed together with the power curve and forecast horizon are the most influential sources explaining the uncertainty [62]. As the forecasts in this study consist of power data, which already incorporates the transformation from wind speed through the power curve, the model is made conditional on the forecast $\hat{y}_t$ and the forecast horizon k. Alternatively, in the case of wind power the model could be applied to wind speed instead of power, as the method is also suitable for stochastic power modelling [56]. However, when using it for stochastic power modelling, binning of data or having variable kernel bandwidths becomes more important, as the spread of power values for varying wind speed values varies greatly [56]. This is less an issue when applying the model directly to power values, as all values fall within a specified domain. A downside to applying it directly to power values is that the uncertainty from power modelling is not captured directly, which seems a likely source for increased sharpness of the probabilistic forecast, which is now not captured.

CKD estimation of the conditional probability density function of the forecast error $\mathscr{E}_t = \epsilon$, given forecast $\hat{Y}_t = \hat{y}$ and time of day $K_t = k$ can be expressed as follows [56] (step 1.2):

$$F(\epsilon|\hat{y}, h) = \frac{\sum_{t=1}^{n} K_{h_\epsilon}(\mathscr{E}_t - \epsilon) K_{h_{\hat{y}}}(\hat{Y}_t - \hat{y}) K_{h_k}(K_t, k)}{\sum_{t=1}^{n} K_{h_{\hat{y}}}(\hat{Y}_t - \hat{y}) K_{h_k}(K_t, k)} \tag{4.2}$$

Where $F(\epsilon|\hat{y}, h)$ is a cumulative density function, n is the sample size and $K_h(\cdot)$ is a kernel function with bandwidth h. Note that for the estimation of the model the forecast horizon index k is excluded, which is because the index is included itself as an independent variable. The kernel used for $\hat{Y}_t$ and $\mathscr{E}_t$ is a Gaussian and the kernel for K_t is the Aitchison-Aitken kernel. The former is also applied successfully in the studies by Jeon and Taylor, while the second is a kernel function specifically designed for discrete variables [2, 29, 56]. The inclusion of the forecast horizon as a discrete variable is an alternative to the binning of data applied by Jeon and Taylor [29, 56]. The former allows for information sharing between time horizons, while binning leads to a strong reduction in data points, which hurts estimation of multidimensional densities, as data

sparsity is an important issue with kernel density estimation. This added dimension from using the forecast horizon as a conditional variable can be seen as applying information sharing between different forecast horizons. For wind power this added dimension does not hurt performance, mostly due to the fact that wind power is a process that has two bounds, 0 and P^{max}. However, in the case of electricity price processes, there are no such bounds, which makes it difficult to achieve reliable predictive densities, as data sparsity is much more an issue with these processes. Hence, for these processes conditional variable K_t may be excluded, resulting in the following model:

$$F(\epsilon|\hat{y}) = \frac{\sum_{t=1}^{n} K_{h_\epsilon}(\mathscr{E}_t - \epsilon) K_{h_{\hat{y}}}(\hat{Y}_t - \hat{y})}{\sum_{t=1}^{n} K_{h_{\hat{y}}}(\hat{Y}_t - \hat{y})} \tag{4.3}$$

The kernel bandwidth was determined by Silverman's rule of thumb, which is common for multivariate kernel density estimation and is given by:

$$\sqrt{h_i} = \left(\frac{4}{d+2}\right)^{\frac{1}{d+4}} n^{-1/(d+4)} \sigma_i \tag{4.4}$$

Where σ_i is the standard deviation of the i^t variable, d is the number of variables, and n is the number of data points.

An alternative to this rule of thumb is to apply cross validation. Using this method, the model is fit to part of the data, after which a metric such as maximum likelihood is computed to determine how well this model fits the remaining data. Such an empirical approach to the selection of model parameters is very flexible, and can be used independent from the underlying distribution of the data, whereas Silverman's rule assumes the variables to be Gaussian. However, using Silverman's rule of thumb, computational gains are very significant while performance is satisfactory. Even in the case where the data are not Gaussian, use of Silverman's rule often results in an accurate representation of the distribution.

4.2.2. Constructing a density forecast

A full description of future values of a stochastic process is given by a density forecast, which is expressed as $\hat{F}_{t+k|t}$, which is issued at time t for forecast horizon $t+k$, conditional on all information up to time t [42]. This means that the model specified in equation 4.2 is estimated using data up to time t, after which it is used to specify future probability density functions for given values for forecast horizon k and forecast value $\hat{y}$. Although the model specified in equation 4.2 can provide a continuous density forecast, for visualization and validation purposes it is common to construct a density forecast as a set of discrete quantiles. To construct the density forecast, the cumulative distribution function $\hat{F}_{t+k|t}$ is used. This is directly provided by the CKD model as the kernel functions are cumulative density functions. A quantile with nominal level α is specified as follows:

$$P[y_{t+k} \leq \hat{q}^{(\alpha)}_{t+k|t}] = \alpha \tag{4.5}$$

Which means that the probability of y_{t+k} being less or equal than quantile $\hat{q}^{(\alpha)}_{t+k|t}$ is α. In order to extract quantiles using the cumulative distribution function $\hat{F}_{t+k|t}$, the following transformation needs to be carried out:

$$\hat{q}^{(\alpha)}_{t+k|t} = \hat{F}^{-1}_{t+k|t}(\alpha) \tag{4.6}$$

Where $\hat{F}^{-1}_{t+k|t}$ is the inverse of the cumulative density function or quantile function, and α is the quantile to be extracted, which in this study range from 0.05 to 0.95 with a 0.05 increment. Using a Beta distribution with $\alpha = 1.5$ and $\beta = 5$, this transformation is visualized in figure 4.2.

To be able to apply the function, the quantile function $\hat{F}^{-1}_{t+k|t}$ needs to be known, which is not the case with the non-parametric CKD model. Hence, this quantile function needs to be approximated. A typical numerical method to do so is by using a root-finding algorithm, which functions based on the following principle:

$$F^{-1}(\alpha) = \inf\{x \in \mathbb{R} : \alpha \leq F(x)\} \tag{4.7}$$

Where the quantile function $F^{-1}(\cdot)$ returns a threshold value for a quantile $0 < \alpha < 1$ below which random draws from the distribution function would fall an α fraction of the time.

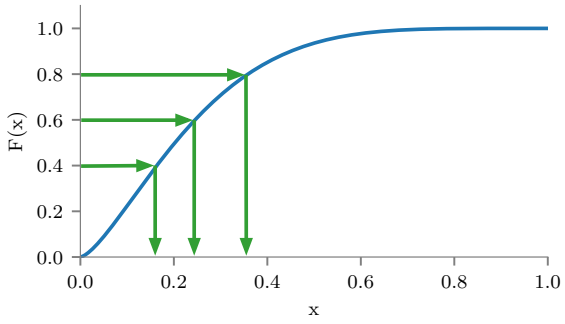


Figure 4.2: Extraction of quantiles for α=0.4,0.6 and 0.8 for a Beta(1.5,5) distribution.

Unfortunately, this method is computationally expensive, especially when applying such a method to the amount of data needed for the construction of scenarios. For instance, in the scenario generation process described later, for each timestep N scenarios are generated, meaning that the method needs to be applied N times, where $N \geq 1000$. Hence, an alternative method is applied, which is based on interpolation and can be evaluated much faster compared to the root finding method. This method approximates $\hat{F}^{-1}_{t+k|t}$ by sampling from $\hat{F}_{t+k|t}$ on a specific interval at a specific resolution. The accuracy of this method is dependent on the resolution of the sampling, which is why a comparison between several different resolutions is made, where the resolution is varied by taking three different fractions of the Silverman's rule of thumb kernel bandwidth h for a sample size of 35,040 samples from the Beta(1.5,5) distribution, which is equivalent to a year of data for the stochastic processes involved in this study. The kernel bandwidth is used as it is a good indicator of the scale and resolution of the distribution estimated by the CKD model. The root finding method used is Brent's method, which finds the root of the function f on the sign changing interval [a , b] and claims guaranteed convergence within the interval, as long as the function can be evaluated [7]. The comparison is depicted in figure 4.3,which shows that varying resolutions yield varying accuracy of the approximation. Although no strict rule exists for the accuracy, given the increase in accuracy between the three resolutions, the resolution chosen is $h/4$. Although still some accuracy is lost, the final accuracy is deemed acceptable, as computational intensity increases strongly with further increases in resolution.

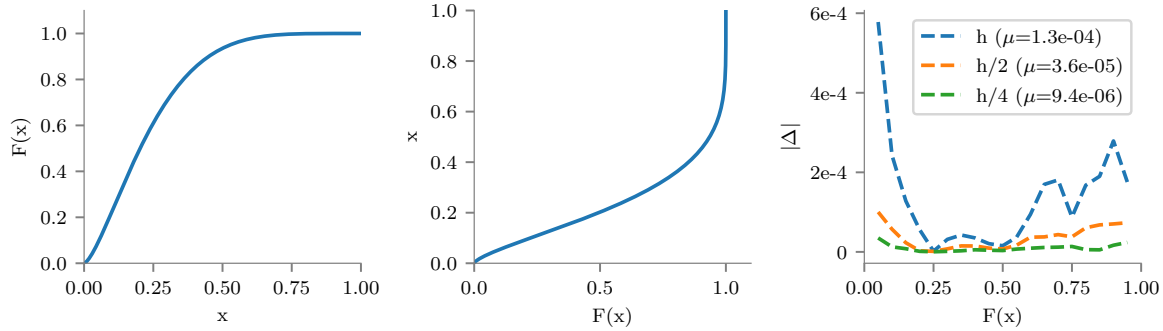


Figure 4.3: a: Cumulative distribution function for Beta(1.5,5).
b: Approximation of quantile function for Beta(1.5,5).
c: Absolute difference between quantiles from root finding and interpolation for varying resolutions as a fraction of the Silverman's rule of thumb kernel bandwidth h. The quantiles range from 0.05 to 0.95 in steps of 0.05.

4.3. Scenario generation

The scenarios used as input to the stochastic optimization models should not only honor the previously determined probabilistic description of the stochastic processes, but also honor the temporal correlation between time steps. This temporal correlation means that if for instance a forecast error is high at time $t+k$,

there is a certain likelihood that it will be high at time $t+k+1$. If this correlation is weak, then forecast errors for future times can be seen as relatively independent from previous forecast errors. However, in the cases of wind power and electricity prices, this correlation is quite high, making it an important characteristic of the stochastic process. The most basic scenarios are in the form of time trajectories. However, one can incorporate dependencies with other stochastic processes as well, e.g. wave-wind, wind-solar or other related trajectories. The function which generates these trajectories is the multivariate random variable Z_t [42]:

$$Z_t = \{Y_{t+k}, k = 1, ..., K\} \tag{4.8}$$

This means that Z_t is a random variable which combines the random variable Y_t at different lead times k and specifies their interdependence structure. J time trajectories can then be issued using model Z_t, at time t for a set of K successive lead times:

$$\hat{z}^{(j)} = [\hat{y}^{(j)}_{t+1|t}, \hat{y}^{(j)}_{t+2|t}, ..., \hat{y}^{(j)}_{t+K|t}]^\top \tag{4.9}$$

This joint predictive density $\hat{F}_{Z_t}$ can be obtained by combining marginal predictive densities with an interdependence structure, where the marginal predictive density is a density for given conditional values of K_t and $\hat{Y}_t$ for the CKD model specified in equation 4.2, or for a given conditional value of $\hat{Y}_t$ for the CKD model specified in equation 4.3. This combining of the marginal densities can be achieved using a Copula model, where a set of marginal predictive density functions $\{\hat{F}_{t+k|t}\}$ for time horizons K are coupled through the Copula model:

$$\hat{F}_{Z_t} = \left\{\{\hat{F}_{t+k|t}\}, C(\delta k)\right\} \tag{4.10}$$

Where the Copula model $C(\delta k)$ provides the interdependence between the different lead times k. This function permits us to define the joint cumulative distribution function $\hat{F}_{Z_t}$ on the condition that each marginal density is calibrated. The specific Copula model applied here is a Gaussian copula, which is the simplest and most convenient one. This model consists of three steps. First, it consists of transforming each variable to uniform space by applying the marginal density functions $F_{t+k|t}$ to the observations of the dependent variable $\epsilon_{t+k|t}$. Second, the uniform random variable $U_{t+k|t}$ is transformed to standard Gaussian using the inverse cumulative density function ϕ^{-1} of a standard Gaussian distribution, which is analytically known. This results in the Copula transformation:

$$U_{t+k|t} = F_{t+k|t}(Y_{t+k}) \sim U[0,1] \tag{4.11}$$

$$N_{t+k|t} = \phi^{-1}(U_{t+k|t}) \sim N(0,1) \tag{4.12}$$

This transformation is visualized in figure 4.4. This shows how using a marginal cumulative distribution function, here represented by a Beta(1.5,5) distribution, can be used to transform observations to their standard Gaussian counterpart. This can be thought of as the removal of the influence of the forecast horizon k and the forecast $\hat{y}$ from the error.

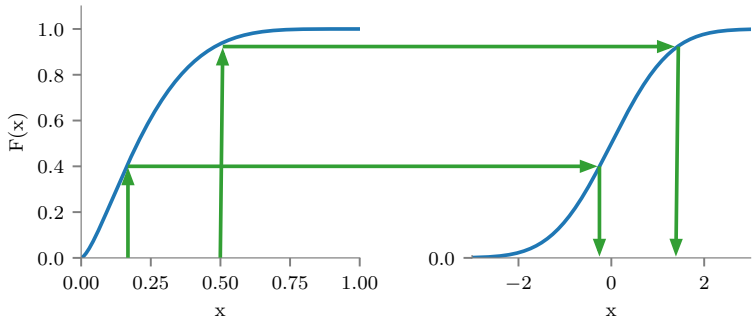


Figure 4.4: a: Cumulative distribution function of Beta(1.5,5).
b: Cumulative distribution function of N(0,1).
The green lines visualize the transformation in equation 4.11

Next, the Gaussian Copula model summarizes the entire interdependence structure in a covariance structure, which is used to define a multivariate normal distribution. This is because a multivariate normal distribution is completely defined by a mean vector and a symmetric covariance matrix. This can also be referred to as a correlation structure, as the variables Y_{t+k} are transformed to $N_{t+k|t}$, which are standard Gaussian variables, which means they all have unit variance. The fact that all variables $N_{t+k|t}$ are standard Gaussian also means that the mean vector of the multivariate normal distribution is a zero vector. This means that the last step in defining the Copula model $\hat{F}_{Z_t}$ is defining the covariance matrix Σ, which describes the relationship between each vector $N_{t+k|t}$:

$$C(\delta k) = \Sigma(N_{t+i|t}, N_{t+j|t}), i, j \in \{1, 2, ..., K\} \tag{4.13}$$

The determination of such a covariance matrix is visualized in figure 4.5. This shows how $N_{t+k|t}$ variables can be related to each other through a covariance structure. On the diagonal, the variance is determined of each standard Gaussian variable, while the other values in the matrix quantify the covariance each standard Gaussian variable has with each other standard Gaussian variable. The figure also shows how the covariance weakens as the difference in forecast horizon increases.

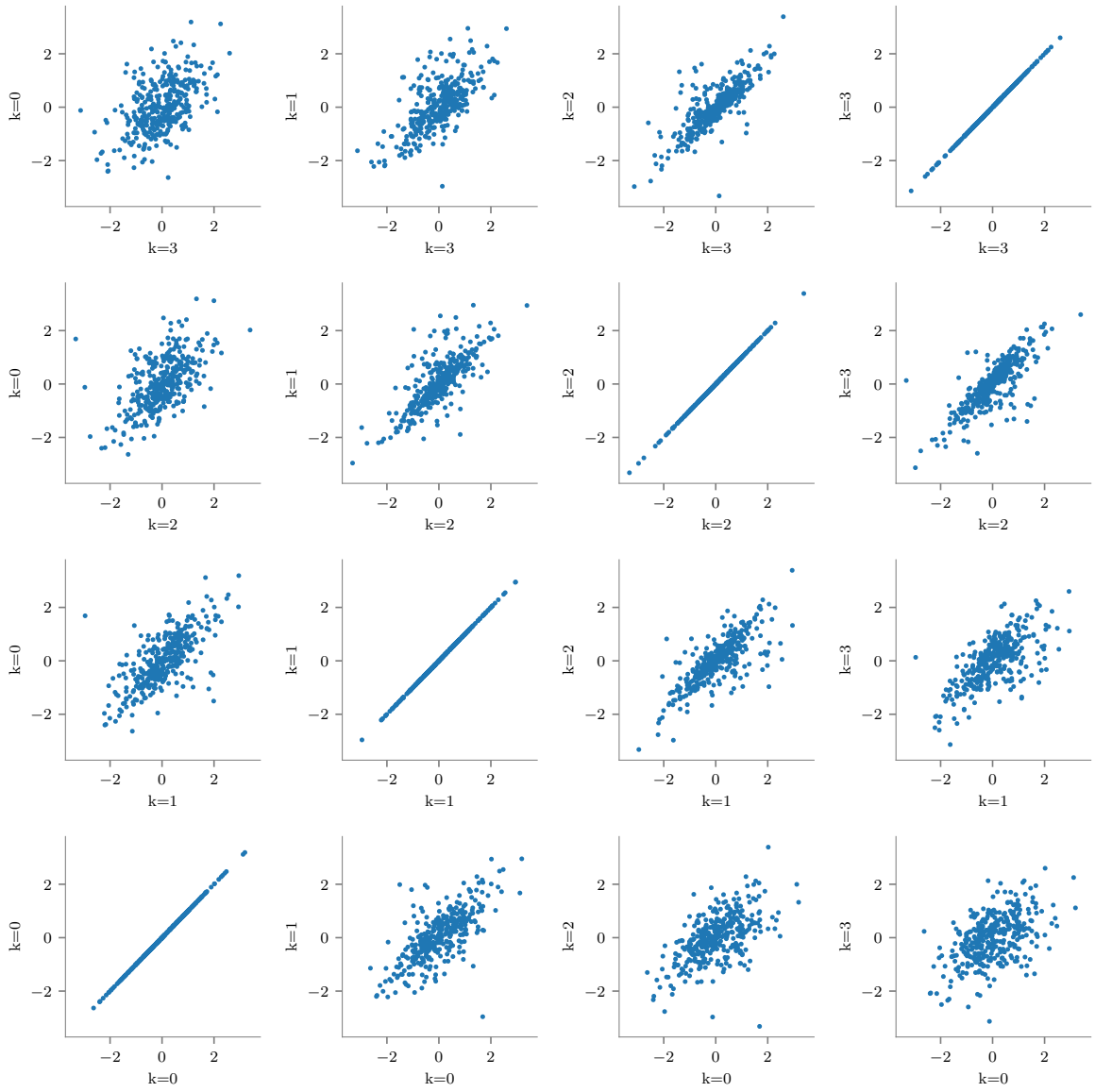


Figure 4.5: Scatter plots of samples of $N_{t+k|t}$ for $k \in 1, 2, 3, 4$. The plots show how the covariance between each variable is related to each other variable.

This correlation structure can easily be extended to cover multiple Gaussian variables, such as different types of renewable generation and different locations, capturing space-time and other dependencies. Moreover, when the dimension is low, this can be done through a single correlation matrix $\Sigma = \rho_{ij}$, where each element ρ_{ij} may inform of interdependence between locations, times and different source of generation. However, as the number of dimensions increases, a more sophisticated covariance model may need to be employed. However, in this study 96 different forecast horizons are used, where the 96x96 matrix consisting of 9126 correlations is computationally insignificant.

Based on the Copula model defined in equation 4.10, scenarios of the stochastic process can be generated. This is done by first defining a multivariate Gaussian random variable Z, with mean vector $\mathbf{0}$ and covariance matrix Σ defined in equation 4.13:

$$Z \sim N(\mathbf{0}, \Sigma) \tag{4.14}$$

This variable Z can then be used to generate J standard normal trajectories, resulting in the set $\{z^j, j = 1, ..., J\}$. Using the standard Gaussian cumulative distribution function ϕ and the marginal inverse predictive cumulative density functions $\hat{F}^{-1}_{t+k|t}$. For every lead time the multivariate Gaussian trajectories can be transformed into trajectories of the dependent variable $\epsilon_{t+k|t}$ through the following transformation:

$$\hat{z}^{(j)} = \hat{F}^{-1}_{t+k|t}(\phi(z^{(j)})), j = 1, ..., J \tag{4.15}$$

Where $\hat{z}^{(j)}$ is a single scenario from Copula model $\hat{F}_{Z_t}$, as defined in equation 4.10. This proces is visualized in figure 4.6, which shows how a single trajectory from a multivariate normal distribution is transformed using marginal densities. The densities shown in the figure are 96 Beta distribution where α varies between 0.5 and 1.5 and β is 5. This transformation needs to be applied to all J trajectories using k approximations of the inverse marginal cumulative density function $\hat{F}^{-1}_{t+k|t}$.

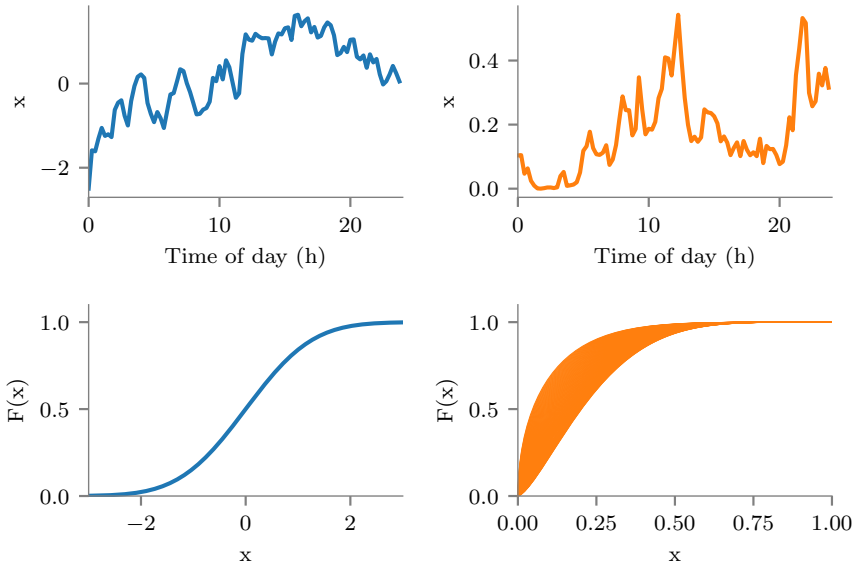


Figure 4.6: a: Gaussian trajectory generated using Z from equation 4.14
b: Transformed trajectory.
c: Corresponding standard Gaussian cumulative density function ϕ.
d: 96 Beta density functions, representing the Marginal density functions. α is varied between 0.5 and 1.5, $\beta = 5$

4.4. Scenario Tree Construction

After scenarios are constructed for all stochastic processes of interest, another step is needed before the scenarios can be used in stochastic optimization. Stochastic processes as input to stochastic optimization are conventionally modeled through a symmetric scenario tree, where each scenario of each process is combined with each scenario of each other process. For S processes with J scenarios, this results in J^S tree scenarios. Conventionally, the first step consists of scenario reduction, where the number of scenarios is reduced to a number that can be optimized in an appropriate execution time, addressing the problem of tractability [43]. After this reduction, the remaining scenarios are structured in a scenario tree. These two steps are discussed in the next section. The section after that presents an alternative method, which preserves the relationships between the four processes using Copula modeling, which does not require scenario reduction or the construction of a symmetric scenario tree.

4.4.1. Scenario Reduction and Tree Construction

In stochastic programming problems, it is possible to reduce the original scenario set to a reduced set, which is close to the original set in terms of a probability distance. Chapter 3 of 'Decision Making Under Uncertainty in Electricity Markets' [10] states that when two sets of scenarios are sufficiently close in terms of probability distance, the optimal value of the solution to the problem from the reduced set is close to the value of the one from the original set. For more information on this relationship and the notion of probability distance, please see [12]. According to Conejo et al. (2010) [10], the most common probability distance used in stochastic programming is the Kantorovich distance. Here a reduced version of the Kantorovich distance is used, which is determined as:

$$D_K(Q, Q') = \sum_{\omega \in \Omega \setminus \Omega_S} \pi_\omega \min_{\omega' \in \Omega_S} v(\omega, \omega') \tag{4.16}$$

Where $D_K(\cdot)$ is the Kantorovich distance, v is a cost function, Q and Q' are two probability distributions, π_ω represents the probability of scenario ω in set Ω, according to the probability distribution Q. The summation in equation 4.16 is performed over the relative complement of set Ω_S with respect to set Ω, also referred to as the difference of sets Ω and Ω_S, which is expressed as $\Omega \setminus \Omega_S$. This is the set of elements in Ω but not in Ω_S. Regarding the Kantorovich distance as defined here, strictly speaking, it can only be called Kantorovich distance if the cost function v is given by a mathematical norm, in this case the l^1 norm. This means that the cost function is specified as:

$$v(\omega, \omega') = \|\omega - \omega'\|_1 \tag{4.17}$$

Where $\|\cdot\|_1$ stands for the l_1 norm, which is computed as:

$$\|\omega - \omega'\|_1 = \sum_{i=1}^{n} |\omega_i - \omega'_i| \tag{4.18}$$

Where n is the number of elements in vector ω. Applying such a metric can be done in several ways, most commonly by applying backward reduction or forward selection, where the former means that the reduced set is built by eliminating scenarios, whereas the latter means that the reduced set is built by selecting scenarios. Due to the amount of variables involved in electricity market problems solved through stochastic programming such as the one here, a strong reduction is needed, in which case forward selection exhibits better performance according to Conejo et al. (2010) [10], which is why the remainder of this section focuses on this heuristic.

This heuristic starts with an empty set, Ω_S, where for each pair of scenarios the cost function $v(\cdot)$ is computed, resulting in a symmetric matrix with a zero diagonal, $\mathbf{V}$. Then a starting scenario is chosen, which is the scenario from which the reduced subset Ω_S is built. The remaining set is depicted by Ω_J. The starting scenario is the scenario which is most equidistant from all other scenarios, in other words: The average scenario. This first step is mathematically obtained for iteration $i = 1$ through evaluating:

$$\omega_i = \arg\{\min_{\omega' \in \Omega} \sum_{\omega \in \Omega} \pi_\omega v(\omega, \omega')\} \tag{4.19}$$

This means that the selected scenario has the smallest summed l_1 norm with all other scenarios as all scenarios are equiprobable, hence the most equidistant in probability distance. After selecting the first scenario, the heuristic continues by iterating over the cost matrix $\mathbf{V}$ containing all values of function v, when for

each candidate scenario ω the corresponding row of the cost matrix is updated, where each existing value of $v(\omega,\omega')$ is compared with its cost to the scenario chosen in the previous iteration, depicted by ω'', thus $v(\omega,\omega'')$, resulting in matrix $\mathbf{V}'$. This means that for each value in the cost matrix the following function is evaluated for the row corresponding to scenario ω:

$$v(\omega,\omega') = \min\{v(\omega,\omega'), v(\omega,\omega''), \forall\omega,\omega' \in \Omega_J \tag{4.20}$$

Then the following equation is evaluated for all columns of the remaining scenarios of set Ω_J:

$$\omega_i = \arg\{\min_{\omega'\in\Omega_J} \sum_{\omega\in\Omega_J} \pi_\omega v(\omega,\omega')\} \tag{4.21}$$

This means that the scenario is chosen which minimizes the Kantorovich distance between the reduced and original set. The reason this technique is a heuristic, is that there is no guarantee that the reduced set is the closest possible set in terms of Kantorovich distance. Also, there is no guarantee that the reduced set gives a good approximation to the optimal value of the optimization problem. The reason this technique is applied is because empirical cases in literature report good performance in practice [13, 23, 44].

This iteration of evaluating equations 4.20 and 4.21 is continued for J iterations. After the orginal set is reduced, the probabilities of the remaining scenarios in set Ω_J need to be transferred to the reduced set. This is accomplished by iterating over the original cost matrix $\mathbf{V}$ and for each remaining scenario's row choosing the chosen scenario's column where the cost is lowest. Doing so for all scenarios can be expressed as:

$$\pi_\omega = \pi_\omega + \sum_{\omega'\in J(\omega)} \pi'_\omega, \forall\omega\in\Omega_S, \text{where} J(\omega) = \{\omega'\in\Omega_J | \omega = j(\omega')\}, j(\omega') \in \arg\min_{\omega''\in\Omega_S} v(\omega'',\omega') \tag{4.22}$$

After scenario reduction, according to convention the scenarios need to be combined in a scenario tree [43]. This means they are arranged in a symmetric tree. The figure below illustrates the way this is done:

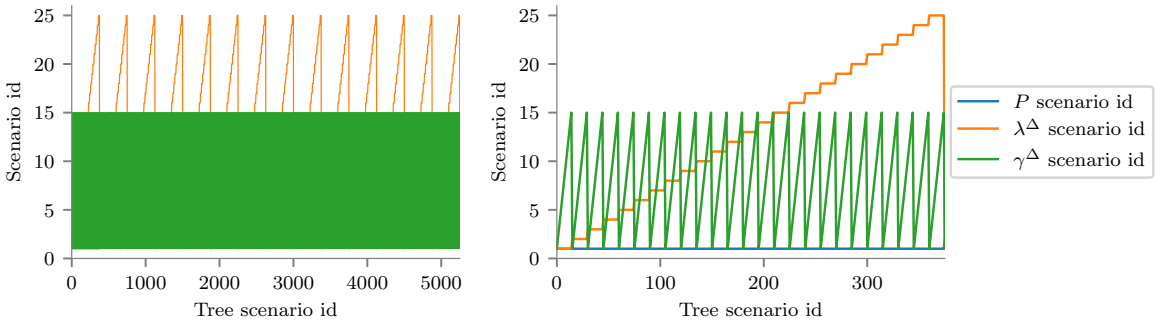


Figure 4.7: a: Scenario tree
b: Zoomed in view of the scenario tree

This shows that for each variable, each scenario is coupled with each scenario of each other variables to form a coupled tree scenario. As each scenario of each variable has a specific probability, these probabilities are recomputed for each scenario in the scenario tree as the product of the three variable scenarios' probabilities in that specific tree scenario. A downside to this way of tree construction is that the underlying assumption is that every combination of the variables is equally probable, where the probability of the tree scenario only says something about the probability of the individual variable scenarios in that tree scenario, not on their combination, thus assuming independence between the processes. Furthermore, as the number of variables dictates the required reduction in order to maintain tractability, much information can get lost from reducing the set. This means that the number of scenarios becomes too small to guarantee a stable solution to the optimization problem. For instance, for the linear version of the optimization model, it was found that the solution to the optimization stabilizes when the set is reduced to 50 scenarios per variable. This was found through experimentally determining the size where the solution from the optimization stabilizes [42]. Lastly, if one would want to incorporate new variables in the optimization model this problem increases further as the number of scenarios in a tree increases exponentially using this technique. For these reasons the next section discusses a method that aims to overcome these problems through the use of an extended Copula model.

4.4.2. Copula Coupled Scenario Tree

Given the reduction of information in the scenario tree, both on individual variables as on their relationship, as independence between processes is assumed, a new method for tree construction is introduced. The assumption of independence seems incorrect as all stochastic processes in the optimization model are coupled through the electricity market. As explained in section 4.2.1, an effective method to couple related processes is through the use of Copula modelling. Chapter 2 of 'Integrating Renewables in Electricity Markets' explains how this coupling is of importance when generating scenarios for multiple sources of interrelated renewable generation, like wind and wave installations, but also when generating scenarios for generation units that have a spatial correlation, like neighbouring wind farms [42]. However, this technique is not limited to these processes. They can also be used to couple other related processes and make sure that their interdependence is respected when generating scenarios. As explained in section 4.2.1, when the dimension of these interdependent processes is low, this can be done through a single correlation matrix. As the price-maker version of the optimization model contains 4 stochastic parameters, of which 3 are of an quarterhourly resolution and one is of an hourly resolution, the resulting matrix is of size $(96+96+96+24)^2 = 97,344$, which is computationally inexpensive to compute. This means the multivariate random variable Z_t is reformulated to include S added random variables for each forecast horizon k:

$$Z_t = \{Y_{s,t+k}, s = 1,...,S, k = 1,...,K\} \tag{4.23}$$

Where index s indicates the stochastic process. For simplicity, all stochastic processes are assumed to have the same number of forecast horizons k. This leads to the following reformulation of the Copula model from equation 4.10:

$$\hat{F}_{Z_{s,t}} = \left\{\{\hat{F}_{s,t+k|t}\}, C(\delta k, \delta s)\right\} \tag{4.24}$$

This means that the extended Copula model only differs from the previously formulated model through the inclusion of all marginal density functions $\hat{F}_{s,t+k|t}$ from all stochastic process s for all forecast horizon k and through the extension of the covariance matrix Σ to include all stochastic processes s:

$$C(\delta k, \delta s) = \Sigma(N_{a,t+i|t}, N_{b,t+j|t}), \text{for } a,b \in \{1,2,...,S\}, \text{for } i,j \in \{1,2,...,K\} \tag{4.25}$$

Where $N_s, t+i|k$ are standard Gaussian variables, obtained through the Copula transformation in equation 4.10. The covariance matrix can then be used to generate J trajectories using Gaussian random variable Z, as defined in equation 4.14. The resulting set of Gaussian trajectories $\{z_{(j)}, j = 1,...,J\}$ can then be transformed through the following transformation:

$$\hat{z}^{(j)} = \hat{F}^{-1}_{s,t+k|t}(\phi(z^{(j)})), j = 1,...,J \tag{4.26}$$

This results in J trajectories where both time dependence as well as dependence between the processes are respected. As J equiprobable scenarios are generated simultaneously with a prespecified ordering through the Copula model, reduction should not be applied, as this would remove the explicit dependencies present in the scenario set. This means no further steps are required, so the scenario set can directly be used as input to the stochastic optimization.

5

Framework for Evaluation of Stochastic Process Modeling

As the solution provided by the stochastic optimization is strongly contingent on the quality of the scenario set, this chapter introduces a framework which can be used to evaluate the inputs, the modeling process and the outputs from the modeling framework introduced in chapter 4. This evaluation framework allows the modeller to make informed decisions on improvements to each stage of the modeling process, with the aim to improve the quality of the modelling. The first section introduces the framework. The second section discusses the quality of point forecasts, which form the inputs to the modelling process. The third section discusses how to assess the reliability of each modeling step, while the final section discusses how to assess the skill of each product from the modeling steps. The tools used for computation in this chapter are the Numpy and Scipy libraries in Python.

5.1. High Level Overview

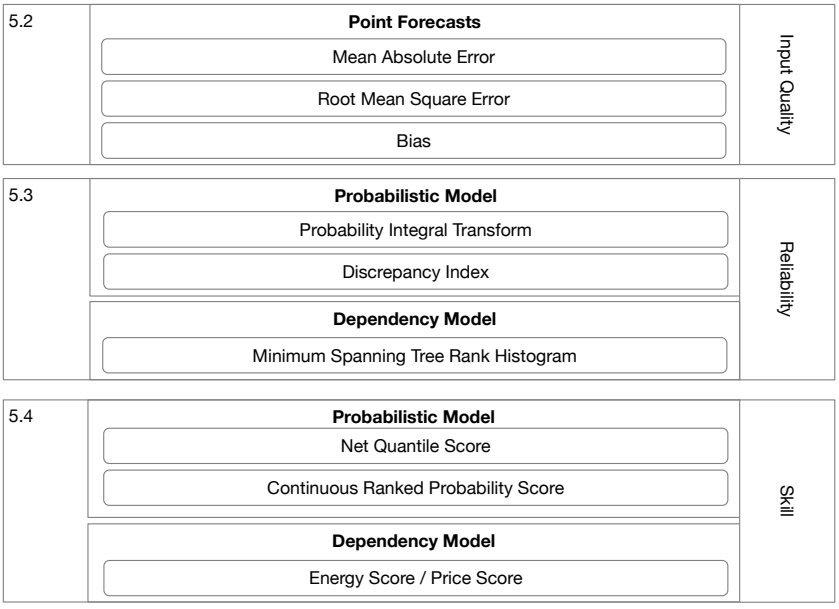


Figure 5.1: Framework for Evaluation of Stochastic Process Modeling. Numbers on the left show which section of chapter 5 discusses the steps. Each step is shown as a rounded rectangle. Where a step is inside a larger rectangle, the model involved is shown in bold. To the right of each step the aspect of evaluation is indicated.

As the previous chapter discussed how to construct scenarios from point forecasts, the first step in evaluating this process starts with evaluating point forecasts. This is a first step which can provide insight in

the statistical properties of each forecast, which can both be used to improve point forecasts and improve the stochastic process modeling. The second step in evaluating the modeling process is the evaluation of its reliability. This needs to be done for each model involved in the process. After the models are found to be reliable, the next step is the assessment of their skill. This gives the modeler the tools to make choices with regards to modeling, as well as which inputs to use for the modeling.

5.2. Point Forecast Statistical Performance

Regarding the statistical evaluation of point forecasts, several research projects have been carried out. One is project ANEMOS, which is a bench-marking exercise carried out in the period 2003-2004, with the aim to analyze the performance of state-of-the-art wind power prediction systems [32]. The main metric used in this study to rank forecasts is the Mean Absolute Error (MAE), which was evaluated as a function of the forecast horizon, enabling evaluation for each time step within the forecast horizon. The second project is the WIRE benchmark, which was carried out in 2015 with the aim to develop a bench-marking platform [54]. Here the main ranking criterion used is also the MAE. Chapter two of the book 'Integrating Renewables in Electricity Markets' presents additional metrics that can be used for the analysis of point forecasts, most notably the Bias and the Root Mean Square Error (RMSE) [42]. As the focus of the analysis in this study lies on performance within the day-ahead window, forecasts are only judged on their performance within this window, even though they are often issued before this window, typically around 9:00AM in the morning before the day-ahead window. This is an important choice, as performance outside this window is not relevant for the optimization. The MAE is specified as follows:

$$\mathrm{MAE}(k) = \frac{1}{T}\sum_{t=1}^{T} |\epsilon_{t+k|t}| \tag{5.1}$$

Which means that for each forecast horizon k, the MAE is determined as the average of the absolute forecast errors over an evaluation set of length T. For instance, for an evaluation set of length 1 year, for each lead time k within the day ahead window, the absolute error is determined, summed and divided by the number of values.

Next is the Root Mean Square Error (RMSE), expressed as:

$$\mathrm{RMSE}(k) = \sqrt{\frac{1}{T}\sum_{t=1}^{T} (\epsilon_{t+k|t})^2} \tag{5.2}$$

Which is defined as the square root of the sum of squared errors over an evaluation set of length T for each forecast horizon k. The RMSE gives higher weight to larger errors, which means it provides insight in the spread of forecast errors, while it is expressed in the original unit.

Next is the bias, which is expressed as:

$$\mathrm{bias(k)} = \frac{1}{T}\sum_{t=1}^{T} \epsilon_{t+k|t} \tag{5.3}$$

Which is defined as the mean of all errors over the evaluation period of length T for each forecast horizon k. This is an important metric, as it provides insight into the existence of a structural bias. Furthermore, for statistical modelling, it is important that in-sample and out-of-sample data are consistent in their statistical properties, especially the bias, as it can introduce a structural error in subsequent transformations. While this also holds for the previous two metrics, it is especially important for the bias. Insight in how these measures evolve over time can thus also help guide modeling choices.

5.3. Reliability of Stochastic Process Modelling

The measure of the performance of a transformed point forecast (to probabilistic form) can be divided into two categories, one related to the probabilistic reliability of the method and one related to the skill of the probabilistic forecast [50]. The former relates to how well the method works, whereas the latter relates to how well the transformed forecast performs overall. The first category is important, as high reliability ensures that the probabilistic forecast is probabilistically calibrated, which is required for preventing sub-optimality when used in optimization. Regarding the second category, high skill means the predictive interval is relatively narrow. This means the interval is more informative. Pinson et al. (2007) [50] point out that reliability

should be the primary requirement of probabilistic forecasts and should be evaluated first. Hence this section firstly discusses how the reliability of predictive densities are determined and secondly how the reliability of scenario forecasts are determined.

5.3.1. Reliability of Predictive Densities

Methods that evaluate reliability of predictive densities are the Probability Integral Transform (PIT), and reliability diagram, which is based on the discrepancy index, which is defined later in this section [45, 50, 58].

Probabilistic calibration or reliability is said to be met if the predictive density $\hat{F}_{t+k|t}$ meets the following requirement:

$$\hat{F}_{t+k|t}(Y_{t+k}) \sim U[0,1] \tag{5.4}$$

Which means that application of the predictive cumulative density function on realizations of the stochastic process it describes should result in uniformly distributed data. This requirement is evaluated through a Probability Integral Transform (PIT), which is typically evaluated using a histogram as depicted below in figure 5.2. The reason the predictive density is applied, which is estimated on data preceding the observation, is because the interest of the modeler should lie on how well the model captures the stochastic characteristics of out-of-sample data, not how well the model describes in-sample data. Ideally the PIT should look like the one in figure 5.2, which shows a sample of size $180 * 24 * 4 = 17,280$ drawn from a standard distribution. The sample has the same number of data points as each stochastic variable in the out-of-sample data set.

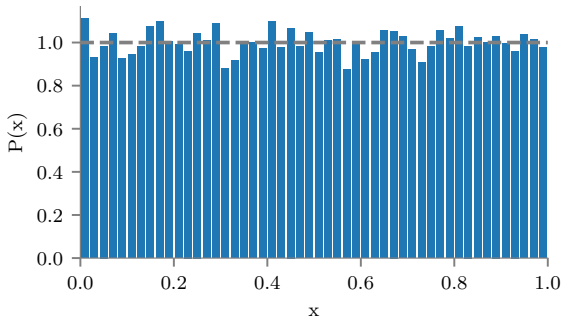


Figure 5.2: Histogram of a sample drawn from a standard uniform distribution.

Although the PIT is a good way of determining overall reliability of a model, it is not easily interpreted. This means it is not easy to use as a tool for identifying where things go wrong. For this purpose, a different metric is introduced: the indicator variable $\xi^{\alpha}_{t,k}$ [42]:

$$\xi^{\alpha}_{t,k} = 1(y_{t+k} < \hat{q}^{(\alpha)}_{t+k|t}) = \begin{cases} 1, & \text{if } y_{t+k} < \hat{q}^{(\alpha)}_{t+k|t} \\ 0, & \text{otherwise} \end{cases} \tag{5.5}$$

Where $\hat{q}^{(\alpha)}_{t+k|t}$ is a specific quantile, as introduced in section 4.2.2, and y_{t+k} is the observation.

This definition means that $\xi^{\alpha}_{t,k}$ is a binary variable, which equates to 1 if the quantile actually covers the measurement or 0 if it does not. This variable allows for the determination of the empirical level $\hat{\alpha}_k$ of quantile α by calculating the mean of the indicator value for a given period:

$$\hat{\alpha}_k = \frac{1}{T} \sum_{t=1}^{T} \xi^{\alpha}_{t,k} \tag{5.6}$$

This allows the calculation of the deviation of each quantile from perfect calibration:

$$\Delta q_m = \alpha_m - \hat{\alpha}_k^{\alpha_m} \tag{5.7}$$

The overall deviation from perfect calibration can then be measures by the discrepancy index $\Delta|q|$ for m quantiles:

$$\Delta|q| = \frac{1}{m}\sum_{i=1}^{m}|\alpha_i - \hat{\alpha}_k^{\alpha_i}| \tag{5.8}$$

Where α_i is between 0 and 1 an indicates what level $\hat{\alpha}_k^{\alpha_i}$ should be. These values are calculated and the empirical versus nominal levels can then be displayed through a reliability diagram, an example of which is shown below in figure 5.3.

