2020

Forecasting the Short-term Value of Wind Power for Risk-aware Bidding Strategies in Single-imbalance Price Electricity Markets

Gianni Goretti
Technological University Dublin, gianni.goretti@tudublin.ie

Funder: Technological University Dublin; European Union

Forecasting the Short-term Value of Wind Power for Risk-aware Bidding Strategies in Single-imbalance Price Electricity Markets

Gianni Goretti (BSc, MSc)

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Supervisors:

Prof. Aidan Duffy, Technological University Dublin, Ireland

Prof. Tek Tjing Lie, Auckland University of Technology, New Zealand

School of Civil and Structural Engineering, Technological University Dublin

May 2020

Abstract

The participation of wind energy in electricity markets and strategic bidding in the day-ahead market has been investigated with growing interest in recent years. However, markets adopting a single-price imbalance settlement — where participants can increase their profits if they help to put the system back into balance — have received very limited attention in the academic literature. In this Thesis, new probabilistic models forecasting the short-term value of wind power are developed and their use in bidding in these types of markets is investigated. The proposed strategies are designed for participants who want to bid strategically in the day-ahead market to increase the value of the energy generated at a wind farm, where value is measured in terms of revenue and exposure to risk. Following an extensive analysis of the available market data, two alternative approaches are developed to generate day-ahead forecasts of the market quantities of relevance for the work. These forecasts are then combined with short-term predictions of wind power in a probabilistic framework. Bids are formulated to reflect the participant's risk profile, conditioned upon the uncertainty in future wind power generation and electricity market conditions. The methodology is applied to a case study where the participation of a real wind farm in the new Irish electricity market is simulated over a test period. The benefits of the proposed models are clearly demonstrated as the strategies successfully improve the value of wind power for the participant by increasing their revenue while reducing the exposure to risk. Moreover, the market quantity forecasts developed in this work prove to be more valuable than a wind power forecast of higher accuracy for a risk-seeking participant.

Declaration

I certify that this thesis which I now submit for examination for the award of Ph.D., is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work.

This thesis was prepared according to the regulations for graduate study by research of the Technological University Dublin and has not been submitted in whole or in part for another award in any other third level institution.

The work reported on in this thesis conforms to the principles and requirements of the TU Dublin's guidelines for ethics in research.

Signature:  Date: 22 May 2020

Gianni Goretti

Acknowledgements

I would like to express my sincere appreciation and gratitude to my supervisor Prof. Aidan Duffy for his guidance throughout this research. I thank him for his consistent support and advice, for all the time, patience and effort he put into helping me with this project.

I would also like to thank all the people I had the pleasure to meet and work with at the Technological University Dublin (formerly Dublin Institute of Technology). In particular, Dr. Marek Rebow for his assistance throughout, John Turner, Dr. Niall Holmes and Una Beagon from the School of Civil Engineering, Prof. Michael Conlon and Dr. Marisa Llorens Salvador, and the Graduate Research School Office for their support. I am grateful to the staff and researchers in FOCAS and ESHI who contributed to making these years an enjoyable experience. A special thanks goes to Prof. Hugh Byrne, Brendan Walsh and Dr. Mohammed Yassin for all the fruitful conversations. I would also like to thank past and present colleagues at the Dublin Energy Lab: Dr. Ronan Oliver, Dr. Fintan McLoughlin, Deirdre Wolff, Dr. Brendan Cleary, Dr. Aritra Ghosh, Dr. Iftekhar Hussain, Daire Reilly, Dr. Philippe Lemarchand, Lee Corcoran, Iordanka Iordanova, Tim O'Leary, Noel O'Neill and Zhengdong Liu.

Thanks to my host supervisor Prof. Tek Tjing Lie at Auckland University of Technology and my colleagues at KEDRI for their support during my time there.

The participation in the IEA Wind Task 36 has allowed me to expand my professional network and to get in touch with experts in the wind energy forecasting field, and I must thank Dr. Gregor Giebel and Dr. Corinna Möhrlen for their

help and support. Thank you to Dr. Paul Leahy from University College Cork for providing me with valuable comments for the present work. I am especially grateful to Dr. Jethro Browell and Ciaran Gilbert from the University of Strathclyde for all their precious advice and the enjoyable conversations. Thank you also to John McCann from the SEAI and Des Farren for all the insightful discussions.

In addition, I would like to thank Dr. Bri-Mathias Hodge and Dr. Zeinab Bedri who acted as the examiners in my viva voce examination. Their comments have been extremely valuable and I greatly appreciate the opportunity to discuss the work presented in this thesis with people of their expertise.

For the provision of data, without which this work would not have been possible, I thank our industry partner (identity covered by NDA) and the ESB for their supply of the wind power and market data.

For the financial support, I gratefully acknowledge the DIT Fiosraigh PhD Scholarship Programme and the Erasmus Mundus Action 2 PANTHER Project.

On a more personal note, I would like to say thank you to all my friends and family. I feel extremely fortunate to have such precious people in my life. In particular, I need to thank all those who have made these years in Dublin so exciting and enjoyable. Thank you to Kevin and Declan, your friendship immediately made Ireland feel like home. Thanks to Eddie, Luke and Sebastian for the countless laughs and stimulating conversations, to my housemates in Clonskeagh, and my teammates in Dublin International Handball Club and CLG Na Fianna.

I would like to say thank you to Niamh for her love and encouragement. She is an extraordinary person who inspires me every day, and without her this would not have been possible.

Finally, I am forever grateful to mamma Lucia for her unconditional love and support, Robi for being my brother and best friend, and babbo Aroldo who is in my heart every day.

Abbreviations

AUC	Area under the ROC curve
BM	Balancing Market
CDF	Cumulative Distribution Function
DAM	Day-Ahead Market
EPF	Electricity Price Forecasting
GB	Great Britain
IEA	International Energy Agency
I-SEM	Integrated Single Electricity Market
ISP	Imbalance Settlement Period (30-minute)
MAE	Mean Absolute Error
MASE	Mean Absolute Scaled Error
MAI	Mean Absolute Energy Imbalance
MLE	Maximum Likelihood Estimation
MOS	Model Output Statistics
NI	Northern Ireland
PDF	Probability Density Function
RMSE	Root Mean Square Error
RMSSE	Root Mean Square Scaled Error
ROI	Republic of Ireland

SNSP	System Non-Synchronous Penetration
TSO	Transmission System Operator

List of Symbols

B	backshift operator
C_{t+h}	balancing costs
$CVaR$	conditional value at risk
d	energy imbalance
δ	probability of positive price difference
$\mathbb{E}[\cdot]$	mathematical expectation
e_{t+h}	forecast error
ε_t	forecast residual
E_{t+h}	energy generated
E_{t+h}^{DAM}	energy bid
E_{max}	wind farm's maximum energy bid
$f(x)$	adjustment function
$g()$	quantile bid function
γ	probability of positive imbalance sign
h	forecast horizon (also look-ahead time)
NIV	net imbalance volume
$\mathbb{P}[\cdot]$	probability
P_{nom}	wind farm's nominal power

π^{DAM}	day-ahead (or spot) price
π^{imb}	imbalance settlement (or balancing) price
π^{diff}	price difference
$\hat{q}^{\alpha}$	quantile forecast with nominal level α
R_{t+h}	revenue
ρ	risk factor
x	market quantity forecast
VaR	value at risk
$\tilde{\cdot}$	random variable
$\hat{\cdot}$	forecast

Contents

List of Figures

List of Tables

In loving memory of Lester Ryan

Chapter 1

Introduction

1.1 Preface

In recent years, wind energy has been playing a leading role in meeting the targets set by the Paris Agreement, aimed at achieving a low-carbon future and fighting global warming [1]. With a total of 650 GW installed as of 2019, wind power is the largest non-hydropower source of renewable electricity by installed capacity in the world [2]. In 2019, wind energy covered almost 5.6% of the world's global electricity production, provided 15% of EU's electricity demand, and was China's third largest energy source [3]. Thanks to its rapidly falling costs per kilowatt-hour, wind energy has become the cheapest source of new power generation across more than two thirds of markets around the world [4]. However, three aspects distinguish wind from traditional generators:

- *variability*, as wind varies in both time and space, with fluctuations occurring over time scales ranging from minutes to more than a year, and variations in space depending on terrain, ground cover and height [5];
- *uncertainty*, due to the limited predictability of the renewable energy source (i.e. wind) and thus, of the power output of wind-based units; and
- *asynchronicity*, as modern wind turbines are interfaced with the grid through power electronics instead of a rotating generator [6], and an increase in wind

generation has the effect of reducing the inertia from synchronous generators.

Wind power's success therefore comes with new challenges. Electricity markets were historically designed for large-scale, dispatchable and centralised fossil power. As the share of renewable but variable and uncertain resources becomes more prevalent in the generation mix, serious issues in terms of security of supply and grid stability arise, driving significant changes in the design and operation of electricity systems and markets. Grid flexibility and transmission systems must be improved, and transmission system operators need to pay increasing attention to operational challenges such as grid imbalances, penetration of non-synchronous sources and curtailment of renewable generation.

It is in this context that wind power forecasting is essential to the renewable energy transition. By reducing the uncertainty of supply, forecasting constitutes a non-generating technology that is able to mitigate the undesirable effects of wind power's variability on system operations. Electricity markets are increasingly incentivising accurate predictions of the generation for the next minutes, hours and days, as their use is considered necessary for the integration of large-scale wind power, as advocated by Holttinen and Hirvonen in [7]. In particular, *short-term* wind power forecasts — where the forecast horizon ranges from 6 to 36 hours [8] — are now commonly employed for unit commitment and energy trading purposes in markets with a significant penetration of wind, as recent work from the International Energy Agency (IEA) Task 36 on Forecasting for Wind Energy has highlighted [9]. Indeed, the value of short-term forecasting at the system level grows as the share of variable renewable energy in power systems increases, as argued by Hodge et al. in [10]. In this study, the authors find that more accurate forecasts improve the system flexibility and lead to larger production cost savings, with the trend becoming more pronounced as the penetration rate increases. Similar findings are obtained by Wang et al. in [11], who also highlight how the economic value of forecasting improvements depends largely on the generation mix of the system. Moreover, wind curtailment is reduced as a result of improved short-term

forecasting, as Wang et al. find in [12] in a related study.

1.2 Background

Ireland[1] has a unique power system, in that it is a small island system with little interconnection and a high penetration of wind energy. The Republic of Ireland (ROI) and Northern Ireland (NI), when compared with the rest of Europe, form a relatively small synchronous system, i.e. a power grid where electricity is generated at a single synchronised frequency of 50Hz. The transmission system shown in Figure 1.1 is linked to European markets through Great Britain (GB) via two 500-MW interconnectors. The project for a new 700-MW interconnector with France, the so-called "Celtic Interconnector", is currently undergoing the planning phase, with the construction phase expected to start in 2022 and to be completed by 2026 [14, 15]. The import dependence on oil and gas is high compared to the European Union (EU) average (see Figure 1.2), resulting in Ireland being a de facto price-taker on these commodities, as noted in the "Energy in Ireland 2019 Report" by the Sustainable Energy Authority of Ireland (SEAI) [16]. Electricity prices to business consumers have been higher than the EU and Euro Area since the second half of 2011 (see Figure 1.3a), with the average price in Ireland in the first half of 2019 being 7.4% and 8.2% above the EU and Euro Area, respectively. To household consumers, on the other hand, prices have mostly been above the EU and close to those in the Euro Area, as shown in Figure 1.3b.

The construction of new wind farms on the island has seen remarkable growth in wind generation capacity over the last two decades, as seen in Figure 1.4, reaching a total installed wind capacity of 5,403 MW in December 2019. In 2019, the total wind generation was 11,958 GWh, accounting for 32.3% of demand and representing the second largest source of electricity generation after natural gas at 50.2%, as reported in Table 1.1. The source of these data is EirGrid's "System and Renewable

[1]When used alone, the term "Ireland" will refer in this manuscript to the all-island of Ireland, a union of the markets of Republic of Ireland and Northern Ireland.