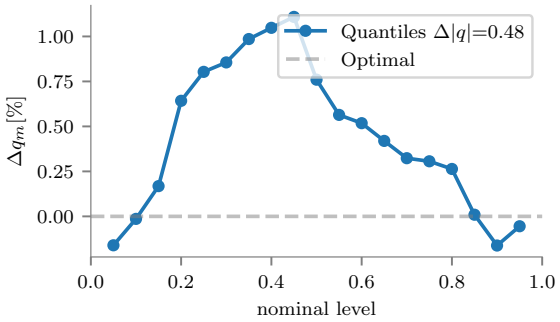


Figure 5.3: Reliability diagram for quantiles of a sample drawn from a standard uniform distribution.

Figure 5.3 displays the empirical quantiles of the same sample used in figure, compared to perfect quantiles for a standard uniform distribution 5.2. Although these are point quantiles, which are determined once for the entire sample, instead of the dynamic quantiles from the predictive densities, where the quantiles consist of one point for each time step in the dataset, it provides the opportunity to illustrate its interpretation: Most quantiles are too high, which means that too many datapoints are below the quantiles. Furthermore, there is a higher deviation for middle quantiles compared to lower and higher quantiles. This indicates that the density is skewed towards higher values.

5.3.2. Reliability of Dependency Models

Next, the reliability of the added modeling step converting densities into scenarios needs to be evaluated. This is done based on the concept of ranks. Gombos (2007) and Gneiting et al. (2008) propose the use of rank histograms for the evaluation of joint probabilities of multiple variables [18, 19]. In the case of the scenarios in this study, the number of variables is 96, as the multivariate model consists of 96 forecast horizons. To judge the calibration of the model, a Minimum Spanning Tree Rank Historgram (MST RH) is used, as introduced by Wilks (2014) [61]. This tool consists of two concepts, the first is the minimum spanning tree. The trees in this context is a vector representing a scenario. To determine the lengths of the trees, the Euclidian distance (l_2 norm) is used, which is the square root of the sum of the squared values of a vector:

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} \tag{5.9}$$

This norm is determined for each scenario $\hat{\mathbf{z}}_t^{(j)}(j = 1,...,J)$ and observation $\mathbf{z}_t$, where each of these vectors contain K elements, where K stands for the forecast horizon. This results in two sets of l_2 norms, those of the scenarios and those of the observations. The next concept in the MST RH is that of the rank histogram.

A rank histogram for the two sets of norms is a histogram where the bins are determined in such a way that the resulting histogram for a sample of observations is perfectly uniform, after which the spread of a sample of *ensembles* can be evaluated against the spread of the observations. In this case this means the sample of l_2 norms of observation vectors $\{\mathbf{z}_t, t = 1,...,T\}$ are used to determine the bins for the rank histogram. This means that the bins are determined based on the empirical quantiles of the observed norms, which are found by evaluating equation 4.7. The rank histogram is then made by evaluating the counts of the *ensemble* within the previously determined bins, where the *ensemble* in this case consists of the sample of l_2 norms of scenario vectors $\hat{\mathbf{z}}_t^{(j)}(j = 1,...,J)$. This concept is visualized in figure 5.4.

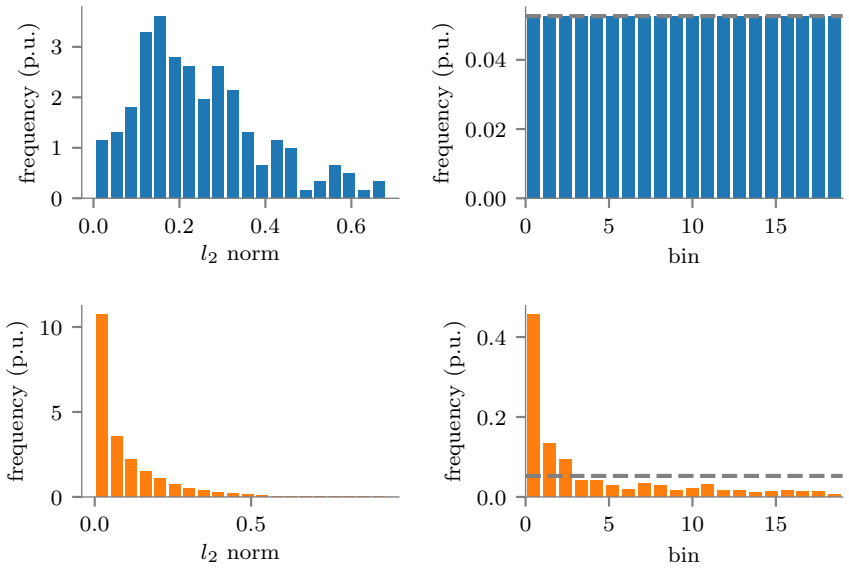


Figure 5.4: a: Histogram of 180 samples from a Beta(1.5,5) distribution, representing observations.
b: Rank Histogram of observed samples.
c: Histogram of 180,000 samples from a Beta(0.5,5) distribution, representing scenarios.
d: Rank Histogram of scenarios.

Figure 5.4a and 5.4c show two distributions, which are distinctly different. In the application here figure 5.4a shows the lengths of observations of a stochastic process, while figure 5.4c shows the lengths of scenarios generated by a stochastic model. Figure 5.4b shows the rank histogram determined by the observation sample shown in figure 5.4a, which shows a perfect uniform distribution, as each bin is determined by the corresponding empirical quantiles for the distribution in figure 5.4a. Figure 5.4d shows the counts of the distribution in figure 5.4c within the bins of the rank histogram. This can then be interpreted as follows: There is a clear downward trend in the rank histogram, which means that on average the scenarios have a lower value for their l_2 norm. This means that on average the scenarios are shorter than the observations, which means that the dependency model *overestimates* the covariance.

After being confirmed reliable, comparative assessment should involve sharpness, which can be evaluated through skill scores.

5.4. Skill of Stochastic Process Modelling

The leading metric for skill is the continuous ranked probability score (CRPS). Another metric often used is the negative quantile score (NQS). The difference between them is that the NQS assesses the quality of the quantiles, which were obtained through equation 4.6, whereas the CRPS assesses the quality of the entire probability function. Hence, the CRPS should be seen as the main ranking criterion. However, for completeness the NQS is also discussed, as some BRP's might only have quantile forecasts available, without the underlying predictive density that generated them.

Hence, the first measure to look at is the negative quantile-based score (NQS). This is defined as [42]:

$$\mathrm{NQS}(k) = \frac{1}{T}\frac{1}{m}\sum_{t=1}^{T}\sum_{i=1}^{m}(\alpha_i - \xi_{t,k}^{(\alpha_i)})(y_{t+k} - \hat{q}_{t+k|t}^{(\alpha_i)}) \tag{5.10}$$

Where $\hat{q}_{t+k|t}^{(\alpha_i)}$ is a set of quantile forecasts with nominals levels $\alpha_i, i = 1,...,m$, y_{t+k} is a set of corresponding measurements and $\xi_{t,k}^{(\alpha_i)}$ is defined in equation 5.5. $\xi_{t,k}^{(\alpha_i)}$ evaluates to one if quantile i at time t,k is above the observations y_{t+k}, so the term within the first set of parentheses is positive if the quantile is lower than the indicator variable. The term within the second set of parentheses is positive if the quantile is lower than the observation. This means that the NQS for a specific horizon is high if many quantiles are wrong, while the extent to which they are wrong is large. This makes it a negatively oriented skill score, where 0 indicates a

perfect quantile forecast.

The CRPS for predictive densities $\hat{F}_t$ and the corresponding measurements y_t is calculated as [42]:

$$\text{CRPS} = \frac{1}{T}\sum_{t=1}^{T}\int_x (\hat{F}_t(x) - H(x - y_t))^2 dx \tag{5.11}$$

Where H is the Heaviside step function, which evaluates to 1 for $x \geq y_t$ and 0 otherwise. The CRPS estimates the area between the predictive cumulative density function and the cumulative density function of the observation, which is represented by the Heaviside function. The computation of the CRPS can be thought of as computing the area between the predictive density function and the density function related to the observation. This is visualized in figure 5.5.

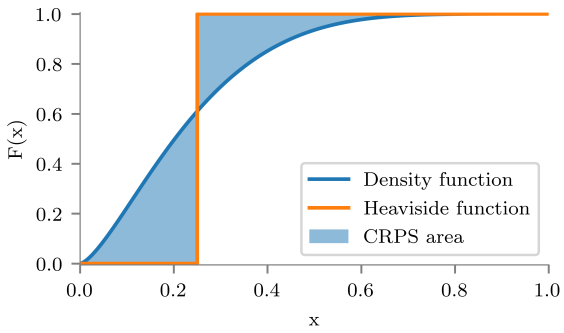


Figure 5.5: Calculation of CRPS for Beta(1.5,5), where $y_t = 0.25$. The red area is is the surface area used to compute the CRPS.

The CRPS is represented in figure 5.5, where the area is indicated in red. This area increases in size when the observation is further from the center of the density function and when the density function itself is less steep. This means that both a good forecast is needed around which densities are constructed, but also a sharp density in order to achieve a good CRPS score. As said, the CRPS assesses the quality of the entire predictive density. As the number of quantiles used to calculate the NQS tends towards infinity, it holds that CRPS=2NQS [42]. This means that if a predictive density has a smooth shape, it can be evaluated well by the NQS. However, when a predictive density has a more complex shape, the difference between the CRPS and NQS can become significant.

To be able to evaluate the skill of the dependency model, a multivariate version of the CRPS is proposed by Gneiting et al. (2008) [18]. This score is referred to as the energy score (Es), which is a direct generalization of the CRPS. The energy score is defined as follows:

For a given set of scenarios $\hat{\mathbf{z}}_t^{(j)}$ $j = 1, ..., J$, issued at time t, the energy score Es_t is given by:

$$Es_t = \frac{1}{J}\sum_{j=1}^{J} \|\mathbf{z}_t - \hat{\mathbf{z}}_t^{(j)}\|_2 - \frac{1}{2J^2}\sum_{i=1}^{J}\sum_{j=1}^{J} \|\hat{\mathbf{z}}_t^{(i)} - \hat{\mathbf{z}}_t^{(j)}\|_2 \tag{5.12}$$

Where $\|.\|_2$ is the Euclidean norm (or l^2 norm), which is often used in the context of probability distance. This score is then calculated and averaged over an evaluation set of length T,

$$Es = \frac{1}{T}\sum_{t=1}^{T} Es_t \tag{5.13}$$

Similar to the CRPS, the energy score is averaged over the T forecast series of the evaluation set. This score is minimal when the trajectories are generated using the true underlying distribution. The Es score is negatively-oriented and it has the same unit as the unit of the trajectories. A lower score on Es means the scenario forecast in question has higher skill. Hence it provides an evaluation tool for improving any part of the forecasting process leading to scenario trajectories. This makes it is a valuable tool in choosing which methods and inputs are to be used to construct the scenarios. In the context of the current study, this means that it can also be used to directly compare different forecasts as a basis for scenarios.

As the current study not only models wind power processes, but also electricity price processes, another metric is introduced. Although the Es was proposed for scenarios of renewable energy processes, there is no

reason why it cannot also be applied to price processes. Hence the price score or Ps is also quantified using equations 5.12 and 5.13.

6

Construction of Price Forecasts

Important inputs to the optimization model are the day-ahead market price λ^D, the balancing market prices λ^+ and λ^- represented by λ^{Imb} and the price-maker effect γ^Δ. For this study no point forecasts were available for these price processes. Hence, the goal of this chapter is to construct point forecasts for each. Firstly, the methodology is discussed based on the literature on forecasting electricity prices. The second section discusses its application to λ^D. The third section discusses its application to λ^{Imb}. The final section discusses its application to γ^Δ. The tools used in this chapter are the Statsmodels, Numpy and Scipy libraries in Python and the forecast package in R.

6.1. Methodology

Different methods exist for forecasting electricity prices, which can be classified into three groups, as explained in a study by Aggerwal, Saini and Kumar (2009) [1], which provides an overview of methods used for this purpose. The first group is based on game theory, which includes models such as Nash equiblibrium, Cournot model, Bertrand model and supply function equilibrium. These models focus on the modeling of strategies of market participants in order to find an optimal solution. The second group is based on simulation methods, which aim to mimic the dispatch of generators in the system, based on the physical state of the system within its physical constraints. The third group consists of time series methods, which utilize past behavior of price series combined with explanatory variables to forecast future prices.

Although each of these methods has its merits, achieving accurate results using the first method requires a deep understanding of the actors and their means and motives. Although the second method can provide detailed insights into the mechanics of the price signal, its implementation is complex and has high computational cost, while it is not able to capture strategic aspects of the market. Lastly, the third group is more straightforward in its implementation, while offering the ability to accurately capture underlying dynamics of the market through explanatory variables. Hence, this study chooses to apply methods from the third group.

Within this third group a distinction can be made between artificial intelligence (AI) and statistical methods. The former have shown to be able to provide an accurate prediction of electricity prices, see for instance the 2014 study by Weron or the 2009 study by Aggerwal et al., which provide an overview of different methods, including indications of their relative performance [1, 59]. Their main drawbacks are that their black-box nature makes them hard to interpret and achieving higher performance compared to their more straightforward statistical counterparts requires considerable effort. As the scope of this study is limited and statistical methods have a proven track record in reliably providing strong performing forecasts, focus is shifted to this second group.

Time series methods have been widely applied in practice and contain among other methods such as autoregressive (AR), moving average (MA), ARMA, and ARMA with exogenous regressors (ARMAX). These method can be extended in multiple ways, including the incorporation of seasonal terms, as well as by combining them with other methods. The advantage of these types of methods is the simplicity and explicitness of their model structure and the proven accuracy of their predictions, see for instance the comparison study by Weron (2014) [59]. Disadvantages are that these type of models require linearity, stationarity, homoscedasticity and normality. The first one relates to the model terms which are restricted to modeling linear relationships, the second refers to the mean of the series, which is required to be stable throughout the time series.

The third relates to the variance of the residuals, which is also required to be stable throughout the time series. The last requirement entails that the residuals should be normally distributed. A typical method to help stabilize series and meet these requirements is by doing an initial transformation before the models are applied. The transformation applied here is the normal inverse transformation using a univariate kernel density model. The kernel density model is constructed as follows [36]:

$$F(y) = \sum_{t=1}^{n} K_h(y - y_t) \tag{6.1}$$

As in section 4.2.1, the kernel bandwidth is determined by Silverman's rule of thumb and the kernel is Gaussian. Using the kernel density model the inverse cumulative distribution function can be determined, as explained in section 4.2.2, after which the following transformation can be applied:

$$U_t = F^{-1}(Y_t) \sim U[0,1] \tag{6.2}$$

$$N_t = \phi^{-1}(U_t) \sim N(0,1) \tag{6.3}$$

After this initial transformation, a method inspired by Joónsson, Pinson, H. Nielsen, Madsen and T. Nielsen (2013) [31] is applied, which is explained in the next section. Although this method is applied to all three series, for balancing prices other statistical models have been applied successfully, for instance through combining SARIMA and Markov proccesses [46]. The study by Weron (2014) [59] contains an extensive overview of different model combinations for different price series. However, because of their proven ability to forecast each of these series successfully also shown in the study by Weron (2014) [59], as well as the goal to provide a single methodology, a single method is applied to all three series.

The model is built up of two different components. Inspired by Jónsson et al. (2013) [31], first a model is constructed that captures the influence of explanatory variables on the price process, which in their model are the day-ahead forecast of aggregate wind and the load. Although the model used by Joónsson et al. [31] captures these influences as nonlinear using a polynomial of order 2, the model applied here is a linear regression model, which is defined as:

$$N_t = \beta_0 + \sum_{0}^{n} \beta_n x_{n,t} + \eta_t \tag{6.4}$$

With intercept β_0, coefficients β_n, explanatory variables $x_{n,t}$ and error term η_t, where a requirement is that $\eta_t \sim N(0,\sigma)$.

Jónsson et al. did not include additional seasonal components in their model, as they state that all seasonality of the price should be captured by the load forecast, which they included as an explanatory variable. However, this study does include additional seasonal components in the regression model in the form of dummy variables for daily seasonality, dummy variables for weekly seasonality and Fourrier series for longer seasonal periods, which are more suitable for longer periods than dummy variables [27]. The rationality behind the Fourrier series is that a series of sine and cosine terms of different frequencies can approximate any periodic function. Hence the following Fourrier terms are included, where m is the seasonal period, which is 1 year:

$$x_{1,t} = \sin(\frac{2 * \pi * t}{m}) \tag{6.5}$$

$$x_{2,t} = cos(\frac{2 * \pi * t}{m}), \tag{6.6}$$

$$x_{3,t} = sin(\frac{4 * \pi * t}{m}) \tag{6.7}$$

$$x_{4,t} = cos(\frac{4 * \pi * t}{m}) \tag{6.8}$$

The inclusion of these seasonal terms allows for identification of the influence of exogenous variables such as the load outside of their seasonal patterns. Several exogenous variables are included in the terms of the linear model. The variables considered are day-ahead forecasts for Belgian aggregate demand, aggregate wind generation, aggregate solar generation and available generation capacity aggregated by fuel type. These forecasts are all published daily on the website of Belgian Transmission System Operator (TSO) Elia. Finally,

dummy variables for holidays and vacation periods are also considered for the regression model. To accommodate such influences these exogenous variables are included in the terms of the linear model. An overview of the explanatory variables considered is given in table 6.1. The daily, weekly, yearly, vacation and holiday variables are considered decomposition variables, which aim to decompose the original signal of seasonal components and trends, revealing the decomposed signal which can then be explained by autocorrelation and exogenous influences [27]. In this study decomposition is included within the regression model with the exogenous variables.

Variable	Number of variables	Domain
Daily seasonality	96	$\in \{0,1\}$
Weekly seasonality	52	$\in \{0,1\}$
Yearly seasonality	4	$\in [-1,1]$
Vacation periods	1	$\in \{0,1\}$
Holidays	1	$\in \{0,1\}$
Aggregate demand forecast	1	$\in \mathbb{R}$
Aggregate wind forecast	1	$\in \mathbb{R}$
Aggregate solar forecast	1	$\in \mathbb{R}$
Generation capacity forecast	1	$\in \mathbb{R}$

Table 6.1: Explanatory variables for regression model

Although the series are stabilized by the transformation by the univariate kernel density model, the explanatory variables are not transformed and contain extreme values, which in ordinary least squares regression can still lead to unstable coefficients. To deal with this issue, a robust linear model is fitted using iteratively reweighted least squares estimation, given a robust criterion estimator, which means that outliers are downweighted, thus reducing their influence. The robust criterion used here is Huber's t, which is based on the median absolute deviation [26]. The advantage of using this metric over the standard deviation is that it is more robust to sample size and the influence of extreme outliers.

Explanatory variables are selected using a process of backward reduction, where the selection criterion is Akaike's Information Criterion (AIC). Significance is not taken into account as a criterion, because statistical significance does not indicate predictive value. Instead, cross-validation or measures such as Bayes Information Criterion (BIC) and AIC are preferred. Here AIC was used as an indicator, which generally is a good indicator whether or not the removal of a variable can increase predictive value, while being computationally advantageous compared to cross-validation [27].

The second component of the model aims to explain the residuals from the previous model, represent by the error term η_t. The error term η_t is modelled as a SARMA process, similar to the approach by Jónsson et al., where the residuals are modelled as a seasonal AR model. Although Jónsson et al. applied robust and adaptive estimation for the model parameters, here a simpler approach was applied: The model is estimated daily with a moving window of one year of data, where model terms are estimated using least squares by minimizing the conditional sum of squares function (CSS), a typical method for this class of model [59]. The model components are chosen once in the initial analysis. An overview of possible model terms is obtained by inspecting the autocorrelation plot and the partial autocorrelation plot. The former shows the correlation the series has with its own lags, the latter shows the correlation series has with its lags, where the influence of the correlation of earlier lags is removed.

After obtaining an overview of possible SARMA model configurations a gridsearch is performed to select final model components. The criterion on which the model components are selected is the bias corrected version of the Akaike Information Criterion (AICc). While cross-validation is generally preferred, the AICc approaches results from using cross-validation while computational gains are significant [27]. The reason no attention was given to statistical significance tests for choosing model orders or components is because statistical significance is generally a poor indicator for out-of-sample performance, as removing terms or variables from linear models based on significance often hurts out-of-sample predictive performance [27]. The equation for this model is as follows:

$$\eta_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t - \sum_{j=1}^{q} \theta_j \epsilon_{t-j} \tag{6.9}$$

With autoregressive parameters ϕ_p and moving average parameters θ_q. ϵ_t is assumed to be an normally

distributed stochastic process with zero mean and variance σ^2, with no significant autocorrelation. Models 6.1, 6.4 and 6.9 are estimated separately, where the kernel density model is estimated first, the robust regression model is estimated second, after which the SARMA-model is estimated on the residuals.

These models are estimated daily using a rolling window of one year of data. Each day the models are estimated, after which one day-ahead forecast is issued at 11:15PM, including data up to 11:00PM. This is because forecasts for aggregate wind and solar are published daily by Elia at 11:00AM. The day-ahead price and the price-maker effect for the current day are known, while the imbalance price is only known up to 11:00PM. The final forecast is obtained by combining the univariate kernel density model from equation 6.1, the regression model from equation 6.4 and the SARMA model from equation 6.9:

$$y_t = F(\phi^{-1}(N_t + \eta_t)) \tag{6.10}$$

This provides the same starting position as for the wind power process, allowing the modelling of the forecast error in chapter 7.

6.2. The Day-Ahead Price

Data for λ^D is publicly available via the website of the European Network of Transmission System Operators (ENTSO-E). The initial in-sample data used for initial estimation of the models consists of 2016 data for all exogenous variables and λ^D. The out-of-sample data consists of values from January 2017 up to and including June 2018. The original time series for the in-sample period for λ^D is depicted below in figure 6.1a. The first thing to note is that although the mean seems relatively stable over time (no overall growth or decrease), the variance seems to be quite heteroscedastic, meaning the variance varies over time. This is a problem for a linear model, as this leads to unstable model parameters. Secondly, there are quite some extreme values, which are also problematic for linear models as this skews results.

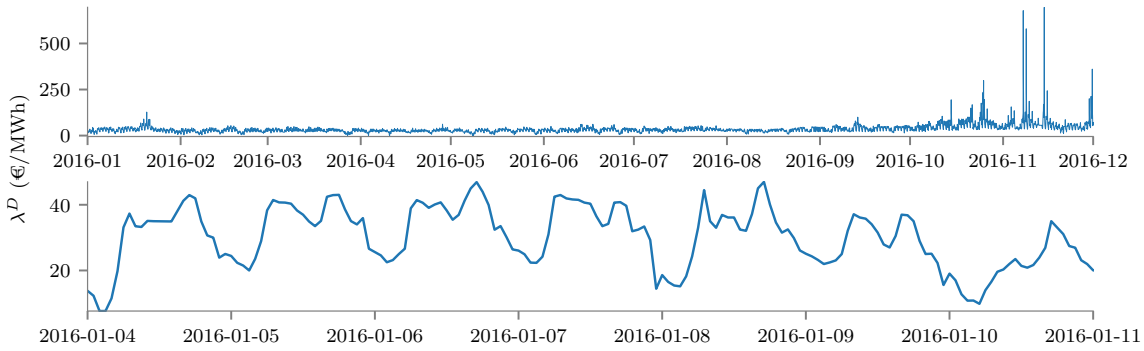


Figure 6.1: a: Belgian λ^D time series for 2016.
b: Zoomed-in view of λ^D

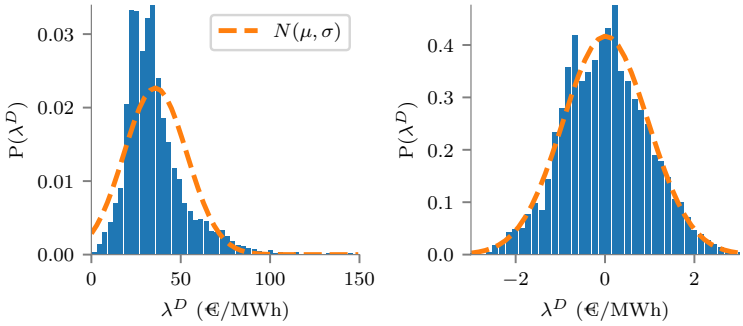


Figure 6.2: a: Histogram of λ^D
b: Histogram of normalized λ^D

Shifting attention to a zoomed-in version of the original time series, visible in figure 6.1b, additional information can be obtained. Clear seasonal patterns can be observed, both daily and weekly. The weekly pattern is best visible in the last two days in this figure, which are part of the weekend, where there is a clear dip in the price. Also, a clear typical peak can be observed for Saturday evening. A daily pattern can also be observed. On average during the night there is a strong decrease in price due to lower demand and relatively high inexpensive wind output. In the morning there is a strong peak and during the day the price levels off only to increase again for the evening peak. This midday leveling off is among other things a combination of inexpensive solar output rising, along with the effect of many people residing together at work, reducing household demand. From these initial observations it seems reasonable to hypothesize that seasonality as well as exogenous variables like aggregate demand and aggregate intermittent renewable output may have an impact on the price.

The first modeling step is to carry out an initial transformation on the data. Looking at the distribution of the values in figure 6.2a, several observations can be made. Firstly, there seem to be long thick tails and signficant outliers, consistent with the series in figure 6.1. Secondly, although not a strict condition for a linear model, the distribution seems skewed. Application of the normal transformation using univariate kernel density estimation corrects for these, as visible in the transformed distribution in figure 6.2b.

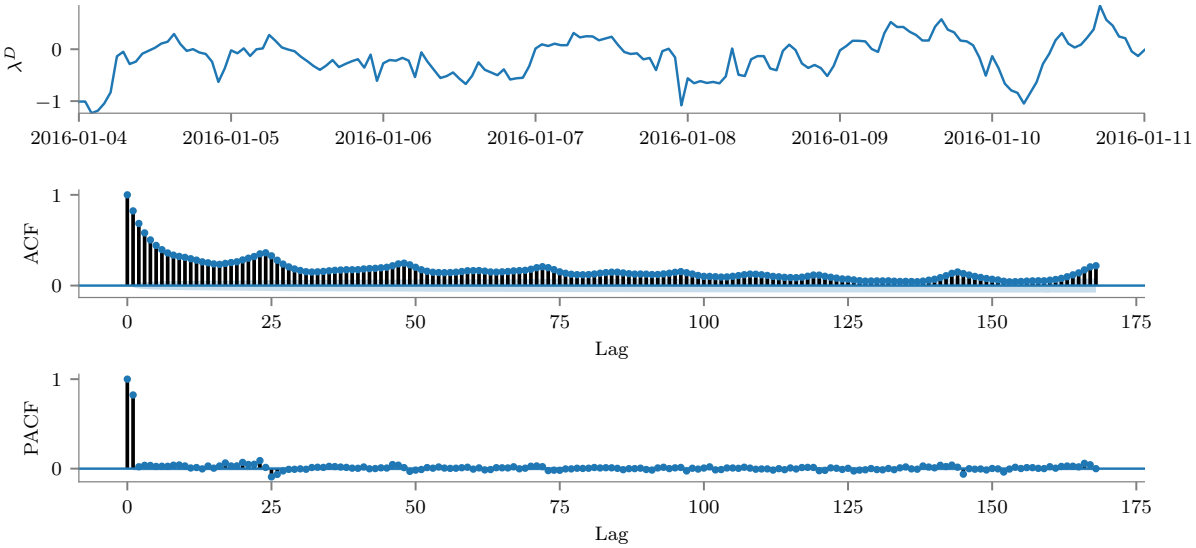


Figure 6.3: a: Residuals from robust regression for λ^D
b: Autocorrelation plot residuals
c: Partial autocorrelation plot residuals

The next modeling step is fitting the robust regression model, as explained in the methodology section. The coefficients for the variables can be found in appendix A.1. This table shows that some variables are not significant. However, as focus is on the predictive quality of the model, backward reduction based on AIC is applied. This heuristic removed three hourly dummy variables. Next, the SARMA model is specified and fitted.

Looking at the series itself in figure 6.3a, there is clearly some structure remaining, which is partly explained by the autocorrelation and partial autocorrelation of the residuals. Starting with the autocorrelation in figure 6.3b, there is a lot of signficant autocorrelation present which tails off during the first day, after which there is a strong peak at the 24-hour mark. Furthermore, many peaks can be observed at every multitude of the 24-hour lag. This suggests either a SARMA or an AR model. Moving to the partial autocorrelation function, depicted in figure 6.3c, a significant peak can be observed at lag 1 and less so around the 24-hour lag. Furthermore, there seem to be multiple significant lags around the 24 hour lag. These observations combined imply at least one AR component combined with at least one seasonal AR and MA component. To select the final model components a gridsearch is carried out, choosing a model by adjusted AIC (AICc). This search provided a SARMA model with specification $(3,0,1)(1,0,1)_{24}$, meaning 3 AR, 1 MA, 1 seasonal AR and 1 seasonal MA component are included in the model.

An example of a forecast is shown in figure 6.4. This figure shows how each step of the model adds predictive ability to the forecast. Firstly, the partial regression model, which only contains the decomposition

variables, already captures much of the price behavior. Secondly, the complete robust regression model provides additional predictive ability by including the influence of explanatory variables. Thirdly, much remaining structure is captured by the SARMA model, which further increases the predictive ability. Although these observations are made for this single day, the specific performance improvements also hold for the whole out-of-sample dataset.

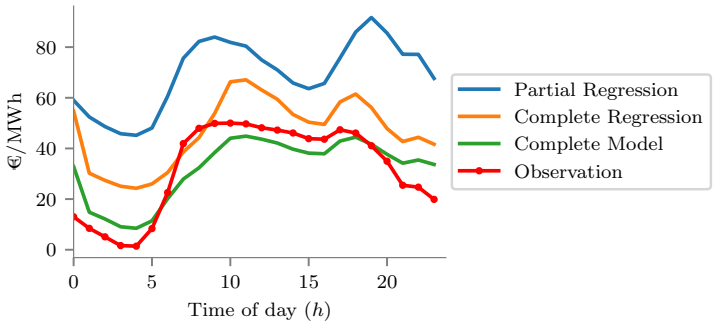


Figure 6.4: Forecast of λ^D for 2018-01-02. The partial regression is a robust regression model using only the decomposition variables. Complete regression includes all variables. Complete model includes the SARMA model.

6.3. The Imbalance Price

As explained in section 3.1 the imbalance price is considered a single price as λ^+ and λ^- hardly ever differ in the Belgian market. Hence the imbalance price without penalty α is chosen to continue the analysis with. As α is never applied to λ^+ in the train or test sample, λ^+ is chosen to represent the imbalance price in this study.

A second choice regarding the imbalance price relates to λ^D. If a market player participates on the day-ahead market and is then asked to provide positive regulation volume, it makes sense that this player would need to be payed a price that is higher than the price it received on the day-ahead price, as it would have offered price quantity pairs according to the marginal prices in his portfolio. In other words: Resources that are left after settling the day-ahead market are more expensive. Thus, if there is a postive NRV, λ^+ would be higher than λ^D and vice-versa. Furthermore, when more expensive resources are activated through the settlement of the day-ahead market, downward regulation bids would require a higher price for not having to generate. Because of this hypothesized shift in the imbalance price by the day-ahead price, the difference price λ^Δ is defined as:

$$\lambda^\Delta = \lambda^+ - \lambda^D \tag{6.11}$$

This difference price is chosen over directly modelling λ^+, because much of the reserve generation is contracted after the day-ahead market is closed, indicating a strong dependency of λ^Δ on λ^D. Data for the imbalance price for Belgium is publicly available through the website of the Belgian Transmission Systems Operator (TSO) Elia. Next, the predictive model is specified for λ^Δ. Again, the in-sample data consists of 2016 data for all explanatory variables and λ^Δ. The test sample consists of values from January 2017 up to and including June 2018. The original time series for the training period for λ^Δ is depicted in figure 6.5a. The first thing to note is that although the mean seems relatively stable over time, the variance seems to be quite heteroscedastic. This is a problem for a linear model, as this leads to unstable model parameters. Secondly, there are quite some extreme values, which are also problematic for linear models as this skews results.

Shifting attention to a zoomed-in version of the original time series, visible in figure 6.5b, additional information can be obtained. Unlike with λ^D, no clear seasonal patterns can be observed, daily nor weekly. However, this does not mean no seasonal pattern exists. For instance, as there are multiple ramps in demand, most notably during the morning and evening, there is often a distinct saw-tooth pattern in the imbalance price. This is caused by the difference in resolution between the day-ahead market and the balancing market, which causes generation profiles of BRPs to be balanced at an hourly resolution but not at a quarter hourly resolution. An example of this effect is shown in figure 6.6a: As electricity is bought to cover demand at an

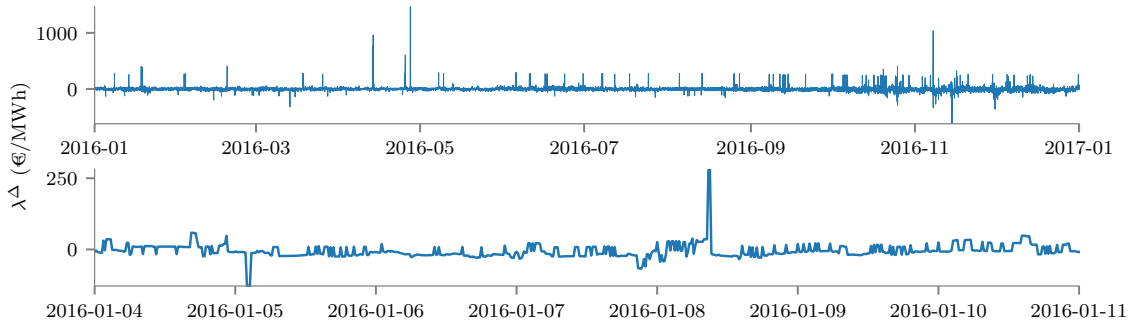


Figure 6.5: a: Belgian λ^Δ time series for 2016. b: Zoomed-in view of λ^Δ.

hourly resolution, due to the upward ramp the market is collectively long during the start of each hour, while the market is collectively short near the end of each hour, which leads to a saw-tooth pattern in the system imbalance, which leads to an opposing NRV and a saw-tooth pattern in the imbalance price, which is shown in figure 6.6b.

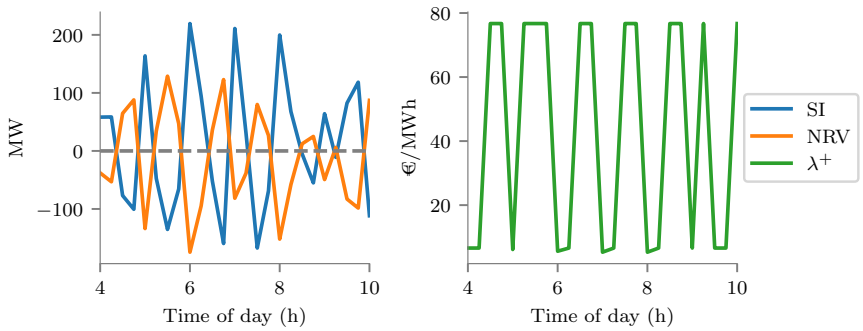


Figure 6.6: Illustration of saw-tooth patterns in the balancing market for 2018-01-03.
a: System Imbalance (SI) and Net Regulation Volume (NRV).
B: λ^+.

Figure 6.6a clearly shows a repeating hourly pattern. Such seasonality can be captured by the robust regression model. Furthermore, although visually not much else can be concluded, the model fitting process will reveal if value can be obtained from seasonal or exogenous variables. Hence it still seems reasonable to hypothesize that seasonality as well as explanatory variables like aggregate demand and aggregate intermittent renewable output have an impact on the price.

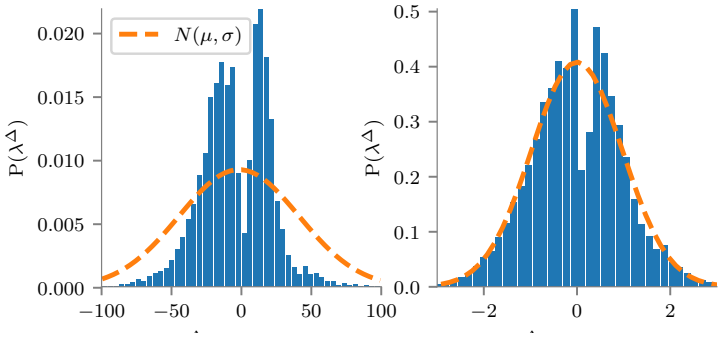


Figure 6.7: a: Histogram of λ^Δ
b: Histogram of normalized λ^Δ

The first step in the modelling process is again to carry out an initial transformation on the data. Looking at the distribution of the values in figure 6.7a, there seems to be a clear bimodal distribution. Furthermore, long tails seems present, which may lead to unstable model parameters. The kernel density model is able to help with the long tails, but some slight bimodality is still present in the transformed data, as shown in figure 6.7. This is mainly due to the fact that Silverman's rule of thumb used to determine the kernel bandwidth assumes a Gaussian distribution, which this clearly is not. Although the resulting distribution is not perfectly Gaussian, it is more suitable for a linear model than the untransformed data.

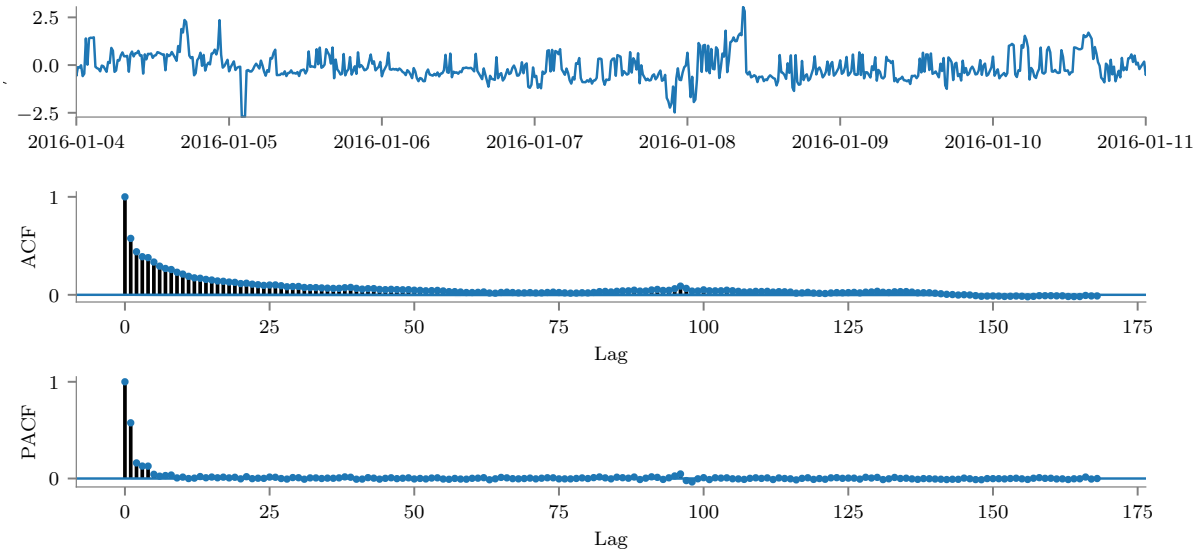


Figure 6.8: a: Residuals from robust regression for $^{\Delta}$
b: Autocorrelation plot residuals
c: Partial autocorrelation plot residuals

The next step is fitting the robust regression model. Based on the backward reduction process seven hourly dummy variables were removed. The model components are listed in appendix A.2. Next, the SARMA model is specified and fitted.

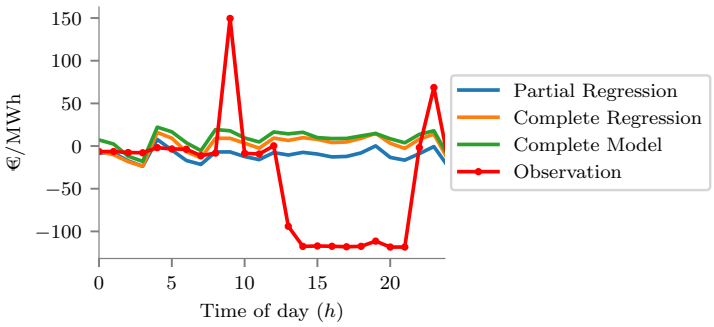


Figure 6.9: Forecast of λ^{Δ} for 2018-01-02. The partial regression is a robust regression model using only the decomposition variables. Complete regression includes all variables. Complete model includes the SARMA model.

Looking at the series itself in figure 6.8a, there is still some structure present, which is partly explained by the autocorrelation and partial autocorrelation of the residuals. Starting with the autocorrelation in figure 6.8b, there is a lot of signficant autocorrelation present which tails off during the first day, after which there is a strong peak at the 24-hour mark (lag 96). Furthermore, one other peak can be observed at lag 192. This suggests either a SARMA or an SAR model. Moving to the partial autocorrelation function, depicted in figure 6.8c, significant peaks can be observed at lags 1,2,3,4,5 and 6 and less so around the 24-hour lag. Furthermore, there seem to be negative lags and multiple significant lags around the 24-hour lag. These

observations combined imply at least one AR component combined with at least one seasonal AR and MA component. Although these plots provide an indication what kind of model may be suitable, for forecasting purposes it is more appropriate to apply a gridsearch, choosing a model by adjusted AIC (AICc). This search provided a SARMA model with specification $(4,0,1)(1,0,1)_{96}$, meaning 4 AR, 1 MA, 1 seasonal AR and 1 seasonal MA components

An example of a forecast is shown in figure 6.9. This figure shows the influence of each of the modelling steps. It also shows how λ^{Δ} can shift strongly between time periods. The forecast is not very similar to the observation, as this shifting is rather symmetrical, as it occurs in both directions and the linear model forecasts the mean value.

6.4. The Price-Maker Effect

As explained in section 2.3, a wind farm of this size is in fact a price-maker and not a price-taker on the Belgian imbalance market. Hence, it makes sense to include a variable which is able to capture the effect a wind farm has on the imbalance price. Conejo et al. (2010) [10] describe the price-maker effect as the slope of the inverse ordered supply curve to the balancing market. To estimate this slope, this supply curve needs to be known. Fortunately, an approximation of this supply curve in 100MW blocks is published by Elia. This curve goes from -800MW up to 800MW in regulation capacity. However, as shown in figure 6.10, the system imbalance rarely exceeds the boundaries of -500MW and 500MW. For the entire dataset this occured only 0.7% of the time. For this reason only the part of the supply curve between -500MW and 500MW was used to determine the slope of the inverse ordered supply curve. This is illustrated through an example using the same supply curve as in figure 3.3, which is shown in figure 6.10b.

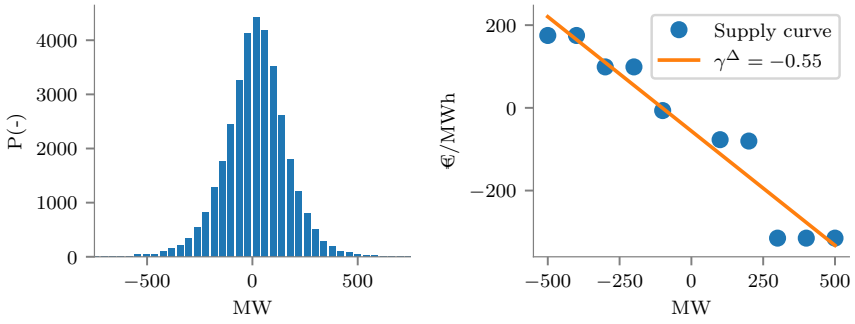


Figure 6.10: a: Histogram of system imbalance for 2016.
b: Inverse ordered supply curve with linear regression for 2018-01-01 0:00-0:15.