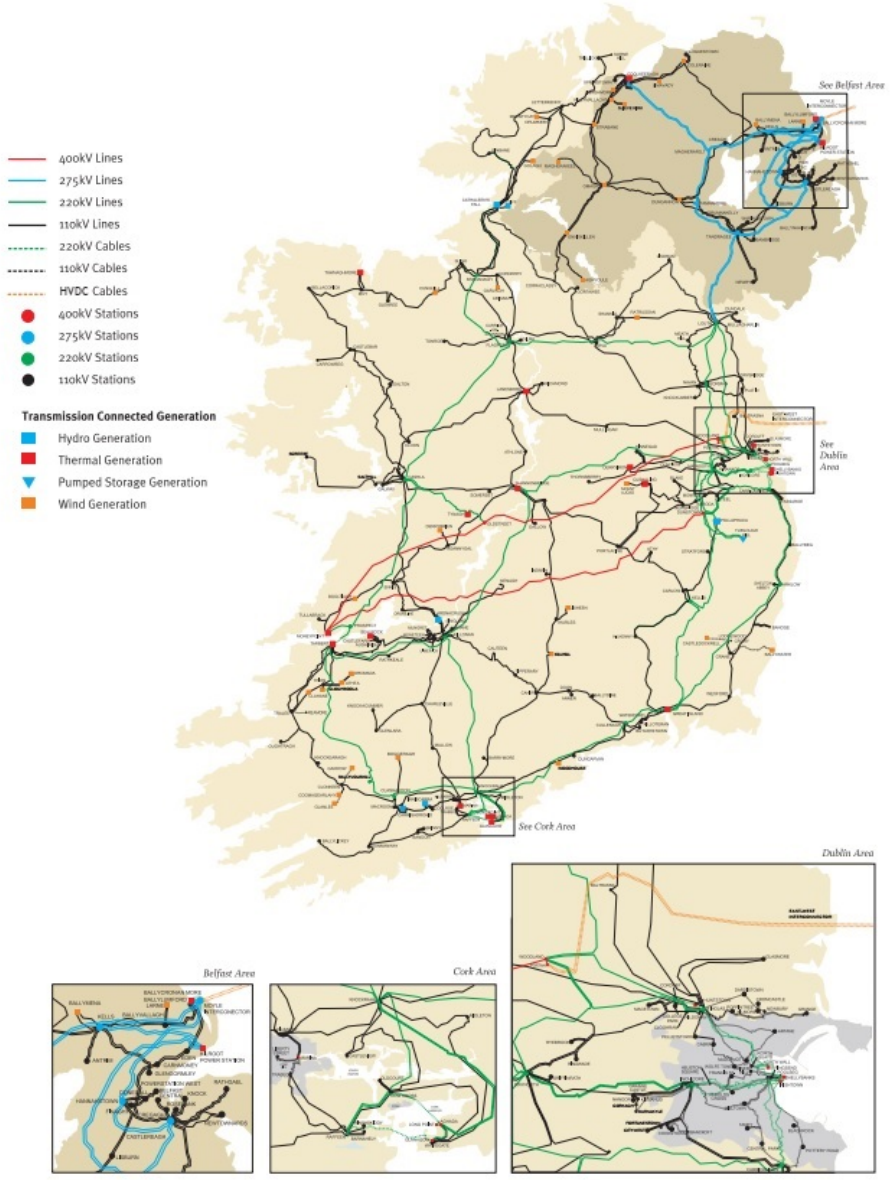


Figure 1.1: All-island of Ireland transmission system. Source: EirGrid [13].

Data Summary Report" for year 2019 [18]. According to the "Renewables 2019 Global Status Report" by REN21 [3], in 2018 Ireland was ranked third in the world by share of electricity generation from variable renewable energy, after Denmark and Uruguay (see Figure 1.5), and the second in the world for wind power capacity per capita after Denmark in 2018.

In 2019, renewables made up 11% of Ireland's total energy consumption, with a target of 16% set for 2020. In order to achieve this, the Irish Government has set a 40% renewable electricity target for 2020 [19], the highest for any synchronous

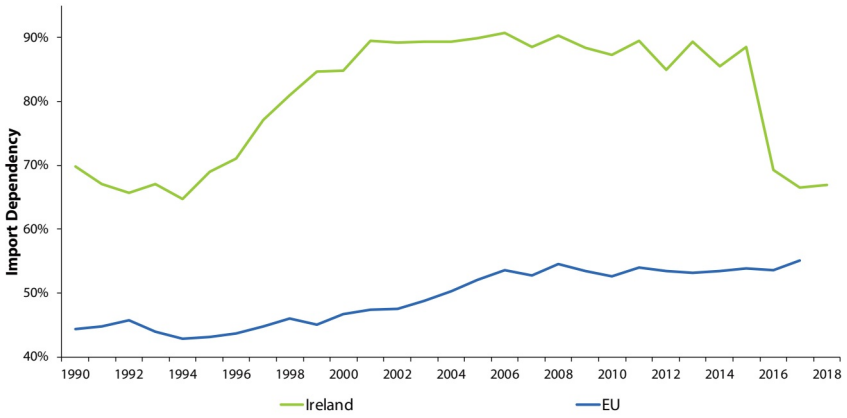
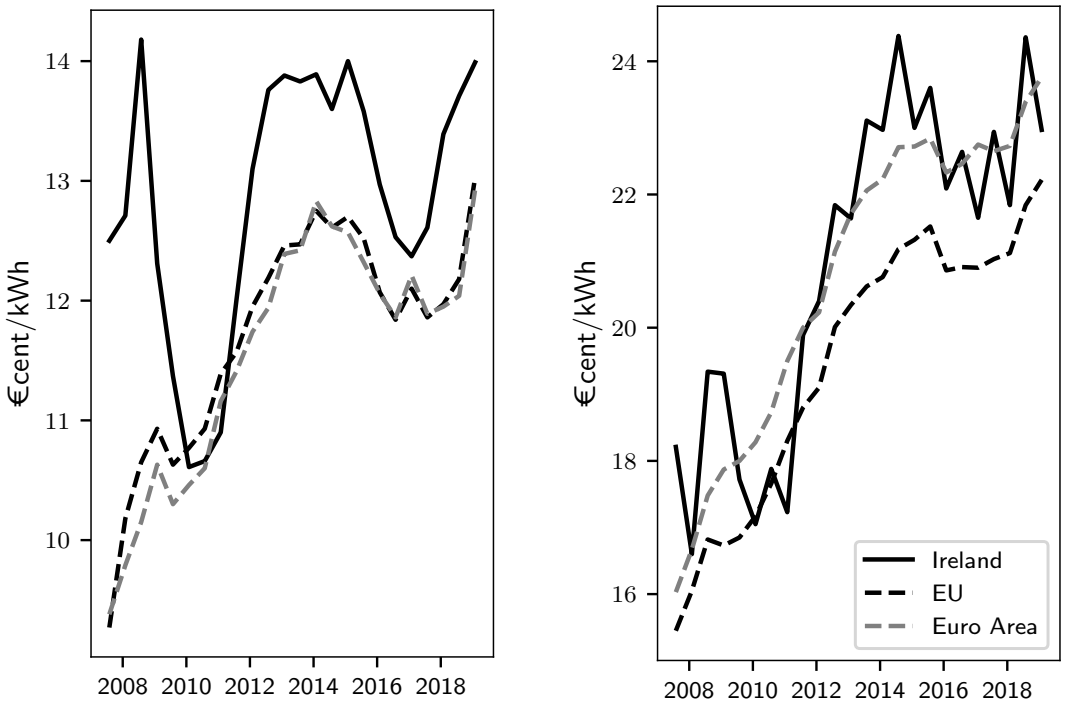


Figure 1.2: Import dependence of Ireland and the EU. Note the sharp drop for Ireland in 2016 following the start of natural gas production from the Corrib gas field. Source: "Energy in Ireland 2019 Report" by the SEAI [16].



(a) Business consumers (ex VAT)

(b) Households consumers (inc VAT)

Figure 1.3: Electricity prices per kWh in Ireland, the EU, and the Euro Area. Source of data: SEAI [17].

system in Europe (see Figure 1.6). To increase the amount of renewable energy in the Irish power system in a secure and reliable manner, the "DS3 Programme – Delivering a Secure, Sustainable Electricity System" has been undertaken [20]. The aim of the programme is to address the factors that influence the System Non-

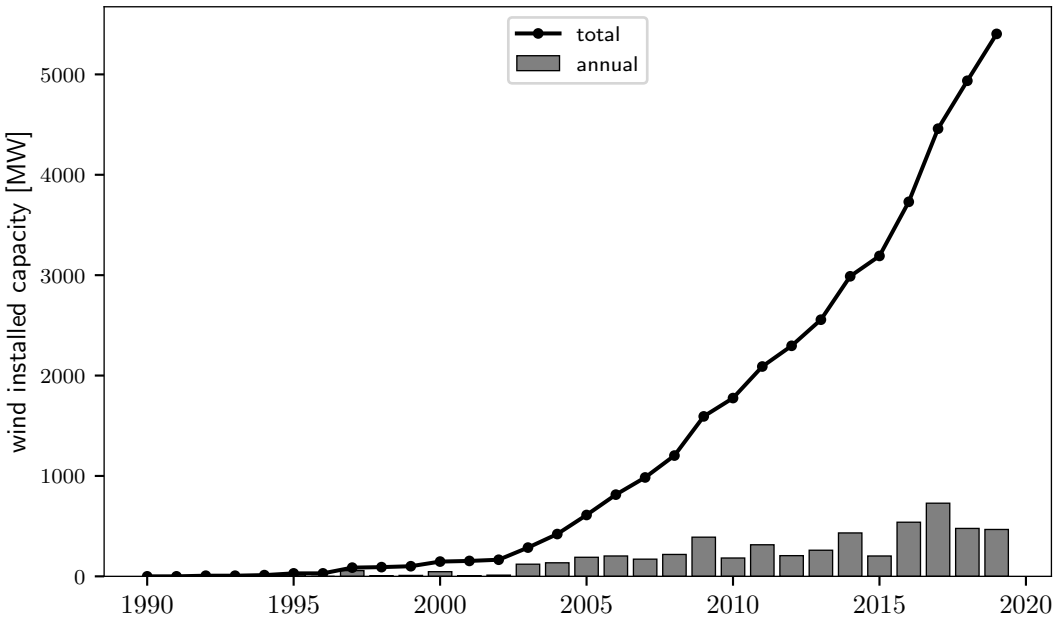


Figure 1.4: Total and annual installed wind generation capacity in Ireland. Source of data: "System and Renewable Data Summary Report" for year 2019 by EirGrid [18].

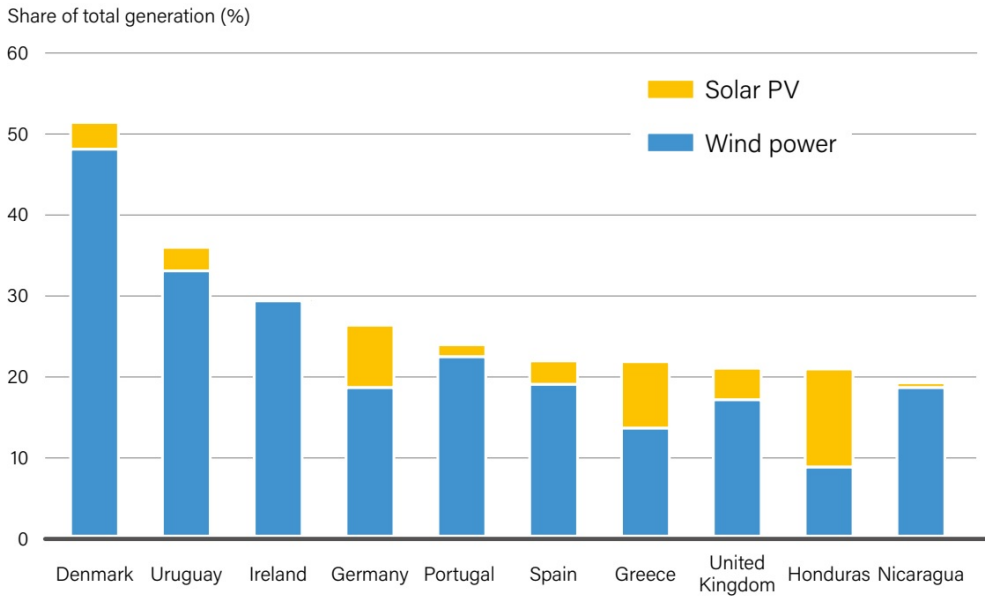


Figure 1.5: Top 10 countries for share of electricity generation from variable renewable energy in 2018. Source: "Renewables 2019 Global Status Report" by REN21 [3].

Synchronous Penetration[2] (SNSP) limit, with the final goal of increasing this limit from 50% to 75%. Since April 2018, the DS3 Programme has enabled the system

[2]This is the real-time measure of the percentage of electricity generation coming from non-synchronous sources relative to the system demand.

Fuel type	%
Gas	50.2
Wind	32.3
Peat	5.6
Coal	3.8
Hydro	2.4
Net Imports	2.1
Other Renewables	1.8
Other Non-Renewables	1.4
Oil	0.4

Table 1.1: Electricity fuel mix in Ireland in 2019 as percentage of demand. Source of data: "System and Renewable Data Summary Report" for year 2019 by EirGrid [18]

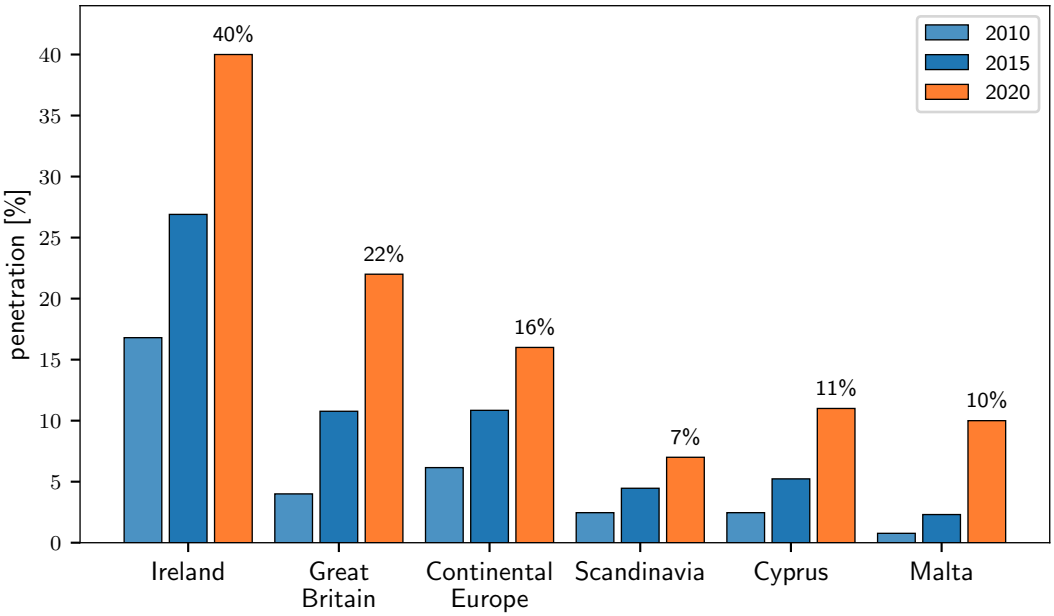


Figure 1.6: Penetration of non-synchronous renewables in each European synchronous system from 2010 to 2020. Source of data: "The National Renewable Energy Action Plans" [19].

operators to handle up to 65% of variable renewable energy on the grid at any given time, making the Irish power system the first in the world to reach this level [21]; this has made it possible to increase the percentage of time where the SNSP is higher than 50%, as seen in Figure 1.7. Nevertheless, part of the wind-generated electricity is still curtailed due to exceeding the SNSP limit, as shown in Figure 1.8.