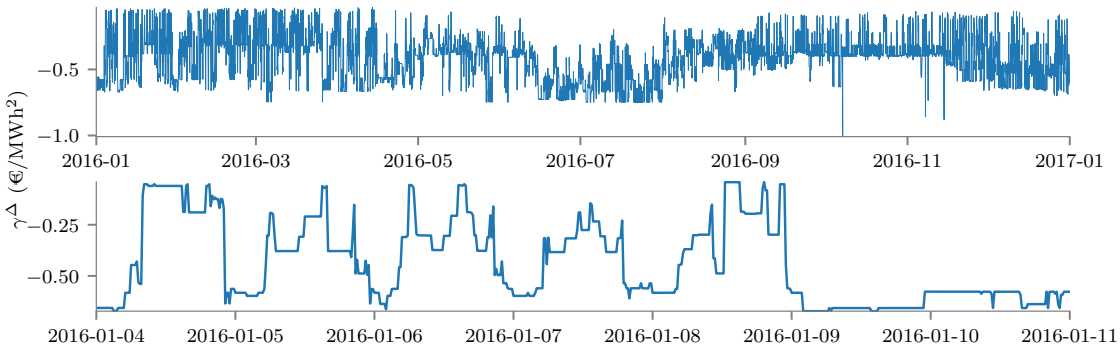


Figure 6.11: a: Belgian γ^{Δ} time series for 2016. b: Zoomed-in version of γ^{Δ} time series.

Next, the predictive model is specified for γ^{Δ}. Again, the initial in-sample data consists of 2016 values for all explanatory variables and γ^{Δ}. The out-of-sample data consists of values from January 2017 up to and including June 2018. The original time series for the training period for γ^{Δ} is depicted below in figure 6.11a. The first thing to note is that the mean seems relatively unstable over time, while the variance seems to be quite

heteroscedastic, which again is a problem for the linear model, as this leads to unstable model parameters. Secondly, there are quite some extreme values, which are also problematic for the linear model as this skews results.

Shifting attention to a zoomed-in version of the original time series, shown in figure 6.11b, additional information can be obtained. A clear seasonal pattern can be observed, mostly daily. The can be explained as on average during the night there is a strong decrease in demand, leaving more power available for balancing services. Although the seasonality seems less pronounced than was the case for λ^D, still it seems reasonable to hypothesize that seasonality as well as exogenous variables like aggregate demand and aggregate intermittent renewable output have an impact on γ^Δ, which will be tested through the regression model.

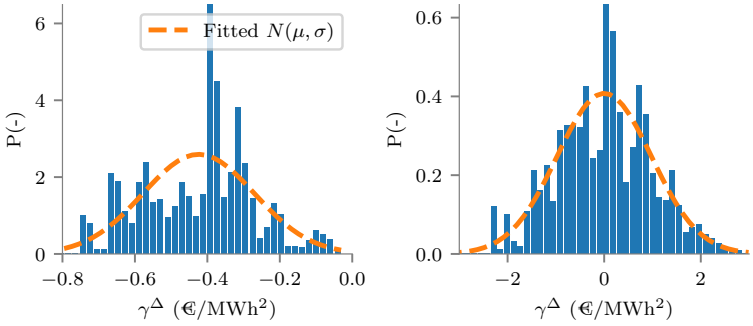


Figure 6.12: a: Histogram of γ^Δ.
b: Histogram of normalized γ^Δ.

The first step is again to carry out an initial transformation on the data. Looking at the distribution of the values in figure 6.12a, there seems to be no clearly defined distribution. As the distribution is also far from a normal distribution, application of the inverse normal transformation using univariate kernel density estimation seems appropriate. As with the imbalance price model, the kernel density model is not able to capture the underlying distribution well, resulting in a non perfect Gaussian distribution in figure 6.12. However, as the transformation results in a more stable mean, more homoscedastic variance and fewer outliers, the transformed dataset is used.

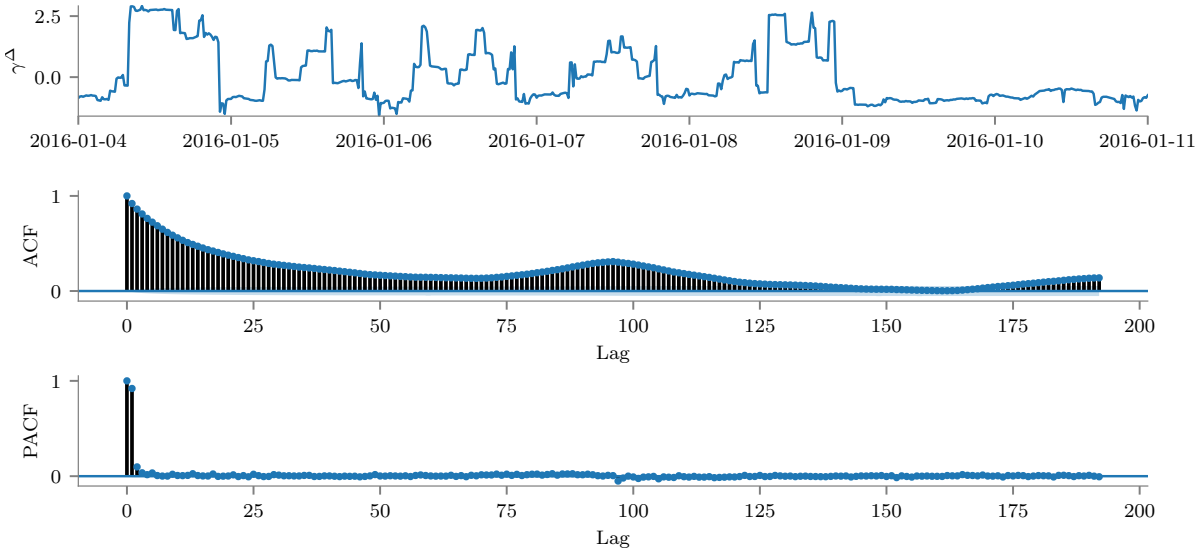


Figure 6.13: a: Residuals from robust regression for γ^Δ
b: Autocorrelation plot residuals
c: Partial autocorrelation plot residuals

The next step is fitting the robust regression model. Based on the backward reduction process two types of variables are removed, which are the wind power capacity forecast and three hourly dummy variables. A

list of the model components can be found in appendix A.3.

Looking at the residual series in figure 6.13a, there is clearly still much structure present, which is partly explained by the autocorrelation and partial autocorrelation of the residuals. Starting with the autocorrelation in figure 6.13b, there is a lot of signficant autocorrelation present which tails off during the first day, after which there is a strong peak at the 24-hour mark (lag 96). Furthermore, one other peak can be observed at lag 192. This suggests either a SARMA or a SAR model. Moving to the partial autocorrelation function, depicted in figure 6.13c, significant peaks can be observed at lags 1, and 2 and less so around the 24-hour lag (lag 96). These observations combined imply at least one AR component combined with at least one seasonal AR and MA component. Although these plots provide a clear indication what kind of model may be suitable, for forecasting purposes it is more appropriate to apply a gridsearch, choosing a model by adjusted AIC (AICc). This search provided a SARMA model with specification $(4,0,1)(1,0,1)_{96}$, meaning 4 AR, 1 MA, 1 seasonal AR and 1 seasonal MA component.

An example of a forecast is shown in figure 6.14. This figure shows how each step of the model adds to the forecast. While less pronounced than with λ^{Δ}, the observation shows a strong shift between one time period to the next. Furthermore, this particular day does not show clear signs of seasonality. Lastly, each model again shows a particular distinct forecast, where the more complex models for this day seems to outperform the less complex models, which is also true for the out-of-sample set.

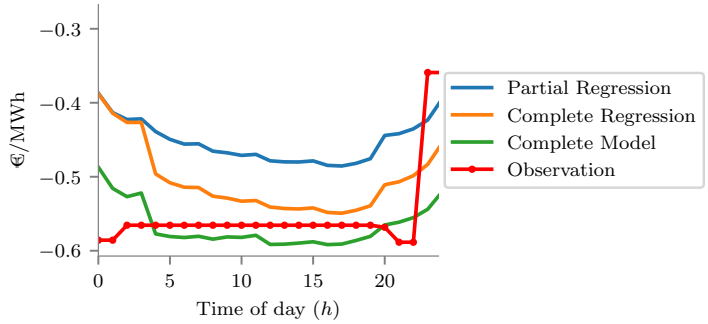


Figure 6.14: Forecast of γ^{Δ} for 2018-01-02. The partial regression is a robust regression model using only the decomposition variables. Complete regression includes all variables. Complete model includes the SARMA model.

7

Results Stochastic Process Modeling

This chapter takes the frameworks from chapters 4 and 5 and applies it to wind power forecasts and the price forecasts generated in chapter 6. First the three wind power forecasts are discussed in section 7.1. Second, the day-ahead price forecast is discussed in section 7.2. Third, the imbalance price forecast is discussed in section 7.3. Fourth, the price-maker effect forecast is discussed in section 7.4. Last, construction of the scenario tree is discussed in section 7.5

7.1. Wind Power

As mentioned in section 1.2, three different wind power forecasts are analyzed and modelled. This provides the opportunity to show how the evaluation framework presented in chapter 5 can be used to make choices with regards to inputs to the stochastic process modelling. The three forecasts being compared are one provided by Essent and two provided by Whiffle, which are forecasts generated by a higher resolution NWP model, based on Large Eddy Simulation (LES) [53]. The first Whiffle forecast consist of the raw output out of their own NWP model, while the second combines the first with information from several different sources to provide a statistically optimized version. Below these three forecasts are analyzed on their statistical performance. Note that power values are normalized by the nominal power of the park, which is 147.6MW (24 turbines of 6.15MW each). This is done because it allows for comparison of forecasts between different park setups, independent of the installed capacity, increasing interpretability. First, the statistical performance of the different point forecasts is analyzed in section 7.1.1. Second, the stochastic models are discussed in section 7.1.2. Last, their performance is evaluated in section 7.1.3.

7.1.1. Wind Power Point Forecast Performance

First, the point forecasts are judged on the main ranking criterion, the Mean Absolute Error (MAE) as presented in equation 5.1. Looking at the results for the three forecasts in figure 7.1, several conclusions can be

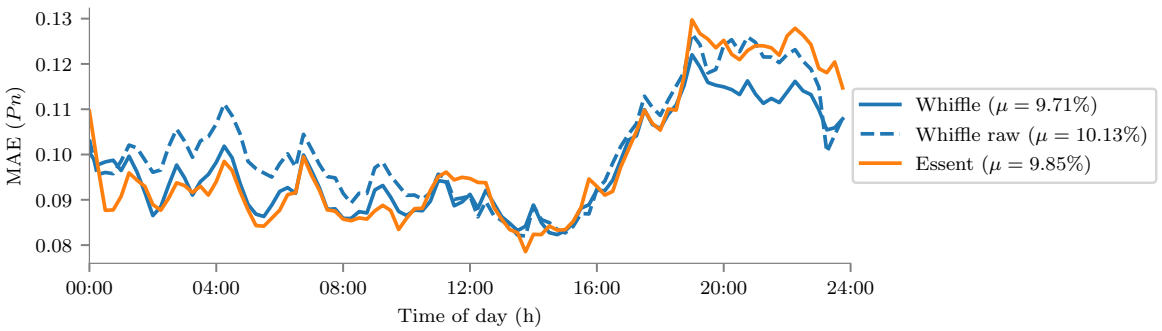


Figure 7.1: Mean Absolute Error for 2018-01-01 - 2018-06-30. The 2017 MAE for Whiffle was 8.98%, for Whiffle raw 9.35% and for Essent 9.17%.

drawn. First, a clear diurnal pattern can be identified in all, which is quite similar in shape between them. This pattern is characterized by a relatively large MAE during the night relative to daytime, which is explained by the fact that at night the wind speed on average is higher, resulting in larger absolute errors [54]. As all analyses are in local time, which is subject to day-light saving, this effect is offset slightly as the atmosphere does not conform to day-light saving. Second, the Whiffle raw forecast performs worse to a varying degree relative to the Whiffle forecast throughout the day. Second, from the 10 hour until the 20 hour mark, the Whiffle and Essent forecast seem very close in performance. Outside of this interval the Whiffle forecast outperforms the Essent forecast, most notably during the night, which implies that the Whiffle forecast better at forecasting higher wind speeds. Last, as the MAE is the main ranking criterion, the Whiffle forecast is ranked highest, the Essent forecast second and the Whiffle raw forecast third.

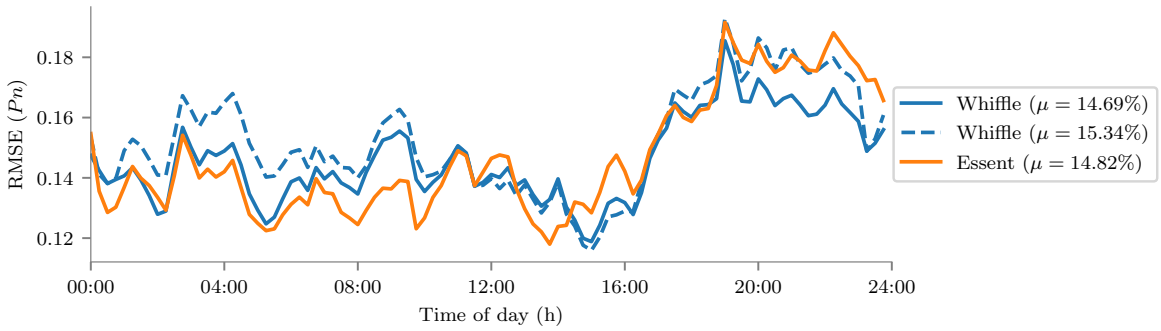


Figure 7.2: Root Mean Square Error for 2018-01-01 - 2018-06-30. The 2017 RMSE for Whiffle was 13.51%, for Whiffle raw 14.30% and for Essent 14.06%.

Second, the forecasts are judged on the Root Mean Square Error, as presented in equation 5.2. Looking at the results in figure 7.2, several conclusions can be drawn. Firstly, as the RMSE is a quadratic loss function, it is more sensitive to a larger spread in errors, giving more weight to larger errors. Figure 7.2 shows that the difference between the Whiffle raw forecast and the others is much larger when compared to the MAE. This means that the Whiffle raw forecast has a large spread in its errors than the others, indicating greater consistency for the other two. Comparing the better performing two, their difference in RMSE earlier in the day seems smaller when compared to the MAE. Also, percentage wise the improvement of the Whiffle forecast over the Essent forecast is only 0.85%, which is smaller than the 1.42% in MAE. This indicates that the performance improvement lies less in the reduction of the variability of its errors, but more in absolute performance.

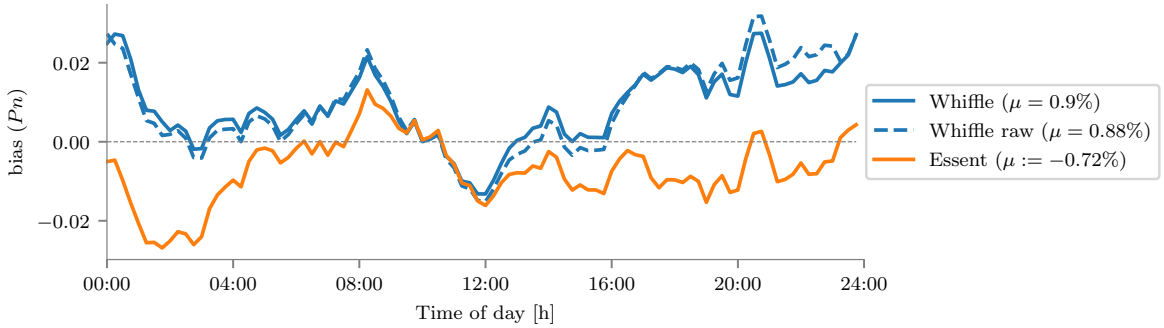


Figure 7.3: Bias for 2018-01-01 - 2018-06-30. The 2017 Bias for Whiffle was -0.49%, for Whiffle raw -0.79% and for Essent -1.69%.

Last, the forecasts are judged on the bias, as presented in equation 5.3. Figure 7.3 does not provide much insight into the absolute performance of the forecast, but it does allow one to gain insight in the existence of systematic errors, which can be corrected in post-processing using statistical methods. Overall the Essent forecast does a better job at processing out the systemic part of the error. Interestingly, both the Essent and

the Whiffle and Whiffle raw forecasts seem to have a similar trend in their bias. This indicates that there are some underlying similarities to the models generating these forecasts. This makes sense, as typically lower level NWP models use the output from higher level NWP models to determine their inititial and boundary conditions, as explained in section 1.1.1.

In conclusion, on all statistical performance measures the Whiffle forecast tends to outperform the other forecasts. When ranking them on statistical performance the initial ranking is derived from looking at the MAE, which means the Whiffle forecast performs best. However, evaluating the statistical performance of point forecasts is not sufficient when considering one forecast over another, as ultimately its performance in its specific use case is most important.

7.1.2. Wind Power Stochastic Process Modelling

First, the models used for generating density and scenario forecasts are discussed. Second, their performance is evaluated using the framework from chapter 5. The conditional kernel density (CKD) model, as presented in equation 4.2, is constructed with forecast error $\epsilon_{t+k|t}$ as the dependent variable and forecast $\hat{y}_{t+k|t}$ and forecast horizon k as the independent variables. In the CKD model the numerator of equation 4.2 is the joint distribution of $F(\epsilon, \hat{y}, k)$. This can be thought of as looking in 3d space and estimating the probability density function based on local densities. This is done by placing Gaussian cylinders (kernels) with a 3-dimensional size defined by the bandwidth parameters on each data point in space, after which the individual kernels are summed to form the kernel density estimate. This estimation in three dimensions does suffer from sparsity, specifically for medium to high values for $\hat{y}_{t+k|t}$, as high power values occur less often than lower power values. This is why increasing dimensionality in kernel density estimation can quickly decrease the validity of the resulting estimated distribution and the performance of the model. Also, with differences in data density, binning the data or applying variable bandwidths for different areas becomes important, where binning increases performance as for each region of the data a specific kernel bandwidth can be determined. However, as the wind power process has clearly defined boundaries between 0 and 1, constant bandwidths work well.

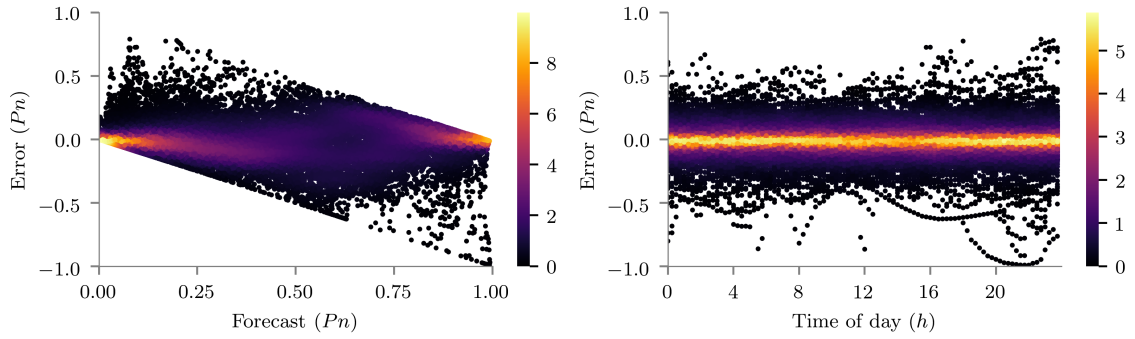


Figure 7.4: CKD models for Essent Wind Power. Lighter colors indicate higher probabilities.
a: Essent Wind Power CKD model $f(\epsilon|\hat{y})$.
b: Essent Wind Power CKD model $f(\epsilon|k)$.

The first relationship being analyzed is that of the probability of the normalized error $\epsilon_{t+k|t}$, conditional on forecast power $\hat{y}_{t+k|t}$. The first thing to notice from figure 7.4a is that the distribution of the errors is very narrow and skewed both for low and high values of $\hat{y}_{t+k|t}$, while it is wide and centered for middle values. The second relationship is that of the probability of the $\epsilon_{t+k|t}$, conditional on forecast horizon k, of which the probability density function is shown in figure 7.4b. This figure shows that the distribution of errors is relatively narrow for early lead times and then becomes wider later in the day, which is the same pattern as for the MAE and RMSE, although it is difficult to notice in figure 7.4b due to it being quite a nuanced difference. This means that later in the day the forecast starts to perform worse, which is consistent with figure 7.1. Although figures 7.4a and 7.4b shows two separate CKD models, each with a single independent variable, the complete CKD model uses two independent variables. This results in shifted probabilities for the data points shown in the figures. Their probabilities are shifted because of the added explanatory conditional variable. The stronger this shift, the more influential an added independent variable is.

For the initial estimation of the CKD model 2017 data is used, after which quantiles are constructed for the following 6 months of data. The CKD model is estimated daily using a rolling window of the past 12 months

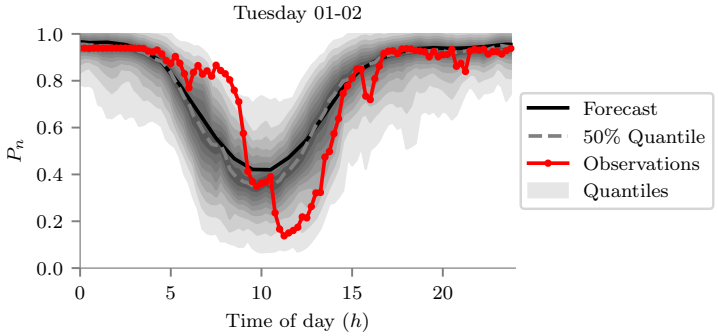


Figure 7.5: Quantile forecast Essent for 2018-01-02. The forecast shows quantiles 5% up to 95% in 5% increments.

of data. Figure 7.5 shows a density forecasts for the Essent forecast. This figure shows much detail in the quantiles, indicating heterogeneous marginal predictive densities. Furthermore, it shows that the observed power is much more volatile than the forecast, indicating that neither the forecast, nor the quantiles are realistic representations of the underlying stochastic process. Last, there is a significant difference between the original forecast, which forecasts the mean, and the 50% quantile, which is due to the skewness in the density, which leads to a difference between the mean and the 50% quantile.

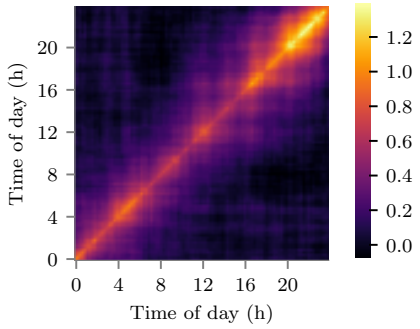


Figure 7.6: Covariance for Gaussian Essent errors for 2017.

The second model is the Copula model, as introduced in equation 4.10. The first part of the Copula model consists of the marginal predictive densities, which are obtained from the model shown in figure 7.4. These marginal predictive densities are used to transform all observations for $\epsilon_{t+k|t}$ by applying equation 4.11. The second part of the Copula model is the covariance matrix, which captures the dependency between the Gaussian variables as expressed in equation 4.13. This matrix is shown in figure 7.6, where lighter colors indicate a higher covariance. This matrix can be interpreted as follows: First, the diagonal represents the covariance between each Gaussian vector with itself, which is its variance. Second, it is a strictly symmetric matrix along the diagonal, where the covariance between each Gaussian vector of each forecast horizon with every other Gaussian vector of each other forecast horizon is represented in the matrix. The matrix shows quite some covariance between different forecast horizons, which makes sense, as the wind power process has limited volatility.

To adequately represent the stochastic process, a sufficient number of scenarios needs to be generated to cover the most plausible realizations of the process. A typical number used for this type of decision making problem is 1000 [10]. An example of a scenario forecast for 1000 scenarios is shown below in figure 7.7. Although this figure has clear similarities with its quantile counterpart, it is not as easily interpretable. As the marginal densities have long tails, the range of scenarios is significantly wider. Furthermore, it is difficult to see the density of scenarios in different areas. This shows that although scenarios are the desired input for optimization purposes, for decision-makers value can also be obtained by visual inspection of more easily interpreted quantile forecasts. Next, the quality of the modeling process is evaluated.

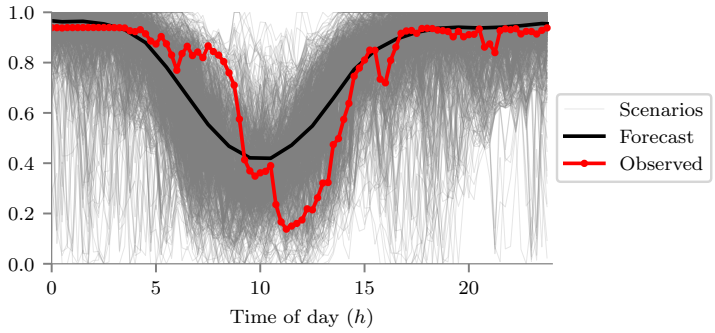


Figure 7.7: Scenario forecast for 2018-01-02

7.1.3. Wind Power Stochastic Process Evaluation

The first modelling step that needs to be evaluated is that of the construction of density forecasts by the CKD model. The first analysis is that of the Probability Integral Transform, which evaluates equation 5.4. The histogram for each density forecast is shown in figure 7.8, which is used to evaluate the PIT. The histograms for all three forecasts in figure 7.8 show that the predictive models are not perfectly calibrated. Although they differ somewhat in shape, the overall trend is similar. The hump shape present in all three histograms indicate on average too wide predictions intervals, which means that the predictive distributions are overdispersed. Although this provides an initial insight in whether or not the predictive density is calibrated, it is not easily interpreted, nor can it directly give insight whether or not the model should be re-calibrated or can be safely applied in an operational context.

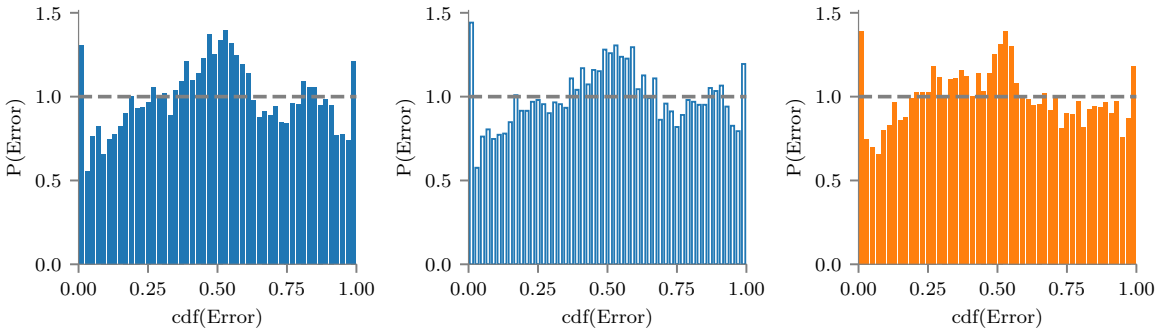


Figure 7.8: a: PIT for Whiffle density forecast.
b: PIT for Whiffle raw density forecast.
c: PIT for Essent forecast.

The reliability diagram is shown in figure 7.8. This shows how the 19 quantiles perform individually. This enable the user to make further inferences on its calibration. First, when quantiles are on average too low, less than the nominal fraction of observations indicated on the x-axis is observed below the specific quantile, resulting in a negative value for Δq_m. All three show this for quantiles 50% and lower. When quantiles are on average too high, more than the nominal fraction of observations is observed below the quantile, resulting in a positive value for Δq_m. All three show this as well for higher quantiles. This pattern is consistent with the conclusion from the PIT histogram, as it indicates overdispersed predictive distributions, where the predictive intervals are too wide on average. Second, although all three show a similar trend, the Essent forecast is most symmetric around the 50% quantile, while the other two show a slightly more asymmetrical trend, especially the Whiffle forecast. This means that the Whiffle predictive density is most skewed, the Whiffle raw is slightly less skewed and the Essent forecasts is quite symmetrical. Thirdly, the Whiffle forecasts both seem shifted downward relative to the Essent forecast, which indicates a bias in the predictive density as on average it predicts too low quantiles. As the model is estimated on observations and it should remove any bias present in the data, which it does for the Essent forecast, this indicates that either the model is not successful at removing the bias for the Whiffle raw and Whiffle forecasts or that there is an inconsistency between the

in-sample and out-of-sample data. Figure 7.3 reports the 2017 bias values for the three forecasts, which are most different for the Whiffle raw and Whiffle forecasts relative to the Essent forecast. As the model is estimated daily on the past year of data, a shift in bias between in-sample and out-of-sample data leads to a bias in the predictive density.

Last, although these detailed observations provide insights that can be used to improve the probabilistic model, the discrepancy index provides a concrete measure on which to judge the reliability of the overall model. For all three it is near 2%, where the Essent predictive density is most reliable, the Whiffle raw is second and the Whiffle predictive density is least reliable. Although there are no clear guidlines on what is an acceptable value, Morales et al. (2014) [42] deem a discrepancy index of 1.2% to be very good. Although 2% is higher than 1.2%, as there are no strong biases present in the predictive density, the predictive densities are deemed reliable.

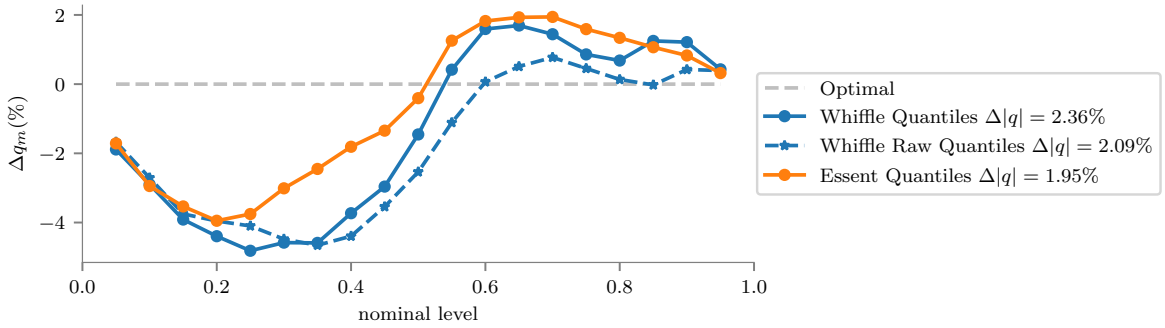


Figure 7.9: Reliability diagram for the Whiffle, WHiffle raw and Essent forecasts. Positive values indicate too high quantiles, as a higher percentage than the nominal fraction is below them and vice-versa.

The final reliability test is that of the Minimum Spanning Tree Rank Histogram (MST RH), as presented in section 5.3.2. The MST RHs for all three forecasts are shown in figure 7.10. Compared to the example MST RH in figure 5.4, all three histograms seem quite irregular. This is due to the fact that the underlying probabilistic models are significantly more complex than the Beta distributions used for sampling the data for figure 5.4. The purpose of the MST RH for complex distributions with a limited sample size is not to test perfect uniformity, but to identify possible trends in the ranking of the MSTs of the ensembles versus the observations, where the scenarios are the ensembles and the observations are the observations of the stochastic process. In the example there was a clear trend present, which in this case would indicate that the dependency model overestimates the covariance. However, all three show no clear downward or upward trend. This indicates that the dependency model accurately captures the covariance between the different forecast horizons, which means that both the predictive density and the dependency model are deemed reliable.

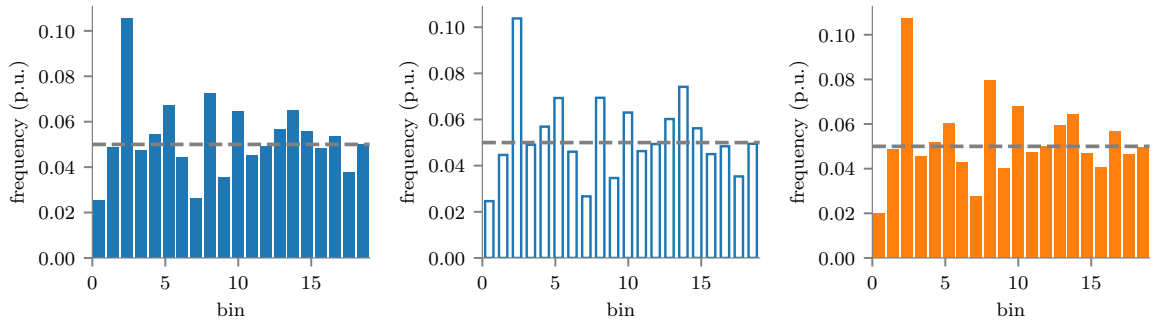


Figure 7.10: Minimum Spanning Tree Rank Histogram for 2018-01 - 2018-06 scenarios. Based on scenario forecast consisting of 1000 scenarios.
a: Whiffle MST RH. b: Whiffle Raw MST RH.
c: Essent MST RH.

As the models for all three forecasts are deemed reliable, focus is shifted to their performance. The first performance metric is the Net Quantile Score (NQS) as discussed in section 5.4. This measure computes a metric from the 19 quantiles generated by the predictive density. The results for this metric are shown in figure 7.11. The overall trend in this figure is very similar to the MAE shown in figure 7.1. Moreover, the ranking between the three forecasts is preserved in their quantile form. However, the percentage wise difference between the Essent and Whiffle forecast is much smaller in their quantile form. This can be explained by the bias shown in the reliability diagram. As the NQS penalizes wrong quantiles as well as the extent to which they are wrong, the added sharpness from the point forecast is hurt by the larger bias present in the predictive density. This effect is also present in the difference between the Essent and Whiffle raw forecast, as the percentage wise different in MAE is smaller than the difference in NQS, again showing the penalty placed on inferior calibration. Overall, the Whiffle raw forecast is still the best performing predictive density, albeit with a smaller margin.

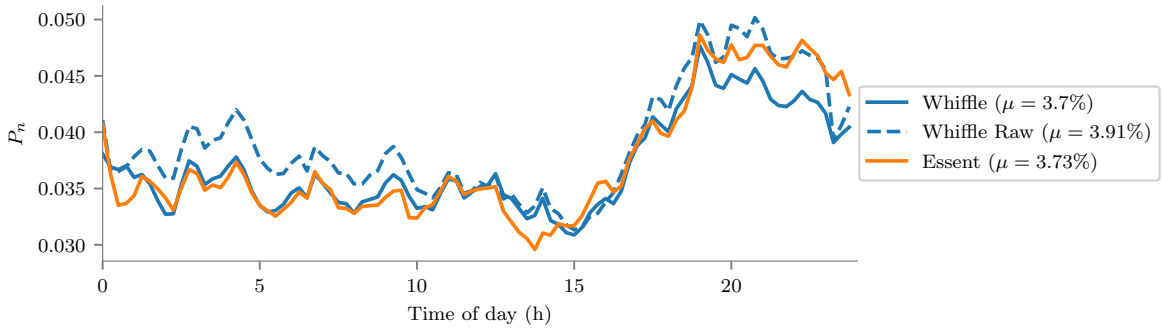


Figure 7.11: Net Quantile Scores as a function of forecast horizon for the Whiffle, Whiffle Raw and Essent forecasts for 2018-01 - 2018-06.

The main ranking criterion of predictive densities, the Continuous Ranked Probability Score is also introduced in section 5.4. As this is computed over the entire predictive density, this should provide the most accurate judgment on which predictive density is superior. The difference between the NQS and CRPS is expected to be large when the underlying predictive density is more complex and of a higher resolution. As the model producing the predictive densities is continuous, the CRPS is considered important, whereas it would not be if the density would have been generated by a discrete model, such as a quantile regression model. The CRPS score, displayed in figure 7.12 again shows a very similar trend to the MAE and NQS. However, the differences between the three forecasts are more similar to the differences in MAE. This indicates that the NQS can result in skewed results, due to a lack of resolution. The conclusion from the CRPS is again that overall the Whiffle forecast is superior, with the Essent forecast coming in second and the Whiffle raw forecast in third. All three predictive densities' CRPS score is lower than its MAE score, which can be interpreted as the predictive densities adding skill to the point forecast [42]. This is because the MAE is the same as the CRPS when the CRPS is computed for a step function representation of the point forecast.

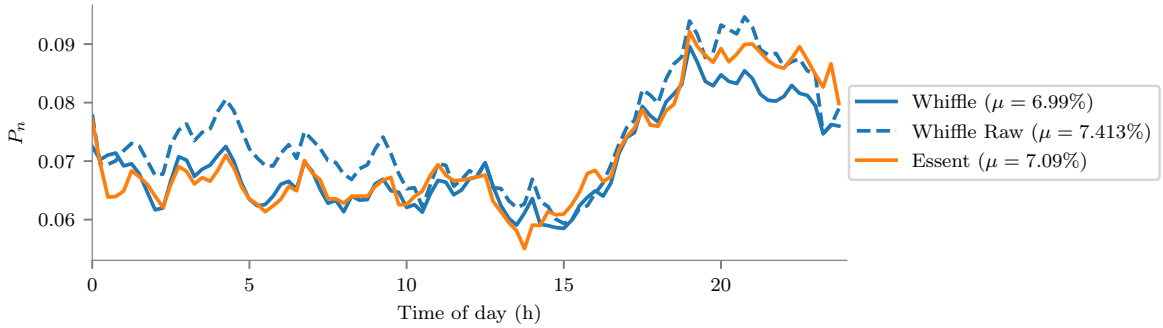


Figure 7.12: Continuous Ranked Probability Score as a function of forecast horizon for the Whiffle, Whiffle Raw and Essent forecasts for 2018-01 - 2018-06.

The last ranking criterion is the Energy Score (Es), as introduced in section 5.4. This is also a strictly negatively-oriented score and evaluates to 0 if the scenarios perfectly capture the underlying distribution. As the energy score evaluates scenarios, which are a result of a point forecast, a predictive density and a dependency model, this can be used for decisions on improving any step in the process of constructing scenarios. Figure 7.13 shows the Es distributions for the different forecasts. On average the Whiffle forecast is again superior to the other forecasts, with the Essent forecast again coming in second and the Whiffle raw forecast in third. However, for this final metric, the difference is smaller compared to the difference in CRPS. As this is the final ranking metric, the Whiffle forecast is considered superior as input to the optimization model, with the Essent forecast slightly behind it and the Whiffle raw forecast a distant third. In an absolute sense the energy scores are acceptable, compared to scenarios analyzed in a 2012 study by Girard and Pinson, where the best scenario forecast achieved a score of 1.130 [48]. However, as that was a different case, it should only serve as an indication that the scenarios perform well, it cannot be used to conclude that these scenarios are superior to those in the Girard and Pinson study.

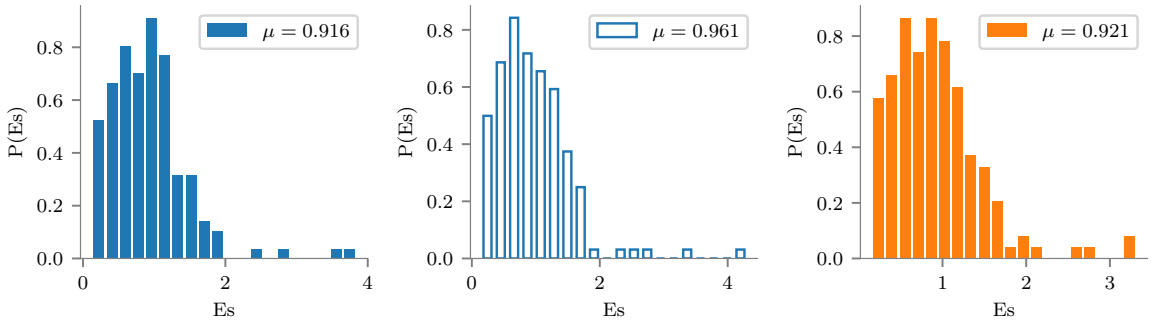


Figure 7.13: Continuous Ranked Probability Score as a function of forecast horizon for the Whiffle, Whiffle Raw and Essent forecasts for 2018-01 - 2018-06.

7.2. Day-Ahead Price

The first step in the modeling process is again judging the performance of the point forecasts. Especially in the case of the price forecasts, which can be easily generated by WPPs themselves, for instance using the method introduced in chapter 6, this initial evaluation can provide insights in how point forecasts can be improved. In figure 7.14 the performance of the forecast of λ^D is presented.

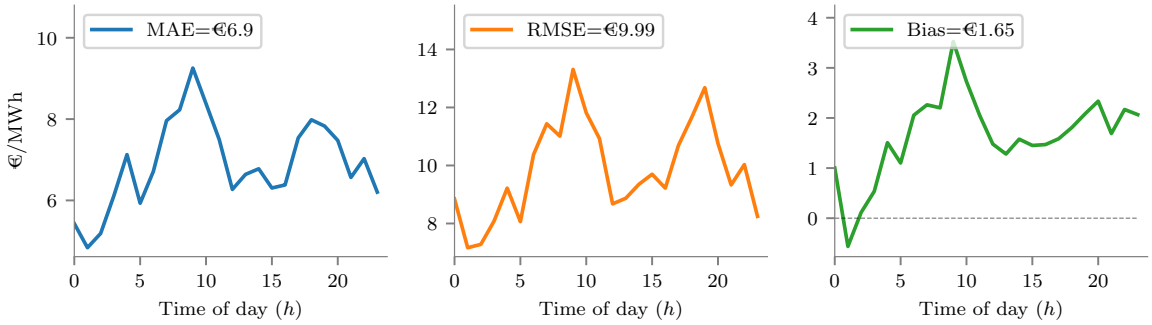


Figure 7.14: Statistical performance of point forecast for λ^D.
a: Mean Absolute Error. 2017 MAE=€6.11.
b: Root Mean Squared Error. 2017 RMSE=€10.36
c: Bias. 2017 Bias=€1.47.

The left panel shows the performance on the main ranking criterion, the MAE. Unfortunately, unlike with the wind power forecasts, no reference is available. However, it does show some interesting characteristics of the underlying process. The forecast error is low early in the day, then rises during the morning ramp.

It decreases slightly during the day, only to sharply increase during the evening ramp. This is analogous to the MAE rising for wind power during the night. As wind speeds are on average higher during the night, the absolute forecast error naturally increases. This is also the case for the day-ahead price, as the pattern can be explained by the price behaviour driven by the morning and evening ramps. This same pattern is also visible for the RMSE in figure 7.14b. Looking at the bias in figure 7.14c, an interesting observation can be made. First, the bias is positive and quite large relative to the MAE. Fortunately, the probabilistic model should be able to correct for this as long as the bias is consistent between in-sample and out-of-sample data. Second, it shows the same pattern as do the MAE and RMSE. This indicates that the model consistently underestimates the price, which scales with the height of the price. This could mean that higher prices are caused by non-linear effects, which the model is not able to capture.

The next step is the application of the stochastic model. The conditional kernel density (CKD) model is again estimated with the forecast error $\epsilon_{t+k|t}$ as a dependent variable. The independent variable for the model for the day-ahead price is limited to the forecast $\hat{y}_{t+k|t}$. Although for wind power forecast errors also the time of day is considered, here it is not. Because the forecast error for the price data is distributed less uniformly and has no strict boundaries, adding the forecast horizon as independent variable hurt the model's ability to reliably capture the uncertainty. This is due to the fact that data is much more sparse at the extremes with this unbounded process than it is with wind forecast errors. Fortunately, the distinct daily pattern of the day-ahead price means that the forecast value itself is already a helpful indicator for how the shape of the distribution is related to the time of day.