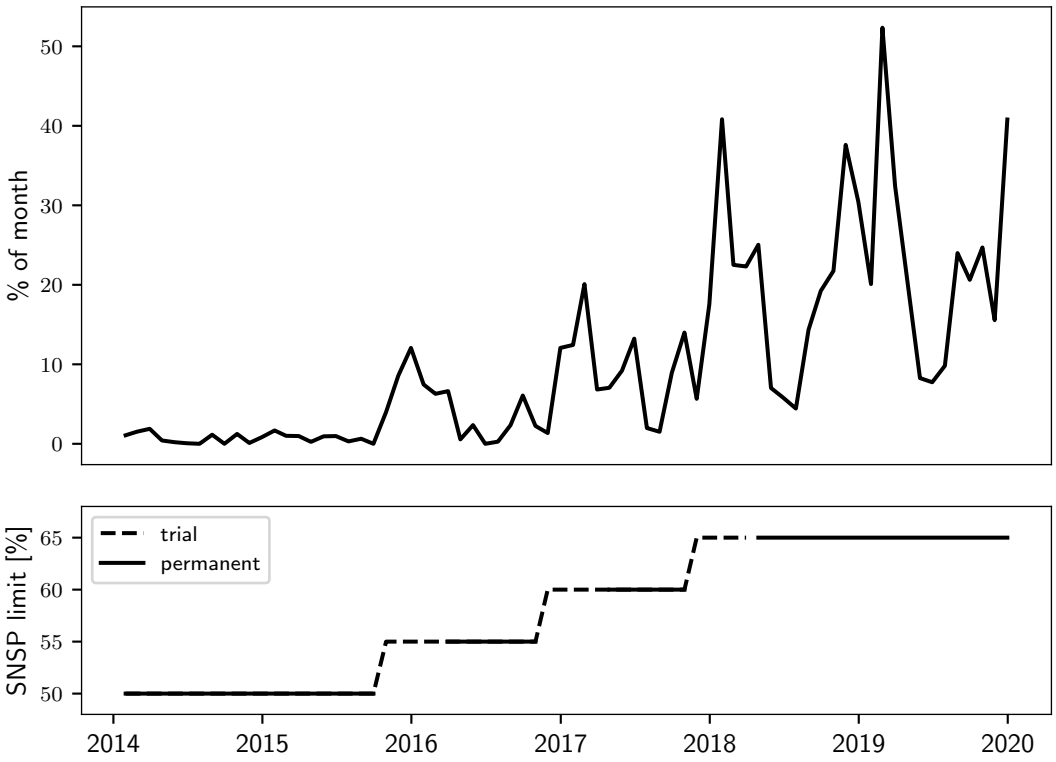


Figure 1.7: Percentage of month with SNSP higher than 50% (top) and SNSP limit (bottom) in Ireland. Source of data: "System and Renewable Data Summary Report" for year 2019 by EirGrid [18].

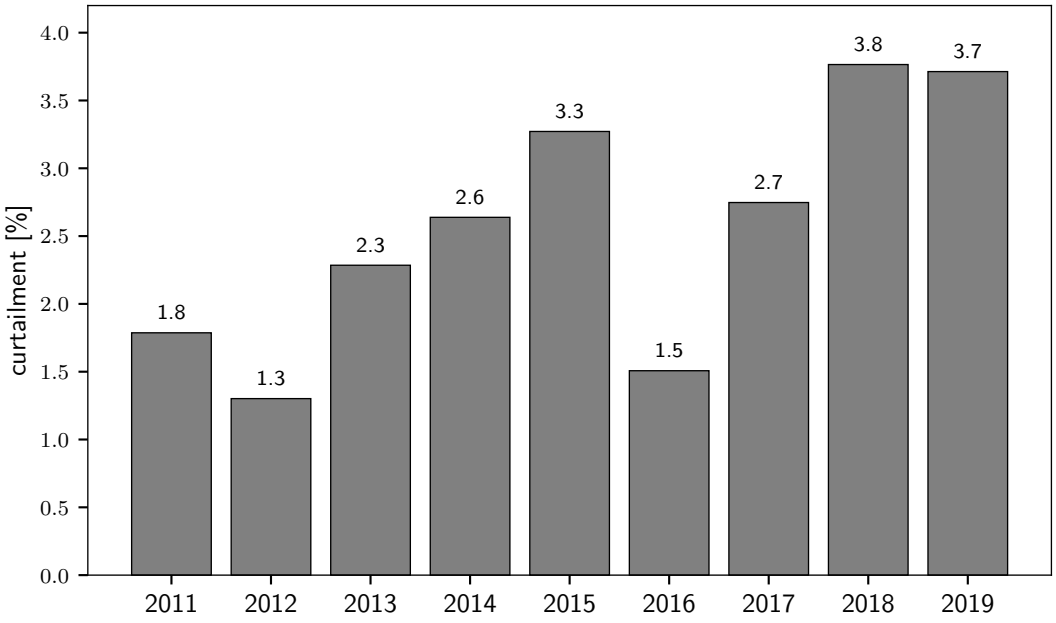


Figure 1.8: Curtailment of wind generation in Ireland. Source of data: "System and Renewable Data Summary Report" for year 2019 by EirGrid [18].

1.3 Motivation

It is in this setting that the new Irish electricity market, the so-called Integrated Single Electricity Market (I-SEM), has come into operation in October 2018. The new market arrangements have been designed to allow the efficient coupling of the all-island electricity market with the wholesale markets across Europe through a single marketplace with common rules and free trade across borders [22]. There are two main objectives: to increase competition in the market in order to apply downward pressure on the cost of electricity to consumers, and to achieve greater levels of security of supply for users [23].

One of the major changes brought in by the new regulations is that *balance responsibility* was introduced for all participants, including wind power generators, to ensure that their notifications of generation or demand best reflect their actual expectations [24]. As participation is at unit level, traders need to establish their position for each wind farm within their portfolio individually. Bids are submitted for every trading period in the day-ahead market and participants are financially responsible for any deviation from their contracted position. Short-term predictions of wind energy generation serve as the basis to formulate these bids. The problem of trading wind energy in the day-ahead and balancing markets will be formulated in mathematical terms in Section 4.1.

The I-SEM is also adopting a single-price imbalance settlement, meaning that deviations are set at the balancing market (or imbalance) price regardless of their direction. This price is higher or lower than the day-ahead (or spot) price if the system is in up- or down-regulation, respectively. This means that participants can increase their profits if the direction of their energy imbalances is opposite to that of the system imbalance, that is, if they help to put the system back into balance. A detailed discussion on price formation mechanisms and deviation penalties will be presented in Section 2.1. It follows that the two variables of primary relevance for a participant who wants to bid strategically are wind power generation and the sign of the difference between day-ahead and balancing market

prices. Directly bidding the point forecasts of wind generation is the most basic approach to trading. Currently, this is common practice among many utilities in Europe and the USA, as the work by Möhrlen et al. in [25] within the frame of the IEA Task 36 has highlighted. However, a more accurate point forecast does not necessarily translate to a higher profit. Indeed, a wind farm operator who bids as closely as possible to the actual production would not obtain the highest possible revenue in a single-price imbalance market. In other words, a wind energy forecast of higher accuracy does not necessarily imply higher profits, and vice versa. For example, in a market where spot prices are on average lower than imbalance prices, a producer would benefit from bidding systematically less than the actual generation, that is, employing a positively biased point forecast. This is demonstrated, for instance, by Bessa et al. in [26], as the results from the case study show that when price penalties are asymmetric, a biased model leads to higher incomes than a neutral model. Moreover, a point prediction of wind energy alone does not provide any information on the uncertainty in future power generation or market conditions. This does not allow the decision-maker to manage the risk of participating in the electricity market. Referring to the types of forecast "goodness" defined by Murphy in [27], rather than a forecast of higher *quality*, wind energy traders are mostly interested in prediction methods that can maximise their benefits; that is, forecasts of higher *value*. However, such advanced offering strategies tailored to the end-user's needs cannot be designed using wind power point forecasts only. For this purpose, it is necessary to move to a probabilistic framework and/or to integrate in the decision-making process information on future market conditions (cf. Pinson et al. in [28]).

The participation of wind energy in electricity markets and strategic bidding in the day-ahead market has been investigated with growing interest in recent years, with studies considering various power systems and prediction methodologies. However, the majority of the published works concentrates on markets adopting a dual-price imbalance settlement, while single-price markets have received very

little attention in the academic literature. The developments in the present Thesis contribute to filling this gap by advancing the discussion on the subject.

1.4 Aim of Research

The aim of this research is to develop new probabilistic models forecasting the short-term value of wind power and investigating their use in bidding in electricity markets adopting a single-imbalance pricing scheme. The perspective adopted here is that of a participant who wants to bid strategically in the day-ahead market to increase the value of the energy generated at a wind farm. Value will be measured in terms of revenue and exposure to risk. The proposed models combine short-term predictions of wind power and of relevant electricity market quantities in a probabilistic framework. Offers are then formulated to reflect the participant's appraisal of future uncertainties in wind power generation and electricity market conditions. The objective of these bidding strategies is to improve the value of wind power for the participant, increasing the revenue while controlling risk exposure. In order to implement these strategies, together with forecasts of the wind farm's power generation, predictive models need to be developed to forecast the electricity market quantities of interest. In turn, this entails the detailed analysis of the market data from the power system under study.

1.5 Contributions to Knowledge

The contribution to knowledge of this research is three-fold. Firstly, novel probabilistic value forecasting models are developed for wind farms operating in electricity markets with single-imbalance price settlement. The bidding strategies enable the participant to improve the value of their energy generation and are designed for wind farms whose size is small relative to the power system . Indeed, value is improved as a result of larger energy imbalances and this is in conflict with the objective of the system operator, who would prefer imbalances to be minimised.

The strategies are novel in that they use probabilistic forecasts of both wind power and market quantities to define the bid while explicitly integrating the participant's risk profile. The methodology is implemented using wind power and market data in Ireland as an example. However, the work is relevant to other markets with similar trading arrangements in countries such as the USA, Great Britain, Germany, and The Netherlands.

Secondly, this work advances the discussion on short-term forecasts of electricity market quantities by proposing a new method to predict the probability of the sign of the difference between spot and imbalance prices. The proposed two-stage methodology employs dynamic regression models and Monte Carlo simulations, and is successfully applied to a new electricity market where the availability of historical data to train the models is extremely limited.

Finally, a detailed description and analysis of market data from the new Irish electricity market are provided. The market commenced trading in October 2018 in a power system characterised by many unique features. The exploratory and statistical analysis performed on the available data give valuable insights into a number of processes for modelling and forecasting purposes, including: day-ahead and balancing prices, direction of system imbalances, the occurrence of anomalous imbalance prices, and the impact of predicted wind power penetration on various market quantities. Such a systematic study on the new Irish market data has not been undertaken previously.

1.6 Thesis Layout

In order to define the general context of the Thesis, wholesale electricity markets and their functioning are described in Chapter 2. Particular attention is dedicated to the new Irish electricity market, with a detailed description and analysis of market data from the first months of its operation. The characteristics of the data and the analysis outcomes provide the motivation for the modelling approaches adopted within the methodology of this work.

Chapter 3 gives a summary of the current methods used for time series forecasting and forecast evaluation, followed by a review of the academic literature of relevance for this work. The state-of-the-art in short-term forecasting of wind power, electricity prices and value of wind energy is presented. Also, the motivation and aim of this research are justified by highlighting the gaps in the literature.

The aim of Chapter 4 is to describe the methodology used in developing the value forecasting models. The problem of trading wind energy in day-ahead and single-price balancing markets is detailed initially, followed by a description of the wind power forecasts available in the study. Then, the two alternative approaches developed to forecast market quantities are explained. Finally, the various bidding strategies are formulated.

The methodology is applied to a case study where the participation of a real wind farm in the Irish electricity market is simulated and the results are presented and discussed in Chapter 5. The performance of the wind power and market quantity forecasts is evaluated and benefits and limitations deriving from the application of the proposed strategies are evaluated.

The overall conclusions from the present work are summarised in Chapter 6 together with perspectives for further research.

Chapter 2

Electricity Markets

This Chapter provides a description of electricity markets and their functioning, paying particular attention to the operation of the Integrated Single Electricity Market (I-SEM), the Irish wholesale electricity market. Firstly, the general features and mechanisms in electricity markets are discussed in Section 2.1, with the focus on the main energy trading platforms. Then, the specific case of Ireland is analysed in Section 2.2: the market's structure is outlined in Section 2.2.1, while the outcomes from the first seven months of its operation are described and analysed in detail in Section 2.2.2.

2.1 Market Organisation

Electricity markets have developed various structures and arrangements across different countries and regions [29–31], but it is still possible to identify some common rules and features that characterise them. The models are classified by increasing levels of competition in: monopoly, purchasing agency, wholesale competition, and retail competition [32]. The focus here will be on wholesale electricity markets, currently the most diffused model in industrialised countries [33]. Here, the activities of generation, transmission, distribution, and retail are separated. Generation and retail are performed in a competitive environment, while the operation of the spot market and of the transmission network remain centralised. The following

entities can be identified:

- *Regulator*: the independent governmental body responsible for the market design and its specific rules; it also carries out monitoring activities to identify misbehaviours in the market.