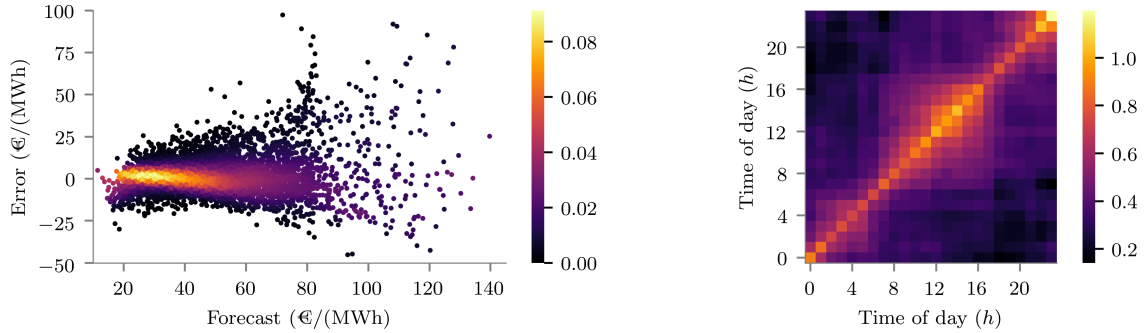


Figure 7.15: a: CKD model for the day-ahead price. Lighter colors indicate higher probabilities.
b: Covariance matrix for Gaussian day-ahead price errors for 2017.

Looking at figures 7.15a, several conclusions can be drawn. First, The distribution has a very distinct shape, as it is very narrow for lower forecasts, while it becomes increasingly wide as forecast values increase. This is in line with the analysis of the point forecast, which indicated that for higher prices the forecast error increases. Secondly, there is a slight skewness present in the model, with a slight negative deviation for higher forecast values. Such particularities are difficult to capture with less flexible methods, while the highly flexible CKD model adequately captures such aspects of a distribution. For higher forecast values data sparsity becomes an important issue. As the kernel bandwidth is estimated on the entire data set, having such sparse data becomes an issue for the model, as it leads to oversmoothing for denser areas and undersmoothing for more sparse areas.

In figure 7.15b the Copula model is visualized. Again, the Gaussian Copula is represented by a single covariance matrix. From it several conclusions can be drawn about the day-ahead price process. First, overall it is quite bright compared to the wind power errors, indicating that there is significantly more temporal dependence present in the process. Second, there are some square structures to be observed within the matrix, most notably between morning and afternoon values. This indicates that there are different regimes in the stochastic process between these two periods, which the Copula model shows to be able to capture.

In figure 7.16 the quantile and scenario forecasts are shown, generated by the models from figure 7.15. Again, the models were estimated daily using the past year of data. For the scenario forecast, each day 1000 scenarios were generated. In the figure, the quantiles seem significantly narrower than the scenarios, which indicates that the marginal predictive densities from the CKD model have quite long tails. Again, the quantile forecast is more easily interpreted, while the scenario forecast is difficult to interpret. For the quantile forecast, the 50% quantile does not seem to deviate much from the forecast, which indicates that overall the

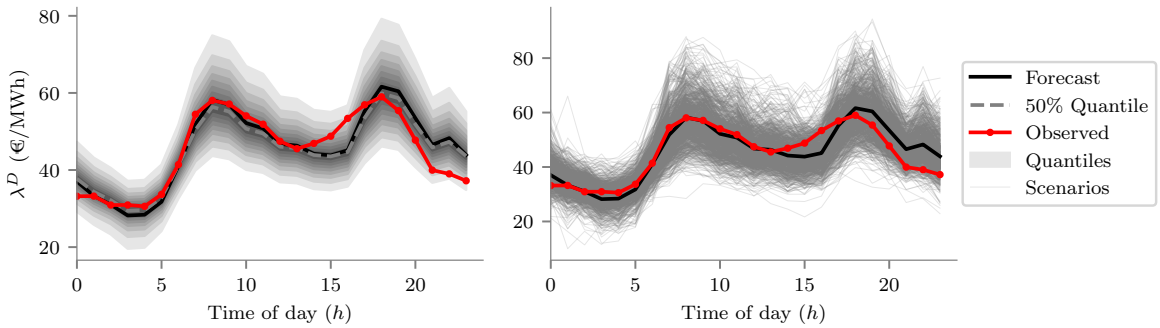


Figure 7.16: a: λ^D Quantile forecast for 2018-02-05.
b: λ^D Scenario forecast for 2018-02-05.

underlying density forecast is not very skewed, which is also observable in figure 7.15a for the model itself. Certainly compared to the wind power model, the model is quite symmetric, resulting in a 50% quantile that is close to the mean forecast.

The next step is verification. To this end the three metrics applied to evaluate reliability are shown in figure 7.17. Firstly, figure 7.17a shows the PIT in a histogram. Contrary to the wind power case, the histogram shows a convex trend. This indicates underdispersed predictive densities, where the predictive intervals are too narrow on average. Figure 7.17b paints a more nuanced picture. It shows that higher quantiles increasingly underestimate the quantiles. This trend in the reliability diagram indicates that the model is also skewed towards lower values, resulting in underestimated predictive quantiles. The reliability diagram also shows a slight upward trend for the two higher quantiles, which makes it relatively wide in the higher quantile region. However, as the overall trend is quite linear the predictive is considered too narrow and skewed towards lower values. However, the discrepancy index $\Delta|q|$ of 2.87% is relatively good. Although no reference cases are found, when compared to the well calibrated wind power models, 2.87% is deemed acceptable, although it must be noted that the model is slightly biased by its skewness. In figure 7.17c, the Minimum Spanning Tree Rank Histogram shows quite a uniform distribution. As it shows no clear indication of a trend, which means that the Copula model is able to capture the underlying dependence structure well, the Copula model is deemed reliable. In conclusion, although the predictive density shows a slight bias, its reliability is deemed acceptable. As the Copula model seems reliable as well, the modelling process for λ^D is deemed reliable.

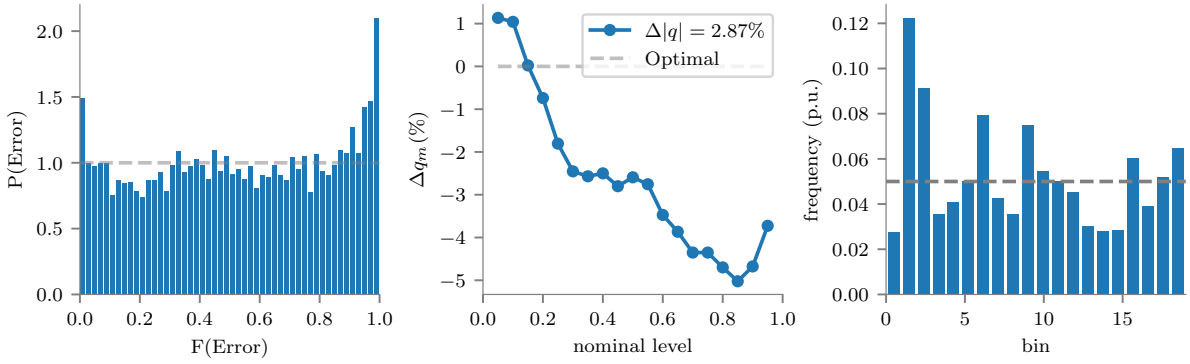


Figure 7.17: Reliability evaluation of forecast for λ^D.
a: PIT histogram.
b: Reliability diagram.
c: Minimum Spanning Tree Rank Histogram.

The performance of the predictive density and scenario forecast is shown in figure 7.18. Figure 7.18a shows the Net Quantile Score for λ^D. Compared to the CRPS shown in figure 7.18b, it seems very similar in overall trend, which is an indicator that the 19 discrete quantiles represent the underlying predictive density well. Second, a slight diurnal pattern can be observed in the NQS and CRPS, although it is quite different from the MAE of the point forecast in figure 7.14, as the large peak has shifted from the evening ramp to the

morning ramp, which indicates that it adds more skill in the time around the evening ramp than the time around the morning ramp. Third, the CRPS score is significantly lower than the MAE score, which indicates that the predictive density adds skill. Unfortunately, direct comparison with other cases is not possible as prices change a lot between countries and periods. So again, this measure serves better as a comparison between methods on a single case. Last, the Price Score (Ps) is shown in figure 7.18, which has the exact same definition as the Es introduced in section 5.4. As with the Es this score carries the same unit as the scenarios from which it is computed. Again, this measure is a negatively oriented score, making it a good skill score to analyze the influence of changes in each step of the modeling process. Here an average value of €29.94 is achieved, which serves as a benchmark for future improvement.

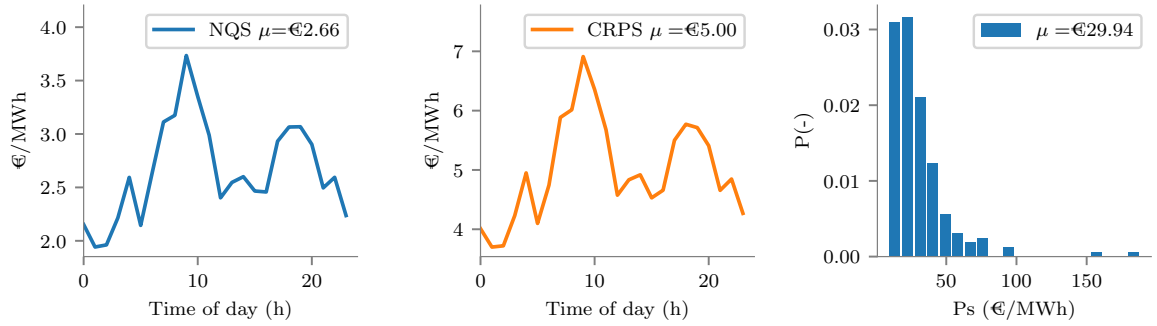


Figure 7.18: Performance evaluation of λ^D process.
a: Net Quantile Score.
b: Continuous Ranked Probability Score.
c: Price Score (Ps).

7.3. Imbalance Price

As with the day-ahead price, the first step is the performance assessment of the point forecasts. The performance metrics for λ^Δ are shown in figure 7.19.

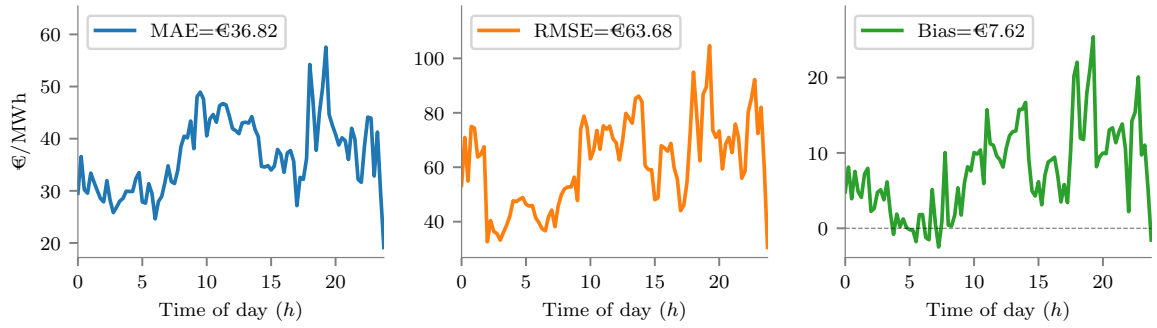


Figure 7.19: Statistical performance of point forecast for λ^Δ.
a: Mean Absolute Error. 2017 MAE=€31.92.
b: Root Mean Squared Error. 2017 RMSE=€52.53
c: Bias. 2017 Bias=€3.28.

First, the MAE in figure 7.19a shows a distinct diurnal pattern which is similar in shape to the MAE of the day-ahead price. The explanation for this is the same as for the day-ahead price. More extreme prices occur during the morning and evening ramps. As absolute errors are naturally higher on average for higher prices, the pattern is to be expected. The RMSE in figure 7.19b is relatively different from the MAE in shape, compared to wind power and the day-ahead price. This is due to the fact that the imbalance price is significantly more volatile, which results in a larger spread in errors. Finally, the bias displayed in figure 7.19c shows to be positive, again with a pattern that is similar to the MAE and RMSE. Compared to the day-ahead price the bias relative to the MAE is relatively small. The caption of figure 7.19 shows the statistics for 2017, which are quite

different from the out-of-sample set. Especially the bias is significantly larger, which could lead to problems for the predictive density.

The next step is the application of the stochastic model. The dependent variable is again the forecast error $\epsilon_{t+k|t}$. As an independent variable the forecast $\hat{y}_{t+k|t}$ was considered. However, this model setup was not able to reliably capture the distinct diurnal uncertainty that is characteristic of the imbalance price, as shown in figure 7.19a. Hence, the forecast for λ^D is used as the independent variable. Again, the forecast horizon is not used as an independent variable, as data sparsity is even more an issue for λ^Δ. Fortunately, the daily pattern of λ^D again means that the forecast value itself is already a helpful indicator for how the shape of the distribution is related to the forecast horizon.

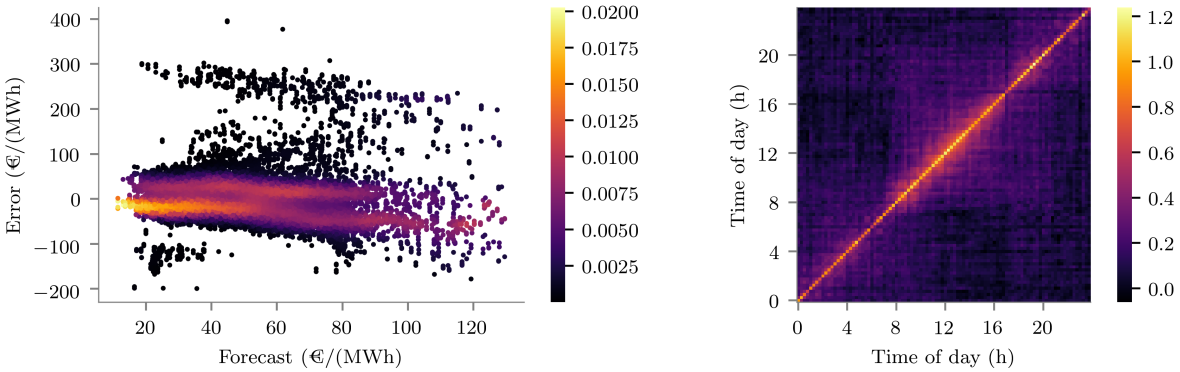


Figure 7.20: a: CKD model for λ^Δ. Lighter colors indicate higher probabilities.
b: Covariance matrix for Gaussian imbalance price errors for 2017.

The CKD model for λ^Δ is shown in figures 7.20a. From it several conclusions can be drawn. First, the distribution has a distinct bimodal shape, which varies strongly for different forecast values. Also, there is a strong negative peak for low forecast values, which means that for low forecast values for λ^D, for 2017 there was a high probability of negative forecast errors for λ^Δ. Lastly, the distribution shows a downward trend, indicating a higher probability for negative forecast errors for higher forecast value of λ^D. Overall, the forecast for λ^D seems well able to predict the uncertainty around the forecast for λ^Δ.

Moving to the Gaussian Copula model, the covariance matrix is shown in figure 7.20b. Firstly, it is much darker when compared to the previous covariance matrices, indicating little temporal dependence, which makes sense as imbalance prices are highly volatile. Lastly, as with λ^D some square structures can be observed, indicating a change in regime between morning, mid-day and night, which the Copula model shows to be able to capture. Figure 7.21 shows examples of quantile and scenario forecast, generated using the models depicted in figure 7.20. For the scenario forecast, each day 1000 scenarios were generated.

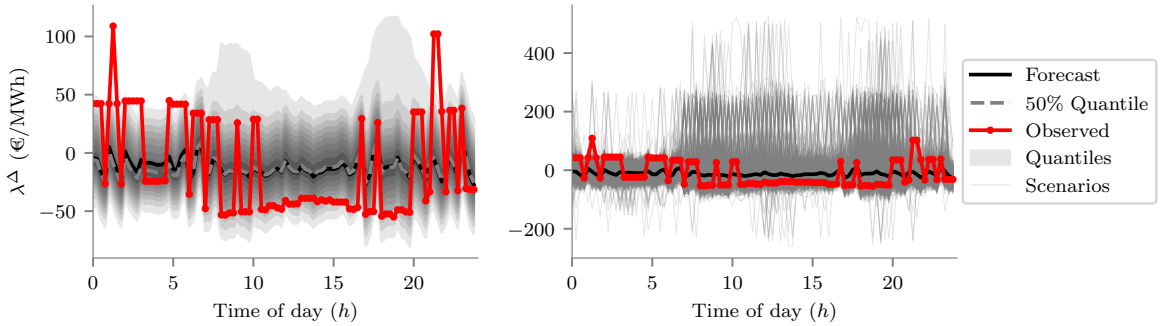


Figure 7.21: a: Quantile forecast for λ^Δ for 2018-02-05.
b: Scenario forecast for λ^Δ for 2018-02-05.

From figure 7.21 several conclusions can be drawn. First, it shows that the model predicts strong upward uncertainty around the morning and evening ramps. This is where the value of the model lies, as the anticipation of such events are important to the optimization. The quantiles seem significantly narrower than

the scenarios, which indicates that the distributions from the CKD model have very long tails, which makes sense as many extreme values can be observed in figure 7.20a. As with wind power and the day-ahead price, the quantile forecast is more easily interpreted, informing on the uncertainty surrounding the price forecast. Next focus is turned to reliability of the models, for which the relevant metrics are displayed in figure 7.22.

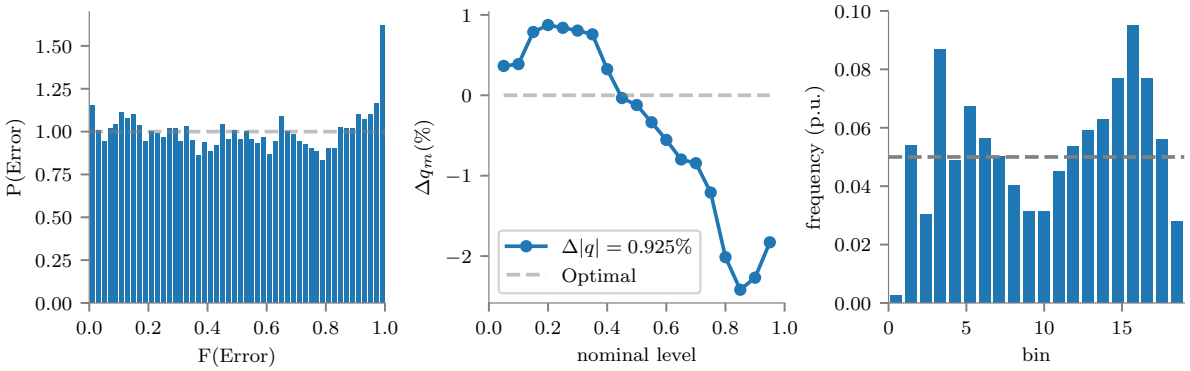


Figure 7.22: Reliability evaluation of forecast for λ^{Δ}.
a: Probability Integral Transform histogram.
b: Reliability diagram.
c: Minimium Spanning Tree Rank Histogram.

Firstly, from figures 7.22a a very slight convex trend can be observed, as the histogram shows slightly higher counts toward the edges. This suggests slightly underdispersed predictive distributions, where the predictive intervals are too narrow on average. This is confirmed through the reliability plot in figure 7.22b. This plot shows that lower quantiles are slightly overestimated, while higher quantiles are slightly underestimated. Overall there is a slight bias in the predictive density, as it is skewed towards higher values. However, the discrepancy index is very good at 0.925%. This deems the model very well calibrated. In the case of prices, a possible explanation for the skewness could lie in the rapid change of the underlying market dynamics, which means that older data may not well represent today's uncertainty, as was shown in the point forecast analysis. The last test for reliability is shown in figure 7.22c, which shows the MST RH. Contrary to the day-ahead price and wind power, this MST RH does show a clear distinct shape. Its bimodal shape is directly related to the behavior of the price signal. As can be observed in figure 7.21a, the price can be stable or it can switch strongly. This is coherent with the explanation in section 2.2.3, which explains that as the sign of the Net Regulation Volume changes, the price changes abruptly. However, when the sign of the NRV stays the same for multiple PTUs, the price is significantly less volatile. As the dependency model applied here only models covariance over time, this switching behaviour is not considered seperately. So while the predictive densities are well able to provide the Gaussian data used to estimate the covariance matrix, the method of covariance estimation is not able to capture the switching behaviour. This means that for days with few switches in the price signal the model underestimates the covariance, while for days with many switches the model overestimates the covariance, resulting in the shape in figure 7.22c. However, the deviation from perfect calibration is not very large, which is why the model is accepted as sufficiently reliable, albeit with the note that a slight bias exists due to the nature of the price signal and the limitations of the Gaussian Copula.

Figure 7.23 shows the performance metrics for the predictive densities and the scenario forecasts. From figure 7.23a and 7.23b several conclusions can be drawn. First, the NQS curve does not seem to approximate the CRPS well, indicating that the 19 quantiles between 0.05 and 0.95 approximate the full distribution poorly. This is important if one wants to model the distribution through a discrete approximation, a higher number of quantiles seems appropriate than the 19 extracted here, in order to accurately capture the distribution. Second, a slight diurnal pattern can be observed, which is quite similar to the forecast performance in figure 7.19a. Direct comparison is again not possible as prices change quite a lot between countries and periods. So again, this measure serves better as a comparison between methods on a single case. Concerning the difference between the MAE and the CRPS, again a significant difference is achieved, which indicates that skill is added to the forecast. Lastly, on the Ps an average value of €379.38 is achieved, which currently cannot be compared to one in literature, thus again serving as a benchmark for future improvement

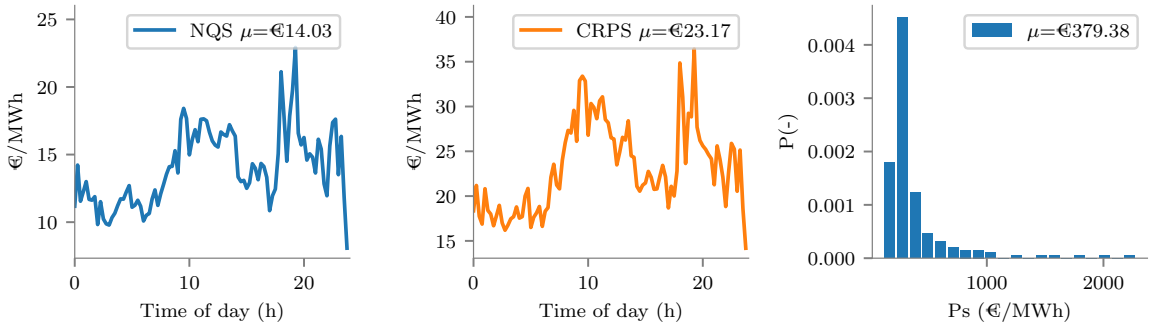


Figure 7.23: a: CRPS and NQS for $^{\Delta}$
b: λ_s for $^{\Delta}$

7.4. Price-Maker Effect

The first step in the process of modelling the price-maker effect, defined as the slope of the inverse ordered supply curve to the balancing market in section 6.4, is analyzing its point forecast performance. First, the Mean Absolute Error in figure 7.24a shows some similarity with the other two processes. This shape can again be related to the aggregate load, which is highest during the morning and evening and lowest during the night. When the load is high, more generation capacity is used to cover the load and less is available for balancing services. This leaves more expensive generation for balancing services and thus a steeper supply curve. A similar trend is present in the RMSE in figure 7.24b. The bias is shown in figure 7.24c, which is relatively large relative to the MAE, indicating a moderately calibrated predictive model. Furthermore, as mentioned in the caption of figure 7.24, the bias for the out-of-sample dataset is significantly larger than for the in-sample dataset, which could be problematic for the probabilistic model.

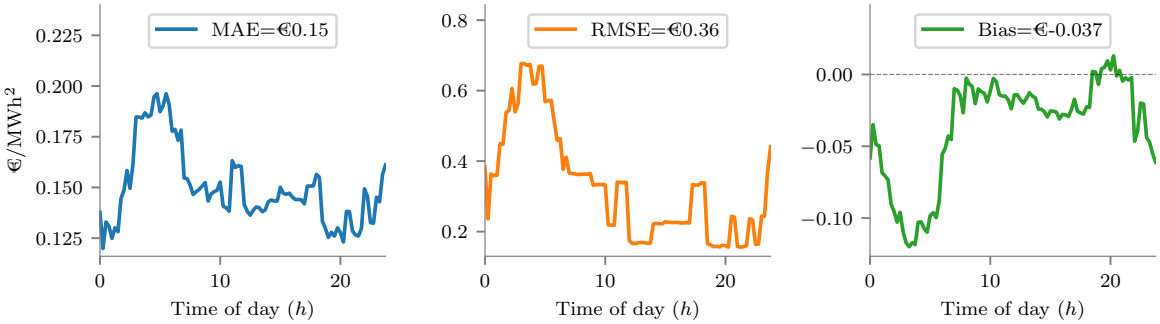


Figure 7.24: Statistical performance of the forecast for γ^{Δ} for 2018-01 - 2018-06.
a: Mean Absolute Error. 2017 MAE=€0.14.
b: Root Mean Squared Error. 2017 RMSE=€0.36.
c: Bias 2017 Bias=€-0.007.

The next step is the application of the stochastic model. The dependent variable is the forecast error $\epsilon_{t+k|t}$. As independent variable the forecast for γ^{Δ} is used. Again, the forecast horizon is not considered as a explanatory variable, due to the same problems as with λ^{D} and λ^{Δ}. Unfortunately, the lack of a daily pattern for the forecast for γ^{Δ} means that the forecast value itself is not a helpful indicator for how the shape of the distribution is related to the time of day.

The CKD model is shown in figure 7.25a. There are some outliers, as 0.48% of the time in the in-sample data, error values fell outside the range from the y-axis in the figure. These are included in the model estimation, but for interpretation purposes, these points are not displayed in the figure. From the figure several conclusions can be drawn. Firstly, the distribution has a distinct shape, which can be related to the wind power case. The similarity is that for high and for low forecast values there is distinct skewness, while for medium forecast values the distribution is much flatter and symmetric. Overall, this distribution shows less of a distinctive shape than the other distributions defined earlier.

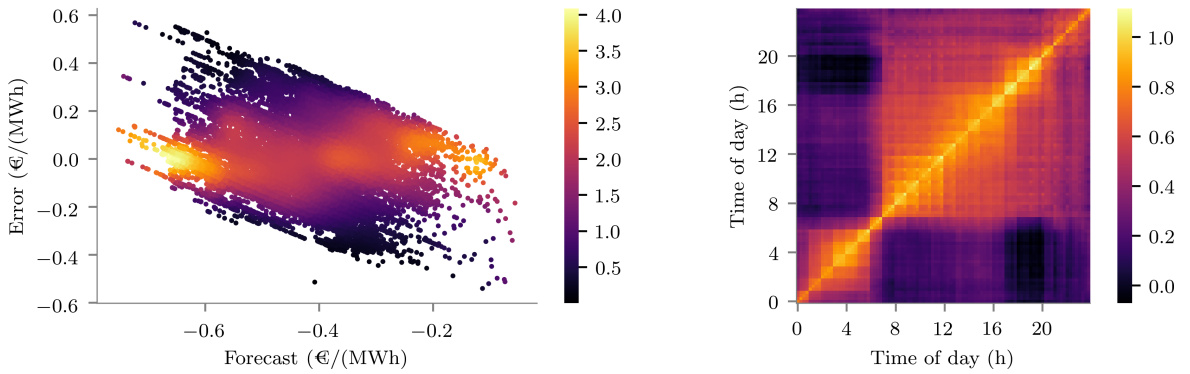


Figure 7.25: a: CKD model for γ^{Δ}. Lighter colors indicate higher probabilities.
b: Covariance matrix for Gaussian price-maker forecast errors for 2017.

Moving to the Gaussian Copula model, the covariance matrix is shown in figure 7.25b. This matrix is much lighter when compared to the other covariance matrices, indicating strong temporal dependence. Again, some square structures can be observed, although this time much more pronounced. This again indicates a change in regime between morning and mid day and less so between mid day and night, which the Copula model again shows to be able to capture. Next, quantile and scenario forecasts are generated, for which an example is shown in figure for 2018-01-02 7.26. For the scenario forecast, each day 1000 scenarios were generated.

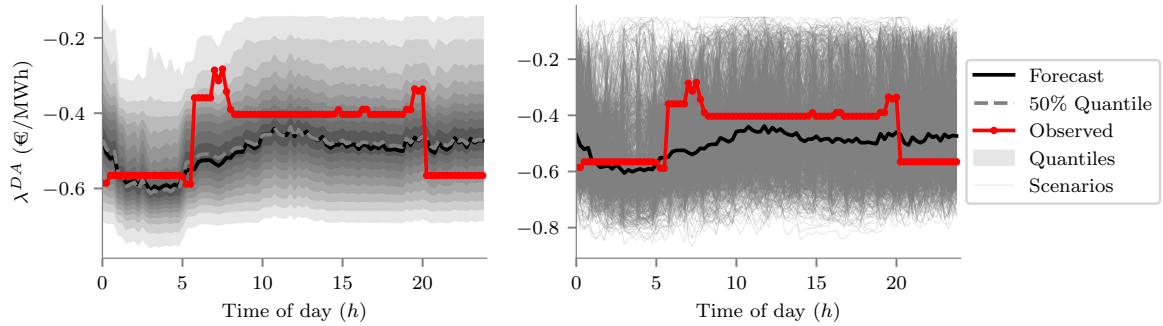


Figure 7.26: a: quantile forecast for γ^{Δ} for 2018-01-02.
b: scenario forecast for γ^{Δ} for 2018-01-02.

From the quantile forecast in figure 7.26a, several conclusions can be drawn. First, the shape of the predictive density is not very distinct, compared to the former predictive densities. This is coherent with the CKD model shown in figure 7.25a, which also does not have a particularly distinct shape. The observation of the process shows to have some similarities to the imbalance price, with distinct switches between time periods, but relative stability between them. However, it is much less volatile than the imbalance price. Compared to the imbalance price, the scale of the y-axis in figure 7.25b is relatively similar to that of the quantile forecast, indicating relatively short tails. However, as there are some occurrences of extreme values, there will sometimes be long tails to the predictive density, resulting in extreme scenarios.

Next focus is turned to reliability of the models, for which the relevant metrics are displayed in figure 7.27. Firstly, looking at figure 7.27a, which displays the histogram for the PIT, a clear downward trend can be observed. This indicates that the predictive density is shifted towards higher values. From figure 7.27b, this observation can be confirmed. It shows that on average quantiles are overestimated, indicating a negative bias in the predictive density. It also shows that lower quantiles are quite accurate, middle quantiles are estimated too high, while higher quantiles again are relatively accurate. This means that there is also skewness towards higher values. Due to its bias its average discrepancy index of 5.46% the model is deemed moderately reliable. The last reliability test is that of the MST RH, shown in figure 7.27c. There is a slight downward trend in the histogram, which indicates that the temporal dependency is overestimated. This can partly be explained by the negative trend in the PIT, which indicates that the transformed Gaussian errors do not meet

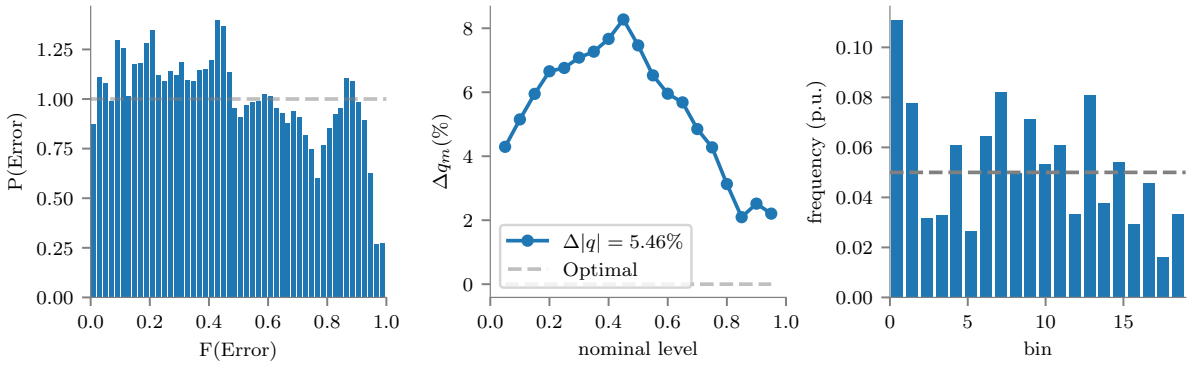


Figure 7.27: a: Quantile-Quantile diagram of forecast for γ^{Δ}
b: Reliability diagram
c: MST rank histogram

the requirement of normality well, which is a strict requirement for the Gaussian Copula model. Last, the Gaussian Copula is not able to capture the switching characteristic of the process well, which also contributes to the non-uniformity of the MST RH. With regards to overall reliability, some cause for improvement seems appropriate before operational implementation.

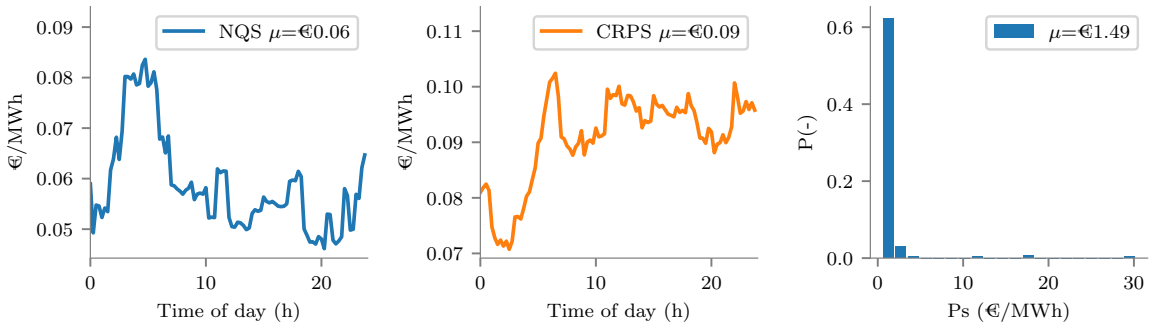


Figure 7.28: a: CRPS and NQS for γ^{Δ}
b: λ_s for γ^{Δ}

Figure 7.28 shows the performance metrics for the predictive densities and the scenario forecasts. From figure 7.28a and 7.28b several conclusions can be drawn. First, again the NQS seems to approximate the CRPS poorly, indicating that the 19 quantiles between 0.05 and 0.95 approximate the full distribution poorly. Second, a slight diurnal pattern can be observed, which is quite similar to the forecast performance in figure 7.24. Third, the CRPS is again significantly lower than the MAE, indicating that again skill is added by the probabilistic model. Last, the score on the Price Score (Ps) is shown in figure 7.28c. An average value of 1.514 is achieved, which currently cannot be compared to one in literature, again only serving as a benchmark for future improvement.

7.5. Scenario Tree Construction

The next step in the modeling process is combining all processes in a scenario tree, as discussed in section 4.4. The modeling framework presents two options for this step. The first is to apply the methods of scenario reduction and tree construction, which are the conventional methods in stochastic optimization, as explained in section 4.4.1. The second method is the method of extending the dependency model, which is explained in section 4.4.2. The next section discusses application of the first method, after which application of the second method is discussed.

7.5.1. Conventional Scenario Tree

Section 4.4.1 explains how to choose scenarios and also how many scenarios are to be chosen in orderto achieve a stable solution to the optimization problem, which for the linear version of the model is 50. However, as the conventional scenario tree is constructed by coupling all scenarios of each variable with all scenarios of each other variable, this means that for S variables and J scenarios the scenario tree will contain J^S scenarios. For the price-taker version of the model this would result in $50^3 = 125,000$ scenarios, making the optimization computationally very expensive. For the price-maker version of the model this would result in $50^4 = 6,250,000$ scenarios, making the optimization nearly intractable. However, one of the assumptions behind this method of scenario tree construction is that all processes are independent. As there is no statistical link between any of the scenarios, the definition of the imbalance price in section 6.4 provides a partial solution to the problem of tractability. As the scenarios for the imbalance price are always relative to the scenarios for the day-ahead price, the solution does not change with the addition of multiple day-ahead price scenarios. This means that the following redefinition can be made for the day-ahead parameter in the definition of the price-taker model in equation 3.1 and the price-maker optimization model in equation 3.5:

$$\lambda^D_{t,\omega} = \hat{\lambda}^D_t \tag{7.1}$$

Where $\hat{\lambda}^D_t$ is the point forecast for the day-ahead price. This means that because of the definition of the imbalance price, no day-ahead price scenarios are needed for the optimizations using the reduced scenario sets. This means that for the price-taker model the scenario tree contains $50^2 = 2,500$ scenarios, which is computationally tractable, and for the price-maker model $50^3 = 125,000$ scenarios are required, which still makes it computationally too intensive. Hence, for the price-maker model the number of scenarios still needs to be lowered below the recommended number of 50 scenarios for each stochastic process. For this purpose a ranking criterion is proposed, which ranks the three remaining processes on the extent to which the stochastic process is captured by the scenarios. The criterion is defined as:

$$\mu_{|y|} = \frac{1}{K}\frac{1}{T}\sum_{k}^{N_k}\sum_{t}^{N_t}|y_{t+k}| \tag{7.2}$$

$$a = \frac{Xs}{\mu_{|y|}} \tag{7.3}$$

Where Xs is either the Energy Score Es or the Price Score Ps and a is the ranking score. This definition means that a represents the ratio of the Es/Ps as a fraction of the mean absolute value of the stochastic process. This allows for the comparison of skill between the mostly positive processes of wind power and the day-ahead price, the switching process of the imbalance price and the mostly negative process of the price-maker effect, as it is corrected for scale by the mean absolute value. These values are shown in table 7.1 for the out-of-sample data set, including the value for the day-ahead price for completeness.

Stochastic Process	a
Whiffle	2.44
Whiffle raw	2.56
Essent	2.44
Day-ahead Price	0.67
Imbalance Price	9.86
Price-maker effect	2.97

Table 7.1: Score of Stochastic Process Sharpness

The table shows the relative skill of the scenarios for the different processes. This study hypothesizes that scenarios with higher skill better describe the uncertainty in the stochastic process and therefore need fewer scenarios to accurately represent the underlying uncertainty. Based on this hypothesis, the day-ahead price would need the lowest number of scenarios, with the three wind power processes coming in second, the price-maker effect coming in third and the imbalance price a distant fourth. As a discrete number of scenarios is needed for each process, the ratio of the price-maker effect process to the wind power process is approximated as 1:1. The ratio of these to the imbalance-price is approximated as 1:3. For tractability, the limit for the total number of scenarios is chosen as 6000. This leaves $j = 12$ for wind power and the price-maker effect and $j = 36$ for the imbalance price, resulting in 5,184 scenarios in the scenario tree.

With the number of scenarios chosen for the scenario reduction of the processes, the three scenario sets can be reduced. Figure 7.29 shows the original scenario set for the Essent wind power process, with the reduced set in colored lines of variable lightness, where brighter scenarios indicate a higher probability. The number of scenarios in the reduced set for this figure is $j = 12$. The figure shows that the reduced set covers much of the range of the original set, but certainly not the entire range. This indicates that although the heuristic has proven performance, the degree of reduction perhaps leads to a too significant loss of information.

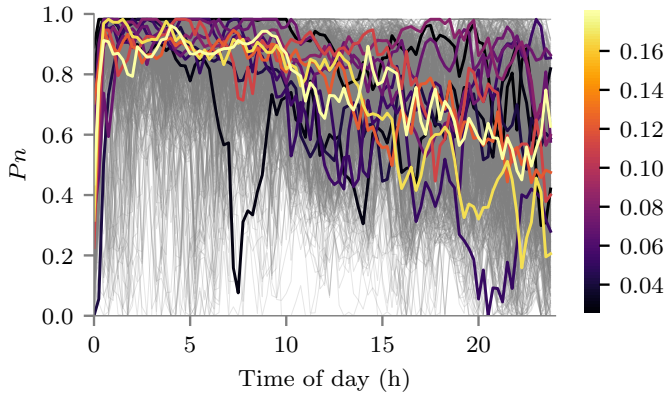


Figure 7.29: Reduced set for Essent wind power scenarios for 2018-01-08 in colored lines, where lighter colors indicate a higher associated probability. Original set is depicted in light grey.

Figure 7.30, shows the original scenario set for the imbalance price process, with the reduced set in colored lines of variable lightness, where the brightest scenario is the most probable scenario in the set. The number of scenarios in the reduced set for this figure is $j = 36$. The reduced set seems to cover much more of the range of the original scenarios. However, again, not the entire range of the original scenario set is covered.

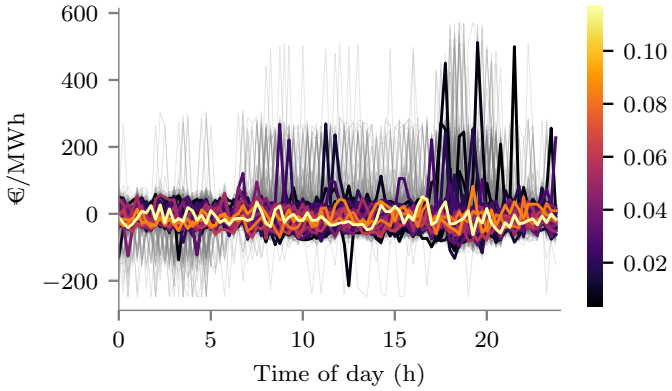


Figure 7.30: Reduced set for Imbalance price scenarios for 2018-01-08 in colored lines, where lighter colors indicate a higher associated probability. Original set is depicted in light grey.

Figure 7.31 shows the original scenario set for the price-maker effect process, with the reduced set in colored lines of variable lightness, where the brightest scenario is the most probable scenario in the set. The number of scenarios in the reduced set for this figure is $j = 12$. Again, the heuristic seems to cover much of the uncertainty of the underlying process, but much information is left out of the underlying set.

Figures 7.29 and 7.31 show that the strong reduction required to obtain a tractable optimization, leads to a lot of loss of information. Although further testing would be required to make any claims on the stability of the specific price-maker optimization result for such a strong reduction, it seems probable that the solution can improve in terms of stability if such a strong reduction can be avoided. Although this is not an issue for the price-taker version of the optimization, there is another limiting factor to this method. Namely, these

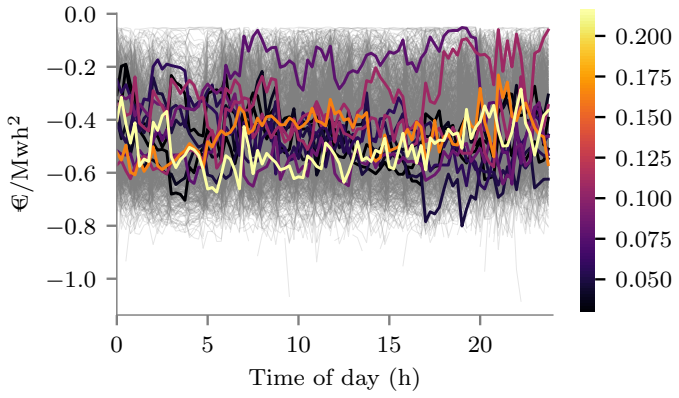


Figure 7.31: Reduced set for the price-maker effect scenarios for 2018-01-08 in colored lines, where lighter colors indicate a higher associated probability. Original set is depicted in light grey.

scenarios are coupled in the scenario tree under the assumption that they are independent, which given their relationship through the electricity market seems improbable, as explained in section 4.4.2. For these reasons, an alternative method is applied in the next section.

7.5.2. Copula Scenario Tree

A method with the aim to overcome the problem of information loss due to scenario reduction and the limiting assumption of independence between processes was introduced in section 4.4.2. This method works by extending the Guassian Copula model, used to model dependencies between forecast horizons, to include the dependencies between processes. The covariance matrix for this model is shown in figure 7.32. The mapping of the colors in this matrix is different from the earlier matrices, in order to highlight dependencies between the processes.

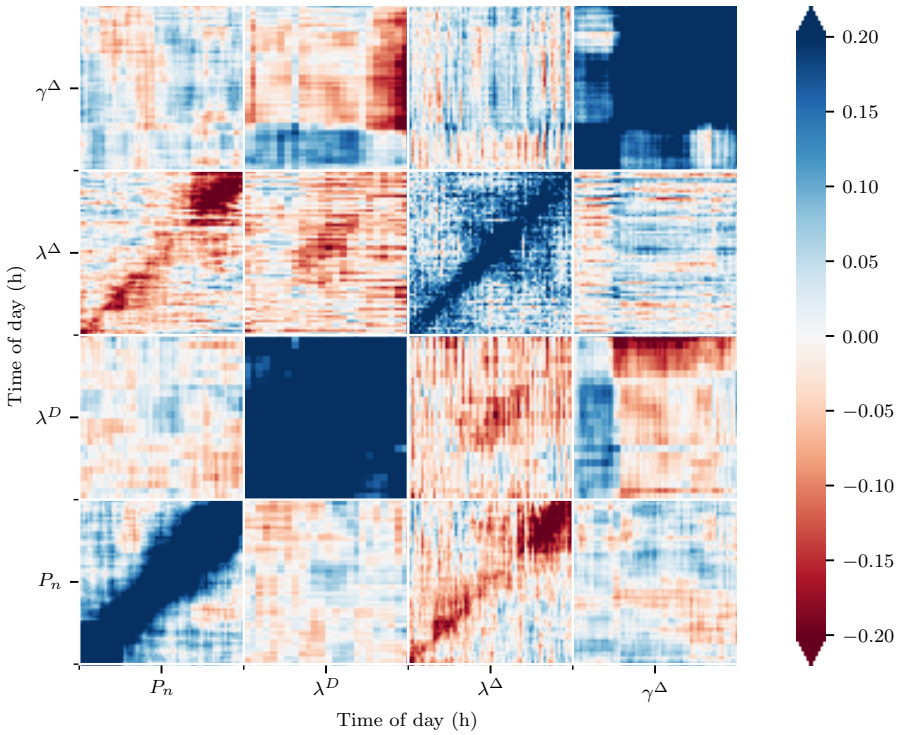


Figure 7.32: Extended covariance matrix for all processes.

From the covariance matrix several conclusions can be drawn. Firstly, there is one strong and distinct

relationship visible, which is between P_n and λ^{Δ}. This negative covariance relationship can be seen in the diagonal of the P_n, λ^{Δ} matrix. This negative covariance exists because when the forecast error for P_n is positive, the wind power producer is long on the balancing market. As the WPP has a strong covariance with all other WPPs within the Belgian market, the likelihood of a positive system imbalance is increased. As a positive system imbalance requires negative balancing action, λ^{Δ} likely decreases. Secondl, a slight negative covariance during the middle part of the day exists between λ^{D} and λ^{Δ}, which is visible on the diagonal of the $\lambda^{D}, \lambda\Delta$ matrix. This implies that when there is a positive forecast error for λ^{D}, a lower forecast error for λ^{Δ} is likely. As λ^{Δ} is defined as $\lambda^{+} - \lambda^{D}$, this negative covariance implies that positive forecast errors for λ^{D} do not lead to equivalent positive forecast errors for λ^{Δ}. Third, there is significant covariance between λ^{D} and γ^{Δ}. This can be observed in the $\lambda^{D}, \gamma^{\Delta}$ matrix. This pattern shows that when there is a positive forecast error for λ^{D} for the entire day, the forecast error for the price-maker effect is decreased during the morning and increased during the rest of the day. This increase in price-maker effect means that the supply curve becomes flatter. This effect is shifted upward for positive forecast errors for λ^{D} later in the day, meaning that the negative covariance with earlier time periods is nearing zero, while the positive covariance is shifted upward. The fact that these relationships are not present on the diagonal indicates that the day-ahead price has a relatively uniform effect for the entire day. Note that these relationships imply no causality, but are simply based on what was observed in the data. The other variables show no clear relationships. This is also valuable, as when trajectories are generated using this covariance matrix, this independence is ensured.

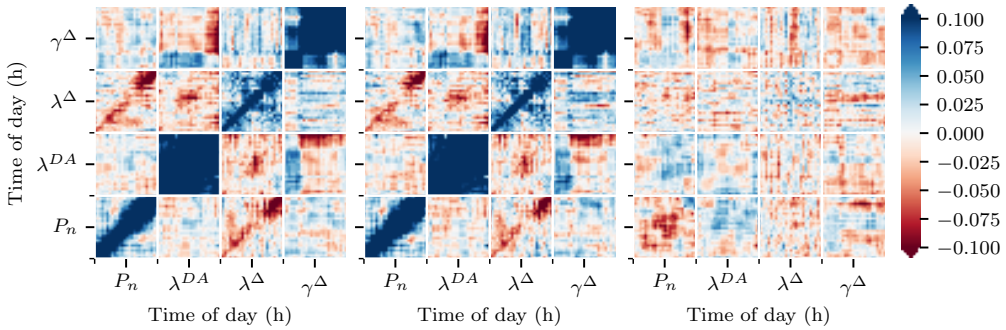


Figure 7.33: a: Original combined covariance matrix.
b: Covariance matrix of 1000 samples.
c: Difference matrix.

In earlier sections the number of scenarios generated were determined by a rule of thumb, which stated that 1000 scenarios for these variables accurately represent their underlying stochastic process. However, this does not provide a concrete measure on how well the scenario set represents the covariance matrix of the observations. To provide an answer on the question on how many samples is sufficient to accurately represent the dependencies found in the covariance matrix, the metrics MAE, RMSE and bias are applied, as introduced in chapter 5. When generating samples using the covariance matrix, a new matrix can be determined on these samples. Using these two matrices, a difference matrix can be computed, which represents how well the samples capture the dependence in the original matrix. Seeing as this matrix contains all sampling errors, the three metrics can be computed. This provides an indicator of the influence of the sample size on the preservation of dependencies, enabling a more informed choice for the sample size. First, in figure 7.33a the combined covariance matrix of the observed Guassian errors is shown, in figure 7.33b the covariance matrix of 1000 samples drawn using the covariance matrix in 7.33a is shown, and in figure 7.33c the difference matrix is shown. The range of the colormap for 7.33a and 7.33b is the same as the one used in figure 7.32, while figure 7.33c uses the scale of the colorbar to its right. This is done to be able to provide insight in the differences.

The difference matrix shows a significant difference in covariance for the wind power errors, which indicates that a sample size of 1000 does not ensure an accurate representation of the covariance for this process. Although the other processes are less affected, other parts of the combined matrix also shows signficiant differences. These differences can also be observed by comparing the interaction matrices between the processes, using the difference matrix as a guide. These significant differences indicate that a sample size of 1000 may not be sufficient to ensure the dependencies are well represented in the scenario set. Using the concept of the difference matrix, for different sample sizes the MAE, RMSE and Bias are computed.

Figures 7.34a, 7.34b and 7.34c show that sample size is important when it comes to the preservation of

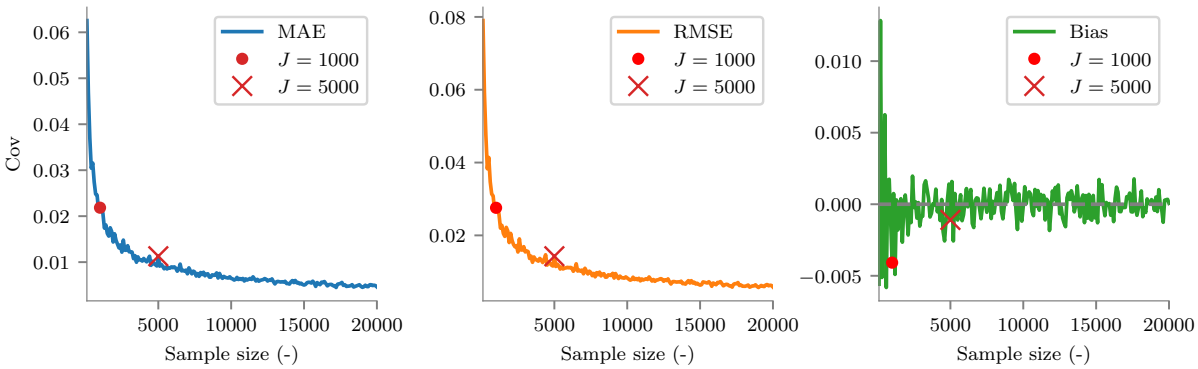


Figure 7.34: Performance metrics for different sample sizes. Scatter points indicate the performance of the reference sample size of $J = 1000$ and for the chosen sample size of $J = 5000$.
a: Mean Absolute Error.
b: Root Mean Square Error.
c: Bias.

the dependence structure. The round markers in figure 7.34 show the MAE, RMSE and Bias for the reference sample size. Although the reference sample size is not meant for such a large dependence model and it may result in adequate performance for a smaller dependency model, for this model it is not adequate. Looking at the performance, it seems the bias stabilizes gradually, while the decrease in MAE and RMSE is initially steep, but on this scale stabilizes around the 10,000 sample size mark. The bias seems to stabilize around the 3000 sample size mark. Although the choice in sample size is arbitrary, here a sample size of 5000 is chosen. This sample size results in a tractable optimization problem and it is of a similar size as the reduced set for the price-maker model, which allows for a fair comparison. This sample size significantly improves the preservation of dependence compared to a sample size of 1000. As no reduction is applied, all samples are equiprobable, which means there is no need to compute probabilities.

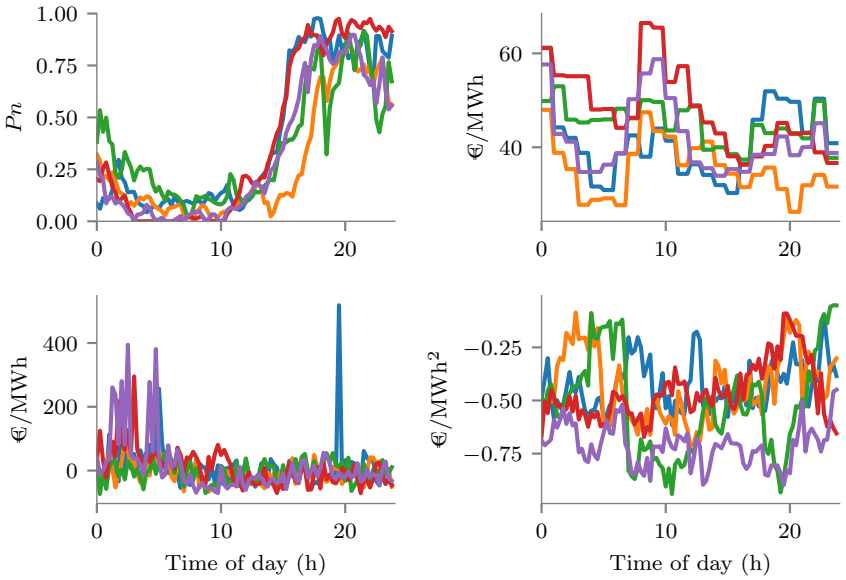


Figure 7.35: Copula Coupled Scenarios for 2018-06-30. Each scenario of equal color is generated concurrently by the dependency model.
a: Scenarios for the Essent wind power forecast.
b: Scenarios for the day-ahead price.
c: Scenarios for the imbalance price.
d: Scenarios for the price-maker effect.

These trajectories are Copula coupled, meaning that each of these is extracted from a single Gaussian

trajectory of length $96 + 24 + 96 + 96 = 312$. This method of scenario generation respects the covariance a variable has with each forecast horizon, as well as the covariance it has with all other variables' forecast for each forecast horizon. As mentioned in the previous section, conventionally scenarios are reduced based on probability distance or solution stability. However, the reasoning behind it is that tractability needs to be achieved, while keeping as much information as possible. Here the covariance matrix ensures that all relationships are respected, meaning the scenarios don't need to be separately coupled through a tree. Hence, tractability is much less an issue.

8

Results Stochastic Optimization

This chapter discusses the results from simulating the bidding strategies from chapter 3, using both the reduced scenario set and the Copula coupled scenario set from chapter 7. Section 8.1 discusses the different strategies that are analyzed. Section 8.2 discusses the performance of all strategies under the price-taker assumption. Section 8.3 repeats the analysis from section 8.2 under the price-maker assumption. Section 8.4 discusses the effects of the strategies on the system imbalance. Section 8.5 discusses the effects of the strategies on the opportunity cost for the market as a whole.

8.1. Optimized Bids

Chapter 3 introduced three different methods for determining optimized bidding strategies. The first, as discussed in section 3.1, consists of a linear optimization model which assumes that the WPP has no influence on the imbalance price. The second, discussed in section 3.2, consists of a quadratic optimization model which assumes the WPP has an influence on the price, which is modelled as a linear relationship. The last, discussed in section 3.3, proposes a closed form reformulation of the model from section 3.2.

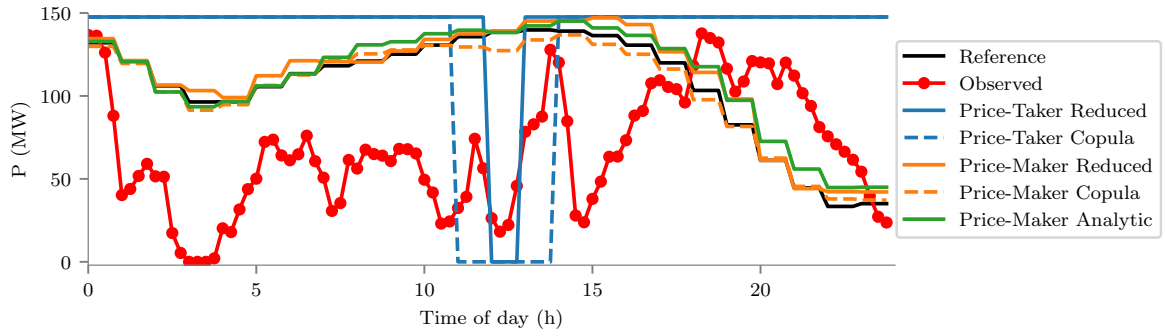


Figure 8.1: Optimized bids from all different strategies for the Essent forecast for 2018-01-05.

An example of the strategies based on the Essent wind power forecast is shown in figure 8.1. This figure shows that the different combinations of methods can result in quite different strategies. First, the price-taker optimized bid using the Copula coupled scenario set, which is referred to as the price-taker Copula strategy, results in a much longer period with a zero-bid compared to the one using the reduced scenario set, which is referred to as the price-taker reduced strategy. Second, the price-maker optimized bid shows periods with large differences between the one using the Copula coupled scenario set and the reduced scenario set, the former of which is referred to as the price-maker Copula strategy and the latter as the price-maker reduced strategy. Third, the optimized bid from the analytic model, referred to as the price-maker analytic strategy, seems significantly different from the two bids from the price-maker optimization model. This indicates that although the study by Bertrand and Papavasiliou [4] stated that the price-maker optimization model can be reformulated as an analytic model, in practice results differ from the stochastic optimization version of the

model. This is due to two assumptions underlying their reformulation. First, they state that the probability weighted average of the scenarios of the stochastic processes in the optimization model equals their respective forecasts. This is not the case, as the scenarios have different statistical properties due to their transformations from point forecasts to scenario forecasts. Furthermore, although the scenario set resulting from scenario reduction is statistically close to the original scenario set, due to the extent of reduction they differ in practice. Second, they assume that the stochastic processes are independent from each other. However, the price-maker Copula strategy assumes dependence between the processes, which results in a significant difference.

Concerning the inputs, three different forecasts are used as inputs for the modeling of wind power as a stochastic process, as discussed in section 7.1. Furthermore, section 7.5 presents two different methods for constructing the scenario set, which form the input to the two optimization models from sections 3.1 and 3.2. This leaves three different models for determining the optimal bid, three different wind power forecasts used as input and two different methods to construct the scenario tree. Furthermore, the reference strategy, as introduced in section 2.3, represents the conventional strategy. This means that for each of the forecasts a reference strategy, a price-taker reduced strategy, a price-taker Copula strategy, a price-maker reduced strategy, a price-maker Copula strategy and a price-maker analytic strategy needs to be analyzed. This means that 18 different strategies need to be tested and compared against each other. The aim of this analysis is to provide insights in the value of the different inputs, but also in the different methods used.

Each of the strategies is simulated for the period from 2018-01-01 up to and including 2018-06-30. Section 8.2 analyzes the bids from the strategies using the method applied in the studies by Rahimiyan et al. (2011) [51] and Chaves Avila et al. (2013) [9], as explained in section 2.3. Section 8.3 analyzes the strategies based on the method introduced in section 3.4. The statistical methods for comparison are explained in section 3.4.2 and 3.4.3.

8.2. Price-Taker Analysis

The first comparison of the results between the different strategies and inputs applies the price-taker analysis. This means that the revenues for all optimized bids are computed using the method from section 2.3, which assumes that the price does not change as a result of a change in strategy. Table 8.1 shows the total revenues for the different strategies and forecasts as a percentage of the reference strategy for the Essent forecast, which was operationally used for this period.

Strategy	Forecast		
	Whiffle	Whiffle Raw	Essent
Reference	100.82	101.62	100.00
Price-Taker Reduced	111.58	111.58	111.58
Price-Taker Copula	111.67	111.79	111.63
Price-Maker Reduced	102.35	103.30	101.71
Price-Maker Copula	101.87	102.77	101.09
Analytic	101.68	102.44	100.88

Table 8.1: Total market revenue for 2018-01 - 2018-06 in % of Essent reference strategy.

Several conclusions can be drawn from table 8.1. Firstly, concerning the strategies, the price-taker strategy seems to outperform all others by a large margin. Within these, the price-taker Copula strategy results in a higher revenue than the price-taker reduced strategy. Concerning the price-maker version of the optimization, the strategies using the optimization model outperform the analytic model, while the price-maker reduced strategy outperforms the price-maker Copula strategy. Concerning the different forecasts, the Whiffle raw forecast outperforms all other forecasts on revenue in all strategies. The Whiffle forecast comes in second in all strategies, while the Essent forecast comes in third in all strategies.

Figures 8.2a and 8.2b show the relationship between the MAE and the revenue. For clarity only the Essent non-reference strategies are shown. This figure shows that the price-taker strategies provide a far worse forecast compared to the others. Figure 8.2b shows that the three forecasts are very similar on the MAE, but are quite distinct in terms of revenue. Of all strategies, the price-maker strategy using the Copula coupled scenario set is most similar to the original forecast on MAE. As all strategies are strongly outperformed by the price-taker strategies, the remainder of this analysis focuses on the price-taker strategies. Results for all strategies can be found in appendix B.1.

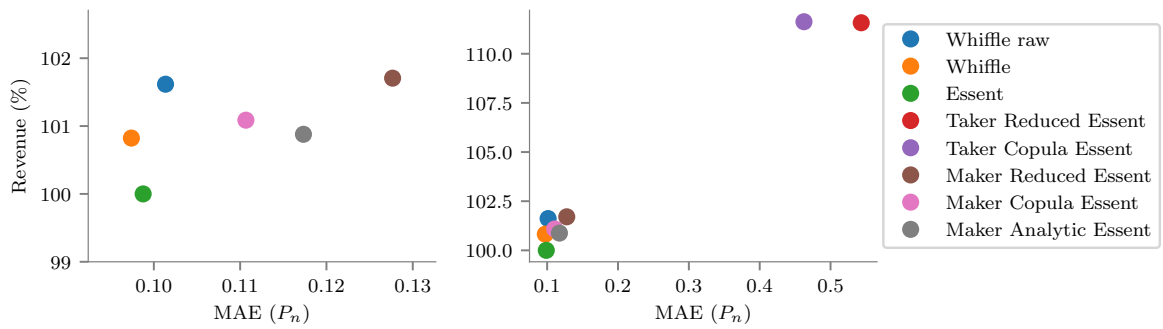


Figure 8.2: Mean Absolute Error (MAE) versus total price-taker revenue as a percentage of the reference strategy using the Essent forecast.
a: Whiffle, Whiffle raw and Essent reference strategies and Essent price-maker strategies.
b: Whiffle and Whiffle raw reference strategies and all Essent strategies.

Figure 8.3 shows results for the expected difference of the mean or shift of the revenue distribution, as introduced in section 3.4.2. The table shows the scale of the shift in € on the colorbar and the relative percentage of this shift in % in the cells. This indicates the difference in expected revenue per PTU, which in this case is a quarter hour. Although the shift in revenue distribution is relatively minor, it shows that the price-maker Copula strategy outperforms all price-maker reduced strategies. Of all forecasts, the Copula scenario set built using the Whiffle raw forecast outperforms all other forecasts. Based on these observations, the strategy from the price-taker optimization model using the Copula coupled scenario set with the Whiffle raw forecast is likely the most valuable. It also shows that there is no difference between the different forecasts used in the price-taker reduced strategy. This is due to the fact that there is no difference in price scenarios between these scenario sets and the wind power forecast plays little role in the price-taker optimization.

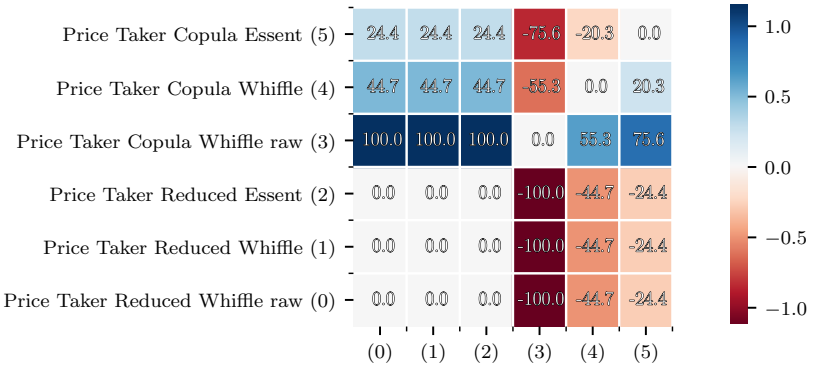


Figure 8.3: Expected price-taker difference of the mean for the price-taker strategies. The colorbar indicates the difference in €. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=€1.13. The expected difference is computed for rows over columns.

Figure 8.4 shows the Conditional Value at Risk (CVaR_α) scores, where $\alpha = 5\%$. As the $\text{CVaR}_{5\%}$ value indicates the expected value for the 5% lowest revenues, it provides insight in the risk of a strategy. Figure 8.4 shows that the difference in $\text{CVaR}_{5\%}$ is much greater than the difference in expected mean. Again, the price-taker strategies using the Copula coupled scenario set outperform the strategies using the reduced scenario set. Although the differences percentage wise are much less significant compared to the expected difference of the mean, again the strategy using the Whiffle raw forecast outperforms the other strategies. Based on these observations, both on expected value and on risk the price-taker strategy using the Copula coupled scenario set with the Whiffle raw forecast is to be preferred.

Figure 8.5a shows the average opportunity price the strategies obtain for their short volume, long volume and total imbalance volume. Several observations can be made. First, the price-taker Copula strategies incur a worse average opportunity price (λ^{Δ^-}) for being short, although the one using the Whiffle raw forecast incurs the least disadvantageous λ^{Δ^-}. Second, the price-taker reduced strategies also incur the most advan-

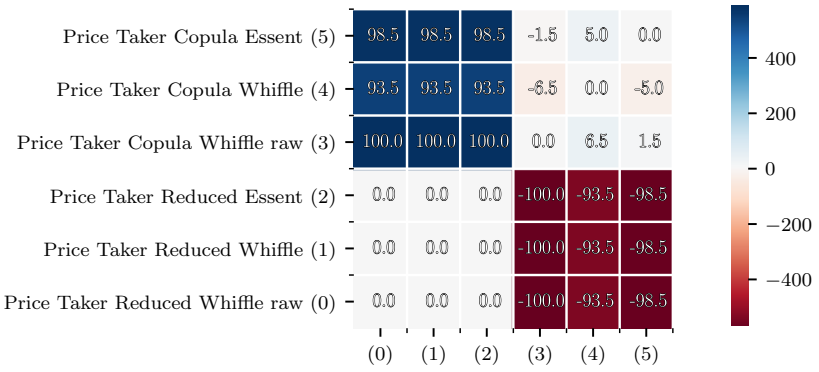


Figure 8.4: Price-taker difference of $CVaR_{5\%}$ for the price-taker strategies. The colorbar indicates the difference in €. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=€576.11. The expected difference is computed for rows over columns.

tageous average opportunity price (λ^{Δ^+}) for being long.

Figure 8.5b shows the short volume, long volume and total volume. First, it shows that the price-taker Copula strategies have a much smaller short volume than the strategies using the reduced set. Second, the strategies using the Copula coupled scenario set have a much larger long volume than the strategies using the reduced set. Thirdly, the strategies using the reduced set are much more short in total than the strategies using the Copula coupled scenario set.

Figure 8.5c shows the total opportunity revenue from being short, long and in total. The first column shows that the price-taker Copula strategy with the Whiffle raw forecast realized the smallest opportunity loss from being short. The second column shows that the price-taker Copula strategy using the Whiffle and Essent forecasts realize the highest opportunity from being long, while the Copula strategy using the Whiffle raw forecast realizes the smallest opportunity from being long. The final column shows that on total opportunity revenue the price-taker Copula strategy using the Whiffle raw forecast realizes the smallest loss, making it the most valuable strategy under the price-taker assumption.

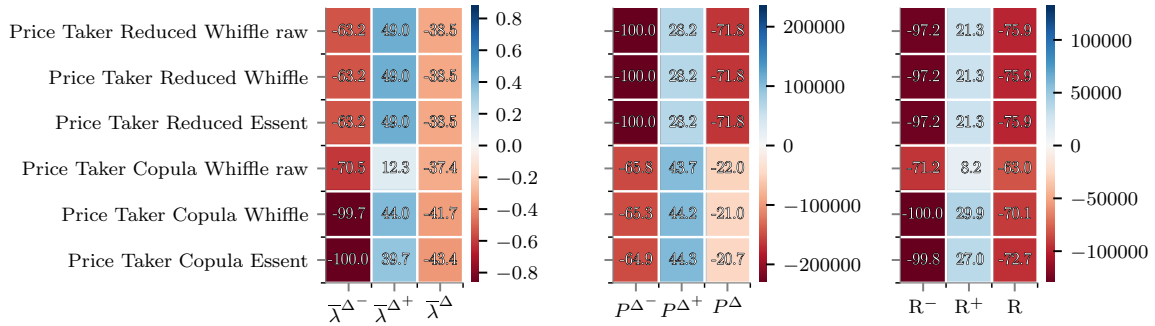


Figure 8.5: Market quantities for the price-taker strategies. The colorbars indicate quantities in the original unit of measurement. The cells indicate quantities as a percentage of the scale of the color bars.
a: Average opportunity price realized for short volume, long volume and total imbalance volume. 100%=€0.86/MWh.
b: Short volume, long volume and total imbalance volume. 100%=231,122MWh.
c: Realized opportunity for being short, long and total. 100%=€130,022.

The observations from figure 8.5 can be used to explain the superior performance of price-taker Copula strategy using the Whiffle raw forecast. First, although it realizes the least favorable opportunity price for being long, because it realizes a relatively favorable short opportunity price and it has a relatively small short volume, on average it realizes the most favorable opportunity price. Its difference with the other strategies using a Copula coupled scenario set on volume is relatively minor, which leads to the conclusion that the Whiffle raw forecast distinguishes itself by its favorable correlation with favorable opportunity prices when short. On the other hand, the other strategies using the Copula coupled scenario set incur relatively unfa-

vorable prices compared to the strategies using the reduced scenario set, but still perform better due to the fact that they have a smaller short volume. This indicates that having a large short volume is unfavorable. This makes sense, as upward regulation is more expensive than downward regulation, resulting in a relatively unfavorable short opportunity price. The reason the strategies using Copula coupled scenario set are so different is due to the dependency model. This dependency model leads to relatively low imbalance price scenarios when the wind power scenarios are relatively high and vice versa. Because it already includes the effect of aggregate wind power on the imbalance indirectly through this dependency model, its scenario set is more realistic, which results in a better optimization.

To explain this influence of the Copula coupling on the result, insight is needed from sections 7.1.3 and 7.3. On average, the Essent forecast for this wind farm for 2018 is 38% of normalized power. Figure 7.4, which shows the conditional kernel density model for the Essent forecast error, shows that this means that the error distribution for wind power on average tends to be skewed towards higher values. On average, the forecast for the day-ahead price λ^D was €42.82/MWh. Figure 7.20, which shows the conditional kernel density model for the imbalance difference price λ^Δ forecast error, shows that this means that on average it is skewed towards higher values. Given the negative correlation between the Gaussian forecast errors of these two processes, which is shown in figure 7.32, this means that relatively small negative forecast errors for wind power on average result in relatively large positive forecast errors for λ^Δ, while relatively large positive forecast errors for wind power on average result in relatively small negative forecast errors for λ^Δ. This makes sense, as upward regulation is more expensive than downward regulation. Combining these two observations, using the Copula coupled scenario set the optimization is less likely to profit from being short relative to the reduced set, while it is also less likely to profit from being long relative to the reduced set, thus placing an empirically derived penalty in the composition of the Copula coupled scenario set. This leads to more conservative strategies with the Copula coupled scenario set. However, given the fact that the price-taker strategy only offers 0 or $P^{\max}$, this conservative bidding does not occur. Hence, the conservative effect is that the price-taker Copula strategy has a smaller short volume. This is because the coupling is non-existent in the reduced set, meaning that negative wind power forecast errors are just as likely to be in a tree scenario with a small positive λ^Δ forecast error as a small negative λ^Δ error, which is not the case in reality. On the other hand, 55% of all scenarios for λ^Δ for the out-of-sample set were negative, which implies that being short pays off. In optimization, on average, moving towards being short is thus penalized more heavily in total revenue for the Copula coupled scenario set, while being long is penalized less, thus resulting in a smaller short position on average for the Copula coupled scenario set.

Table 8.2 shows the correlations with the system imbalance. It shows that the price-taker strategies have a distinctly different correlation compared to the other strategies. Furthermore, the price-taker strategy with the Copula coupled scenario set has a lower correlation with the SI compared to the reduced price-taker strategy, which indicates that is more likely to provide passive balancing successfully, although it is not significantly rewarded for its increased passive balancing.

Strategy	Forecast		
	Whiffle	Whiffle Raw	Essent
Reference	0.232	0.205	0.240
Price-Taker Reduced	0.037	0.037	0.037
Price-Taker Copula	-0.013	-0.014	-0.016
Price-Maker Reduced	0.174	0.156	0.191
Price-Maker Copula	0.177	0.155	0.197
Analytic	0.196	0.175	0.204

Table 8.2: Correlations between price-taker strategy imbalances and the system imbalance (SI) for 2018-01 - 2018-06.

In conclusion, under the price-taker assumption the price-taker strategies strongly outperformed all other strategies, where the strategy with the Copula coupled set slightly outperform the strategy with the reduced coupled set. However, as explained in chapter 3, the imbalance price is influenced by the WPP, which is why the price-taker assumption does not hold. For this reason the next section discusses the price-maker analysis.

8.3. Price-maker analysis

Before the results from the price-maker analysis can be discussed, first the performance of the price reconstruction described in section 3.4 is discussed.

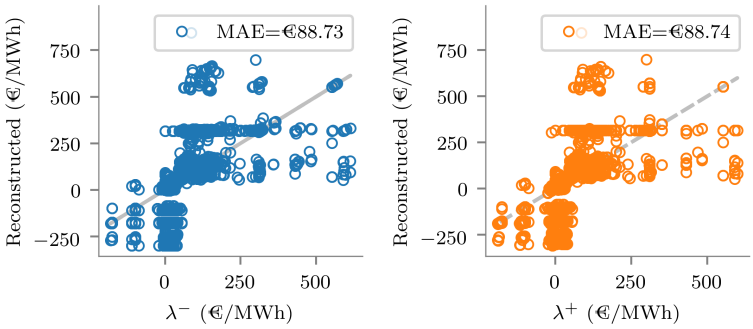


Figure 8.6: When NRV<0, Bias=€51.95. When NRV>0, Bias=€-120.03.
a: Scatter of λ^- versus reconstructed λ^-. Bias=€-28.02.
b: Scatter of λ^+ versus reconstructed λ^+. Bias=€-28.01

In figure 8.6 two scatter plots are shown, which show the performance of the price reconstruction algorithm for 2018-01 - 2018-06, assuming no change in strategy. It shows that the price is not reconstructed particularly well and results in a MAE of 28.02€/MWh for the λ^--reconstruction and -28.01€/MWh for the λ^+-reconstruction. This is due to two things. First, as the published supply curve is a crude approximation to the actual supply curve, it results in plateaus in the scatter plot, which means that prices are assumed not to change, while in practice they do change because of a smaller step size in the actual supply curve. Second, the MDB and MIB are assumed to be the minimum and maximum values of the NRV signal at a one-minute resolution for each PTU, which they are not in practice, due to higher resolution bids and due to grid constraints.

Unfortunately, the reconstructed price is biased very differently when the NRV is negative, versus when the NRV is positive. This would strongly favor strategies that are more often long than short, as being long is overly rewarded. Because of the poor performance of reconstructing the price using the approximated supply curve with the approximated MDB and MIB, the reconstructed price is not used. However, as both the MDP and MIP are published together with the SI and NRV, one can still recreate the effect of sign changes of the NRV on the imbalance price. As shown in figure 3.5b, the sign change of the NRV is one of the strongest influences on the price, as it causes the switching behavior. Furthermore, the imbalance price is characterized by relative stability when the sign of the NRV does not switch between PTUs, which is due to the the supply curve being a step function. As the price-maker bids are relatively similar to the reference strategy, as can be observed in figures 8.1 and 8.2, it seems fair to assume no change in price other than the possible switch caused by a change in NRV. Hence, revenues are reconstructed based on the sign of the NRV, which should allow a more accurate depiction of real-world performance. In table 8.3 the revenues using the price-reconstruction are shown.

Strategy	Forecast		
	Whiffle	Whiffle Raw	Essent
Reference	100.53	100.74	100.00
Price-Taker Reduced	37.24	37.24	37.24
Price-Taker Copula	55.42	55.42	55.42
Price-Maker Reduced	99.93	100.21	99.55
Price-Maker Copula	100.97	101.17	100.75
Analytic	100.32	100.54	99.65

Table 8.3: Total market revenue for 2018-01 - 2018-06 in % of Essent reference strategy.

Several conclusions can be drawn from table 8.3. First, the price-taker strategies perform significantly worse. This is due to the mechanism by which the price is influenced: When the strategy provides too much balancing power, the sign of the NRV changes, which turns an opportunity gain into an opportunity loss,

thus hurting the revenue for the strategy. Of the price-maker strategies, only the price-maker Copula strategy leads to a consequent increase in revenue compared to the reference strategies, while the other two price-maker strategies lead to a decrease in revenue. The only successful non-reference strategy is the price-maker Copula strategy, where the Whiffle raw forecast obtains the most significant increase in revenue. All price-maker strategies perform worse compared to the price-taker analysis. This is due to the fact that when a strategy provides too much balancing power, in the price-taker analysis, the sign of the NRV does not change, while it does in the price-maker analysis. Whenever such a change occurs, the price changes strongly due to the large step in the supply curve between decremental and incremental bids.

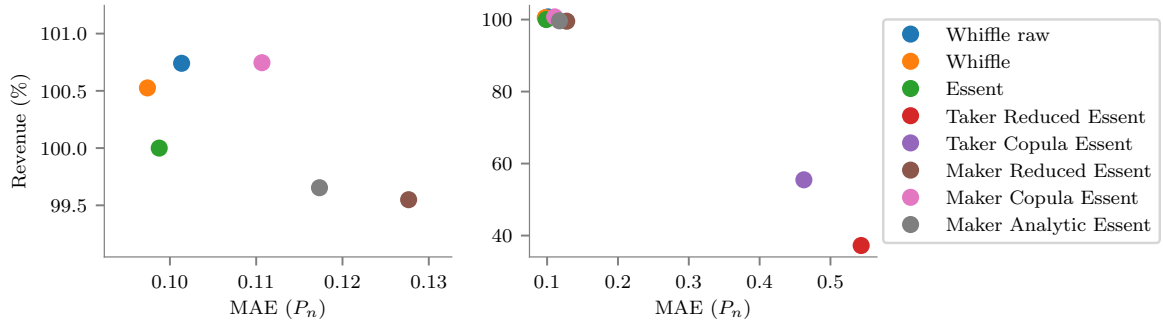


Figure 8.7: Mean Absolute Error (MAE) versus total price-maker revenue as a percentage of the reference strategy using the Essent forecast.
a: Whiffle, Whiffle raw and Essent reference strategies and Essent price-maker strategies.
b: Whiffle and Whiffle raw reference strategies and all Essent strategies.

Figures 8.7a and 8.7b shows the relationship between the MAE and the revenue. For clarity only the Essent non-reference strategies are shown. Figure 8.7b again shows that the price-taker strategies provide a far worse forecast compared to the others, which in the price-maker analysis leads to a strong decrease in revenue. Of all non-reference strategies, the price-maker Copula strategy is most successful both in terms of revenue and MAE. As the price-taker strategies are strongly outperformed by the reference and price-maker strategies, the remainder of this analysis focuses on the reference and price-maker strategies. Results for all strategies can be found in appendix B.2.

Figure 8.8 shows the expected differences of the mean for the reference and price-maker strategies, which confirms the results from the table. Several observations can be made. First, it shows that only the price-maker Copula strategy is able to consequently realize a positive shift of the revenue distribution from the reference strategies. Second, the two Whiffle reference strategies already provide a strong increase in revenue. Third, the price-maker Copula strategies are all ranked higher than the highest ranked reference strategy. Based on these observations, the preferred strategy would be the price-maker Copula strategy using the Whiffle raw forecast.

Figure 8.9 shows the Conditional Value at Risk scores, where $\alpha = 5\%$. Figure 8.9 shows that the difference in $\text{CVaR}_{5\%}$ is again much greater than the expected difference in mean. Unlike the expected difference in mean, the price-maker Copula strategy does not always outperform the reference strategies on risk. Although the one using the Essent forecast does outperform its reference strategy, the one using the Whiffle forecast underperforms its reference strategy relatively strongly and the one using the Whiffle raw forecast slightly underperforms its reference strategy. Based on these observations, the choice of strategy depends on the risk preference of the WPP, as the price-maker Copula strategy using the Whiffle raw forecast outperforms the Essent one on expected mean value, whereas the Essent one performs better on risk.

Figure 8.10a shows the average opportunity price the strategies obtain for their short volume, long volume and total imbalance volume. From it, several observations can be made. First, the reference strategies incur a worse λ^{Δ^-} for being short compared to the optimized strategies, of which the price-maker Copula strategy incurs the least advantageous price. Second, the reference strategy also incurs the least advantageous λ^{Δ^+} for being long, where the price-maker reduced and the price-maker Copula obtain the most favourable prices.

Figure 8.10b shows the short volume, long volume and total volume. First, it shows that the price-maker Copula strategies have a significantly smaller short volume than the other optimized strategies, although larger than the volume for all reference strategies. Second, the price-maker Copula strategies have a significantly larger long volume than the other strategies. Third, the other optimized strategies are significantly

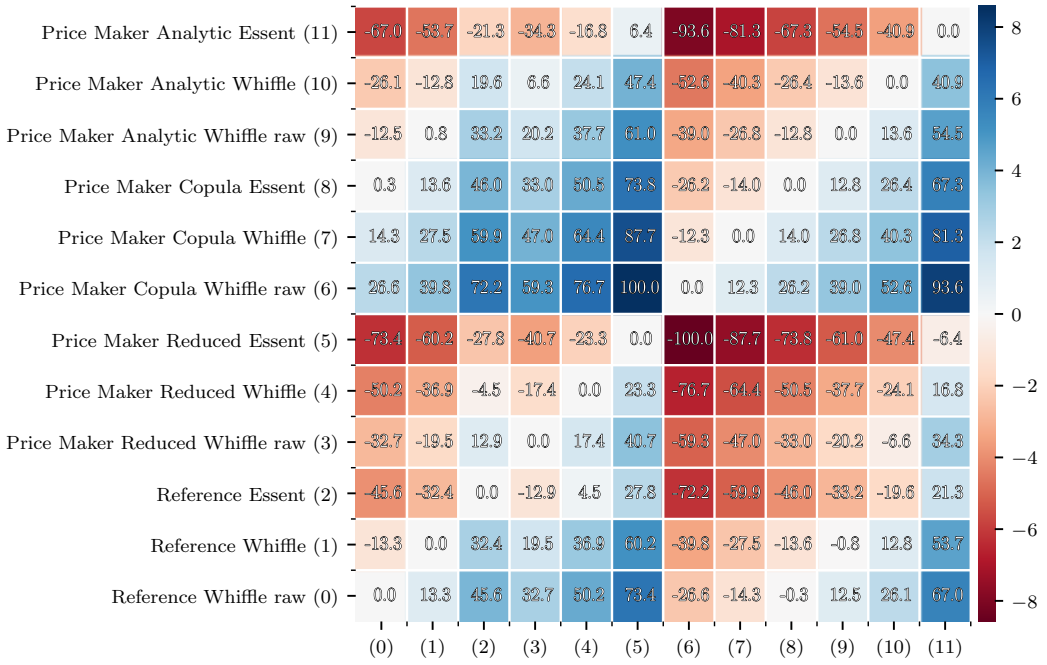


Figure 8.8: Expected price-maker difference of the mean for the price-maker strategies. The colorbar indicates the difference in €. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=€8.57. The expected difference is computed for rows over columns.

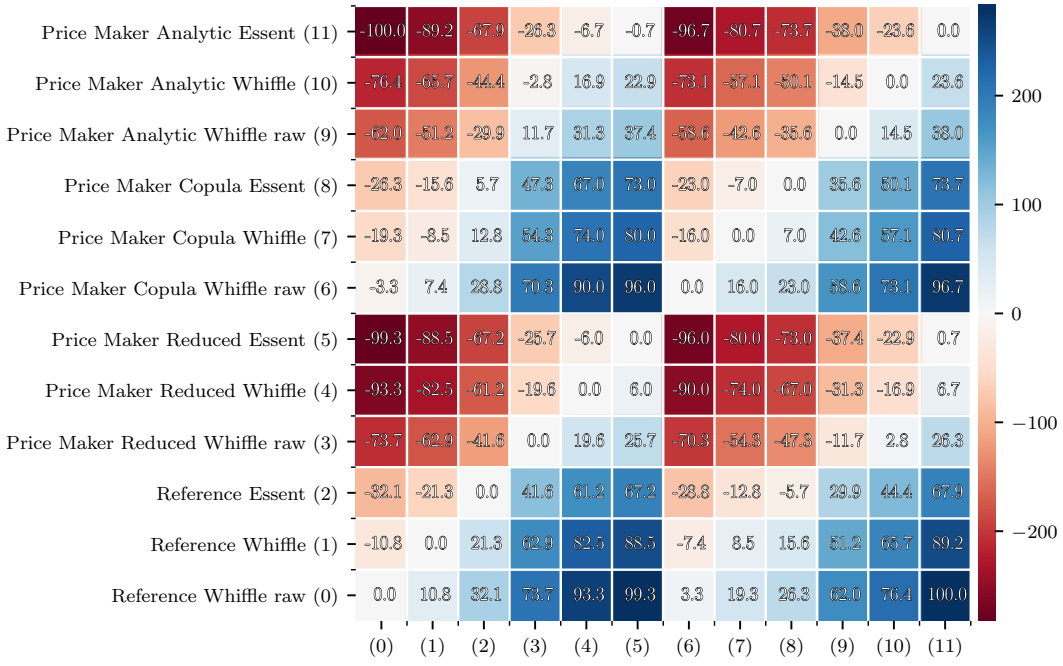


Figure 8.9: Price-taker difference of $CVaR_{5\%}$ for the price-maker strategies. The colorbar indicates the difference in €. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=€282.41. The expected difference is computed for rows over columns.