- *Market Operator* (MO): organises and operates the market place, matching bids and offers submitted by buyers and sellers, and taking care of the settlement of the accepted bids and offers.

- *System Operator* (SO): is responsible for system security by ensuring that the system is always in balance. In Europe, this takes the form of the *Transmission System Operator* (TSO), who also owns and operates the network. In the USA, it can take the form of the *Independent System Operator* (ISO), where SO and MO are within the same body, or of the *Regional Transmission Organisation* (RTO), for non-market locations which usually encompass multiple states.

- *Distribution System Operator* (DSO), or distribution companies: owns and operates the distribution grid and often acts as a retailer as well.

- *Retailers*: buy electricity in bulk on the wholesale market in order to sell it to consumers.

- *Generating Companies*: own the production assets (from a single generator to a portfolio of plants of various technologies) and sell the electrical energy generated from them; they can also sell ancillary services like regulation, voltage control and reserve.

- *Consumers*: buy the electricity from a retailer (or directly from the wholesale market, in the case of large consumers) and use it for any purpose.

An electricity market generally comprises various markets, which can be grouped as:

- *Capacity* market, a mechanism developed by the SO to ensure that the generation capacity in the system is sufficient to meet the demand in future years.

- *Energy* markets, which constitute the main place for the optimal scheduling and settlement of energy exchanges; here, participants pay or get paid in proportion to the actual amount of energy drawn from or injected into the grid.

- *Ancillary services*, which include all the services purchased directly by the SO to maintain the integrity and stability of the electricity supply and the power quality, such as frequency control, voltage control, and system restart services.

Depending on their timeline, energy markets are divided into:

- *Futures*, where financial instruments (future contracts) are traded with time horizons up to a few years with the aim of reducing risk and hedging against volatile prices.

- *Day-ahead*, the central, highly liquid energy trading platform for scheduling and matching of electricity bids and offers; the auction submission gate usually closes at midday of the day before delivery.

- *Intraday*, a continuous trading platform between the closure of the day-ahead market and real time allowing traders to adjust their positions following changes in market conditions; for example, an unscheduled plant outage or an updated forecast of wind power generation.

- *Balancing*, a managed spot market for the SO to balance load and generation close to real time.

The last three markets are the most relevant for wind power trading and therefore will be presented in more detail in the next Sections. The focus will be on the European type of power exchanges.

2.1.1 Day-ahead Market

The day-ahead market (DAM) is an electricity pool based on a two-sided auction mechanism which provides a centralised form of system management. This is also referred to as *spot market* in Europe and as *forward market* in the USA. Trading always refers to a certain amount of megawatt-hours to be delivered over a fixed period of time, also referred to as *trading period*. The length of such period depends on the specific market but is usually set at one hour, 30 minutes or 15 minutes. For each trading period, all bids and offers are placed at the same time and participants are unaware of others' behaviour. Note that in this text, a generation (or supply) offer is referred to as a *bid* and a consumption offer as an *offer*. Bids and offers consist of a set of quantity-price pairs, specifying the amount of energy to be sold or bought and the corresponding price. Bids submitted by generating companies to supply a certain volume of electricity at a certain price are ranked in order of increasing price, following the *merit order* principle; this way, the market *supply curve* is constructed. On the other hand, offers from consumers are ranked in order of decreasing price forming the *demand curve*; at times, this passage is omitted and the value of demand is determined directly through a forecast. If constraints on the transmission grid are not considered, the intersection between these two curves represents the market equilibrium and determines the scheduled generation and consumption levels and the market clearing price, or *day-ahead price* (π^{DAM}), as illustrated in Figure 2.1. This is also referred to as *spot price* or as *system marginal price* (SMP), since it represents the price of supplying one additional megawatt-hour of energy. However, if constraints on the lines are considered, a different price is associated with each node in the grid and is called the *locational marginal price*. This *nodal* price rationale is adopted in most of the markets in the USA, as opposed to the *zonal* pricing rationale which is popular in Europe. It is worth mentioning at this stage that in many markets, participants are allowed to submit bids with negative prices. This, for example, can happen when the cost of shutting down and ramping back up a unit is higher than the loss from receiving a negative price for

the generation. When the demand is very low or the production from renewable sources is very high (or both), negatively priced bids can potentially lead to the formation of negative clearing prices.

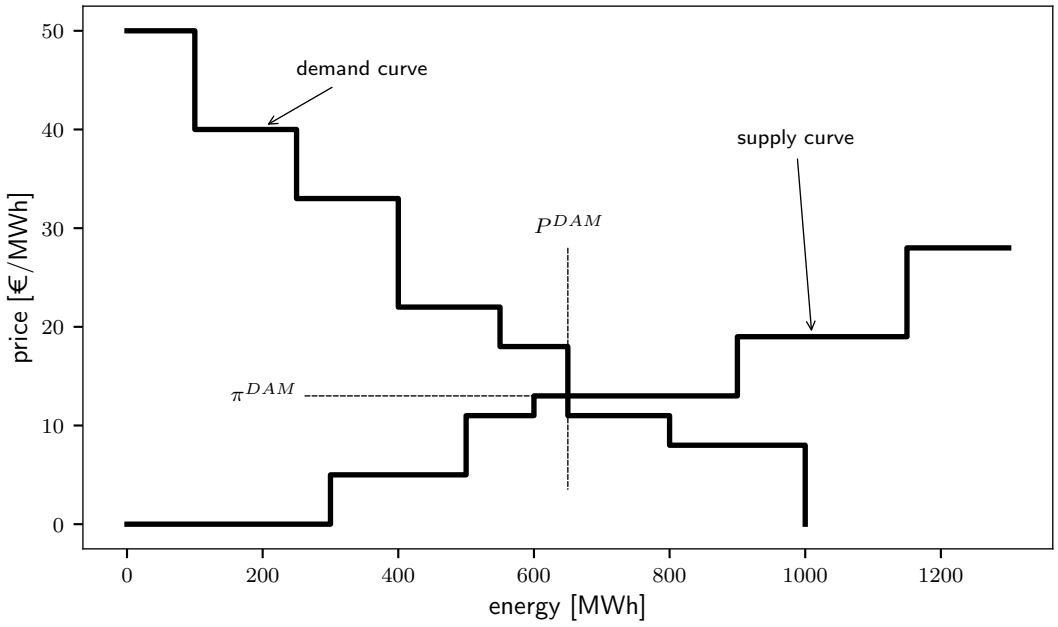


Figure 2.1: Day-ahead market clearing where the equilibrium is found at level P^{DAM} and day-ahead price π^{DAM}.

The clearing of the market is usually determined through the running of a centralised algorithm that determines which bids and offers are retained to calculate the SMP. Once this clearing process is complete, the SO is informed about the trades that occurred. In European systems, participants can usually self-schedule their units after clearing according to their portfolio-based position. In the USA, on the other hand, schedules are financially binding at the unit level.

After the market clearing, the *settlement* process occurs, that is, establishing who should pay what, and who should get paid and what amount. The two settlement options are:

- *pay-as-bid*, where everyone gets paid or pays the price they submitted before market clearing; or
- *uniform pricing*, where generators are paid the SMP for every MWh they produce and consumers pay the SMP for every MWh they consume, irrespective

of the bids and offers that they had submitted.

The latter is generally preferred and is currently more diffused in the markets [34].

As mentioned earlier in this Chapter, the DAM gate closure occurs several hours before delivery. Therefore, deviations will often arise between the amount of energy contracted and the amount that is actually needed or that can be produced. These energy deviations are called *imbalances* and can be handled by each market participant in different ways:

- re-dispatch of own units, by compensating with other generation/consumption means within their portfolio;
- adjust their position in the intraday market through agreements with other players; or
- let the SO put the system back to balance in the balancing market.

If, for example, a generator fails to produce the amount of energy that it has contracted to sell, the TSO — who has the ultimate responsibility to keep its transmission system in balance — buys replacement energy on the balancing market. Similarly, if a large user or retailer consumes less than it has bought, the TSO sells the excess on the balancing market. These balancing activities carry a cost and the parties that are responsible for the imbalances should pay this cost. Normally, market participants are *balance responsible*, meaning that they assume the financial responsibility for keeping their own position balanced over a given time frame.

2.1.2 Intraday Market

The *intraday* or *adjustment market* gives participants the possibility to modify their position between day-ahead market gate closure and delivery. The market is still centrally organized but is based on bilateral contracts. These are contracts for a direct exchange of power between a buyer and a seller (who may both be producers and/or consumers) where the system operator is informed about the trades that occurred. The types of bilateral trading are:

- customised long-term contracts, very flexible contracts with typically large transaction costs and large amounts of energy traded over long periods of time;

- over-the-counter trading, involving standard contracts with lower transaction costs, smaller amounts of energy and shorter lead times; and

- electronic trading, fast and cheap trading mechanism based on a computerised platform that consistently matches supply and offer bids.

The need for corrective actions may vary significantly depending on how new information is disclosed between day-ahead market clearing and actual operation. It may be difficult to foresee the actual imbalance that would eventually need to be fixed and decision-making in such adjustment markets can be very complex. A practical consequence is that, in general, traded volumes and liquidity in intraday markets are low — meaning that participants do not find counter-parties to trade with easily — as shown by Weber in [35] for the main European markets.

2.1.3 Balancing Market

The balancing market (BM) (also referred to as "real-time imbalance market" or "regulation market") takes place shortly before energy delivery (generally one hour or 30 minutes before, depending on the market) to ensure that energy supply meets demand at all times. Its function is to match the residual load and generation by adjusting the production of flexible generators and curtailing the demand of willing consumers. The energy adjustment needed is called *net imbalance volume* (NIV). There are three possible imbalance situations for both the system as a whole and for a single supply (or demand) participant. For the system, there can be:

- $NIV < 0$ when there is a power surplus, as in the overall supply is larger than the overall demand. In this case, the system is said to be *long* and there is a need for downward regulation.

- $NIV > 0$ when there is a power deficit, as in the overall demand is larger than the overall supply. In this case, the system is said to be *short* and there is a need for upward regulation.
- $NIV = 0$ when there is no imbalance, i.e. supply matches demand. Then, there is no need for regulation.

Similarly, for a single generator there can be three imbalance cases:

- the actual generation is larger than the contracted quantity;
- the actual generation is smaller than the contracted quantity; and
- actual and contracted generation match.

Let us assume that from the previously cleared day-ahead market, the equilibrium for a certain trading period is found at generation and consumption level P^{DAM} and day-ahead price π^{DAM}. When reaching the balancing market (i.e. real time), there are N_B dispatchable generators, called *balancing generators*, able to move both up ($\uparrow$) and down ($\downarrow$). They offer:

- upward regulation $P_j^{\uparrow}$ at price $\pi_j^{\uparrow}$, for $j = 1, \ldots, N_B$; and
- downward regulation $P_j^{\downarrow}$ at price $\pi_j^{\downarrow}$, for $j = 1, \ldots, N_B$.

where one necessarily has:

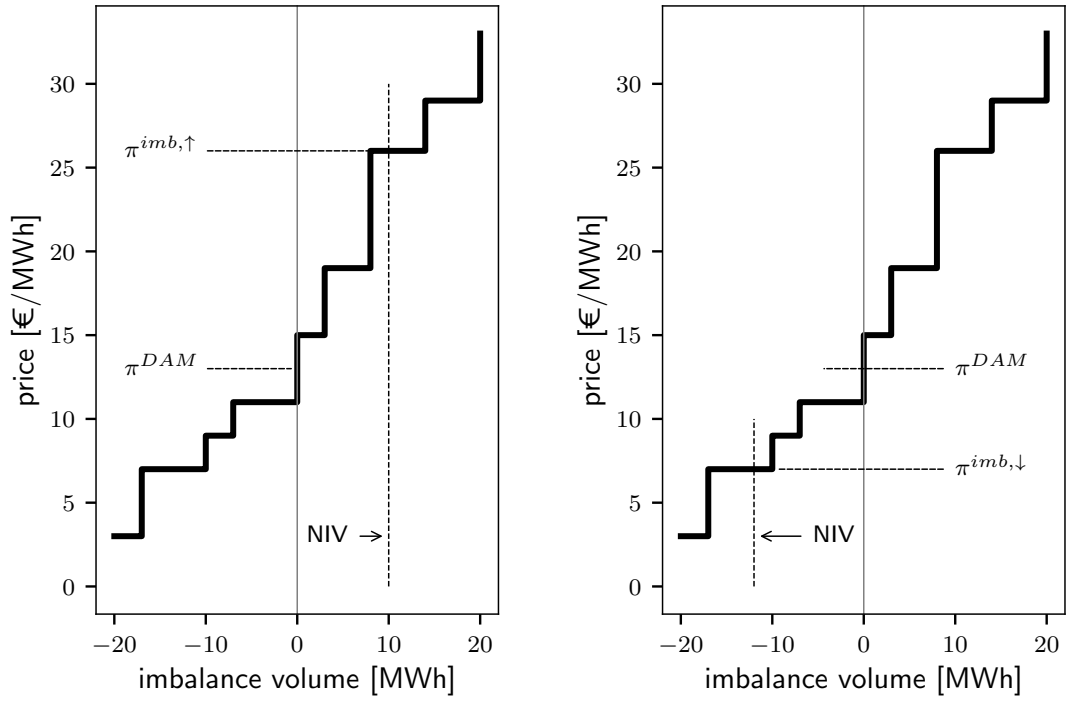
- $\pi_j^{\uparrow} > \pi^{DAM}$ for $j = 1, \ldots, N_B$; and
- $\pi_j^{\downarrow} < \pi^{DAM}$ for $j = 1, \ldots, N_B$.