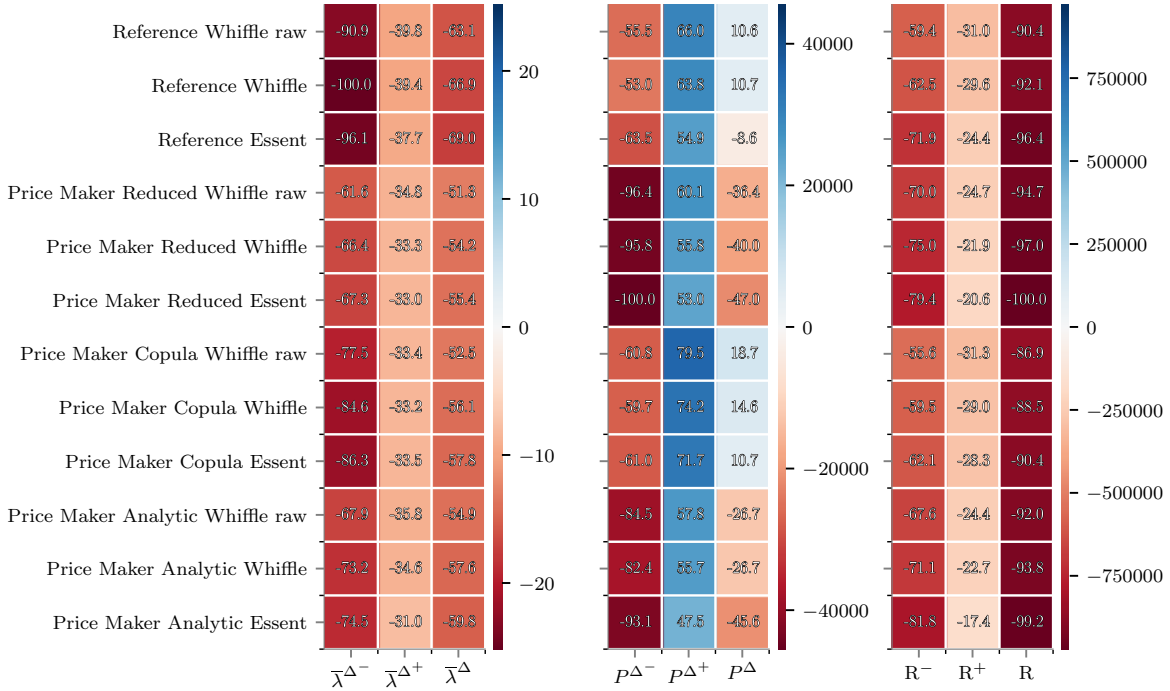


Figure 8.10: Market quantities for the reference and price-maker strategies. The colorbars indicate quantities in the original unit of measurement. The cells indicate quantities as a percentage of the scale of the colorbars.
a: Average opportunity price realized for short volume, long volume and total imbalance volume in €/MWh. 100%=€25.16/MWh.
b: Short volume, long volume and totale imbalance volume in MWh. 100%=45,505.65MWh.
c: Realized opportunity for being short, long and total in €. 100%=€970,972.00.

more short in total than the reference strategy or the price-maker Copula strategy.

Figure 8.10c shows the total opportunity revenue from being short, long and in total. The first column shows that the price-maker Copula strategy using the Whiffle raw forecast realized the smallest opportunity loss from being short. The second column shows that the price-maker analytic strategy realizes the smallest opportunity from being long, while the price-maker Copula strategies realize the highest opportunity loss from being long. The final column shows that on total opportunity revenue the strategy using the Copula coupled scenario set with the Whiffle raw forecast realizes the smallest loss, making it the most valuable strategy under the price-maker assumption.

The reason the price-maker Copula strategies outperform the other strategies seems to lie with the fact that the former avoid opportunity costs from being short. This is in contrast with the other non-reference strategies as these choose to increase their short volume relative to their long volume. This is due to the fact that the more realistic Copula scenario sets already include a penalty on being short, which was explained in section 8.2.

Table 8.4 shows the correlations with the reconstructed system imbalance. The computation of the effect a strategy has on the system imbalance is straightforward, as the imbalance of the operationally used Essent reference strategy is known. The imbalance of this strategy in MW is subtracted from the SI, after which the imbalance of the other strategies can be computed. The table shows that the price-taker strategies have a distinctly different correlation compared to the other strategies, as they to a large extent determine the sign and amplitude of the imbalance, which helps explain their poor eonomic performance. Concerning the price-maker strategies, the strategy using the Copula coupled scenario set shows the smallest correlation with the system imbalance.

In conclusion, it seems more risky and more difficult to profit from λ^{Δ^-} than from λ^{Δ^+}. The Copula strategies recognize this and are the only ones that avoid the relatively unfavorable short opportunity prices successfully. This leads to a smaller opportunity loss from being short while maintaining a relatively favourable opportunity loss from being long. Of the forecasts, the Whiffle raw forecast is superior to the others, both

Strategy	Forecast		
	Whiffle	Whiffle Raw	Essent
Reference	0.250	0.240	0.240
Price-Taker Reduced	0.489	0.489	0.489
Price-Taker Copula	0.458	0.457	0.456
Price-Maker Reduced	0.232	0.229	0.231
Price-Maker Copula	0.215	0.209	0.215
Analytic	0.233	0.226	0.224

Table 8.4: Correlations between price-maker strategy imbalances and the system imbalance (SI) for 2018-01 - 2018-06.

when used as a reference strategy and when used as input to the price-maker optimization. Based on all observations, it seems that the Copula based optimization model is best able to reliably provide passive balancing and is therefore deemed most likely to perform well when applied operationally.

8.4. System Effects

This section discusses the physical effects of the strategies on the grid. This effect is captured by the price-maker system imbalance (SI). As explained in 3.4.3, interest lies not with the sum of all imbalances, but with the sum of all absolute values of the SI, as the SI is centered around zero.

Strategy	Forecast		
	Whiffle	Whiffle Raw	Essent
Reference	100.36	100.36	100.00
Price-Taker Reduced	118.92	118.92	118.92
Price-Taker Copula	112.78	112.75	112.75
Price-Maker Reduced	100.08	100.13	99.98
Price-Maker Copula	99.89	99.95	99.77
Price-Maker Analytic	100.03	100.03	99.72

Table 8.5: Reconstructed total absolute system imbalance for 2018-01 - 2018-06 in % of Essent reference strategy.

Table 8.5 show the values for the reconstructed sum of the absolute system imbalance for all strategies. Several conclusions can be drawn from the table. First, the price-maker Copula strategy is on average most successful in the provision of balancing power. Given how the imbalance price is modelled in the optimization model, the price seems to provide a good indicator on the need for passive balancing. Surprisingly, the price-maker analytic strategy using the Essent forecast performs best in terms of total SI. Of the three forecasts, the Essent forecast seems to consequently result in the largest decrease in imbalance. The price-taker strategies show a significant increase in absolute imbalance, which indicates that these are not capable of providing balancing power. As the price-taker strategies are strongly outperformed by the reference and price-maker strategies, the remainder of this analysis focuses on the latter two. Results for all strategies can be found in appendix B.3.

Figure 8.11 shows the expected difference in mean system imbalance per PTU. It confirms the results of table 8.5, where the price-maker Copula strategies outperform all other strategies on average, except for the price-maker analytic strategy using the Essent forecast. It also shows that the price-maker Copula strategy manages to realize a greater reduction in imbalance using the Whiffle and Whiffle raw forecast than with using the Essent forecast relative to their reference strategies. This may be caused by the specific bias remaining in the Whiffle and Whiffle raw scenarios or by the different correlations present in its Copula model. Last, the figure shows that based on the reduction of absolute system imbalance, the price-maker analytic strategy using the Essent forecast is deemed superior.

Figure 8.12 shows the conditional value at risk at $\alpha = 5\%$ ($CVaR_{5\%}$), which is computed differently from the previous $CVaR_{5\%}$, according to equation 3.20. This means it shows the difference in expected value of the 5% highest absolute system imbalances for each strategy. Several observations can be made. First, the reference strategy for the Whiffle and Whiffle raw forecast show a significant decrease compared to the operationally used Essent reference strategy. Of all strategies, the price-maker strategies using the Copula coupled scenario

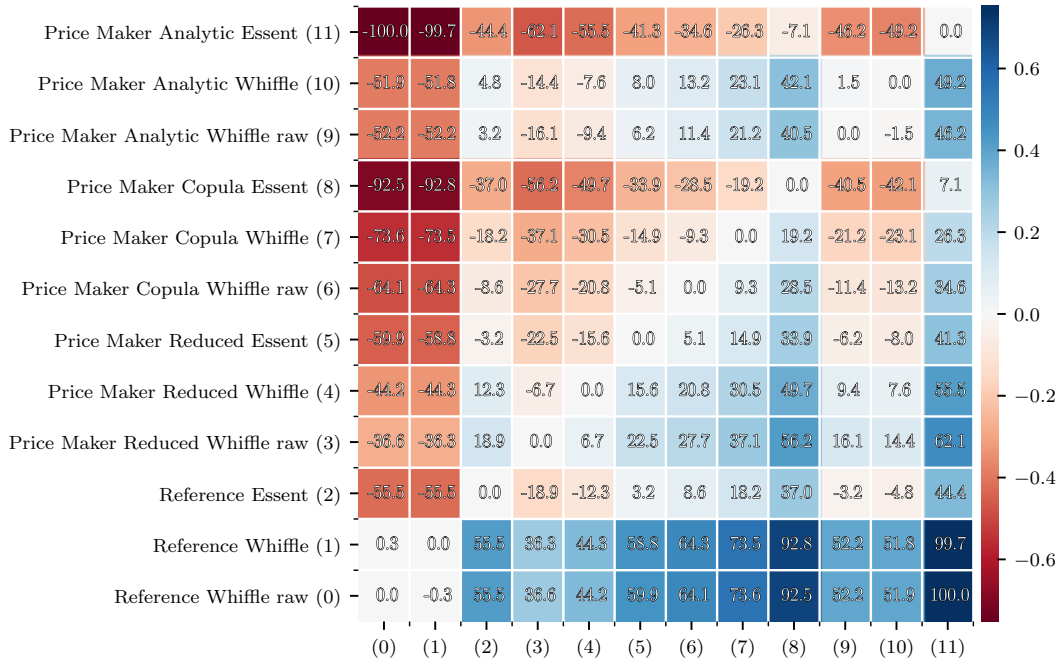


Figure 8.11: Expected price-maker difference of the mean system imbalance for the reference and price-maker strategies. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=0.75MW. The expected difference is computed for rows over columns.

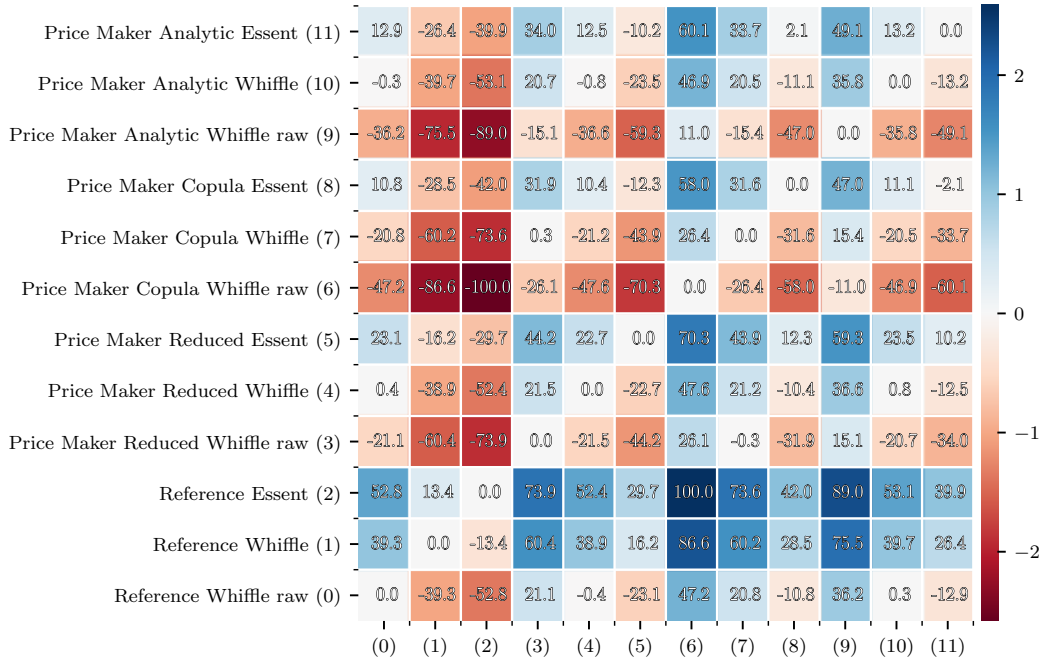


Figure 8.12: Expected price-maker difference of the $CVaR_{5\%}$ of the system imbalance for the reference and price-maker strategies. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=2.58MW. The expected difference is computed for rows over columns.

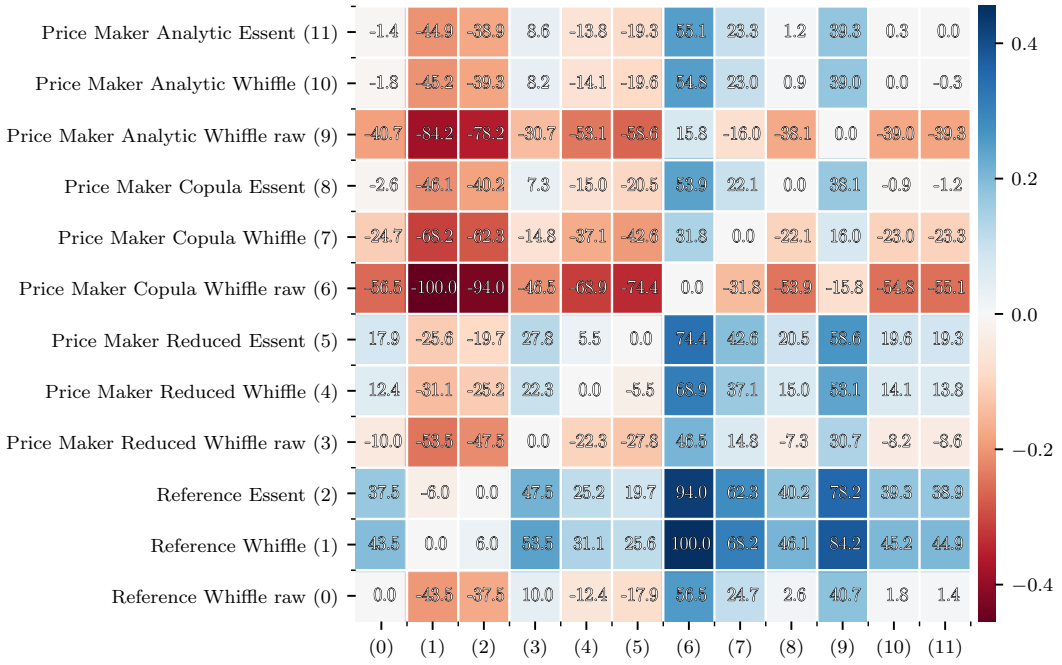


Figure 8.13: Expected price-maker difference of the volatility of the system imbalance for the reference and price-maker strategies. The colorbar indicates the difference in MW. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=0.45MW. The expected difference is computed for rows over columns.

set show the largest decrease relative to all other strategies, including the price-maker analytic strategy using the Essent forecast. Also, the version using the Whiffle raw forecast significantly outperforms those using the other forecasts, which was to be expected given its lower correlation with the system imbalance shown in table 8.4.

Figure 8.13 shows the volatility of the strategies in MW, as computed through equation 3.21. First, it again shows that the Whiffle and Whiffle raw reference strategies already provide an improvement over the Essent reference strategy. Second, of all strategies, the price-maker strategy using the Copula coupled scenario set again outperforms all other strategies.

In conclusion, the price-maker Copula strategy consistently performs best, except on expected difference of the mean system imbalance, where it is outperformed by one specific strategy, the price-maker analytic strategy using the Essent forecast. These results make sense because of how the price-maker optimization model works: The model anticipates prices rather than imbalances, where high volatility and extreme imbalances are more likely to coincide with extreme prices. This is due to the fact that in situations with high volatility more expensive generators need to be activated, sometimes even out of merit order due to ramp-rate constraints. Furthermore, extreme imbalances are always more expensive to balance. Of forecasts used for the Copula coupled scenario set, the Whiffle raw outperformed the other scenario sets, which makes sense, given its lower correlation with the system imbalance, as presented in table 8.4. In conclusion, as the optimization model does not aim to counter system imbalances, but aims to anticipate expensive imbalances, the effect of the optimization on the average system imbalance may not be beneficial, while the effect on the cost of countering imbalances will be, given a successful optimization, such as the price-maker Copula strategy using the Whiffle raw forecast.

8.5. System Costs

This section discussed the effects of the strategies on the total opportunity revenue of the market as a whole, as discussed in section 3.4 and computed through equation 3.22. This revenue represents the value the market as a whole obtains from trading on the balancing market. For each strategy this value is different, as strategies both influence the total system imbalance and the imbalance price, through the switching of the

sign of the NRV.

Strategy	Forecast		
	Whiffle	Whiffle Raw	Essent
Reference	100.00	99.81	100.00
Price-Taker Reduced	112.38	112.38	112.38
Price-Taker Copula	106.70	106.76	106.85
Price-Maker Reduced	99.85	99.60	99.78
Price-Maker Copula	99.47	99.17	99.41
Price-Maker Analytic	99.90	99.70	99.95

Table 8.6: Reconstructed total system revenues for 2018-01 - 2018-06 in % of Essent reference strategy.

Table 8.6 shows the values for the sum of the total system opportunity revenue. As the total system opportunity revenue is negative, lower values indicate a positive effect. Several conclusions can be drawn from it. Firstly, the price-taker strategies lead to significant increases in total system cost. Concerning the price-maker strategies, the strategy using the Copula coupled scenario set is able to reduce opportunity cost for the system as a whole the most. While this is partly due to its own increase in opportunity revenue, it is the only one that also increases opportunity revenue for the system as a whole outside of its own increase. Of the three forecasts, the reference strategies of the Whiffle raw forecast already has a positive impact on the market as a whole, which increases when it is used in the price-maker Copula strategy. As the price-taker strategies are significantly outperformed by the reference and price-maker strategies, the remainder of the analysis only discusses the latter two. Results for all strategies can be found in appendix B.4.

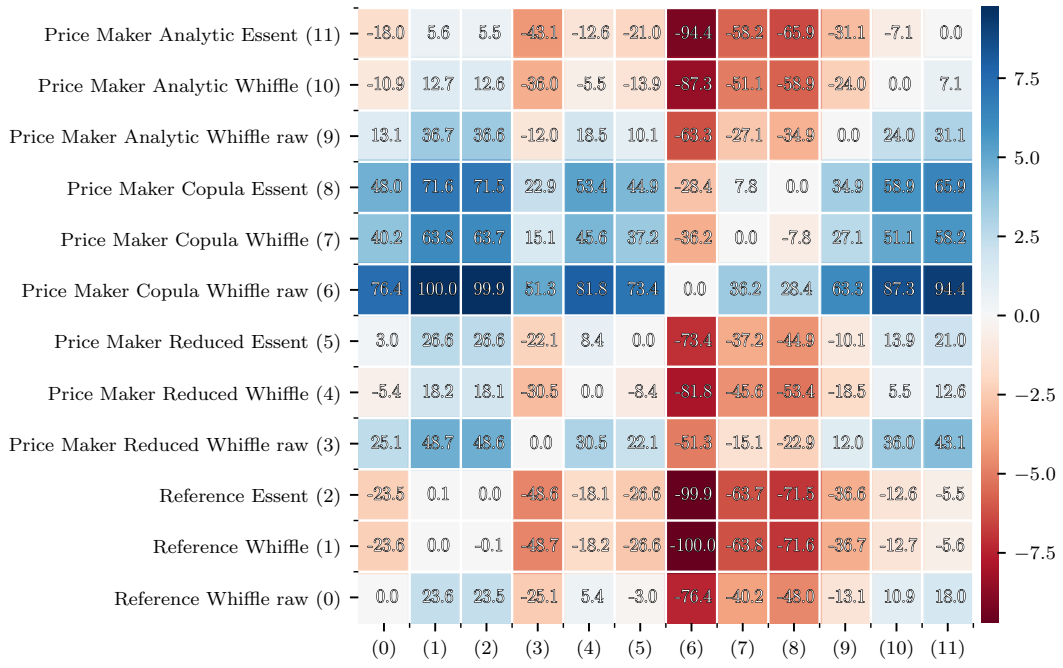


Figure 8.14: Expected price-maker difference of the mean system opportunity for the reference and price-maker strategies. The colorbar indicates the difference in €. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=€9.71. The expected difference is computed for rows over columns.

Figure 8.14 shows the expected difference of the mean between the relevant strategies. It shows that of the reference strategies the one using the Whiffle raw forecast is superior in terms of opportunity revenue. Furthermore, it shows that the application of the price-maker Copula strategy increases it further. Last, it shows that the other price-maker strategies lead to a significant decrease in opportunity revenue.

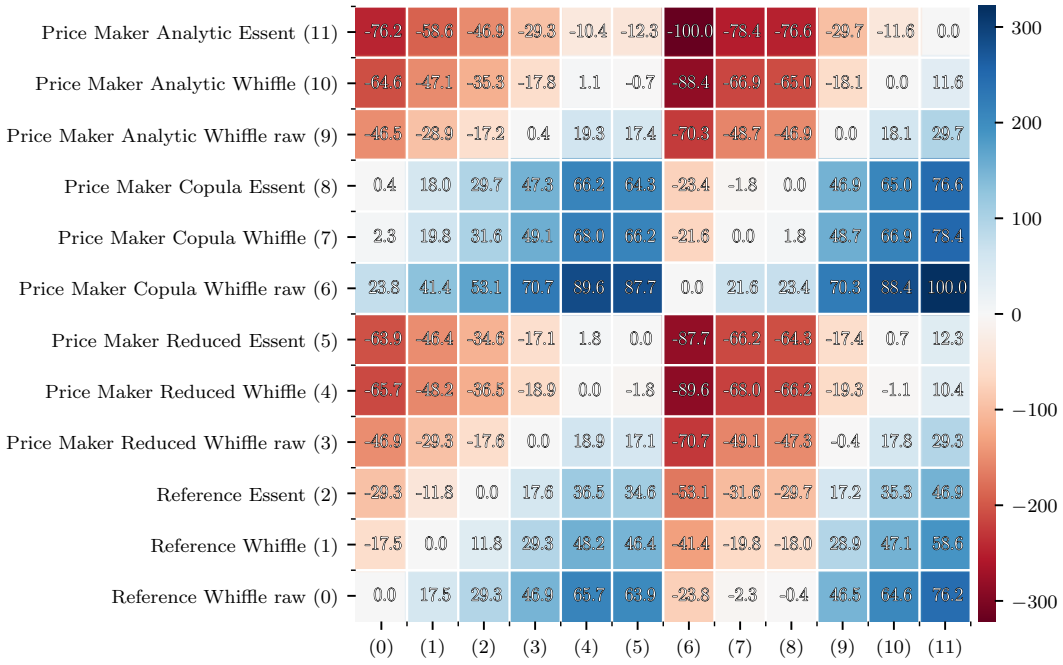


Figure 8.15: Expected price-maker difference of the $CVaR_{5\%}$ for the reference and price-maker strategies of the system opportunity. The colorbar indicates the difference in €. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=€321.49. The expected difference is computed for rows over columns.

Figure 8.15 shows the difference of the conditional value at risk at α=5% for the system opportunity revenue. This indicates the difference in expected value of the 5% lowest revenues. First, it shows that the Whiffle and Whiffle raw reference strategies already lead to a significant decrease in risk. Second, it shows that the price-maker Copulastrategy is again able to further reduce risk.

In conclusion, the price-maker Copula strategy results in a significant increase in opportunity revenue for the system as a whole. This indicates that these strategies are successful at their interaction with the imbalance price. As explained in section 3.4.4, strategies can be successful in four ways. All of these ways entail that the imbalance price is dampened, which has a positive effect on the system as a whole. Unfortunately, this dampening could not be quantified in this analysis. However, as the optimized strategies try to provide balancing power when the imbalance price is favorable, without regarding the effect of the change in sign of the NRV, the decrease in revenue from the price-taker to the price-maker assumption can be explained as follows: As the switch in NRV does not result in a switch of the imbalance price in the optimization, the optimization leads to a too high volume of balancing power when relatively high imbalance prices are predicted. This results in an unwanted sign change of the NRV, which changes the price from the MIP to the MDP and vice-versa. As this overly aggressive provision of balancing power occurs when either the MIP is high of the MDP is low, the switch in NRV leads to a dampened opportunity price, thus lowering total system opportunity cost beyond the lowering of its own opportunity cost. As the price in practice will be dampened by a strategy that is successful at providing passive balancing power, the system cost is expected to be lowered further than shown here.

9

Conclusions, Discussion and Recommendations

This research set out to identify the extent to which the value of wind power can be increased through improved offering strategies. For this purpose, multiple strategies were constructed and simulated for the period 2018-01-01 – 2018-06-30. Chapter 3 introduced three different methods for determining optimized offering strategies. The first, as discussed in section 3.1, consists of a linear optimization model which assumes that the wind power producer (WPP) has no influence on the imbalance price. The second, discussed in section 3.2, consists of a quadratic optimization model which assumes the WPP does have an influence on the price, which is modeled as a linear relationship. The last, discussed in section 3.3, is a closed form reformulation of the model from section 3.2. Concerning the inputs, three different forecasts are used as inputs for the modeling of wind power as a stochastic process, as discussed in section 7.1. Furthermore, section presents two different methods for constructing the inputs to the optimization models in 7.5. These inputs in the cases found in literature are constructed in a symmetric scenario tree using scenario reduction. This study introduces an alternative method, which applies an extended Gaussian Copula model and compares it to the method using the reduced scenario set. This leaves three different models for determining the optimal bid, three different wind power forecasts used as input and two different methods to construct the scenario tree. Furthermore, the reference strategy, as introduced in section 2.3, represents the conventional strategy, which consists of offering the wind power forecast on the day-ahead market. Hence, for each of the forecasts this study researched a reference strategy, a price-taker reduced strategy, a price-taker Copula strategy, a price-maker reduced strategy, a price-maker Copula strategy and a price-maker analytic strategy. The aim of this research was to provide insights in the value of the different inputs, but also in the different methods used. This chapter presents conclusions on the methods themselves, the performance of the strategies and how to assess performance, the impact these strategies have on the system and the market as a whole and how market characteristics influence the strategies and their performance. First, 9.1 presents conclusions on the sub research question. Second, the conclusions are tied together in order to answer the main research question in section 9.2. Third, the research is discussed. Last, recommendations are made in 9.4.

9.1. Sub conclusions

(1) *What model is best suited for constructing bidding strategies for the Belgian case?*

Section 8.2 shows that the price-taker model strongly outperforms all other models under the price-taker assumption, resulting in a strongly increased revenue compared to the reference strategy. Table 8.1 shows that both the price-taker as well as the price-maker models successfully anticipate opportunities, as both win more than they lose, both resulting in an increase in revenue, where the price-taker strategies result in an overall larger expected increase in revenue. However, the price-taker strategies are characterized by more extreme bidding, which means that the wins and losses are larger, which leads to a substantial increase in risk. Section 2.3 concludes that a wind farm can in fact influence the imbalance price, for which the mechanism is explained in section 3.4, while 3.1 showed that the price-taker model results in extreme bidding strategies, which leads to large volumes being traded on the imbalance market. Section 8.3 shows that this leads to a

strong decrease in revenue compared to the reference strategy, which is because of its negative impact on the imbalance price. This can mainly be explained by the fact that the successful bids under the price-taker assumption result in shifts in the sign of the NRV, resulting in fewer and smaller wins and more and larger losses.

Section 8.3 shows that of the price-maker strategies, only the price-maker Copula strategy is able to realize positive expected difference in mean, compared to the reference strategy. Furthermore, the price-maker analytic strategies and the price-maker reduced strategies lead to a substantial increase in risk. Although all price-maker strategies take the effect their strategy has on the imbalance price into account, both the price-maker analytic and the price-maker reduced strategies assume that the processes of wind power, day-ahead price, imbalance price and price-maker effect can be seen as independent from each other. The implication of this assumption is discussed in the conclusion on the third sub question.

In conclusion, concerning specifically the difference between the price-taker and price-maker model, the difference between them in performance, both on the expected difference in revenue as well as on risk is substantial. This shows that in a single-price imbalance market it is crucial to capture the effect on the market. The only strategy with a positive expected difference in revenue is the price-maker Copula strategy. However, there is a strong negative shift in performance for the winning strategy shift from the price-taker analysis to the price-maker analysis, which requires further explanation. This shift in performance can be explained by the fact that the shift in price by change of sign of the NRV is not incorporated in the optimization. This means that the optimization model expects the price to behave according to the price-maker formulation of the effect in chapter 3. This means that it expects the price to decrease or increase up to a certain optimum, which does in fact help the system, but in practice often lead to an unfavorable sign change of the NRV. Due to the exclusion of the influence of the sign change of the NRV, the strategy results in too aggressive passive balancing, which hurts its own revenue. As the model does not anticipates this sign change, the price in reality shifts, which makes the price very undesirable for the WPP. This means that the optimization model needs to endogenize the NRV, MIP and MDP, which is expected to improve the results of the optimization. Furthermore, this would allow the modelling of the price-maker effect in two sections, and the modelling of two imbalance prices separately, which would significantly add resolution to the optimization. Finally, if there is a large ramp of the NRV signal within a PTU, a switch in the sign of the mean NRV within that PTU will result in a large effect on the price. This can be captured stochastically by separately forecasting the NRV, MIP and MDP. Hence, the final conclusion is that for a single price market a price-maker model is needed that is able to captures the dependencies between processes as well as its effect on the NRV.

(2) *What methods can be applied to provide price forecasts?*

Although many different methods can be used to generate point forecasts for electricity prices, the method introduced in chapter 6 shows to be able to be generally applicable to all three electricity price processes. Second, section 7.2,7.3 and 7.4 show that these forecasts can be used to construct reliable scenarios. Third, section 8.3 shows that when used to construct offering strategies, the resulting strategies consequently lead to both an increase in expected value and a decrease in risk for the price-maker Copula strategy. Hence, the price forecasts generated by the method form a valuable input when used in advanced bidding strategies. Unfortunately, the forecasts generated by the model cannot be compared to forecasts from other methods as the comparison with other methods is left out of scope and the performance of other methods reported in literature cannot be compared to the method applied, as market conditions differ strongly between markets and periods. In conclusion, although no comparison is made to other methods, the method introduced shows to be able to forecast all relevant price series, which form a valuable basis for improved offering strategies.

(3) *What methods can be applied to transform point forecasts to the form required for stochastic optimization?*

The first modeling step of transforming point forecasts to predictive densities is carried out using a conditional kernel density model. Chapter 7 shows that this method is able to reliably generate densities for all series, making it generally applicable, although it is less successful for the price-maker effect. The second modeling step of transforming predictive densities to scenario forecasts is carried out using a Gaussian Copula model. This model shows that it is capable of reliably generating scenarios for wind power and the day-ahead price. However, it is not able to capture switching behavior, which is present in the imbalance

price and the price-maker effect. Section 4.4.2 presents an innovation in the form of constructing a scenario set by extending the dependency model. Section 8.3 shows that this final step of the modeling framework is crucial for real-world performance, as the only strategy that improves the expected revenue compared to the reference strategy is the price-maker Copula strategy. The performance evaluation of the price-maker reduced strategy showed that the modeling of the dependencies between the processes is an important step in constructing a successful bidding strategy, as its performance both on expected value and on risk is significantly worse than the price-maker Copula strategy.

This can be explained by the fact that the covariance matrix in section 7.5.2 shows significant covariance between the Guassian forecast errors of all series. This means that the assumption of independence underlying both the reduced set as well as the price-maker analytic strategy does not hold. Furthermore, the covariance matrix shows very specific relationships between the processes, which are empirically derived. Hence, when these dependencies are not captured by a model, the resulting scenario set is unrealistic. This means that for instance a scenario with a high error for the WPP is just as likely to be in a tree scenario with a high error for the imbalance price as with a low error for the imbalance price. The Copula coupled scenario set on the other hand has more tree scenarios with high errors for wind power with low errors for the imbalance price than with high errors for the imbalance price, due to their negative covariance. As this leads to a significant difference in the scenario set, the outcome of the optimization changes as well. For instance, even though the scenarios of the individual processes in the scenario tree are quite similar between the reduced set and the Copula coupled set, in the price-maker analysis the price-maker reduced strategy is significantly outperformed.

The coupling between the wind power forecast error and the imbalance price forecast error can be seen as indirectly capturing the effect of the aggregate wind power forecast error on the imbalance price. As the 'correlation effect' (chapter 1) is known to be high, one can expect that a scenario where the WPP's forecast error is low, the imbalance price is likely to be high. Even though the price-maker optimization model does capture the effect the wind farm itself has on the market, it assumes that this is the only effect on the imbalance price, while the 'correlation effect' teaches us that this is not the case. The Copula coupled scenario set corrects for such assumptions by modeling all these relationships stochastically through the covariance between their Gaussian errors.

In conclusion, the transformation of point forecasts to predictive forecasts can reliably be achieved for all processes by conditional kernel density estimation. The modeling of dependencies between time steps can be modelled by a Gaussian copula, although less successfully for processes that show switching behaviour. However, in the case of strategic bidding, it is important to capture the dependencies between the processes, otherwise no improvement in expected revenue or risk is to be expected.

(4) *What measures are best able to judge forecasts on performance?*

Concerning statistical performance, different types of forecasts should be judged differently. First, when interest lies with the performance of a point forecast, its statistical performance can be ranked on the mean absolute error. Concerning the more advanced predictive densities and scenario forecasts, judgment becomes complex. Concerning the former, an important factor is the point forecast which is used as input to the predictive densities. Section 7.1.3 shows that the lower reliability of the Whiffle and Whiffle raw predictive densities hurts their relative performance on the continuous ranked probability score relative to their point forecast performance. This is mainly due to a relatively large difference in bias between the in-sample and out-of-sample data for these two forecasts compared to the Essent forecast. This means that when models are made to generate predictive densities, reliability is an important indicator of performance. Furthermore, having the knowledge that the change over time of the statistical properties of the point forecast leads to a lowered reliability means a modeler may choose for a modelling method that allows for older data to be gradually forgotten. Concerning the second, section 7.3 shows that the predictive density for the imbalance price is reliable. However, it also shows that the dependency model is less reliable in capturing the dependency between different forecast horizons. This means a modeler may choose to experiment with different methods to capture this dependency, after which its impact can be evaluated by the price score. In conclusion, judgment of the statistical performance is dependent on the form in which the forecast is to be used. Point forecasts may be ranked on the mean absolute error, predictive densities on the continuous ranked probability score and scenario forecasts on the energy or price score. However, when one is responsible for constructing these, judgment of its inputs should not be overlooked, while assessing reliability can help guide modeling choices.

Although statistical performance is an important aspect of the different types of forecasts, sections 7.1

and 8.3 show that when point forecasts are used with the purpose of directly determining the amount to bid, which is the reference strategy, the performance on the main ranking criterion the mean absolute error (MAE) is a poor indicator of economic performance. The Whiffle raw forecast performs worst on the MAE, but strongly outperforms the other forecasts on expected difference in revenue as well as on the ability to reduce risk. This can be explained by the fact that it has a more favorable correlation with opportunity prices. This shows that when choosing a forecast, its performance for its specific use case should be evaluated before a decision is made. This also applies to more advanced use cases. In scenario form the Whiffle raw scenarios also perform worst on energy score, but in all advanced uses, the Whiffle raw scenarios lead to the highest performance ánd to the lowest risk.

In conclusion, while statistical performance evaluation provides tools for making choices between inputs and modeling methods, performance in its use case should always be considered when making choices, as statistical performance is a poor indicator for performance in a market application.

(5) *How can the different strategies be ranked on value?*

The conclusion from sub question (1) is that a price-maker model is a requirement for achieving an improvement in expected revenue. Second, sub question (4) concludes that the price-maker optimization using the Copula coupled scenario set outperforms the reduced scenario set. Third, sub question (4) concludes that the Whiffle raw forecast outperforms all other forecasts in all strategies. Finally, the price-maker analytic strategy is significantly outperformed by the price-maker Copula strategy, both on expected revenue as well as on risk. The analysis in section 8.3 shows that the largest improvement on the expected revenue as well as on reducing risk can be made by using the Whiffle raw forecast, while applying the price-maker Copula strategy further enhances performance.

Unfortunately, this is not the final ranking, as a final ranking should incorporate the complete influence a strategy has on the price. The method for determining such an influence applied in this study proved insufficiently reliable. Hence, a new method is required which is able to reliably reconstruct the imbalance price.

(6) *What are the effects of strategic bidding on the imbalance costs of the market as a whole and the total system imbalance?*

First, section 8.4 shows that the price-maker strategies lead to a decrease in expected absolute system imbalance, of which the price-maker analytic strategy using the Essent forecast led the largest decrease. However, the price-maker Copula strategies are best able to reduce the risk of extreme system imbalance as reduce the volatility of the system imbalance. As more extreme system imbalances require larger balancing action, while high volatility require high-ramp rate, it is expected that these aspects of the system imbalances represent a higher cost of balancing. As the improved offering strategies anticipate imbalance prices rather than imbalances, it seems the price-maker Copula strategy is most successful at alleviating costly system imbalances.

Concerning the total system cost of imbalances, the price-maker strategy with the Copula coupled scenario set manages to reduce the expected cost and to reduce risk. As the total costs are actually decreased beyond the decrease in cost of the WPP itself, strategic bidding shows promise for the system as a whole. Although the dampening effect on the imbalance price was not quantified outside of the effect of the sign change of the NRV, this effect is expected to be larger in an operational context. This is due to the fact that successful passive balancing influences the imbalance price in two ways. First, by providing passive balancing when the cost of active balancing is high. Second, it is successful in reducing its negative influence on the imbalance price when it loses. This means that if successful it has a dampening effect on the imbalance price, thus resulting in lower imbalance costs for the system as a whole. The reason the price-maker Copula strategy results in a positive influence on the opportunity cost for the market as a whole, is because the optimization aims to take advantage of high imbalance prices, which means it often causes a sign change of the NRV in the case of a higher likelihood of either high or low imbalance prices, which hurt its own profits, but helps the market as a whole. If one were to incorporate the NRV in the optimization, this loss is expected to be reduced for the WPP, while the price is still dampened, which should result in an opportunity gain for the market as a whole without the WPP losing out.

In conclusion, the price-maker Copula strategy results in an increase in opportunity revenue for the system as a whole, by providing passive balancing when the cost of active balancing is highest. This indicates that these strategies are successful at their interaction with the imbalance price. As explained in section 3.4.4,

strategies can be successful in four ways. All of these ways entail that the imbalance price is dampened, which has a positive effect on the system as a whole. Unfortunately, this dampening could not be quantified in this analysis.

(7) *What are the effects of market specifics on the composition and performance of bidding strategies?*

Section 2.2.4 discussed the main design characteristic of the short-term electricity markets in Europe and concluded that the main differentiating characteristic lies with the pricing mechanism of the balancing market. Section 2.2.4 discussed the specifics of imbalance price mechanisms and provided several insights. First, for WPPs a single pricing system seems most appealing, as it provides the WPP an opportunity to compete on the provision of balancing power and be rewarded for it. Regarding second place, it depends on the likelihood of two prices to occur in the mixed-pricing system. If there are hardly ever two prices, while it requires much balancing power before an opposite regulation bid is activated, a mixed-pricing system is more attractive, as it allows for an opportunity gain from the provision of passive balancing. If there are often two prices, while it requires little balancing power before an opposite regulation bid is activated, a dual pricing system seems more attractive.

In terms of optimization, for the mixed-pricing system one needs to generate scenarios on the lower bound and upper bound of the NRV within a PTU, as it indicates whether one can expect two prices. As the NRV is the average volume of activated regulation, it provides a clear indication whether or not downward or upward regulation can be expected, and thus whether or not one needs to consider two prices. As it is common for TSOs to only publishes the upward NRV and downward NRV, but not how close it was to activating the bids within a PTU, one should approximate the MDB and MIB from a higher resolution NRV signal. This means that if a TSO wants to incentivize WPPs to effectively provide passive balancing power it should provide this data at a sufficient resolution or publish the lower and upper bounds of activated bids per PTU. This implies that for WPPs it is more beneficial to be in a single pricing system rather than a mixed pricing system, as it simplifies operations and it can still profit when there is both upward and down regulation within a PTU, thus incentivizing it to provide passive balancing. The mixed-pricing system on the other hand punishes all those who have an imbalance when there are two prices, while it is more likely to cause two prices to occur when one bids strategically, thus realizing the unfavorable shift in price sooner, making it harder for wind power producers to profit. This means that the mixed pricing system provides less of an incentive to provide balancing power. This could be seen as negative for the TSO, as the successful provision of balancing power at the quarterhour resolution means that the imbalance price is dampened. Last, the dual pricing system never rewards the provision of passive balancing as one is payed the day-ahead price when one does so successfully. This means that one is incentivized to provide balancing power, but not rewarded for doing so.

9.2. Main conclusion

The main research question is:

To what extent can BRP's improve the market value of their wind power portfolio?

The best estimate for this case is 1.17%, which is the price-maker increase in expected mean revenue from switching from the currently used Essent reference strategy to the price-maker Copula strategy using the Whiffle raw forecast. This result is contingent on multiple factors. First, the forecasts used as input to the modeling. Secondly, the modeling methods applied and third, the currently applied strategy. However, as most WPPs currently apply the reference strategy, the framework presented is likely to be able to help WPPs increase the value of their wind power portfolio.

9.3. Discussion

Section 1.1.3 discusses the main research gaps. The first gap relates to the choice of model. This research contributes to this aspect in multiple ways. First, it shows that the price-taker version of the stochastic optimization model is entirely unsuited for a single-price market. This extends to the mixed-pricing market studied in [9], as the Dutch market often has one-sided balancing within one PTU. For PTUs with one-sided balancing a mixed-pricing market is equivalent to a single-pricing market, which is when results are similar for that market. Hence, a price-taker model seems inappropriate for such a market. Second, this study fur-

ther adds to this discussion by showing that although price recreation is difficult for the imbalance price, the simple method of incorporating the effect of a sign change of the NRV signal in the performance evaluation already shows that improved offering strategies often do not improve economic performance. Also, this study emphasizes the mechanism determining the imbalance price, as it has such a large influence on the price. In hindsight it seems inappropriate not having endogenized this effect inside the price. Not only does it explain the switching behavior on the imbalance price, it is likely it would also alter the results from the optimization.

The second gap relates to the methods used to model the stochastic processes, used as input to the stochastic optimization. While the studies discussed in 1.1.2 all apply simplistic methods to construct them, this study introduces the method of extending the Gaussian Copula to directly generate a scenario tree and it is the first case study to apply it. Although the price-taker analysis showed no improvement over conventional methods, specifically the price-maker analytic version and the price-maker reduced version, the price-maker analysis showed a significant difference. Furthermore, the price-maker Copula strategy showed to be the only offering strategy to consequently improve the economic performance relative to the reference strategy. Further still, analysis on a system level showed that it is the only one helping reduce expensive balancing action. Hence, this study shows the empirical dependence through inspection of the covariance between Gaussian forecast errors and shows through simulating offering strategies that respecting these dependencies is crucial for strategy performance. Furthermore, application of conventional scenario reduction illustrated that the conventional method of constructed a symmetric scenario tree dictates the number of stochastic processes being incorporated in the optimization, which is highly limiting. The introduction of the method of the Copula coupled scenario tree enables the incorporation of many more stochastic process inside the optimization. Without this method, the recommendation to incorporate the effect of the sign change within the model would not be possible, as the number of processes inside the optimization increases by two. Without the method introduced in this study the exponential increase of scenario in a symmetric scenario tree would lead to an intractable optimization.

Finally, it is important to emphasize that this study has important operational implications. First, it shows that the method of evaluating strategy performance is crucial to prevent significant cost to the individual WPP and the market as a whole. This is especially true for the price-taker strategies, but also for the other strategies. Although cases in literature showed improved economic performance for the price-maker reduced strategies, the price-maker analytic strategies and the price-taker strategies, this study found a negative impact on economic performance for these strategies. This means that without a price-maker effect inside the optimization specifically combined with the Copula coupled scenario set, negative performance is likely, again emphasizing the importance of proper dependency modeling.

9.4. Recommendations

Balance responsible parties tend to have demand in their portfolio, which often includes decentralized solar power generation, which is also a stochastic process. Although demand can be forecast relatively well, the solar power component that is included in the demand is more difficult to forecast. Furthermore, wind power and solar power are shown to be negatively correlated, an indirect correlation is present between wind power and demand [60]. This means that if demand with solar power either directly or indirectly were to be included in the optimization, further gains in revenue are likely to be made. Also, conventional generation can be used to offset imbalances of the stochastic parts of a balance responsible parties portfolio. These additions can be realized through the concept of a virtual power plant [42]. A recommendation is not to separately optimize bids for different parts of a balance responsible parties portfolio, although it may help to anticipate strategic offers from optimizing demand through incorporating aggregate demand in the Copula coupled scenario set, thus capturing the effect of demand on the imbalance price. However, the gain of optimizing over both is likely to be larger as they have some correlation through decentralized solar power generation. As one of the main conclusions of this research is that capturing dependencies in the scenario set increases value, this is a clear recommendation.

To help improve the outcome of the optimization, it may be beneficial to include Gaussian forecast errors for aggregate wind and aggregate solar in the dependency model, as this allows one to directly model the effect of the market as a whole on the imbalance price, instead of modeling it indirectly through the 'correlation effect'. The correlation between aggregate wind and solar and the imbalance price should also to some effect capture strategic bids from competing market parties. This is due to the fact that this correlation should decrease when they do so successfully, as it would have a dampening effect on the imbalance price. This dampening effect indicates that even when other parties successfully bid strategically, all other parties tend

to profit from it, as the imbalance price is increasingly dampened by it. However, a better alternative would be to incorporate the NRV in the optimization combined with both the MDP and MIP. This would allow the WPP to anticipate strategic bids through the NRV, which is shown to be crucially important in determining the revenue of a strategy. Furthermore, the current model is betting on predictions of prices in which the Essent forecast had an influence. Ideally, one should remove the influence of the Essent bid from the imbalance bid. As one makes the influence on the imbalance price endogenous in the optimization model, one should use an imbalance price without the influence of the park. This would also allow the Copula model to better estimate the influence of other parks. Furthermore, if the NRV is to be incorporated in the optimization, the influence of the operational strategy should be removed from the NRV, as this would allow the WPP to model the influence of its competitors.

Concerning the methods applied in this study, several directions for improvement are identified. Firstly, section 7.1.3 mentioned that a difference in bias between the in-sample data and out-of-sample data for some forecasts led to decreased reliability of their resulting predictive densities. A solution to this problem would be to incorporate a forgetting factor, which places a weight on each sample which decreases over time. This way, new information has a higher weight on determining the shape of the probabilistic model, thus ensuring that changes in the statistical properties of the underlying stochastic error process are learned faster. This also helps with the stochastic processes related to the electricity market, as the electricity market is especially dynamic, where relationships tend to change over time.

Secondly, the kernel density model could be improved by researching the impact of cross-validation of the kernel bandwidths, different kernel functions, variable kernel bandwidths or Copula kernel density estimation. Another option for improving the predictive densities lies with the application of other methods. One promising method for estimating predictive densities consists of quantiles from a generalized additive model, which is discussed in a study by Gaillard, Goude and Nedellec (2016) [14].

Thirdly, improvements can be made to the dependency model. Inclusion of all processes in the dependency model is shown to be crucial for the performance of the optimized strategy. When it is preferable to control risk, which is discussed below, scenarios of different processes cannot be decomposed by timestep in the optimization as risk control is modeled using an inter-temporal constraint, which means that realistic trajectories for all processes become crucially important [10]. As shown in sections 7.3 and 7.4, the switching behaviour of both the imbalance price and the price-maker effect is not captured well by the Gaussian Copula model, which leaves room for improvement. However, when choosing a model with the aim to generate more realistic trajectories, the need for modeling all dependencies should be respected, as it is shown to be key for real-world performance.

Fifth, improvements can be made to the optimization model. Straightforward additions are the incorporation of risk control and bidding curves, as introduced by Morales, Conejo and Pérez-Ruiz (2010). These can be modeled simply by adding additional constraints to the existing optimization model. For the modeling of the bidding curve, it is crucial however to have scenarios for the day-ahead price. This means that strong scenario reduction would be required without a dependency model. This combined with possible incorporation of other stochastic processes such as demand again shows the value of the dependency model. Another possible valuable addition is the incorporation of the supply curve to the balancing market as a stochastic process. This is done for the day-ahead market by Mitridati and Pinson who incorporate the supply curve as a piece-wise function in their optimization model (2018) [41]. One would then also need to forecast and model the marginal decremental bids, marginal incremental bids and the NRV as stochastic processes. As shown in the price-maker analysis, the NRV is crucially important for the revenue of a strategy. Incorporating it into the price-maker analysis should significantly help increase revenue. Furthermore, the price-maker effect as a linear term is an overly simplifying assumption, as the imbalance price is quite stable when the NRV does not switch sign. Hence, a better approximation of the supply curve should help improve revenue as well, while also providing a better tool to determine the real-world performance of a strategy. Recreating prices using the slope of the supply curve overly influences everything a strategy does, thus resulting in an unrealistic result for the strategy. This has important implications for the optimization, as it too uses this principle. The main aim of this is to result in a more realistic optimization. However, this author hypothesizes that changing the linear effect to a realistic piece-wise supply curve should result in much better results. Also, an important factor determining the profit of a strategy is the sign of the NRV. As this is not taken into account in the optimization, results are rather underwhelming. Using two prices instead of one with the NRV in the optimization is believed to lead to significant improvements to the optimization.

Sixth, as was shown by the two Whiffle forecast, much value can be obtained by different point forecasts as input to the modeling process. While no comparison was made on different forecasts for the electricity price

processes, it seems likely that additional gains in revenue can be obtained by choosing different methods for generating these forecasts. Promising methods include long short-term memory neural networks, which is a machine learning method specifically powerful in forecasting time-series and is capable of capturing non linearities as well as delayed effects between multiple time series, or generalized additive models, which are capable of capturing non linear effects while remaining highly interpretable due to their additive nature.

Finally, the price-maker strategy using the Copula scenario set showed in section 8.4 that it is able to decrease the highest 5% of absolute system imbalances, while reducing the volatility of the system imbalance. As the model optimizes over the imbalance price, which is an important indicator of the system imbalance, it seems likely that performing a root-finding optimization over the system imbalance without taking revenue into account would provide a guide to anticipate system imbalances. This could provide TSOs with a tool to trade power on the day-ahead market to anticipate system imbalances, reducing the need of short-term balancing actions and avoid steeper ramps at the quarterhourly resolution, helping prevent wear and tear of the generators providing balancing action. Especially considering the strain short term balancing requirements place on conventional generators and the system as a whole, this could prove an important tool for TSOs. Furthermore, the price-maker strategy using the Copula scenario set proved particularly capable of decreasing the expected value of the highest 5% of absolute system imbalance, which may prove valuable as the 'correlation effect' entails that wind power producers are often wrong simultaneously, increasing the risk of extreme imbalances.

On a macro-level it would be interesting to research the effects on individual revenues if multiple market parties apply optimized bidding independently or through centralized optimized bidding through an external agent. A study on this was conducted by Guerrero-Mestre, de la Niete, Contreras and Catalão for the Spanish (dual pricing) market in 2016, who found that centralized optimization by an external agent signficantly improved revenue for all [20]. In the Belgian single price case, when processes in the optimization are Copula coupled and a forgetting factor is implemented, this author expects that when significance of the covariance between the wind power forecast error and the imbalance price forecast error reduces due to successful strategic bidding, the gain from strategic bidding reduces for individual market parties, but the dampening effect on the imbalance price leads to an increase in revenue for all. Hence, if successful, the optimal passive balancing from wind farms would reduce the impact of wind power forecast errors on the system imbalance.

Concerning the operational implementation of the algorithms, several recommendations are formulated. First, the tools used to implement the algorithms are mentioned at the start of each chapter, which should enable experts to implement them. Furthermore, although the frameworks in this research are formulated with the goal of empowering WPPs to implement the concepts themselves, further details and practical aspects can be found in the referenced literature, specifically [42] and [10]. However, the framework structures the lessons from these and multiple sources in such a way that it lends itself well for use in an operational context, enabling both the tools to model the stochastic processes, as providing evaluation tools to improve said modeling. Furthermore, the price-maker stochastic optimization model provides a starting point, from which the operational user can depart. The concept of a virtual powerplant can then be applied to incorporate other parts of its portfolio. This author does wish to emphasize several concrete points of improvement, before operational implementation. First, a forgetting factor seems crucial for obtaining satisfactory reliability and fast learning for all models, both predictive and stochastic. Second, although the price recreation algorithm was not sufficiently reliable, the difference between the price-taker and the price-maker performance analysis shows that proper performance analysis is crucial before making any decision with regards to strategy. Hence, before any operational implementation, a method must be found to recreate the impact a strategy has, preferably on all price processes. One way to do so, would be to implement a stochastic method, which incorporates the uncertainties surrounding the impact. Third, all explanatory variables used for the predictive models are public and can be used operationally. Unfortunately, the aggregate wind power and aggregate solar power forecasts are published at 11:00AM, which for some WPPs is considered too late to formulate their bid. In this case these can be replaced by the irradiation and wind power forecasts used operationally, which due to the 'correlation factor' already capture much of the influence of their aggregate counterpart. Fourth, continuous monitoring of performance seems in order, to keep evaluate the value of strategies. It may be appropriate to test strategies in parallel, in order to improve decision making. Fifth, in the recommendation section, multiple methods for improvement are mentioned. One important one is the implementation of the NRV signal as endogenous to the optimization, which allows for the WPP to anticipate strategic bidding from competitors. Also, using the frameworks presented in chapters 4 and 5, the improvements to the stochastic process modeling can be continuously tested, which should empower the process of continuous improvement. Sixth, the addition of risk control should go hand in hand with the easily interpreted quantile forecasts

combined with the experience of traders, to identify risks and apply appropriate risk control. For instance, a daily dashboard could be constructed, which allows traders to visually examine the stochastic processes on a daily basis, which helps them to gradually develop a feel for the algorithm and its worth, while also developing appropriate levels of trust. This is an important benefit of the offering strategy, as this distinguishes it from a black box method, making it explicitly white box. Finally, concerning professionals less familiar with the methods applied in this thesis, it may be appropriate to acquire knowledge by consulting statistical modelers or mathematical modelers with both experience in stochastic process modeling as in stochastic optimization, which are the main fields of expertise required for operational implementation.

10 Reflection

The purpose of this chapter is to reflect on some of the implications of the findings. First, the quality of the decision-making is discussed in section 10.1. Second, the interests of the actors introduced in chapter 2.1.2 are discussed in section 10.2. Third, the implications on system optimality are discussed in section 10.3. Finally, the implications on market design are discussed in section 10.4

10.1. Quality of Decision-Making

The decision-maker in the context of this research is the WPP. The decision to be made is on which amount to be offered on the day-ahead market. The conventional strategy for this type of decision is to offer the expected amount. However, due to the stochastic nature of wind power, the WPP is forced to participate on the balancing market, where it experiences negative market outcomes, decreasing its revenue. The fact that this strategy of participation on the balancing market results in a negative market outcome for WPPs calls for a more informed decision on this participation. Visual inspection of the stochastic processes involved in determining the strategy, presented as quantile forecasts in sections 7.1, 7.2, 7.3 and 7.4, already provides the WPP with additional information, allowing for a more informed decision. Its decision will then be based on its knowledge on the mechanics of the two electricity markets involved. As these mechanics can be summarized mathematically and are expressed as such in the price-maker optimization model discussed in section 3.2, the ability of such a model to honor the considerations of the WPP should inspire confidence in the model. Furthermore, as the WPP would base its strategy on the information presented visually through the quantile forecasts and the optimization model uses the same data similarly to derive its strategy, this again should inspire confidence in the strategy it produces. A third factor the WPP should take into account are the correlations between the stochastic processes, as this leads to a superior strategy, as was shown in section 8.3. This can be visually presented as scatter plots, providing the WPP with a feel of all the interactions it should take into account when determining its strategy. Seeing as the model takes these relationships into account as well, confidence should be inspired in the strategy it produces.

Concerning the measurable quality of the strategy the model produces, this study shows that the strategy produced by the model can be simulated, after which the quality of a strategy can be quantified, making the confidence of the WPP in the model an informed confidence. Section 8.3 provides two measures on which this quality can be evaluated. Unfortunately, section 8.3 also shows that the imbalance price could not be reconstructed accurately, while the difference in results between section 8.2 and 8.3 shows that capturing the effect a strategy has on the price strongly influences the evaluation of the quality of the decision. As the recreated price only captures two factors that determine the price, which already strongly changes the quality of the decision, for a definitive quality evaluation the ability to accurately recreate the imbalance price is crucial. However, given the fact that the price-maker strategy using the Copula coupled scenario set on average only deviates by a small amount from the Essent reference strategy and imbalance prices are relatively stable when the sign of the NRV is stable, it does seem likely that there are multiple strategies that increase the quality of the decision, both on risk and on expected revenue. Finally, the extension of the model to include the control of risk through the conditional value of risk, should provide the decision-maker with an added tool to further enhance the quality of its decision-making, as it allows the decision-maker to use its experience on when risk-aversion is appropriate.

In conclusion, the white-box nature of the model producing the strategy allows the decision-maker to validate each step of the process. This means it aids the decision-maker in the formulation of intelligent complex strategies without trading-off on interpretability. Unfortunately, although measures are introduced that enable the decision-maker to evaluate the quality of the decision, without an accurate recreation of the imbalance price, no definitive evaluation of said quality can be given.

10.2. Strategic Interests of Market Actors

Concerning the strategic interests of market actors, which are introduced in section 2.1.2, several implications are made by the research.

The first regards the day-ahead market operator, which is interested in ensuring an effectively functioning market. Seeing as not much changes for the day-ahead market operator, other than slightly altered bids, which does not impact its goal of an effectively functioning market, the impact on this actor is neutral.

Second, the TSO, which is also a market operator, is responsible for maintaining grid stability at low cost. Seeing as the price-maker Copula strategy helps stabilize the grid, while reducing the imbalance costs for the market as a whole, the TSO is expected to welcome such a strategy. Especially since the penetration of intermittent stochastic generation is expected to increase towards the future, which places additional strains on the grid and balancing power.

Third, the BRP is discussed. Seeing as the BRP can have a varying composition, it makes sense to discuss the impact on each actor separately. First, there are conventional power producers, which are interested in obtaining the highest possible profit from their generation. In this sense, it will profit from an increase of the day-ahead electricity price, irrespective of its marginal costs or its place in the merit order. A higher price would result in its offer either being accepted more often, or it receiving a higher price for its generation. This means that it stands to benefit if the WPP were to offer less on the day-ahead market, due to the price decreasing effect of wind power (see chapter 1). Current market conditions lead to this effect, seeing as the best strategy results in the wind power producer being long more often, as it is more expensive to be short on average. This is due to the asymmetry in the supply curve to the balancing market. This strategy of being long could change in the future if the supply curve would become more symmetric, e.g. through an increase in electricity storage and demand-response.

The second type of power producers are the stochastic producers, with uncertain generation. When these maintain a conventional bidding strategy, these stand to profit from the increase in the day-ahead market price. Furthermore, as such a producer is forced to participate on the balancing market, where it incurs a decrease in revenue, it also benefits from the dampened imbalance price, although the extent to which would depend on its specific correlation with the imbalance price. When it also successfully bids strategically, it should help increase this dampening effect, while it would also increase its own revenue through the success of its strategy. If it were to bid strategically without success, this would mean it would not properly anticipate strategic bids from other stochastic producers, which means it would hurt revenues for all participants to the balancing market.

Third, there is the supplier with demand in its portfolio. When only demand is considered, there are several expected effects. First, the increased price on the day-ahead market would increase its costs. Second, demand is stochastic, which means it is also forced to participate on the balancing market, where it would profit from the dampened imbalance price. Seeing as the day-ahead market is highly liquid and the balancing market is relatively illiquid and has a steep supply curve, it is expected that the benefit from lower balancing costs will outweigh the disadvantage incurred on the day-ahead market. This same reasoning applies to large industrial consumers.

Last, there are the end-users, or consumers, which do not directly interact with the electricity market, but have contracts with suppliers. As mentioned in section 2.1.2, consumers are able to switch suppliers to obtain a more favorable price for their electricity consumption. Assuming a well-functioning market with a high degree of competition between suppliers, some of the economic gains should eventually result in a decreased price for consumers, meaning they incur a positive impact as well.

Concerning the regulation market and actors providing active balancing power, some decrease in its price and the resulting revenues are to be expected. However, as the societal goal of the electricity market is to ensure a reliable and affordable electricity supply, both of which are addressed by the bidding strategy, the interests of the few actors that profit from increased imbalances fall behind those of the majority of market participants and society as a whole. Furthermore, often producers providing balancing power are part of a larger supplier whose profits are hurt by increased imbalance costs. Finally, balancing power is mostly

provided by fossil powered conventional generators. As this implies that such carbon intensive generators' profit is decreased at the benefit of low carbon intensive wind power, this could be considered a societal gain.

10.3. System Optimality

Concerning the optimality of the solution for the system as a whole as it currently functions, several implications can be derived from this research. To discuss the optimality, the economic theory of second best is used [35]. This theory concerns situations where any number of optimality conditions cannot be satisfied. These optimality conditions are what according to general equilibrium theory define a perfect market [3, 40]. These conditions are collectively referred to as perfect competition, where perfect competition should result in a long-run equilibrium where social welfare is maximized. These conditions include but are not limited to:

- A large number of buyers and sellers
- Perfect information
- Homogeneous products
- Every participant is a price-taker

The balancing market does not seem to satisfy multiple such conditions, as not all information is available and participants are price-makers. The theory of second best states that if such market distortions cannot be removed, other distortions to the market may in fact lead to a more efficient market. Hence, focus should not only lie with removing such distortions, but rather with making rational situation specific interventions.

Maximizing welfare in the case of the balancing market entails minimizing the cost from participating on the balancing market. First, the effects of the strategies in this research are discussed. Second, possibilities for improvement with respect to system optimality are discussed.

Focusing on the price-maker analysis of opportunity cost for the system as a whole (see section 8.5, there are clear indications which strategies move the system away from optimality and which help the system move towards optimality. Specifically, the price-maker strategies all help the system move towards system optimality. Of the price-maker strategies only the Copula strategy actually realizes an increase in revenue, according to the price-maker analysis in section 8.3. As this is also the strategy that is most optimal in the aggregate, it is likely that a strategy that is successful for the individual market participant is also successful for the system as a whole. This is due to the fact that a successful strategy dampens the imbalance price. However, it is important to emphasize that without a proper method for evaluating price-maker performance, system optimality is reduced strongly. Section 8.2 shows that the price-taker strategies perform best by far, which according to the analysis in section 8.5 results in a significantly reduced system optimality. As the common method of the case studies in literature is to apply the price-taker analysis, there is a risk that some market participants will in fact hurt system optimality. Hence, a first improvement would be to enable market participants to accurately reconstruct imbalance prices, which would help their strategy and help system optimality.

Although it seems likely that the price-maker Copula strategy will help the system towards optimality, there are multiple ways in which optimality can be increased further. First, multiple studies researched the impact of information sharing and coordinated bidding between WPPs. In terms of information sharing, the benefit is that the information can be used to obtain strategic information from it, or that it can help improve aggregate forecasts. A recent study by Guerrero-Mestre, Sánchez de la Nieta, Contreras and Catalão (2016) [20] found that jointly offering to the day-ahead market can lead to an increase in overall revenue compared to individual offering. However, this model assumes a price-taker analysis and is carried out for the Spanish dual-pricing market. Furthermore, strategic-bidding is simulated in a boolean fashion, where it either is applied or not at all for the same time period, while in real-world applications, advanced strategic bidding is expected to evolve over time, while aggregate wind power forecasts enable wind power producers to anticipate strategic bids from other wind farms, as these can be used to help forecast the NRV signal. However, both from an economic and governance perspective it does make sense that as information scarcity decreases through information sharing, while collective action enables more effective decisions for the aggregate, system optimality should increase. However, in the case of the WPP the attainment of such optimality would require a central agent orchestrating the collective action or distributed governance with complete information sharing. Unfortunately, competing market parties are not likely to share information, while collective action would entail that wins and losses are shared equally, which also does not seem realistic.

To help alleviate some of the constraints and empower market participants to be more successful at their bidding, member states of the European Union are required to publish data of aggregate wind power generation for each ptu of the following day, following EU regulation No 543/2013 [57]. This data is also used in this study to forecast the price series. This seems like a likely source of improvement for system optimality, as it allows WPPs to better anticipate rival bids, to improve forecasts of prices and of the NRV signal.

Still, the main conclusion is that simultaneous offering for multiple wind farms seems reasonable, as separate wind farms are highly correlated over space and time, which allows for a higher quality bid. Optimizing bidding strategies for a large share of the market would likely lead to a higher total revenue. However, the reality is that market actors are not willing to do so, as they compete and an advantage in information provides a competitive edge. However, this research showed that strategic bids can be anticipated both through the indirect correlation of wind power with the imbalance price and through the NRV signal, as aggregate level forecasts are available, anticipating rival bids may be modeled accurately through the NRV signal. Lastly, as successful strategic bids should result in a dampened electricity price, increased system optimality should be possible without information sharing.

Furthermore, as more market players start bidding strategically, the effect on the NRV will be gradual. Recursive learning models predicting the NRV should pick up this effect, which means that strategic bidding of competitors should be picked up. When the market as a whole is not successful in strategic bidding, the system imbalance and on opportunity cost for the system in will be negatively affected. However, if market players compete on NRV, the expectation is that the effect will be beneficial for the market as a whole, as all market players aim to optimize the NRV with regards to their specific portfolio.

Last, the anticipation of the price signal should be seen as rational behavior of the BRP and as an incentive from the market to prevent expensive balancing action. As the successful price-maker Copula strategy is long on average hence seems rational. The improved offering strategy concept could be seen as an increase of rationality, for which the balancing market communicates a clear signal through the imbalance price. However, if grid constraints are increasingly binding in the activation of balancing services, it may prove more efficient for the TSO to ban passive balancing. Especially for a large offshore WPP, which has a strong impact on system imbalance and grid congestion. In such a case implementing locational prices would seem more efficient to incentivize a balanced grid, which is discussed in the next section.

10.4. Short-Term Electricity Market Improvement

There are certain future developments which should increase economic efficiency for electricity markets. First, the main market is currently the day-ahead market, which operates at a limited resolution. This research showed that the limit in resolution of the day-ahead market leads to specific imbalances on the balancing market, which can be alleviated by an increased resolution. Research on the future requirements of the electricity market indicates that higher resolution markets should help with a more efficient market [52]. Second, as some generation assets require longer term planning of their schedule, while some can deal with very short-term schedules, research on the future needs for restructuring electricity markets [11] suggests that multiple markets should be implemented with different horizons. For instance, a week ahead market for base-load generators, a day-ahead market for intermediate load generators and a one-hour ahead market for peak load markets. Ideally, these markets should have no price caps and should be energy-only markets, where price signals reflect the true price of electricity, thus providing proper incentives to market parties.

Concerning a more fundamental aspect of the current market design in Europe, zonal pricing is believed to decrease market efficiency as grid constraints become increasingly binding. As markets in zonal pricing systems are cleared without taking physical constraints into account, after which generators are redispatched to alleviate possible grid congestion, the true underlying price of grid constraints is not revealed. Having a clear price on grid constraints would help with the spatial allocation of generation and demand assets, as well as help identify efficient investments in grid capacity.

Combining locational pricing with short term high frequency markets should also increase market efficiency. First, because the utilization of the grid is currently quite low. This is due to the process by which grid constraints are taken into account after clearing the day-ahead market. As high frequency optimal dispatch including grid constraints can computationally become very intensive, it is not currently carried out at a high frequency. This means that the grid cannot be utilized efficiently, as it in practice is always utilized up to a level where a certain likelihood of an extreme event can still be managed by the grid. This means the grid must be managed as if a network outage has already occurred [5]. If locational real-time prices from the high frequency optimal dispatch were a reality, market parties would respond to these prices, thus efficiently

alleviating any such grid constraints, which should increase the utilization of the grid.

A

Appendix

A.1. Predictive Linear Model Day-Ahead Price

	coef	std err	z	P>|z|	[0.025	0.975]
const	-3.9077	0.112	-34.908	0.000	-4.127	-3.688
0	-0.1872	0.019	-9.796	0.000	-0.225	-0.150
1	-0.2358	0.020	-11.794	0.000	-0.275	-0.197
2	-0.2080	0.021	-9.933	0.000	-0.249	-0.167
3	-0.2970	0.021	-13.877	0.000	-0.339	-0.255
4	-0.3446	0.021	-16.192	0.000	-0.386	-0.303
5	-0.3060	0.020	-14.988	0.000	-0.346	-0.266
6	-0.2024	0.019	-10.691	0.000	-0.239	-0.165
7	-0.0754	0.019	-4.027	0.000	-0.112	-0.039
8	-0.0654	0.019	-3.389	0.001	-0.103	-0.028
9	-0.0681	0.020	-3.347	0.001	-0.108	-0.028
10	-0.0795	0.022	-3.690	0.000	-0.122	-0.037
11	-0.1054	0.023	-4.609	0.000	-0.150	-0.061
12	-0.0995	0.023	-4.327	0.000	-0.145	-0.054
13	-0.1502	0.023	-6.630	0.000	-0.195	-0.106
14	-0.2185	0.022	-10.038	0.000	-0.261	-0.176
15	-0.2951	0.021	-14.199	0.000	-0.336	-0.254
16	-0.3218	0.020	-16.113	0.000	-0.361	-0.283
17	-0.2605	0.020	-12.888	0.000	-0.300	-0.221
18	-0.1079	0.021	-5.260	0.000	-0.148	-0.068
19	0.0216	0.020	1.075	0.283	-0.018	0.061
20	-0.0078	0.020	-0.399	0.690	-0.046	0.030
21	-0.0760	0.019	-3.977	0.000	-0.113	-0.039
22	-0.0728	0.019	-3.805	0.000	-0.110	-0.035
23	-0.1445	0.019	-7.623	0.000	-0.182	-0.107
mon	-0.5132	0.019	-26.635	0.000	-0.551	-0.475
tue	-0.5347	0.020	-26.512	0.000	-0.574	-0.495
wen	-0.5318	0.021	-25.745	0.000	-0.572	-0.491
thu	-0.5162	0.020	-25.555	0.000	-0.556	-0.477
fri	-0.5030	0.020	-25.610	0.000	-0.541	-0.464
sat	-0.5446	0.017	-32.894	0.000	-0.577	-0.512
sun	-0.7642	0.016	-47.458	0.000	-0.796	-0.733
sin1	-0.6211	0.013	-46.534	0.000	-0.647	-0.595
cos1	-0.2211	0.016	-13.620	0.000	-0.253	-0.189
sin2	-0.1689	0.012	-14.126	0.000	-0.192	-0.145
cos2	0.0873	0.008	10.613	0.000	0.071	0.103
holiday	0.0251	0.019	1.301	0.193	-0.013	0.063
vacation	-0.3230	0.021	-15.442	0.000	-0.364	-0.282
Aggregate_Wind	-0.0004	9.19e-06	-46.567	0.000	-0.000	-0.000
Aggregate_Solar	-0.0004	1.35e-05	-26.438	0.000	-0.000	-0.000
Aggregate_Load	0.0005	7.76e-06	58.535	0.000	0.000	0.000
LF	0.0008	0.000	2.427	0.015	0.000	0.001
NG	0.0001	1.62e-05	8.852	0.000	0.000	0.000
NU	-0.0001	8.31e-06	-12.084	0.000	-0.000	-8.42e-05
WA	0.0001	3.17e-05	4.362	0.000	7.62e-05	0.000
WI	0.0011	7.05e-05	16.160	0.000	0.001	0.001
Other	-0.0002	4.35e-05	-4.852	0.000	-0.000	-0.000

Table A.1: Table of model coefficients for the Robust Linear Regression Model. LF = Liquid Fuel, NG = Natural Gas, NU = Nuclear, WA = Water, WI = Wind.

A.2. Predictive Linear Model Imbalance Price

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.6595	0.151	4.358	0.000	0.363	0.956
0.0	0.1906	0.049	3.913	0.000	0.095	0.286
0.25	0.0772	0.049	1.579	0.114	-0.019	0.173
0.5	-0.1285	0.049	-2.620	0.009	-0.225	-0.032
0.75	-0.3307	0.049	-6.712	0.000	-0.427	-0.234
1.0	0.3973	0.049	8.037	0.000	0.300	0.494
1.25	0.0817	0.050	1.643	0.100	-0.016	0.179
1.5	-0.1095	0.050	-2.192	0.028	-0.207	-0.012
1.75	-0.2933	0.050	-5.844	0.000	-0.392	-0.195
2.0	0.2164	0.050	4.295	0.000	0.118	0.315
2.25	0.0744	0.051	1.471	0.141	-0.025	0.174
2.5	-0.0018	0.051	-0.036	0.971	-0.101	0.098
2.75	-0.0825	0.051	-1.622	0.105	-0.182	0.017
3.0	0.1244	0.051	2.443	0.015	0.025	0.224
3.25	0.0860	0.051	1.687	0.092	-0.014	0.186
3.5	0.0947	0.051	1.854	0.064	-0.005	0.195
3.75	0.0647	0.051	1.266	0.205	-0.035	0.165
4.0	0.0586	0.051	1.148	0.251	-0.041	0.159
4.25	0.1309	0.051	2.567	0.010	0.031	0.231
4.5	0.1767	0.051	3.472	0.001	0.077	0.276
4.75	0.3152	0.051	6.211	0.000	0.216	0.415
5.0	-0.0577	0.050	-1.143	0.253	-0.157	0.041
5.25	-0.0223	0.050	-0.443	0.658	-0.121	0.076
5.5	0.1294	0.050	2.582	0.010	0.031	0.228
5.75	0.2672	0.050	5.363	0.000	0.170	0.365
6.0	-0.3578	0.049	-7.263	0.000	-0.454	-0.261
6.25	-0.1415	0.049	-2.894	0.004	-0.237	-0.046
6.5	0.2697	0.049	5.541	0.000	0.174	0.365
6.75	0.4490	0.049	9.257	0.000	0.354	0.544
7.0	-0.2111	0.048	-4.362	0.000	-0.306	-0.116
7.25	-0.1084	0.048	-2.243	0.025	-0.203	-0.014
7.5	0.2565	0.048	5.309	0.000	0.162	0.351
7.75	0.3063	0.048	6.340	0.000	0.212	0.401
8.0	-0.0279	0.048	-0.578	0.563	-0.123	0.067
8.25	-0.1659	0.048	-3.429	0.001	-0.261	-0.071
8.5	0.0427	0.048	0.880	0.379	-0.052	0.138
8.75	0.0443	0.049	0.913	0.361	-0.051	0.139
9.0	0.0699	0.049	1.436	0.151	-0.026	0.165
9.25	-0.0298	0.049	-0.610	0.542	-0.126	0.066
9.5	0.0589	0.049	1.201	0.230	-0.037	0.155
9.75	-0.0938	0.049	-1.903	0.057	-0.190	0.003
10.0	0.1696	0.049	3.430	0.001	0.073	0.266
10.25	0.0068	0.050	0.138	0.891	-0.090	0.104
10.5	0.0020	0.050	0.040	0.968	-0.096	0.100
10.75	-0.1271	0.050	-2.538	0.011	-0.225	-0.029
11.0	0.0490	0.050	0.974	0.330	-0.050	0.148
11.25	0.0401	0.051	0.794	0.427	-0.059	0.139
11.5	0.0834	0.051	1.642	0.101	-0.016	0.183
11.75	0.0027	0.051	0.054	0.957	-0.097	0.103
12.0	0.2735	0.051	5.380	0.000	0.174	0.373
12.25	-0.0326	0.051	-0.642	0.521	-0.132	0.067
12.5	-0.0232	0.051	-0.456	0.648	-0.123	0.076
12.75	-0.0563	0.051	-1.108	0.268	-0.156	0.043
13.0	0.1907	0.051	3.757	0.000	0.091	0.290

	coef	std err	z	P>|z|	[0.025	0.975]
13.25	0.0611	0.051	1.207	0.227	-0.038	0.160
13.5	0.0428	0.051	0.847	0.397	-0.056	0.142
13.75	-0.1337	0.050	-2.652	0.008	-0.233	-0.035
14.0	-0.0929	0.050	-1.848	0.065	-0.191	0.006
14.25	-0.0457	0.050	-0.913	0.361	-0.144	0.052
14.5	-0.0153	0.050	-0.305	0.760	-0.113	0.083
14.75	-0.0199	0.050	-0.401	0.689	-0.117	0.078
15.0	-0.1126	0.050	-2.269	0.023	-0.210	-0.015
15.25	-0.0323	0.049	-0.653	0.514	-0.129	0.065
15.5	0.0730	0.049	1.482	0.138	-0.024	0.170
15.75	0.0210	0.049	0.427	0.669	-0.075	0.117
16.0	-0.2138	0.049	-4.367	0.000	-0.310	-0.118
16.25	-0.0712	0.049	-1.459	0.145	-0.167	0.024
16.5	0.1001	0.049	2.055	0.040	0.005	0.196
16.75	0.3035	0.049	6.229	0.000	0.208	0.399
17.0	-0.5145	0.049	-10.558	0.000	-0.610	-0.419
17.25	-0.1779	0.049	-3.646	0.000	-0.274	-0.082
17.5	0.2955	0.049	6.039	0.000	0.200	0.391
17.75	0.4898	0.049	9.990	0.000	0.394	0.586
18.0	-0.1985	0.049	-4.049	0.000	-0.295	-0.102
18.25	-0.0969	0.049	-1.975	0.048	-0.193	-0.001
18.5	0.0557	0.049	1.137	0.256	-0.040	0.152
18.75	0.0061	0.049	0.124	0.901	-0.090	0.102
19.0	0.1435	0.049	2.931	0.003	0.048	0.239
19.25	0.0677	0.049	1.386	0.166	-0.028	0.163
19.5	-0.1602	0.049	-3.282	0.001	-0.256	-0.065
19.75	-0.2864	0.049	-5.876	0.000	-0.382	-0.191
20.0	0.3010	0.049	6.181	0.000	0.206	0.396
20.25	0.1699	0.049	3.494	0.000	0.075	0.265
20.5	-0.1209	0.049	-2.490	0.013	-0.216	-0.026
20.75	-0.3967	0.049	-8.175	0.000	-0.492	-0.302
21.0	0.1078	0.048	2.222	0.026	0.013	0.203
21.25	0.0268	0.048	0.552	0.581	-0.068	0.122
21.5	-0.2671	0.048	-5.511	0.000	-0.362	-0.172
21.75	-0.4143	0.048	-8.548	0.000	-0.509	-0.319
22.0	-0.2735	0.048	-5.641	0.000	-0.368	-0.178
22.25	0.0732	0.049	1.508	0.131	-0.022	0.168
22.5	0.0537	0.048	1.107	0.268	-0.041	0.149
22.75	-0.1388	0.048	-2.863	0.004	-0.234	-0.044
23.0	0.4995	0.048	10.301	0.000	0.404	0.595
23.25	0.1507	0.048	3.110	0.002	0.056	0.246
23.5	-0.4362	0.049	-8.994	0.000	-0.531	-0.341
23.75	-0.5626	0.049	-11.581	0.000	-0.658	-0.467
mon	-0.0498	0.026	-1.924	0.054	-0.100	0.001
tue	-0.0081	0.027	-0.298	0.765	-0.061	0.045
wen	0.0389	0.028	1.405	0.160	-0.015	0.093
thu	0.0453	0.027	1.674	0.094	-0.008	0.098
fri	0.1302	0.026	4.939	0.000	0.079	0.182
sat	0.2389	0.022	10.724	0.000	0.195	0.283
sun	0.2640	0.022	12.198	0.000	0.222	0.306
sin1	0.1048	0.018	5.950	0.000	0.070	0.139
cos1	-0.0373	0.021	-1.745	0.081	-0.079	0.005
sin2	-0.1310	0.016	-8.305	0.000	-0.162	-0.100
cos2	0.0403	0.011	3.711	0.000	0.019	0.062
holiday	-0.1394	0.025	-5.484	0.000	-0.189	-0.090
vacation	0.3133	0.028	11.354	0.000	0.259	0.367

	coef	std err	z	P>|z|	[0.025	0.975]
Aggregate_Wind	9.843e-05	1.21e-05	8.124	0.000	7.47e-05	0.000
Aggregate_Solar	-0.0001	1.78e-05	-7.892	0.000	-0.000	-0.000
Aggregate_Load	-9.504e-05	1.02e-05	-9.352	0.000	-0.000	-7.51e-05
LF	-0.0007	0.000	-1.531	0.126	-0.001	0.000
NG	6.481e-05	2.13e-05	3.042	0.002	2.3e-05	0.000
NU	8.254e-05	1.1e-05	7.528	0.000	6.1e-05	0.000
WA	-0.0002	4.18e-05	-5.267	0.000	-0.000	-0.000
WI	-0.0004	9.3e-05	-4.206	0.000	-0.001	-0.000
Other	-0.0003	5.74e-05	-4.369	0.000	-0.000	-0.000

Table A.2: Table of model coefficients for the Robust Linear Regression Model. LF = Liquid Fuel, NG = Natural Gas, NU = Nuclear, WA = Water, WI = Wind.

A.3. Predictive Linear Model Price-Maker Effect

	coef	std err	z	P>|z|	[0.025	0.975]
const	-4.1406	0.136	-30.525	0.000	-4.406	-3.875
0.0	0.1286	0.044	2.946	0.003	0.043	0.214
0.25	0.0718	0.044	1.638	0.101	-0.014	0.158
0.5	0.0658	0.044	1.497	0.134	-0.020	0.152
0.75	0.1221	0.044	2.765	0.006	0.036	0.209
1.0	0.0658	0.044	1.484	0.138	-0.021	0.153
1.25	0.0345	0.045	0.775	0.439	-0.053	0.122
1.5	0.0569	0.045	1.271	0.204	-0.031	0.145
1.75	0.1104	0.045	2.455	0.014	0.022	0.199
2.0	0.0970	0.045	2.146	0.032	0.008	0.185
2.25	0.0935	0.045	2.063	0.039	0.005	0.182
2.5	0.1170	0.045	2.573	0.010	0.028	0.206
2.75	0.1410	0.046	3.093	0.002	0.052	0.230
3.0	0.1421	0.046	3.112	0.002	0.053	0.232
3.25	0.1607	0.046	3.515	0.000	0.071	0.250
3.5	0.1752	0.046	3.828	0.000	0.086	0.265
3.75	0.1780	0.046	3.887	0.000	0.088	0.268
4.0	0.2020	0.046	4.415	0.000	0.112	0.292
4.25	0.2090	0.046	4.573	0.000	0.119	0.299
4.5	0.2069	0.046	4.534	0.000	0.117	0.296
4.75	0.2074	0.045	4.559	0.000	0.118	0.297
5.0	0.2671	0.045	5.902	0.000	0.178	0.356
5.25	0.2351	0.045	5.213	0.000	0.147	0.324
5.5	0.1912	0.045	4.258	0.000	0.103	0.279
5.75	0.1818	0.045	4.071	0.000	0.094	0.269
6.0	0.2186	0.044	4.949	0.000	0.132	0.305
6.25	0.1861	0.044	4.244	0.000	0.100	0.272
6.5	0.1340	0.044	3.072	0.002	0.048	0.220
6.75	0.0883	0.043	2.032	0.042	0.003	0.174
7.0	0.1192	0.043	2.747	0.006	0.034	0.204
7.25	0.0358	0.043	0.827	0.408	-0.049	0.121
7.5	-0.0083	0.043	-0.192	0.848	-0.093	0.077
7.75	-0.0618	0.043	-1.426	0.154	-0.147	0.023
8.0	-0.1165	0.043	-2.689	0.007	-0.201	-0.032
8.25	-0.1250	0.043	-2.882	0.004	-0.210	-0.040
8.5	-0.1434	0.043	-3.300	0.001	-0.229	-0.058
8.75	-0.1400	0.044	-3.216	0.001	-0.225	-0.055
9.0	-0.1648	0.044	-3.773	0.000	-0.250	-0.079
9.25	-0.1664	0.044	-3.799	0.000	-0.252	-0.081
9.5	-0.1837	0.044	-4.175	0.000	-0.270	-0.097
9.75	-0.1717	0.044	-3.887	0.000	-0.258	-0.085
10.0	-0.1296	0.044	-2.924	0.003	-0.216	-0.043
10.25	-0.1302	0.044	-2.928	0.003	-0.217	-0.043
10.5	-0.1528	0.045	-3.420	0.001	-0.240	-0.065
10.75	-0.1548	0.045	-3.450	0.001	-0.243	-0.067
11.0	-0.1762	0.045	-3.907	0.000	-0.265	-0.088
11.25	-0.1690	0.045	-3.728	0.000	-0.258	-0.080
11.5	-0.1817	0.046	-3.988	0.000	-0.271	-0.092
11.75	-0.1766	0.046	-3.865	0.000	-0.266	-0.087
12.0	-0.1142	0.046	-2.507	0.012	-0.204	-0.025
12.25	-0.1448	0.046	-3.182	0.001	-0.234	-0.056
12.5	-0.1403	0.046	-3.081	0.002	-0.230	-0.051
12.75	-0.1430	0.046	-3.140	0.002	-0.232	-0.054
13.0	-0.1806	0.045	-3.969	0.000	-0.270	-0.091

	coef	std err	z	P>|z|	[0.025	0.975]
13.25	-0.1611	0.045	-3.548	0.000	-0.250	-0.072
13.5	-0.1414	0.045	-3.120	0.002	-0.230	-0.053
13.75	-0.1268	0.045	-2.805	0.005	-0.215	-0.038
14.0	-0.1551	0.045	-3.443	0.001	-0.243	-0.067
14.25	-0.1460	0.045	-3.251	0.001	-0.234	-0.058
14.5	-0.1567	0.045	-3.501	0.000	-0.244	-0.069
14.75	-0.1249	0.045	-2.799	0.005	-0.212	-0.037
15.0	-0.1308	0.044	-2.940	0.003	-0.218	-0.044
15.25	-0.1289	0.044	-2.907	0.004	-0.216	-0.042
15.5	-0.1370	0.044	-3.102	0.002	-0.224	-0.050
15.75	-0.1403	0.044	-3.188	0.001	-0.227	-0.054
16.0	-0.1270	0.044	-2.894	0.004	-0.213	-0.041
16.25	-0.1381	0.044	-3.155	0.002	-0.224	-0.052
16.5	-0.1453	0.044	-3.326	0.001	-0.231	-0.060
16.75	-0.1540	0.044	-3.525	0.000	-0.240	-0.068
17.0	-0.1923	0.044	-4.402	0.000	-0.278	-0.107
17.25	-0.2510	0.044	-5.739	0.000	-0.337	-0.165
17.5	-0.2800	0.044	-6.385	0.000	-0.366	-0.194
17.75	-0.2840	0.044	-6.462	0.000	-0.370	-0.198
18.0	-0.3192	0.044	-7.262	0.000	-0.405	-0.233
18.25	-0.3195	0.044	-7.269	0.000	-0.406	-0.233
18.5	-0.3208	0.044	-7.300	0.000	-0.407	-0.235
18.75	-0.3033	0.044	-6.906	0.000	-0.389	-0.217
19.0	-0.2746	0.044	-6.257	0.000	-0.361	-0.189
19.25	-0.2559	0.044	-5.843	0.000	-0.342	-0.170
19.5	-0.2481	0.044	-5.671	0.000	-0.334	-0.162
19.75	-0.2278	0.044	-5.213	0.000	-0.313	-0.142
20.0	-0.1314	0.044	-3.010	0.003	-0.217	-0.046
20.25	-0.1031	0.044	-2.365	0.018	-0.189	-0.018
20.5	-0.1104	0.044	-2.536	0.011	-0.196	-0.025
20.75	-0.0592	0.044	-1.360	0.174	-0.144	0.026
21.0	-0.0079	0.043	-0.182	0.856	-0.093	0.077
21.25	0.0046	0.043	0.106	0.916	-0.081	0.090
21.5	0.0130	0.043	0.299	0.765	-0.072	0.098
21.75	0.0456	0.043	1.050	0.294	-0.040	0.131
22.0	0.0834	0.043	1.920	0.055	-0.002	0.169
22.25	0.0395	0.043	0.909	0.364	-0.046	0.125
22.5	0.0379	0.043	0.872	0.383	-0.047	0.123
22.75	0.0621	0.043	1.429	0.153	-0.023	0.147
23.0	0.0602	0.043	1.386	0.166	-0.025	0.145
23.25	0.0759	0.043	1.747	0.081	-0.009	0.161
23.5	0.1025	0.043	2.357	0.018	0.017	0.188
23.75	0.1690	0.044	3.881	0.000	0.084	0.254
mon	-0.5112	0.023	-22.053	0.000	-0.557	-0.466
tue	-0.5117	0.024	-21.110	0.000	-0.559	-0.464
wen	-0.5028	0.025	-20.257	0.000	-0.551	-0.454
thu	-0.6431	0.024	-26.488	0.000	-0.691	-0.595
fri	-0.6167	0.024	-26.099	0.000	-0.663	-0.570
sat	-0.7598	0.020	-38.039	0.000	-0.799	-0.721
sun	-0.5954	0.019	-30.693	0.000	-0.633	-0.557
sin1	-0.1149	0.016	-7.279	0.000	-0.146	-0.084
cos1	-0.1189	0.019	-6.206	0.000	-0.156	-0.081
sin2	-0.3304	0.014	-23.362	0.000	-0.358	-0.303
cos2	-0.5697	0.010	-58.553	0.000	-0.589	-0.551
holiday	0.0450	0.023	1.975	0.048	0.000	0.090
vacation	0.4560	0.025	18.435	0.000	0.408	0.505

	coef	std err	z	P>|z|	[0.025	0.975]
Aggregate_Wind	-0.0001	1.09e-05	-13.544	0.000	-0.000	-0.000
Aggregate_Solar	-6.844e-05	1.59e-05	-4.299	0.000	-9.96e-05	-3.72e-05
Aggregate_Load	0.0003	9.11e-06	28.400	0.000	0.000	0.000
LF	0.0136	0.000	35.307	0.000	0.013	0.014
NG	0.0002	1.91e-05	8.821	0.000	0.000	0.000
NU	-6.968e-05	9.83e-06	-7.089	0.000	-8.89e-05	-5.04e-05
WA	-0.0010	3.75e-05	-26.803	0.000	-0.001	-0.001
WI	-0.0006	8.34e-05	-6.829	0.000	-0.001	-0.000
Other	0.0010	5.15e-05	18.926	0.000	0.001	0.001

Table A.3: Table of model coefficients for the Robust Linear Regression Model. LF = Liquid Fuel, NG = Natural Gas, NU = Nuclear, WA = Water, WI = Wind.