The market is cleared through an auction in a similar way to the day-ahead market (cf. Figure 2.1), except that:

- offers are for adjustments (both upward and downward) from the day-ahead quantity P^{DAM}; and
- demand is here seen as inelastic (that is, the demand curve is near-vertical).

The clearing of the auction determines the NIV and the balancing market marginal price, also called *imbalance price* (π^{imb}). As shown in Figure 2.2, it will result that:

- when the system is long, $NIV < 0$ and $\pi^{imb} < \pi^{DAM}$; and
- when the system is short, $NIV > 0$ and $\pi^{imb} > \pi^{DAM}$.



(a) Positive imbalance volume. **(b)** Negative imbalance volume.

Figure 2.2: Balancing Market clearing: imbalance prices when the system is short (left) and long (right), corresponding to upward and downward regulation, respectively.

While balancing generators simply sell or buy at π^{imb}, there are different consequences on settlement for stochastic generators (such as wind power plants) dispatched through the DAM, depending on whether a one-price or two-price imbalance settlement is adopted in the specific system.

One-price imbalance settlement

In the *one-price* (or single-price) imbalance settlement, the balancing price π^{imb} is used directly for the settlement. For a stochastic generator j, let us indicate with y_j^{DAM} the quantity contracted in the DAM and with y_j the quantity actually

produced. Depending on the direction of the system imbalance, the following scenarios are possible:

A) $NIV > 0$:

1. generator j producing less than scheduled must buy $y_j^{DAM} - y_j$ at price π^{imb};
2. generator j producing more than scheduled must sell $y_j - y_j^{DAM}$ at price π^{imb}.

B) $NIV < 0$:

1. generator j producing less than scheduled must buy $y_j^{DAM} - y_j$ at price π^{imb};
2. generator j producing more than scheduled must sell $y_j - y_j^{DAM}$ at price π^{imb}.

Meanwhile, in-merit balancing generators simply sell or buy at π^{imb} regardless of their offer prices $\pi_j^{\uparrow}$ and $\pi_j^{\downarrow}$ for upward and downward regulation, respectively. The total payment/revenue of day-ahead market participants for deviations from schedule equals the revenue/payment of the balancing generators.

The rationale behind one-price balancing markets implies that if one's deviation contributes to setting the system off-balance, this leads to a revenue loss; this is the case for scenarios A.1 and B.2 described above. On the contrary, if one's deviation is helping the system go back to balance — regardless if such deviations are wanted or not — this leads to additional profit; this is the case for scenarios A.2 and B.1. In fact, if the system has a shortage of production, generators producing more power than contracted in the DAM (and loads consuming less than contracted) receive a price higher than the DAM price for their sales in the BM. Conversely, when the system has an excess of production, generators that produce less than contracted (and loads consuming more than contracted) pay a balancing price that is lower than the DAM price.

Two-price imbalance settlement

In the *two-price* (or dual-price) imbalance settlement, there are different consequences on settlement for participants dispatched through the DAM.

A) $NIV > 0$:

1. generator j producing less than scheduled must buy $y_j^{DAM} - y_j$ at price π^{imb};
2. generator j producing more than scheduled must sell $y_j - y_j^{DAM}$ at price π^{DAM}.

B) $NIV < 0$:

1. generator j producing less than scheduled must buy $y_j^{DAM} - y_j$ at price π^{DAM};
2. generator j producing more than scheduled must sell $y_j - y_j^{DAM}$ at price π^{imb}.

Meanwhile, balancing generators still simply sell or buy at price π^{imb}.

Under the two-price imbalance settlement, those putting the system off-balance are to be penalised (cases A.1 and B.2). However, those supporting the system unintentionally — that is, stochastic producers like wind units — do not get extra rewards (cases A.2 and B.1); only wanted deviations from dispatchable producers opposite in sign from the system imbalance are rewarded financially with an imbalance price that is more favourable than the day-ahead one.

2.2 I-SEM, the Irish Electricity Market

2.2.1 Market Structure

The Integrated Single Electricity Market (I-SEM) is the all-island wholesale electricity market for Ireland that came into operation on the 1st October 2018 to integrate the Irish electricity market with the European ones. Recall that in this

text, the term "Ireland" is used to refer to the all island of Ireland, a union of the markets of Republic of Ireland (ROI) and Northern Ireland (NI). The SEM Committee is the governing body for the I-SEM, while the regulatory authorities are the Commission for Regulation of Utilities (CRU) for ROI and the Utility Regulator for NI. The distribution system operator is ESB Networks and the transmission system operators (TSOs) are EirGrid for ROI and SONI for NI. EirGrid and SONI are also the Nominated Electricity Market Operators: they operate as the Single Electricity Market Operator Power Exchange (SEMOpx) for the Day-ahead and Intraday markets, and as the Single Electricity Market Operator (SEMO) for the Balancing market. The synchronous transmission system is linked to European markets through Great Britain via two interconnectors:

- the East-West Interconnector (EWIC), a high-voltage direct current (HVDC) submarine and subsoil power cable linking Republic of Ireland to Wales, with a capacity of 500 MW and owned by EirGrid Interconnector DAC; and
- the Moyle Interconnector, an HVDC submarine and subsoil cable linking Northern Ireland and Scotland, with a total capacity of 500 MW and owned by Mutual Energy.

The I-SEM comprises:

- a Capacity Market, where capacity is traded up to five years in advance of the trading day;
- two markets for financial instruments, the Forward Market and the Financial Transmission Right auction, where trading occurs from over a year to one month ahead of the trading day;
- two physical *ex-ante* (or spot) markets for energy trading, the Day-Ahead Market (DAM) and the Intraday Market (IDM);
- a Balancing Market (BM), where energy and non-energy services (e.g. voltage control, energy reserves) are offered for system balancing before and into real

time.

Participation in the DAM occurs through coupling with the European market; participants can then refine their position in the IDM before entering the BM, where participation is mandatory for most generators. Typically, generators and suppliers will establish a physical position in the non-mandatory ex-ante markets to reduce their exposure in the BM. While participants can own a portfolio of resources, participation in the I-SEM is at unit level. A participant unit's net energy position is the accumulated volume of all its trades in the physical markets and any energy balancing action taken by the TSO in the balancing market. All physical trades in the ex-ante markets are *firm*, meaning the participant is financially exposed in the balancing market if it cannot adhere to its commitments. Wind power generators are not allowed to trade quantities above their firm access, i.e. the wind farm's installed capacity.

Generator units are classified as either dispatchable, non-dispatchable but controllable, or non-dispatchable and non-controllable. Wind power units are considered non-dispatchable but controllable, in that they can limit their output to set-point instructions issued by the TSO; they are also given priority dispatch, whereby the TSO is obliged to take energy from these units ahead of other generators, conditional on system security considerations. At times, in fact, the TSO can require reducing the output of controllable wind power units to maintain system security. If the reduction is required for a system-wide reason, for example if the SNSP limit is reached, this is referred to as *curtailment*. If the reduction is required to manage a local issue, for example a grid constraint or network congestion, this is referred to as a *constraint*.

The trading day (D) in the I-SEM is from 23:00 local time the day before (D–1) to 23:00 local time on the day (D), which corresponds to midnight in Central European Time (CET). Note that daylight saving is observed in these regions, so 23:00 local time corresponds to 23:00+00:00 UTC in the winter (GMT), but to 22:00+00:00 UTC in the summer (IST). Where not specified otherwise, time in

this work is local and expressed in 24-hour format.

2.2.1.1 Ex-ante Markets

The Day-Ahead Market (DAM) is a highly liquid, pan-European energy trading platform which constitutes the core of European market integration. Bids and offers and interconnector flows across participating regions of Europe are scheduled through a single centralised price-coupling algorithm (EUPHEMIA). In Ireland, participants submit their bids and offers to SEMOpx through the EPEX Spot trading system. In turn, SEMOpx interacts with the Market Coupling Operator, who runs the EUPHEMIA price-coupling algorithm.

Trades from one bidding zone to another are only restricted by cross-border capacity. In theory, with market coupling, there will be a single price as long as energy can flow freely, where as when the network is congested prices diverge. Physical flows on Moyle and EWIC Interconnectors are linked to the day-ahead markets in Ireland and in Great Britain (GB) and the price difference between the two. Where the DAM price is higher in the I-SEM than in GB, the interconnectors will import power into the I-SEM. Where on the contrary the I-SEM price is lower, the interconnectors will export power to GB. In general, if there are high levels of wind in Ireland, then interconnectors flow out of I-SEM; if the level of wind is low, on the contrary, interconnectors flow into I-SEM.

Participants can submit orders in the DAM for each of the 24 (one-hour) trading periods in trading day D. The submission window opens 19 days before D and closes at 11:00 on D–1. The market is then cleared and schedules are published at 13:00 on D–1. The market is settled daily by SEMOpx and for each hourly trading period k, all transactions are settled at the DAM marginal price, or *day-ahead price*, π_k^{DAM}.

The long-term model for a single European intraday trading platform is based on the continuous trading across interconnectors known as Cross Border Intraday (XBID). However, as an interim solution, intraday trading is continuous only within Ireland at the moment. Three cross-border intraday auctions are also run allowing

trades between I-SEM and BETTA, the bidding zone of GB.

The Intraday Market (IDM) trading day is divided into 48 (30-minute) trading periods. The submission window for the continuous trading opens at 11:45 on D–1 and closes one hour before real time (t–1). The submission window for cross-border trades opens on D–19 and closes:

- for auction 1, at 15:30 D–1 for all 48 trading periods on the trading day;
- for auction 2, at 08:00 on day D for the 24 periods from 11:00 to 23:00 D; and
- for auction 3, at 14:00 on day D for the 12 periods from 17:00 to 23:00 D.

2.2.1.2 Balancing Market

The Balancing Market (BM) reflects the actions taken by the TSO to keep the system balanced and secure. The market determines the imbalance settlement price to settle the TSO's balancing actions and any uninstructed deviations from a participant's contracted ex-ante position. The TSOs are responsible for the safe, secure and reliable operation of the power system and have the obligation to maximise priority dispatch generators while minimising the cost of deviation from physical notifications of participants (i.e. the expected output of generating units that reflects the participant's best estimate of its intended level of generation). SEMO is responsible for the settlement, billing, and credit risk management.

The BM trading day is divided into 48 (30-minute) imbalance settlement periods (ISP). Within each ISP, there are six (5-minute) imbalance pricing period. Note that the BM imbalance settlement period and IDM trading period are aligned and will be referred to simply as *trading period* in the remainder of this text. The submission window opens on D–19 and closes one hour before the start of each 30-minute ISP (t–1). The timing of the three markets in the Irish system discussed above is shown in Figure 2.3, where duration, trading period resolution and gate closing times of each market are also illustrated.

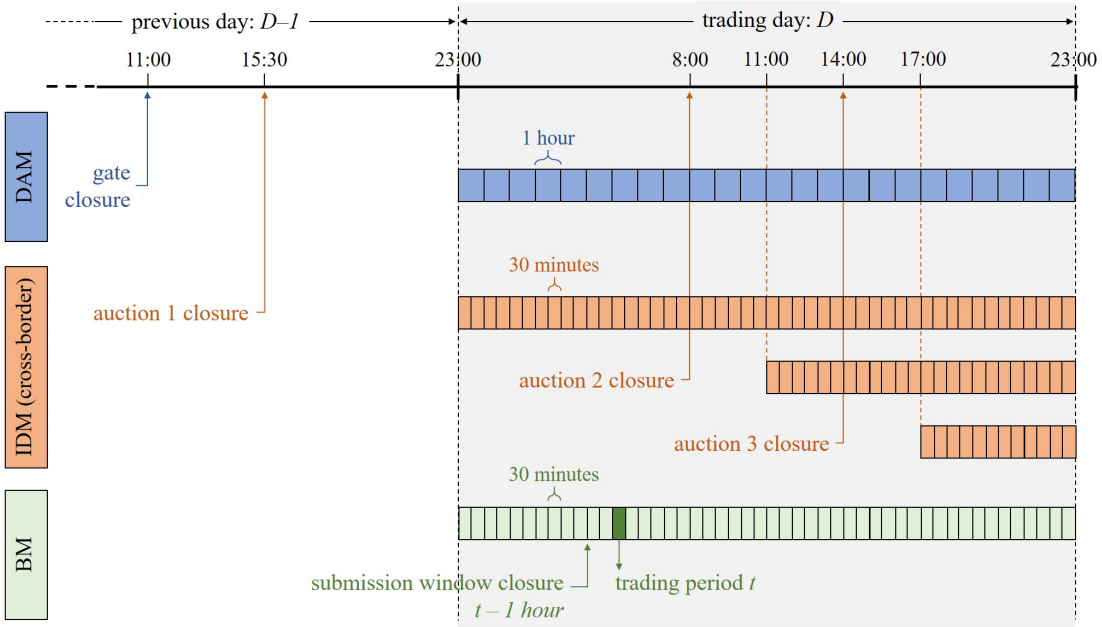


Figure 2.3: Timeline of day-ahead, intraday, and balancing markets in the Irish system.

Although not used directly in settlement of the BM, the imbalance prices for each 5-minute imbalance pricing period are used to calculate the *imbalance settlement price*, or *balancing price*, π^{imb} for each 30-minute trading period. A rules-based, *flagging-and-tagging* process is used to determine the initial imbalance price in each 5-minute period. The process prevents bids and offers that were scheduled due to system constraints or where units were operating at a unit constraint from influencing the imbalance price. The imbalance settlement price for each 30-minute trading period is then found as the average of the six corresponding imbalance pricing period prices. A detailed description of the imbalance pricing mechanisms is given in Appendix A.