B

Appendix

B.1. Price-Taker Analysis

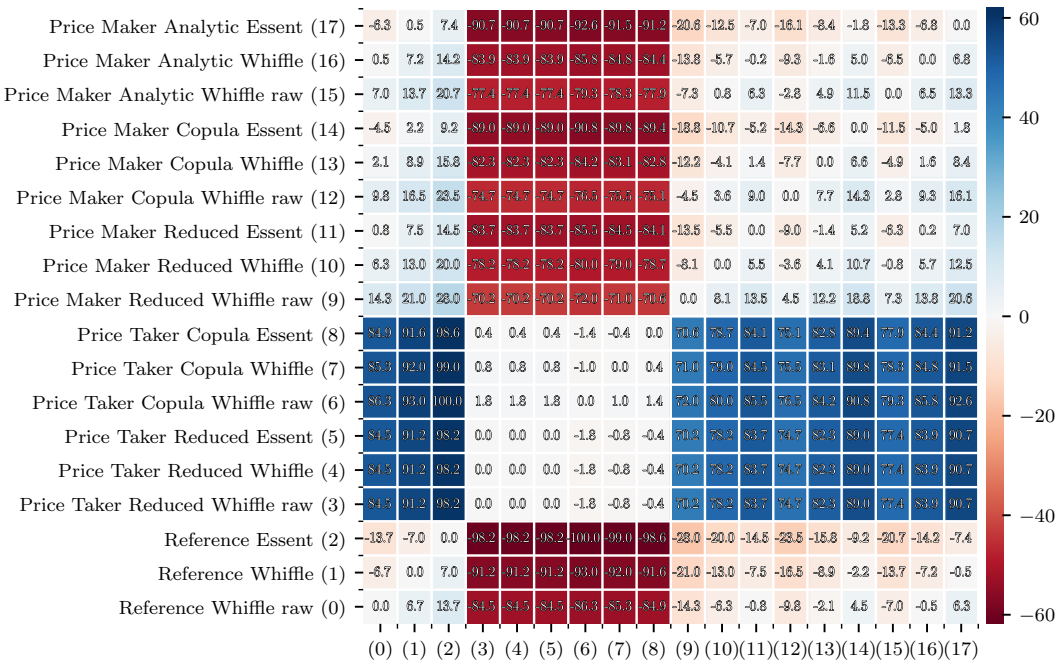


Figure B.1: Expected price-taker difference of the mean for all strategies. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=€61.83. The expected difference is computed for rows over columns.

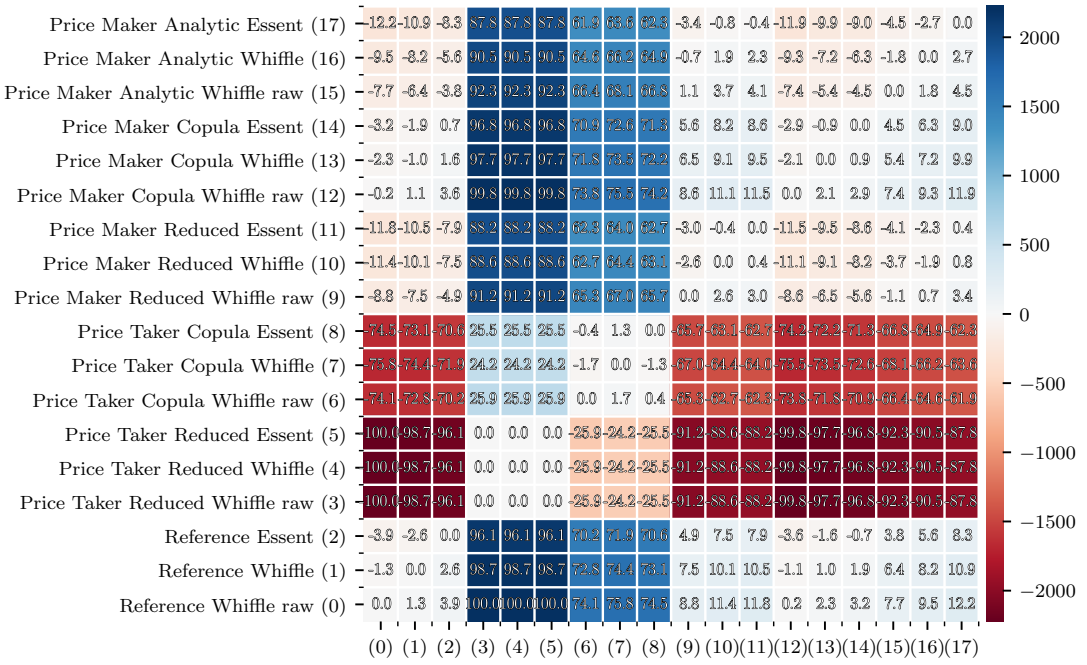


Figure B.2: Price-taker difference of $\text{CVaR}_{5\%}$ for all strategies. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=€2222.75. The expected difference is computed for rows over columns.

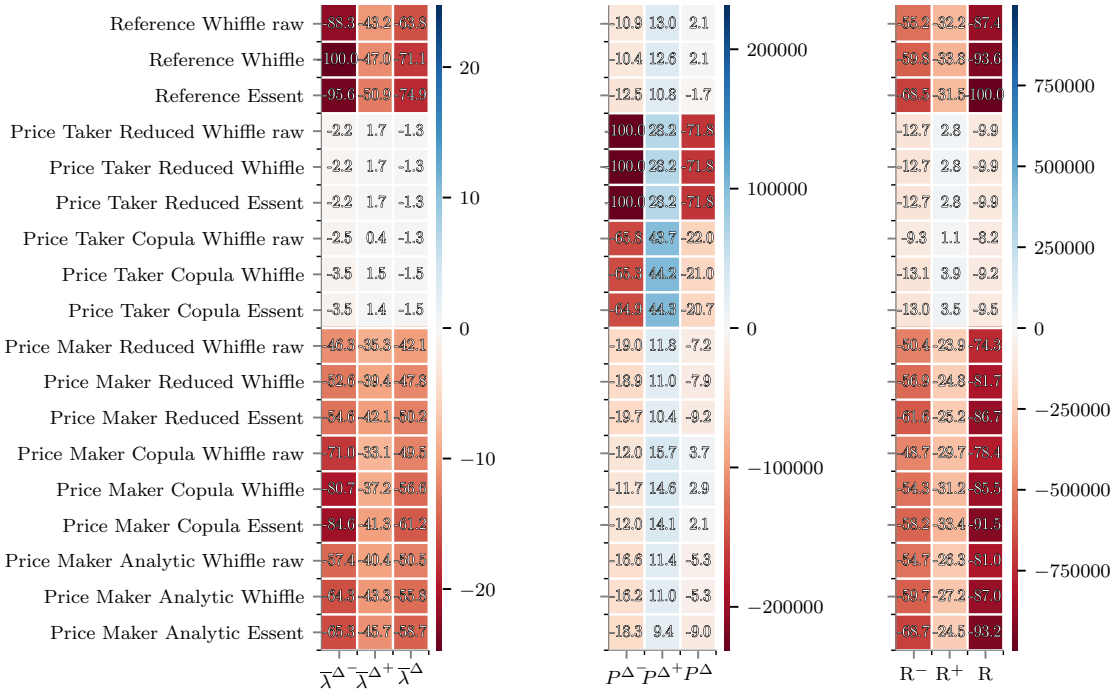


Figure B.3: Market quantities for all strategies. The colorbars indicate quantities in the original unit of measurement. The cells indicate quantities as a percentage of the scale of the colorbars.
a: Average opportunity price realized for short volume, long volume and total imbalance volume in €/MWh. 100%=€24.69/MWh.
b: Short volume, long volume and totale imbalance volume in MWh. 100%=231,122MWh.
c: Realized opportunity for being short, long and total in €. 100%=€996,021.

B.2. Price-Maker Analysis

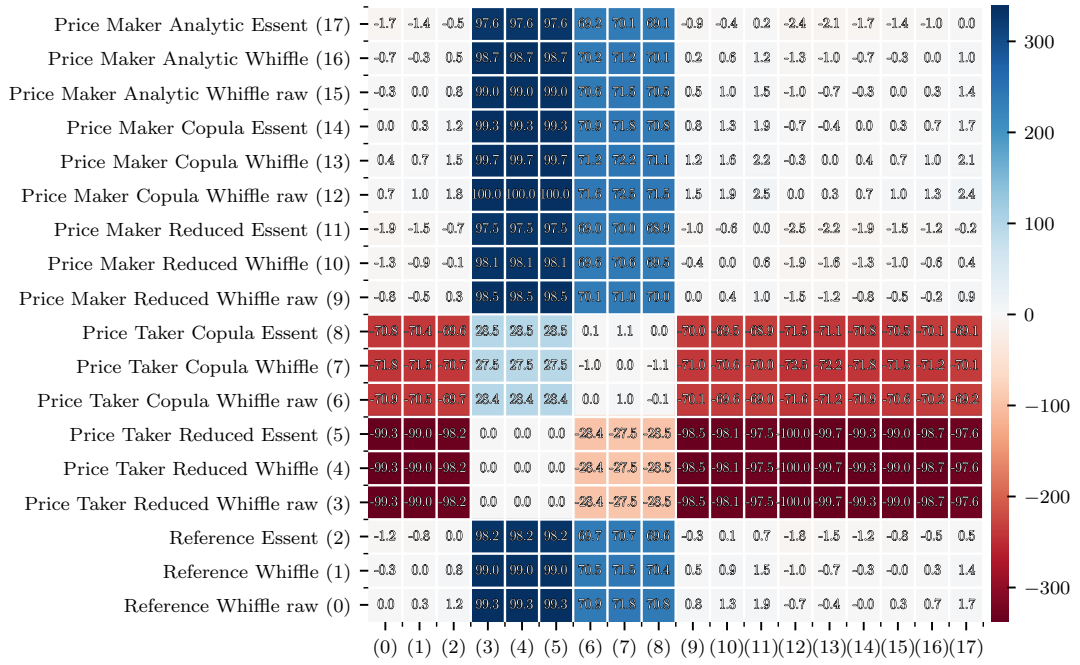


Figure B.4: Expected price-maker difference of the mean for all strategies. The colorbar indicates the difference in €. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=€337.83. The expected difference is computed for rows over columns.

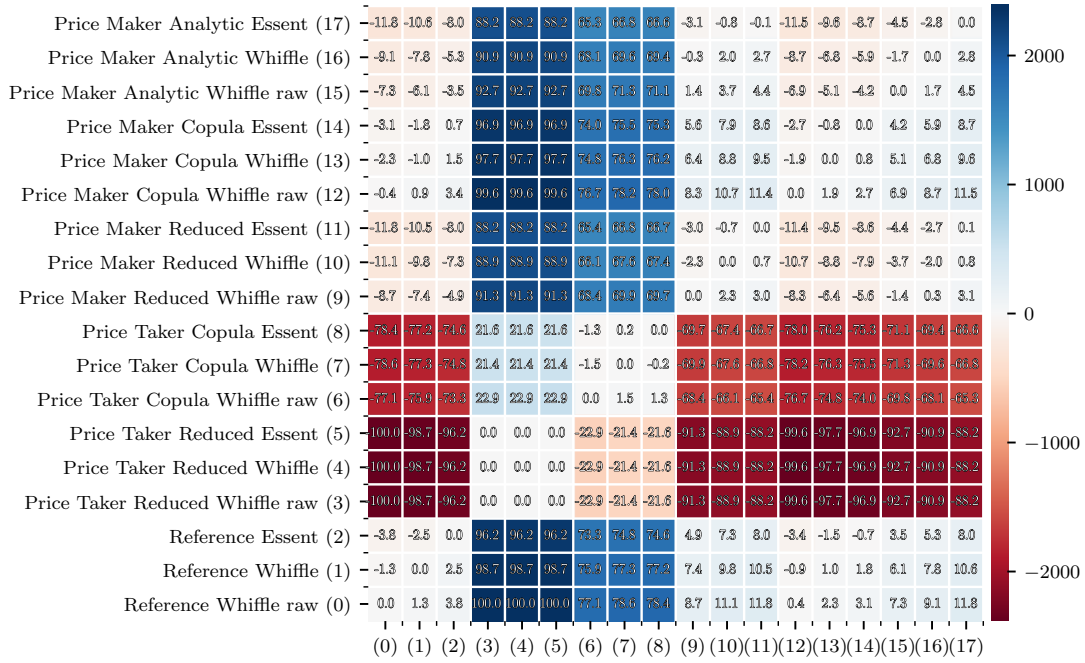


Figure B.5: Price-taker difference of $CVaR_{5\%}$ for all strategies. The colorbar indicates the difference in €. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=€2,383.95. The expected difference is computed for rows over columns.

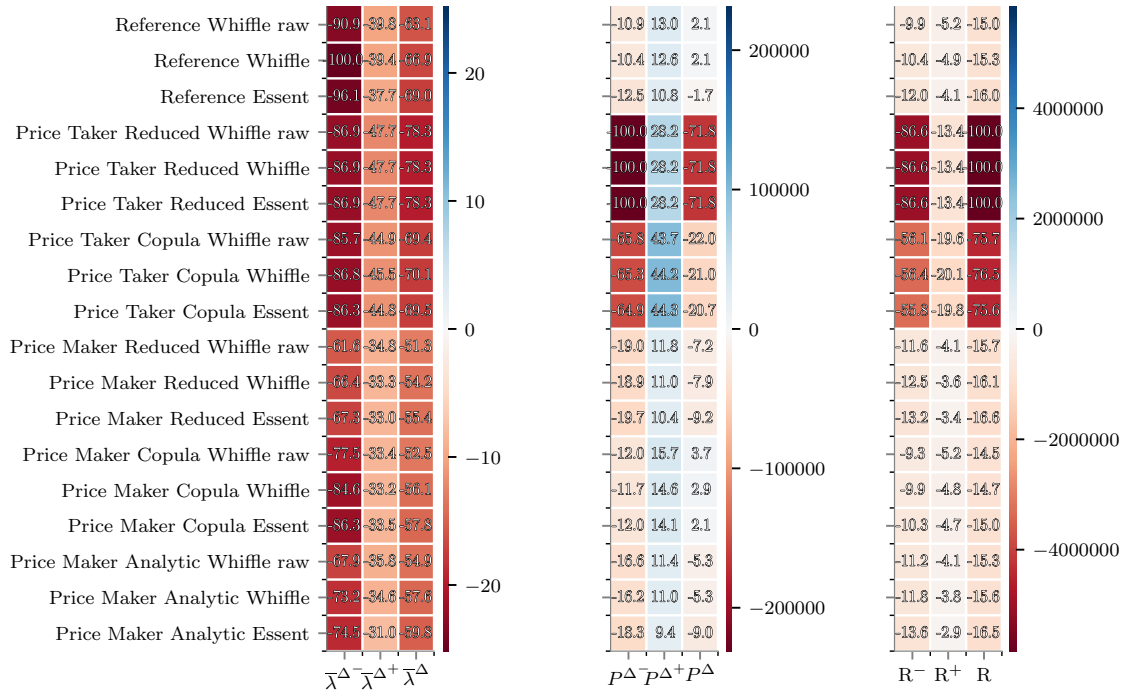


Figure B.6: Market quantities for all strategies. The colorbars indicate quantities in the original unit of measurement. The cells indicate quantities as a percentage of the scale of the colorbars.
a: Average opportunity price realized for short volume, long volume and total imbalance volume in €/MWh. 100%=€25.16/MWh.
b: Short volume, long volume and totale imbalance volume in MWh. 100%=231,122MWh.
c: Realized opportunity for being short, long and total in €. 100%=€,5838,701.28.

B.3. System Effects

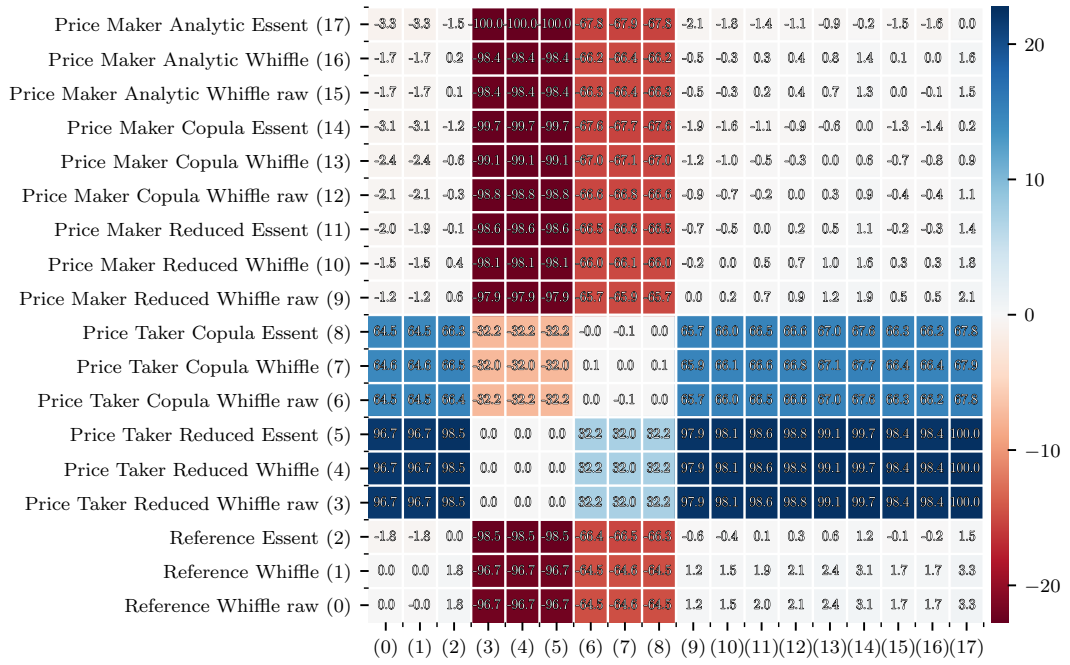


Figure B.7: Expected price-maker difference of the mean system imbalance for the reference and price-maker strategies. The colorbar indicates the difference in MW. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=22.72MW. The expected difference is computed for rows over columns.

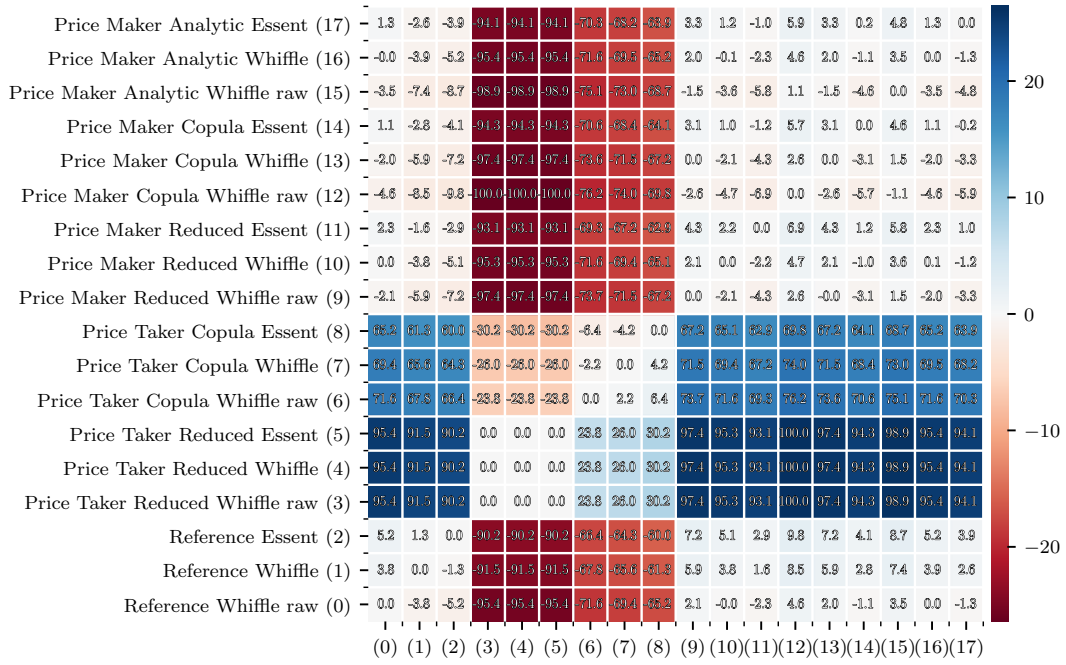


Figure B.8: Expected price-maker difference of the $\text{CVaR}_{5\%}$ of the system imbalance for the reference and price-maker strategies. The colorbar indicates the difference in MW. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=26.41MW. The expected difference is computed for rows over columns.

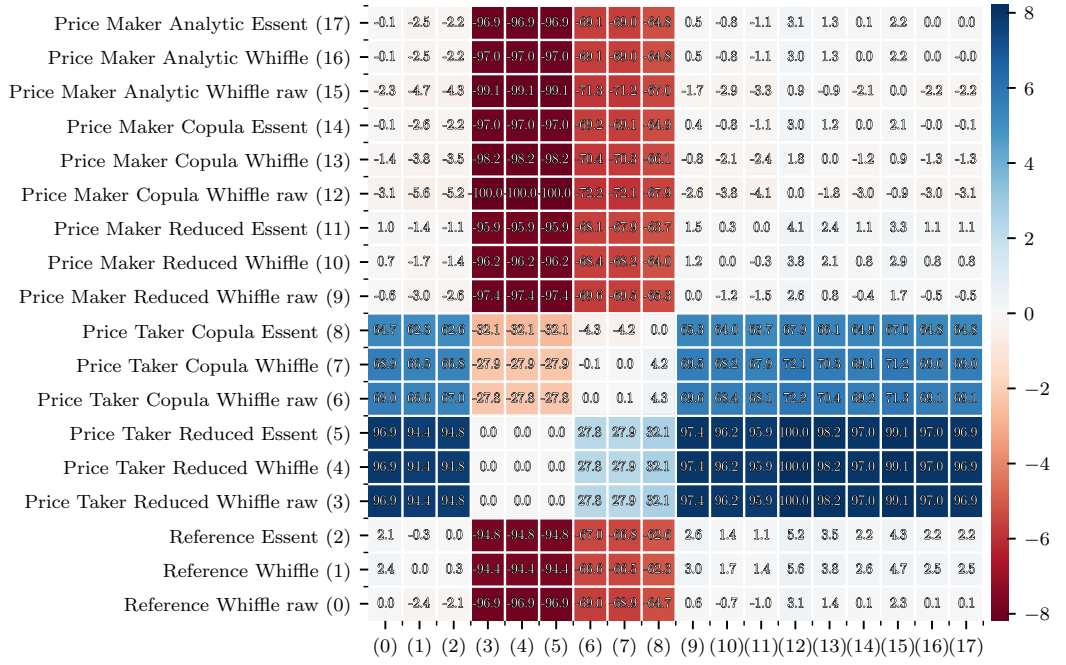


Figure B.9: Expected price-maker difference of the volatility of the system imbalance for the reference and price-maker strategies. The colorbar indicates the difference in MW. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=8.19MW. The expected difference is computed for rows over columns.

B.4. System Costs

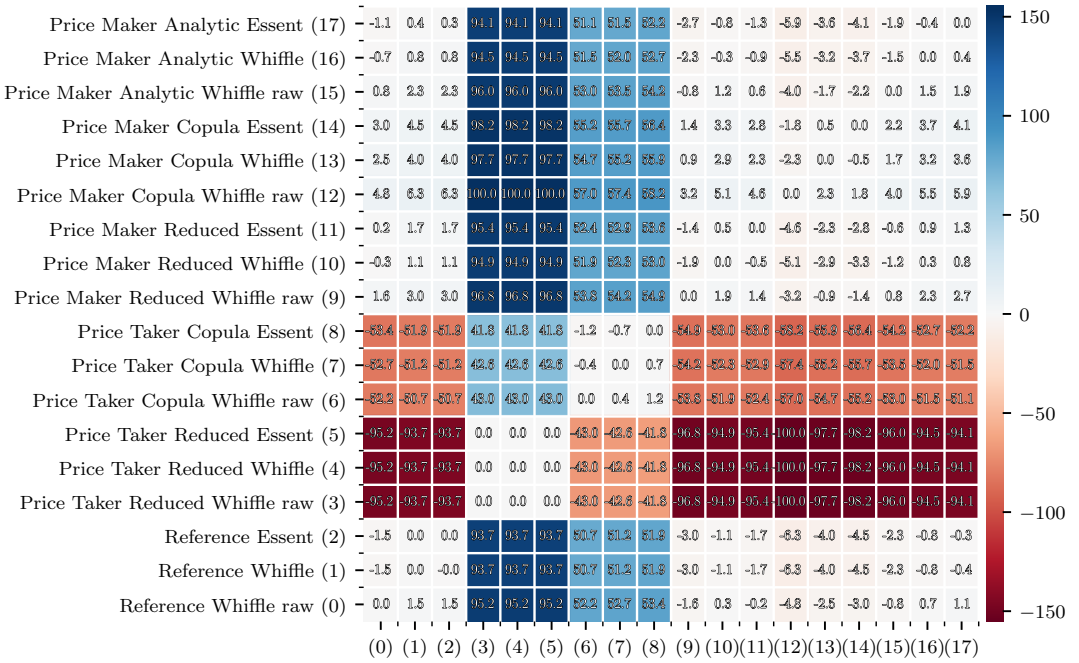


Figure B.10: Expected price-maker difference of the mean system opportunity for the reference and price-maker strategies. The colorbar indicates the difference in €. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=€155.20. The expected difference is computed for rows over columns.

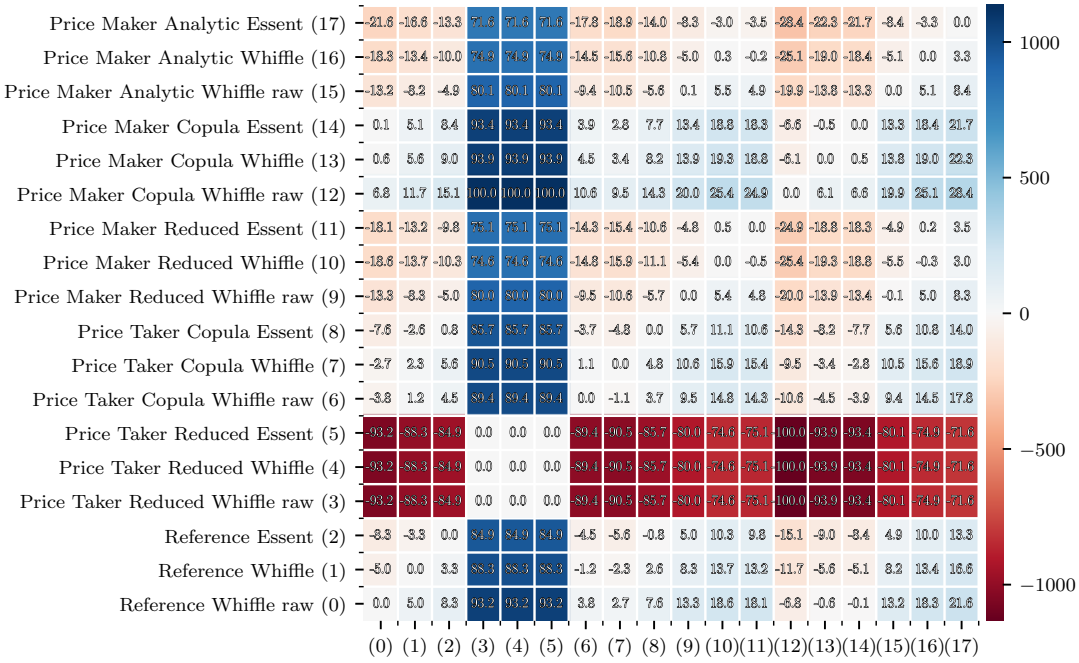


Figure B.11: Expected price-maker difference of the $CVaR_{5\%}$ for the reference and price-maker strategies. The colorbar indicates the difference in €. The cells indicate the value as a percentage of the scale of the colorbar, where 100%=€1,133.41.

Bibliography

[1] Aggarwal S. K., Saini, L. M. and Kumar, A. (2009). Electricity price forecasting in deregulated markets: A review and evaluation. *International Journal of Electrical Power & Energy Systems*, 31(1):13–22. 10.1016/J.IJEPES.2008.09.003.

[2] Aitchison, J. and Aitken, C. G. G. (1976). Multivariate binary discrimination by the kernel method. *Biometrika*, 63(3):413–420.

[3] Arrow, K. J. and Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society*, pages 265–290.

[4] Bertrand, G. and Papavasiliou, A. (2017). Optimal dispatch of wind farms facing market prices. In *2017 14th International Conference on the European Energy Market (EEM)*, pages 1–6. 10.1109/EEM.2017.7981871.

[5] Biggar, D. R. and Hesamzadeh, M. R. (2014). *The economics of electricity markets*. John Wiley & Sons.

[6] BP (2018). Statistical Review of World Energy. Technical report.

[7] Brent, R. P. (2013). *Algorithms for minimization without derivatives*. Courier Corporation.

[8] Chaves Avila, J. P. (2014). European short-term electricity market designs under high penetration of wind power.

[9] Chaves-Ávila, J. P., Hakvoort, R. A., and Ramos, A. (2013). Short-term strategies for Dutch wind power producers to reduce imbalance costs. *Energy Policy*, 52:573–582. doi.org/10.1016/j.enpol.2012.10.011.

[10] Conejo, A. J., Carrión, M., and Morales, J. M. (2010). *Decision making under uncertainty in electricity markets*, volume 1. Springer.

[11] Conejo, A. J. and Sioshansi, R. (2018). Rethinking restructured electricity market design: Lessons learned and future needs. *International Journal of Electrical Power & Energy Systems*, 98:520–530. https://doi.org/10.1016/j.ijepes.2017.12.014.

[12] Constantain, G. (2018). Probability Metrics and the Stability of Stochastic Models. *Bulletin of the London Mathematical Society*, 27(1):87–89. 10.1112/blms/27.1.87.

[13] Dupačová, J., Gröwe-Kuska, N., and Römisch, W. (2003). Scenario reduction in stochastic programming. *Mathematical programming*, 95(3):493–511.

[14] Gaillard, P., Goude, Y., and Nedellec, R. (2016). Additive models and robust aggregation for gefcom2014 probabilistic electric load and electricity price forecasting. *International Journal of Forecasting*, 32(3):1038 – 1050.

[15] Giebel, G., Brownsword, R., Kariniotakis, G., Denhard, M., and Draxl, C. (2011). The state-of-the-art in short-term prediction of wind power: A literature overview. *ANEMOS. plus*.

[16] Giebel, G. and Kariniotakis, G. (2017). 3 - Wind power forecasting—a review of the state of the art BT - Renewable Energy Forecasting. In *Woodhead Publishing Series in Energy*, pages 59–109. Woodhead Publishing. https://doi.org/10.1016/B978-0-08-100504-0.00003-2.

[17] Glahn, H. R. and Lowry, D. A. (1972). The use of model output statistics (MOS) in objective weather forecasting. *Journal of applied meteorology*, 11(8):1203–1211.

[18] Gneiting, T., Stanberry, L. I., Grimit, E. P., Held, L., and Johnson, N. A. (2008). Assessing probabilistic forecasts of multivariate quantities, with an application to ensemble predictions of surface winds. *TEST*, 17(2):211. 10.1007/s11749-008-0114-x.

[19] Gombos, D., Hansen, J. A., Du, J., and McQueen, J. (2007). Theory and Applications of the Minimum Spanning Tree Rank Histogram. *Monthly Weather Review*, 135(4):1490–1505. doi:10.1175/MWR3362.1.

[20] Guerrero-Mestre, V., de la Nieta, A. A. S., Contreras, J., and Catalão, J. P. S. (2016). Optimal bidding of a group of wind farms in day-ahead markets through an external agent. *IEEE Transactions on Power Systems*, 31(4):2688–2700.

[21] Hart, W. E., Laird, C. D., Watson, J.-P., Woodruff, D. L., Hackebeil, G. A., Nicholson, B. L., and Siirola, J. D. (2012). *Pyomo-optimization modeling in python*, volume 67. Springer.

[22] Haupt, S. E., Jiménez, P. A., Lee, J. A., and Kosović, B. (2017). 1 - Principles of meteorology and numerical weather prediction A2 - Kariniotakis, George BT - Renewable Energy Forecasting. In *Woodhead Publishing Series in Energy*, pages 3–28. Woodhead Publishing. https://doi.org/10.1016/B978-0-08-100504-0.00001-9.

[23] Heitsch, H. and Römisch, W. (2003). Scenario reduction algorithms in stochastic programming. *Computational optimization and applications*, 24(2-3):187–206.

[24] Hirth, L. (2013). The market value of variable renewables: The effect of solar wind power variability on their relative price. *Energy Economics*, 38:218–236. https://doi.org/10.1016/j.eneco.2013.02.004.

[25] Hirth, L., Ueckerdt, F., and Edenhofer, O. (2015). Integration costs revisited – An economic framework for wind and solar variability. *Renewable Energy*, 74:925–939.

[26] Huber, P. J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.*, 35(1):73–101. 10.1214/aoms/1177703732.

[27] Hyndman, R. J. and Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.

[28] IRENA (2018). Wind energy. http://www.irena.org/wind.

[29] Jeon, J. and Taylor, J. W. (2012). Using conditional kernel density estimation for wind power density forecasting. *Journal of the American Statistical Association*, 107(497):66–79.

[30] Jónsson, T., Pinson, P., Madsen, H., and Nielsen, H. A. (2014). Predictive densities for day-ahead electricity prices using time-adaptive quantile regression. *Energies*, 7(9):5523–5547.

[31] Jónsson, T., Pinson, P., Nielsen, H. A., Madsen, H., and Nielsen, T. S. (2013). Forecasting electricity spot prices accounting for wind power predictions. *IEEE Transactions on Sustainable Energy*, 4(1):210–218.

[32] Kariniotakis, G., Marti, I., Casas, D., Pinson, P., Nielsen, T. S., Madsen, H., Giebel, G., Usaola, J., Sanchez, I., and Palomares, A. M. (2004). What performance can be expected by short-term wind power prediction models depending on site characteristics. In *Proceedings of the European Wind Energy Conference EWEC*, pages 22–25.

[33] Koenker, R. (2005). *Quantile Regression*. Econometric Society Monographs. Cambridge University Press.

[34] KU Leuven Energy Institute (2015). The current electricity market design in Europe. Technical report. https://set.kuleuven.be/ei/factsheets.

[35] Lipsey, R. G. and Lancaster, K. (1956). The general theory of second best. *The review of economic studies*, 24(1):11–32.

[36] Liu, P.-L. and Der Kiureghian, A. (1986). Multivariate distribution models with prescribed marginals and covariances. *Probabilistic Engineering Mechanics*, 1(2):105–112.

[37] Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130–141.

[38] Matevosyan, J. and Soder, L. (2006). Minimization of imbalance cost trading wind power on the short-term power market. *IEEE Transactions on Power Systems*, 21(3):1396–1404.

[39] Mazzi, N. and Pinson, P. (2017). 10 - Wind power in electricity markets and the value of forecasting A2 - Kariniotakis, George BT - Renewable Energy Forecasting. In *Woodhead Publishing Series in Energy*, pages 259–278. Woodhead Publishing. https://doi.org/10.1016/B978-0-08-100504-0.00010-X.

[40] McKenzie, L. W. (1959). On the existence of general equilibrium for a competitive market. *Econometrica: journal of the Econometric Society*, pages 54–71.

[41] Mitridati, L. and Pinson, P. (2018). A Bayesian inference approach to unveil supply curves in electricity markets. *IEEE Transactions on Power Systems*, 33(3):2610–2620.

[42] Morales, J., J. Conejo, A., Madsen, H., Pinson, P., and Zugno, M. (2014). *Integrating Renewables in Electricity Markets - Operational Problems*.

[43] Morales, J. M., Conejo, A. J., and Pérez-Ruiz, J. (2010). Short-term trading for a wind power producer. *IEEE Transactions on Power Systems*, 25(1):554–564.

[44] Morales, J. M., Pineda, S., Conejo, A. J., and Carrion, M. (2009). Scenario reduction for futures market trading in electricity markets. *IEEE Transactions on Power Systems*, 24(2):878–888.

[45] Nielsen, H. A., Nielsen, T. S., Madsen, H., Giebel, G., Badger, J., Landberg, L., Sattler, K., Voulund, L., and Tofting, J. (2006). From wind ensembles to probabilistic information about future wind power production–results from an actual application. In *Probabilistic Methods Applied to Power Systems, 2006. PMAPS 2006. International Conference on*, pages 1–8. IEEE.

[46] Olsson, M. and Soder, L. (2009). Estimating real-time balancing prices in wind power systems. In *2009 IEEE/PES Power Systems Conference and Exposition*, pages 1–9. 10.1109/PSCE.2009.4840069.

[47] Pinson, P., Chevallier, C., and Kariniotakis, G. N. (2007a). Trading Wind Generation From Short-Term Probabilistic Forecasts of Wind Power. *IEEE Transactions on Power Systems*, 22(3):1148–1156. 10.1109/TPWRS.2007.901117.

[48] Pinson, P. and Girard, R. (2012). Evaluating the quality of scenarios of short-term wind power generation. *Applied Energy*, 96:12–20.

[49] Pinson, P. and Kariniotakis, G. (2010). Conditional prediction intervals of wind power generation. *IEEE Transactions on Power Systems*, 25(4):1845–1856.

[50] Pinson, P., Nielsen, H., Moller, K., Madsen, H., and Kariniotakis, G. (2007b). Non-parametric probabilistic forecasts of wind power: required properties and evaluation. *Wind Energy*, 10(6):497–516. 10.1002/we.230.

[51] Rahimiyan, M., Morales, J. M., and Conejo, A. J. (2011). Evaluating alternative offering strategies for wind producers in a pool. *Applied energy*, 88(12):4918–4926.

[52] Regulatory Assistance Project (RAP) (2014). Power Market Operations and System Reliability: a contribution to the market design debate in the Pentalateral Energy Forum. Technical report.

[53] Schalkwijk, J., Jonker, H. J. J., Siebesma, A. P., and Bosveld, F. C. (2015). A year-long large-eddy simulation of the weather over Cabauw: An overview. *Monthly Weather Review*, 143(3):828–844.

[54] Sperati, S., Alessandrini, S., Pinson, P., and Kariniotakis, G. (2015). The "weather intelligence for renewable energies" benchmarking exercise on short-term forecasting of wind and solar power generation. *Energies*, 8(9):9594–9619.

[55] Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M. (2013). Climate change 2013: The physical science basis. *Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5)(Cambridge Univ Press, New York)*, 25.

[56] Taylor, J. W. and Jeon, J. (2015). Forecasting wind power quantiles using conditional kernel estimation. *Renewable Energy*, 80:370–379. https://doi.org/10.1016/j.renene.2015.02.022.

[57] The European Commission (2013). COMMISSION REGULATION (EU) No 543/2013.

[58] Tilmann, G., Fadoua, B., and E., R. A. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268. 10.1111/j.1467-9868.2007.00587.x.

[59] Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030–1081. 10.1016/J.IJFORECAST.2014.08.008.

[60] Widén, J. (2011). Correlations between large-scale solar and wind power in a future scenario for Sweden. *IEEE transactions on sustainable energy*, 2(2):177–184.

[61] Wilks, D. S. The Minimum Spanning Tree Histogram as a Verification Tool for Multidimensional Ensemble Forecasts. *Monthly Weather Review*, (6):1329–1340. 10.1175/1520-0493(2004)132<1329:TMSTHA>2.0.CO;2.

[62] Yan, J., Liu, Y., Han, S., Wang, Y., and Feng, S. (2015). Reviews on uncertainty analysis of wind power forecasting. *Renewable and Sustainable Energy Reviews*, 52:1322–1330. https://doi.org/10.1016/j.rser.2015.07.197.

[63] Zhang, Y., Wang, J., and Wang, X. (2014). Review on probabilistic forecasting of wind power generation. *Renewable and Sustainable Energy Reviews*, 32:255–270. https://doi.org/10.1016/j.rser.2014.01.033.