The BM is settled weekly based on a single-price approach. Any imbalance which is not due to a balancing action (i.e. uninstructed) is settled at the imbalance settlement price, while any imbalance due to a balancing action (i.e. instructed) is settled at the better of the imbalance settlement price and the bid/offer price. For wind power generators, if the unit is constrained due to local issues, it is compensated based on the market's decremental curve. However, if it is curtailed due to system-wide reasons, the unit is compensated at a curtailment price determined

for each unit based on their undelivered day-ahead market commitments. Further information on curtailment prices and payments can be found in Section E and F of the I-SEM Trading and Settlement Code [36].

2.2.2 Market Operation

This Section provides a description and detailed analysis of I-SEM market data from the first seven months of its operation, starting on the 1st October 2018 until the 25th April 2019, to give a comprehensive picture of the electricity market data available at the time this work was conducted.

The value forecasting models that will be developed in this thesis require the modelling and forecasting of a number of market processes, namely the formation of day-ahead prices, imbalance prices, and system imbalance direction. Therefore, a wide range of exploratory analyses were performed on the available market data to gain insight into these processes and improve the modelling quality and forecasting performance. Day-ahead prices are explored in Section 2.2.2.1, while the time series of imbalance volumes and prices from the Balancing Market are studied in Section 2.2.2.2. The substantial occurrence of anomalous imbalance prices during the first months of operation of the I-SEM is investigated in more detail in Section 2.2.2.3. Finally, a range of statistical analyses are carried out in Section 2.2.2.4 to investigate the impact of wind power penetration on a number of market quantities, namely: day-ahead prices, system imbalance direction, and anomalous imbalance prices.

The outcomes of these analyses and the characteristics of the data presented and discussed in the remainder of this Section will provide the motivation and justification for the modelling approaches adopted within the methodology of this work presented in Chapter 4.

2.2.2.1 Day-Ahead Market

The day-ahead market (DAM) was the dominant energy market in terms of liquidity, with more than 93% of the total ex-ante volumes being traded there. The average price was 64.35 €/MWh, the lowest –10.29 €/MWh and the highest 365 €/MWh. The average volume was 4.56 GW, with a minimum of 2.72 GW and a maximum of 6.84 GW. The time plots of cleared prices and volumes are shown in Figure 2.4. The Figure shows the presence of zero or even negative prices as well as price spikes; their occurrence will be analysed in more detail later in this Section.

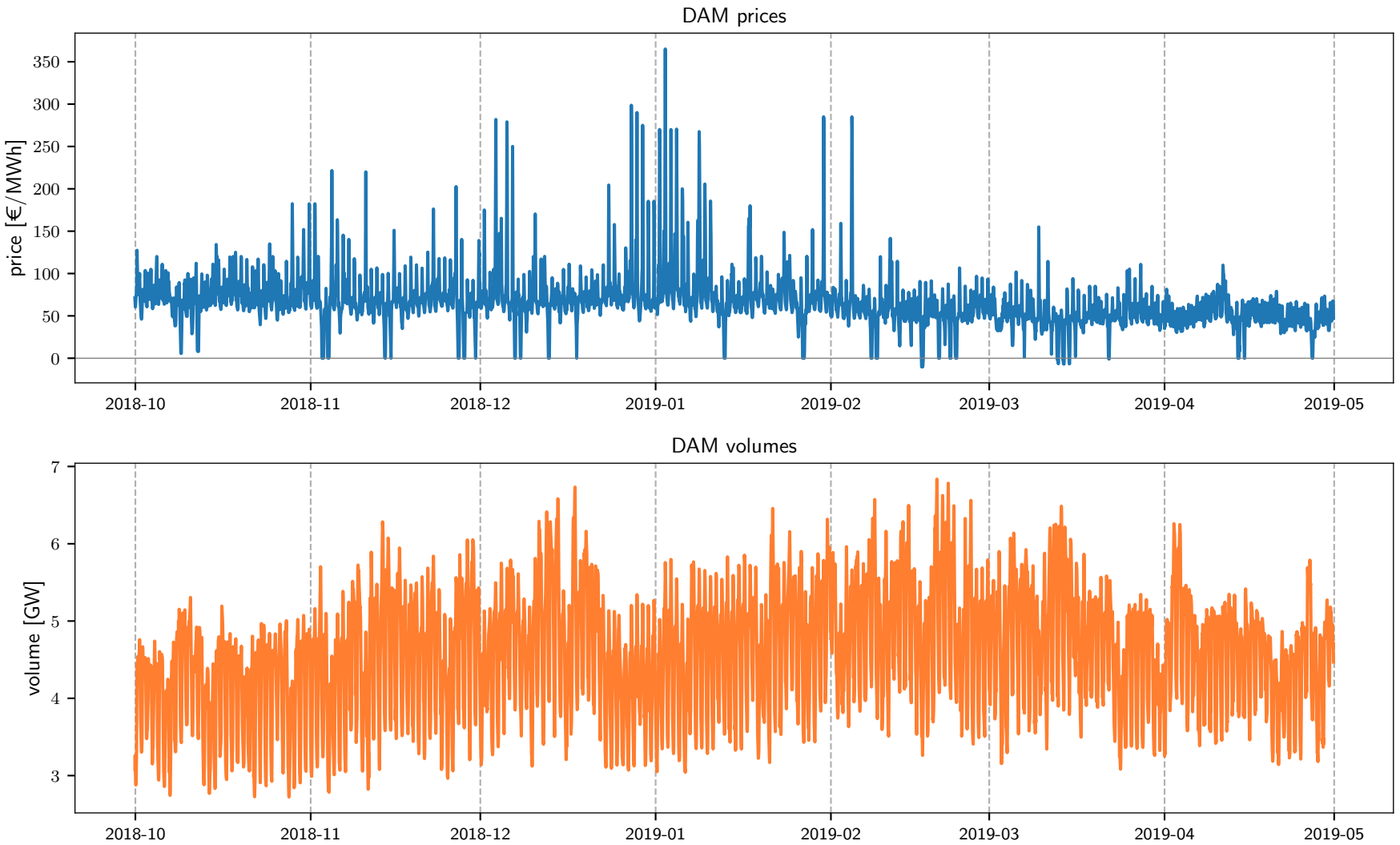


Figure 2.4: Day-ahead cleared prices (top) and volumes (bottom).

The cumulative and 30-day rolling average of prices and volumes are shown in Figure 2.5; it can be observed that prices increased during the month of January, decreased consistently in February and then stabilised from mid-March onwards; volumes, on the other hand, had a generally increasing trend, although there was a marked dip around January. The cumulative and 30-day rolling standard deviation of prices and volumes are shown in Figure 2.6, showing how the volatility of prices increased constantly during the first three months and then experienced two sharp drops at the start of February and March, followed by a decreasing trend; the volatility of volumes increased steeply in the first two months, with marked drops in mid-January and mid-April.

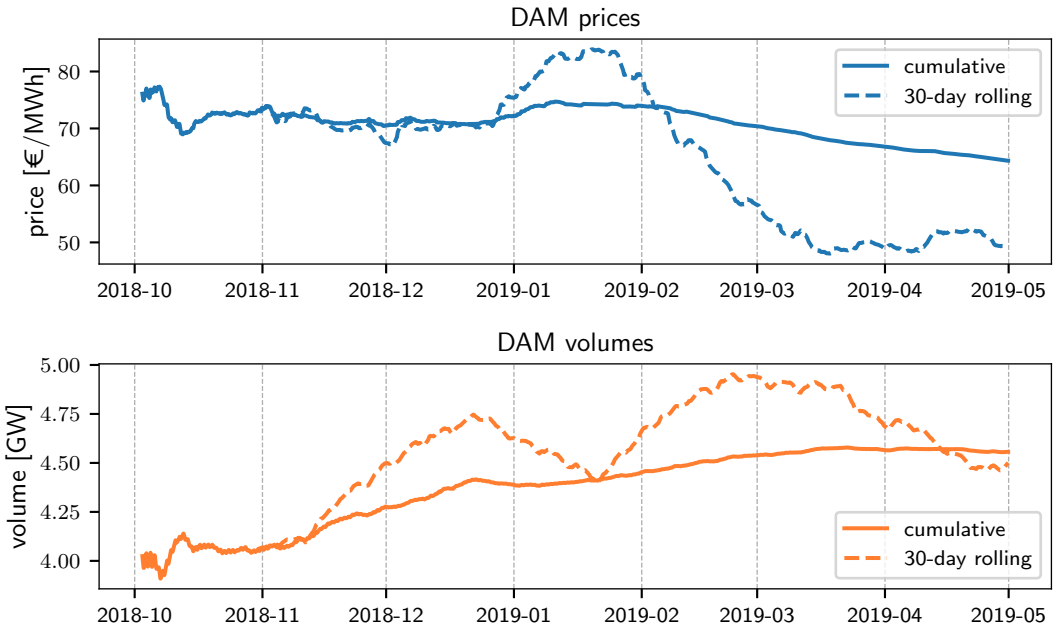


Figure 2.5: Cumulative (solid line) and 30-day rolling (dashed line) average of day-ahead prices (top) and volumes (bottom).

The time series of price and volume are hourly data (recall that the trading period is one hour in the DAM) with two underlying seasonality: a daily pattern and a weekly pattern. The presence of some seasonal components can already be noticed in Figure 2.4 but it is hard to distinguish their periodicities at this resolution, which are instead illustrated clearly by the seasonal plots in the following graphs. The weekly seasonal plots of price and volume shown in Figures 2.7 and

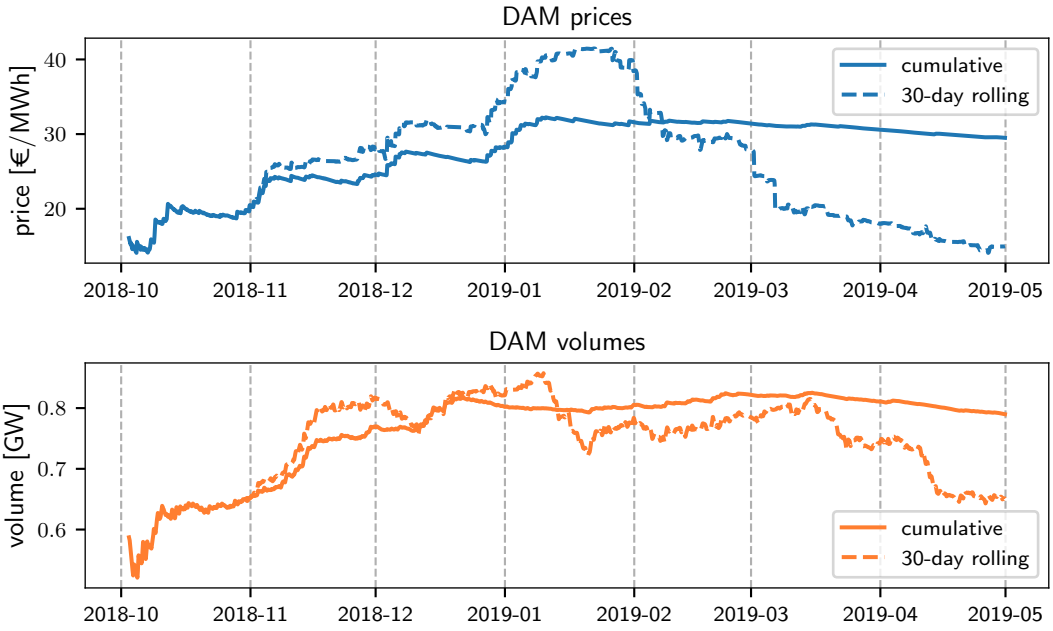


Figure 2.6: Cumulative (solid line) and 30-day rolling (dashed line) standard deviation of day-ahead prices (top) and volumes (bottom).

2.8 highlight the weekly patterns of both variables. Prices and more markedly volumes were in general lower during the weekend; zero or negative prices occurred with higher incidence in the last three days of the week, while price spikes occurred regardless of the day of the week but were lower in value during the weekend.

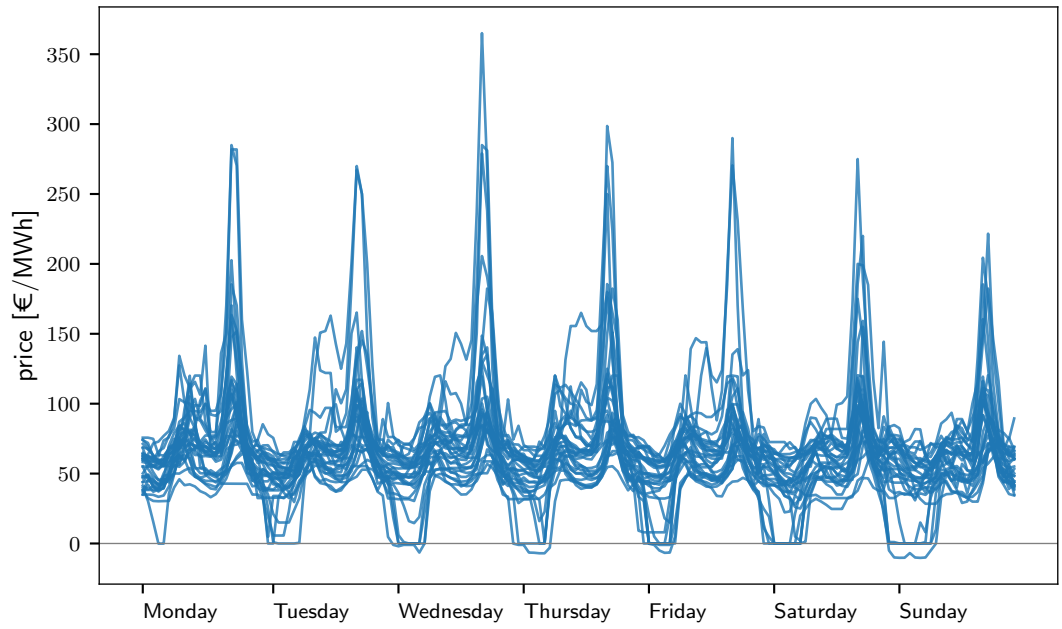


Figure 2.7: Weekly seasonal plot of day-ahead prices.

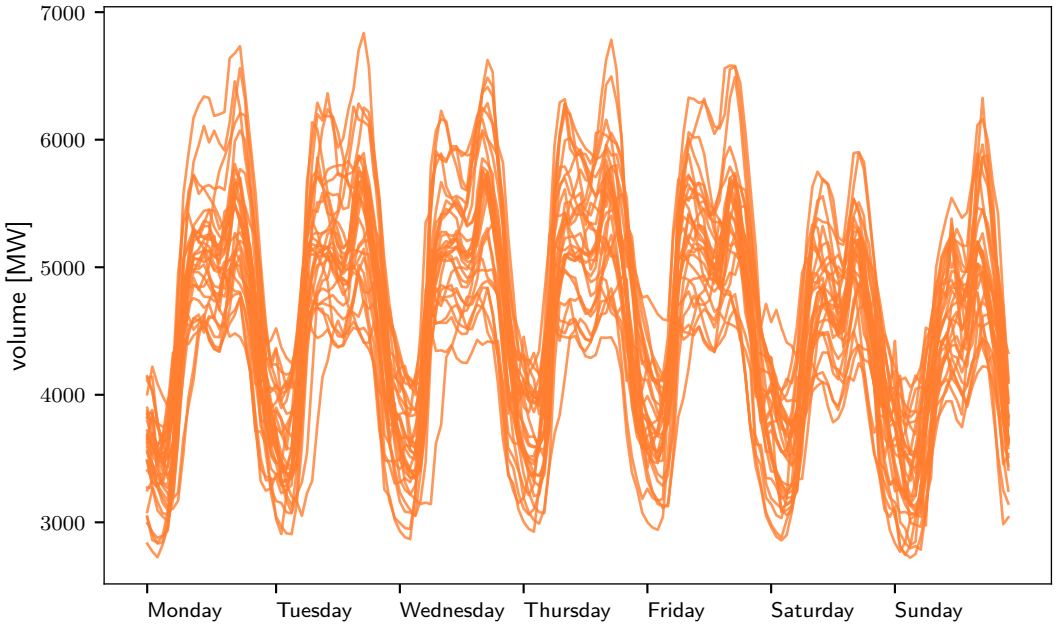


Figure 2.8: Weekly seasonal plot of day-ahead volumes.

Figures 2.9 and 2.10 show the daily seasonal plots of price and volume, highlighting a marked daily pattern for both variables. The daily seasonality of prices is further analysed in Figure 2.11, where box-and-whisker plots are used to show the distribution of prices for each hour of the day. From Figures 2.9 and 2.11, one can see how zero and negative prices occurred only in the early hours of the day (when demand is low), while price spikes occurred between 16:00 and 19:00 (when demand is high). This is displayed with more clarity in Figures 2.12 and 2.13, where the number of instances with zero or negative prices and with prices above 150 €/MWh are shown, respectively. To further investigate the hourly behaviour of prices, the time plot of each one-hour trading period is shown separately in Figures 2.14 and 2.15. The plots show how the variance of the series decreases when considering each trading period separately (cf. Figure 2.4).

The average values of price and volume for each hour of the day are shown in Figure 2.16. The daily pattern of volumes is clear: values are low during the night, increase during the morning with a first peak around 09:00 and then again during the evening with a second, larger peak around 18:00. Prices have a more accentuated peak in the evening but overall follow a parallel trend. This is a

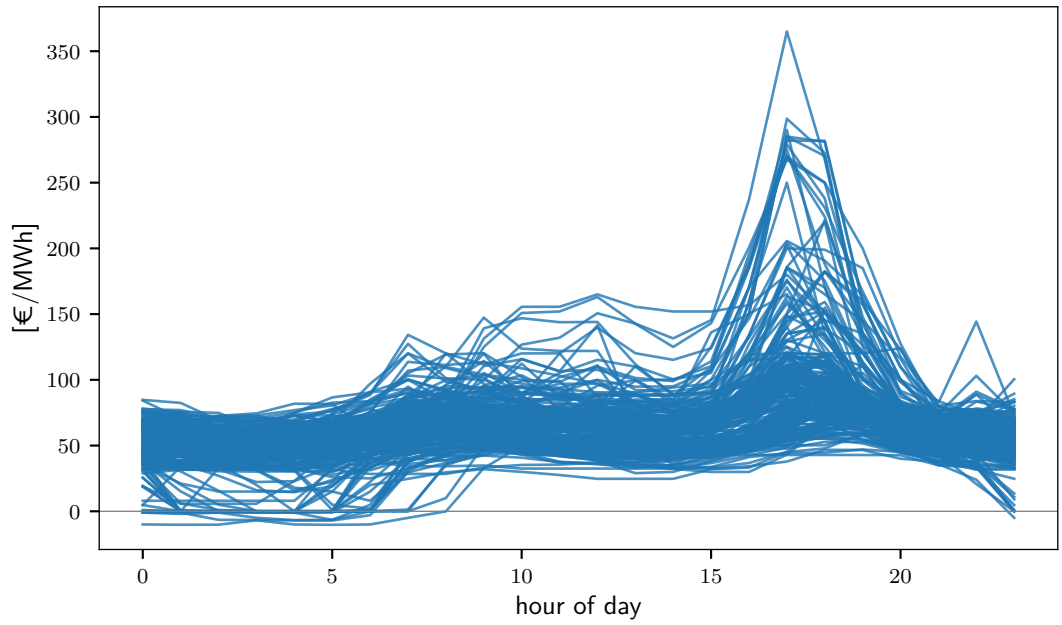


Figure 2.9: Daily seasonal plot of day-ahead prices.

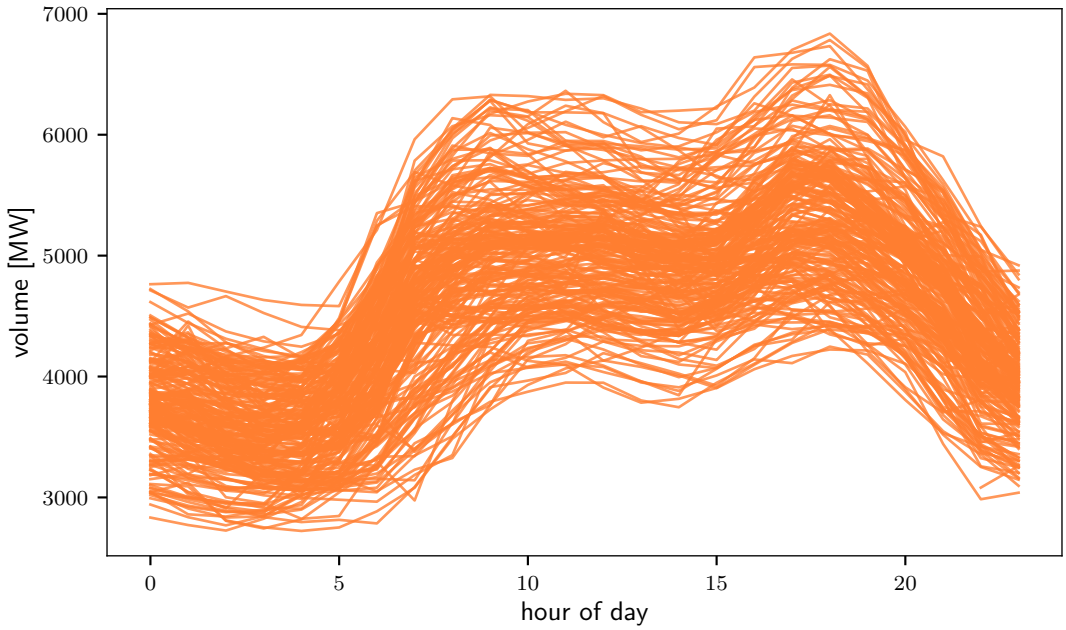


Figure 2.10: Daily seasonal plot of day-ahead volumes.

consequence of more expensive generation being required to meet higher demands and the application of the merit order principle employed in the formation of the supply and demand curves in the day-ahead market clearing (cf. Figure 2.1).

Prices and volumes however do not have similar trends in terms of standard deviation, as shown in Figure 2.17. The variance of prices is constant during the

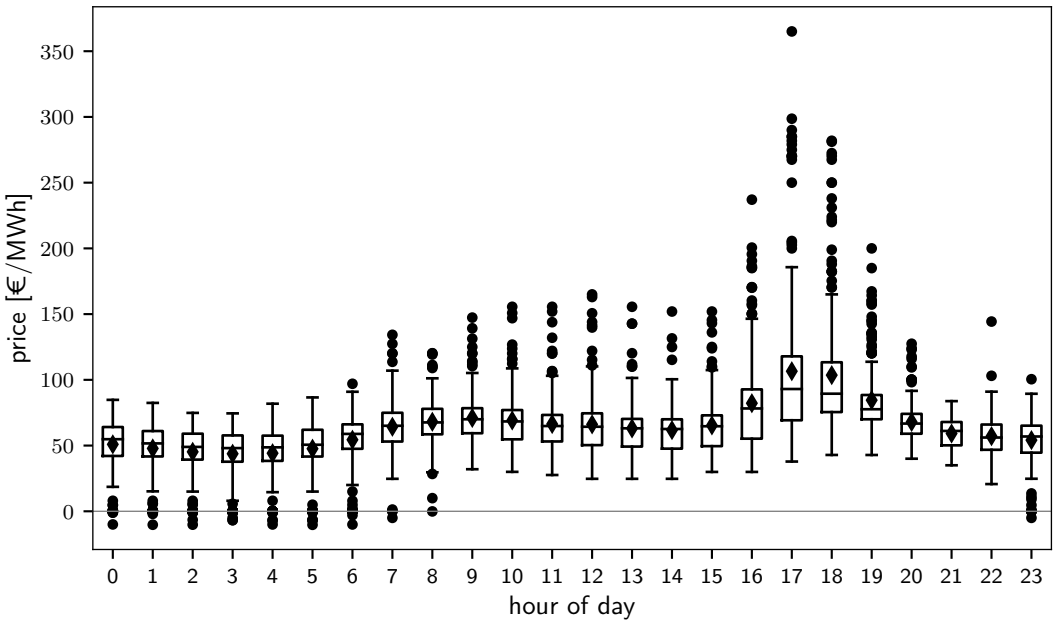


Figure 2.11: Boxplots of day-ahead prices for each hour of the day. The diamond marker indicates the mean of the distribution.

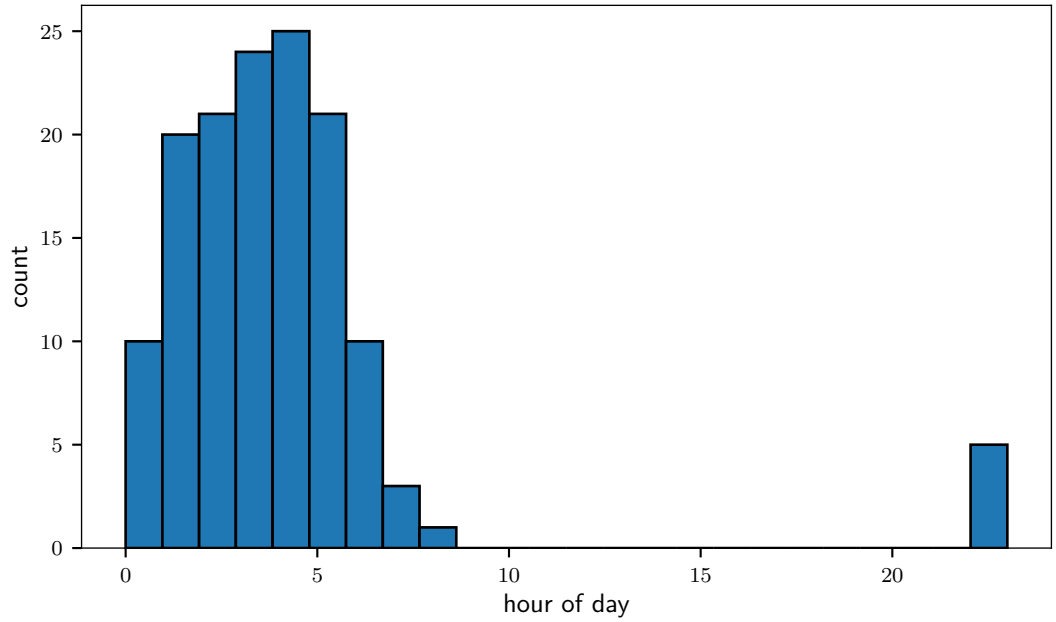


Figure 2.12: Number of instances with zero or negative prices for each hour of the day.

day but becomes noticeably larger in the evening (as visible also in Figures 2.9 and 2.11); volume displays more variability and has the greatest variance during the morning load rise.

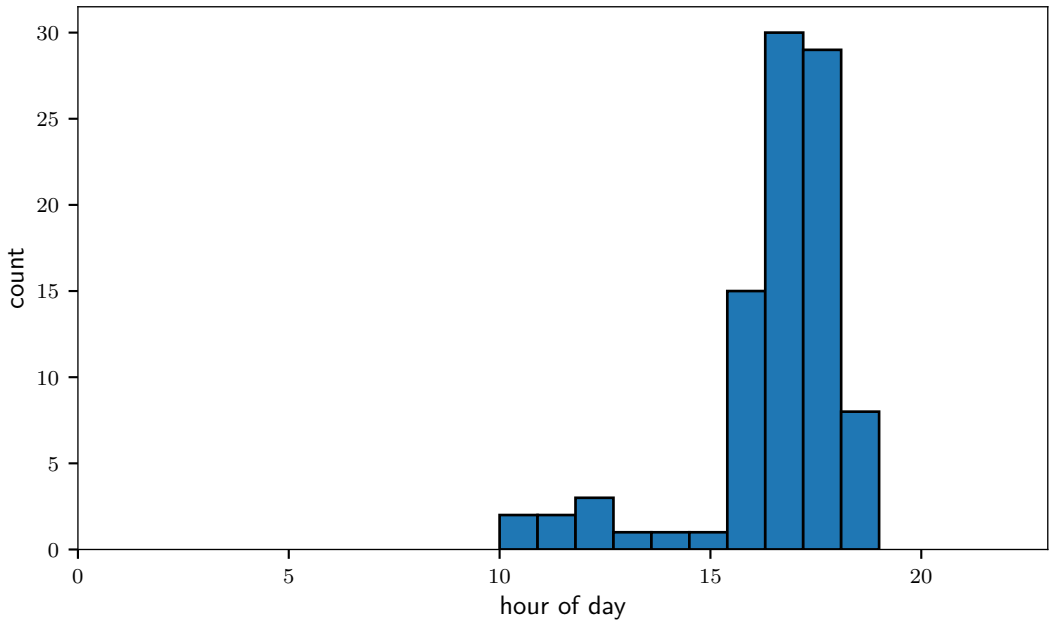


Figure 2.13: Number of instances with prices above 150 €/MWh for each hour of the day.

2.2.2.2 Balancing Market

During the first seven months of operation, the Balancing Market (BM) has been characterised by extremely volatile and often counter-intuitive imbalance prices and the tendency of participants to leave the system systematically short. As the analysis in the "Modification Proposal Mod_06_19" [37] has shown, balancing market outcomes in the I-SEM have been considerably more volatile than in Great Britain, where the market structure is comparable to Ireland. In the document, it was argued that the higher volatility could be a consequence of the smaller size of the system, the more variable generation within it, grid constraints and limited market experience of the participants.

In this Section, a detailed analysis of some key outcomes from the operation of the BM is performed. First, it is recalled that the trading period in the BM — called Imbalance Settlement Period (ISP) — has a duration of 30 minutes. A positive imbalance volume corresponds to the system being short, thus leading to up-regulation; conversely, a negative imbalance volume corresponds to the system being long, thus leading to down-regulation.

The time plot of imbalance volumes is shown in Figure 2.18. The mean absolute volume was 21.56 MWh and the mean volume 2.94 MWh, indicating a positive bias in the series. This bias is also visible from the histogram in Figure 2.19 showing the distribution of the series. Up-regulation occurred 51.9% of the times, with an average upward imbalance volume of 23.62 MWh and a maximum of 134.11 MWh. Down-regulation occurred 42.5% of the times, with an average downward imbalance volume of –21.88 MWh and a minimum of –108.75 MWh. The system was in balance and thus no regulation was needed 5.6% of the times.

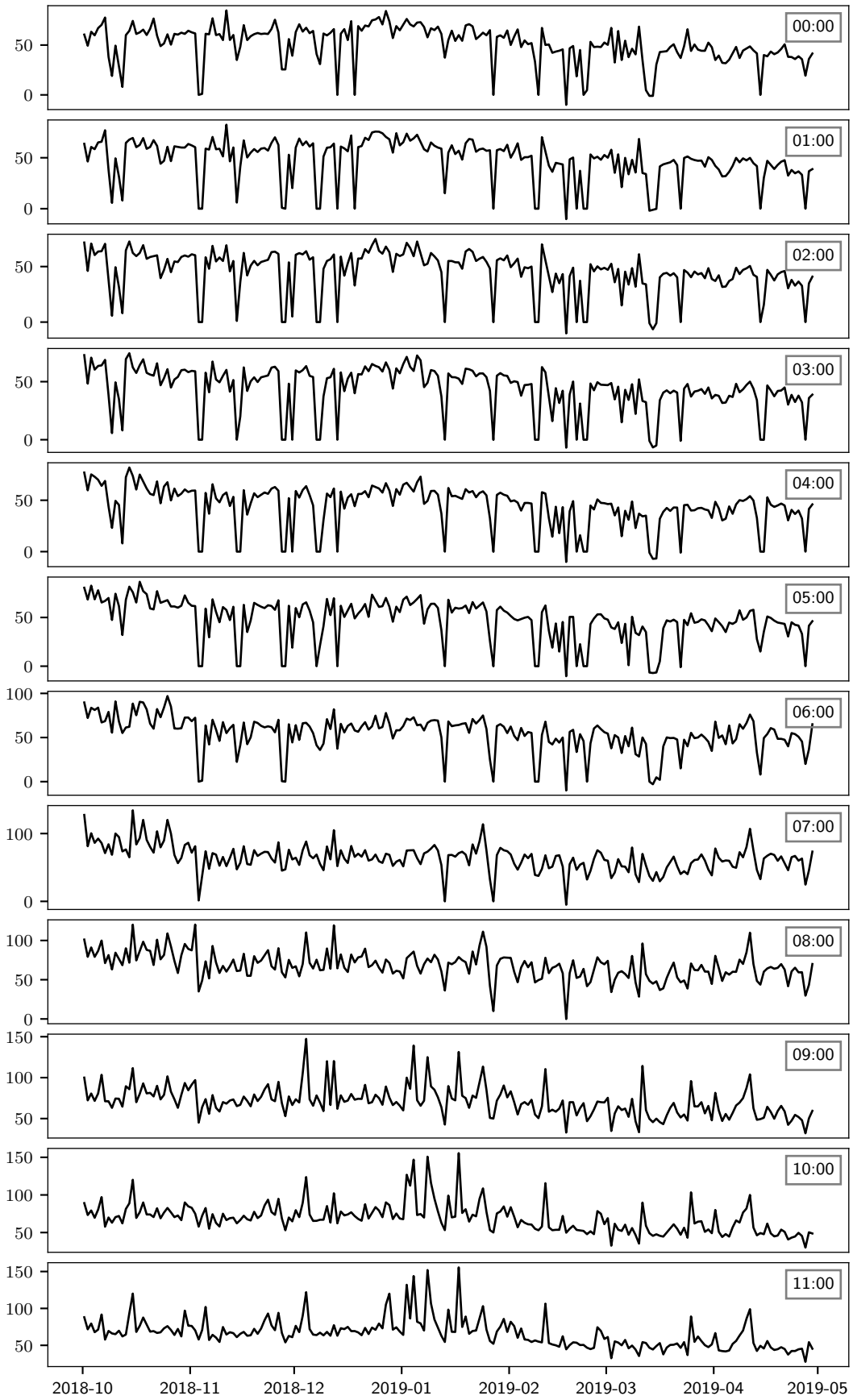


Figure 2.14: Time plot of DAM prices (expressed in €/MWh) for each one-hour trading period (part 1).

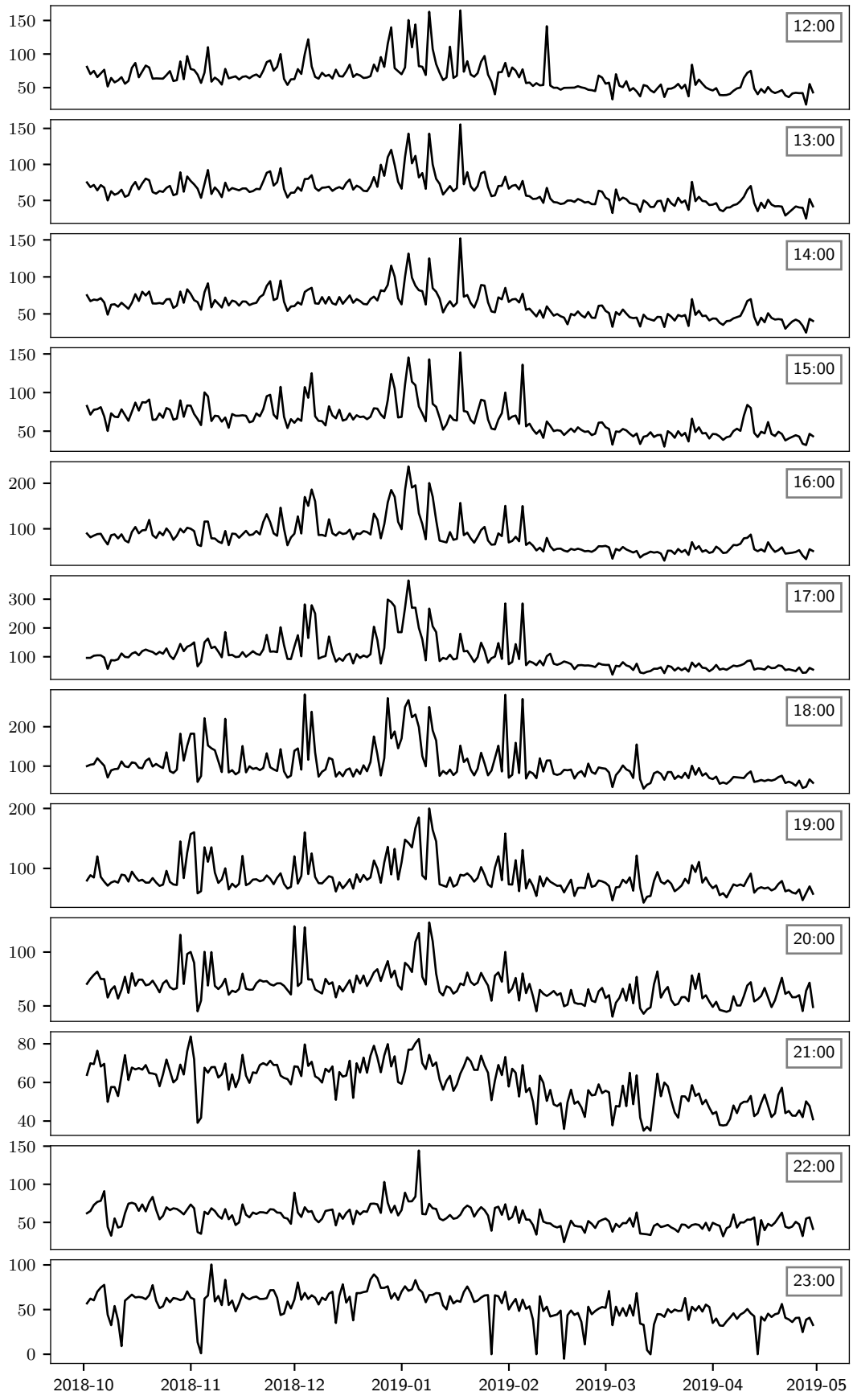


Figure 2.15: Time plot of DAM prices (expressed in €/MWh) for each one-hour trading period (part 2).

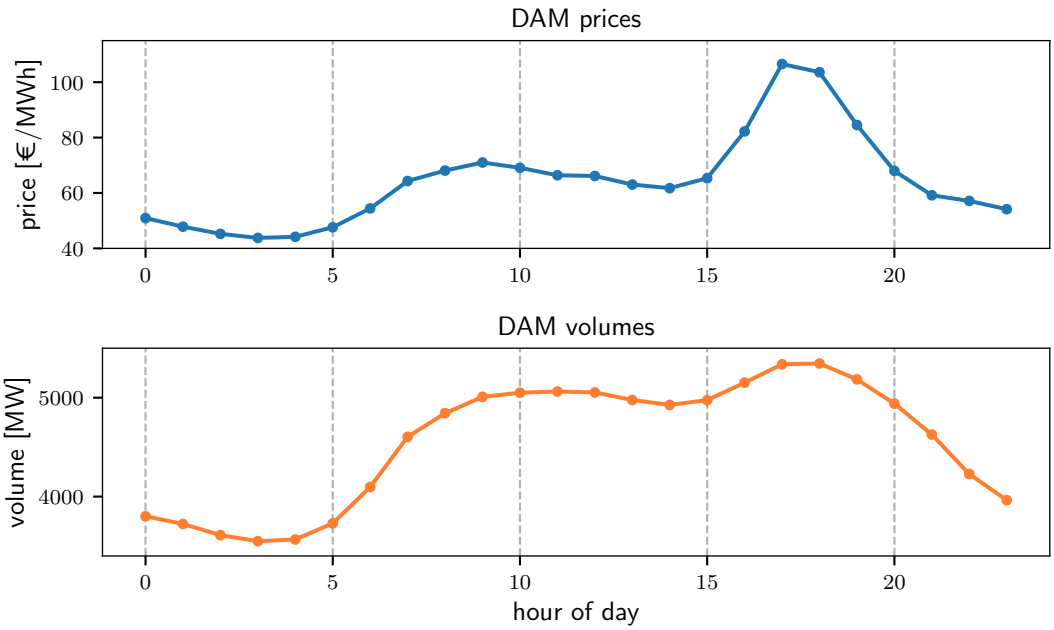


Figure 2.16: Average day-ahead price (top) and volume (bottom) for each hour of the day.

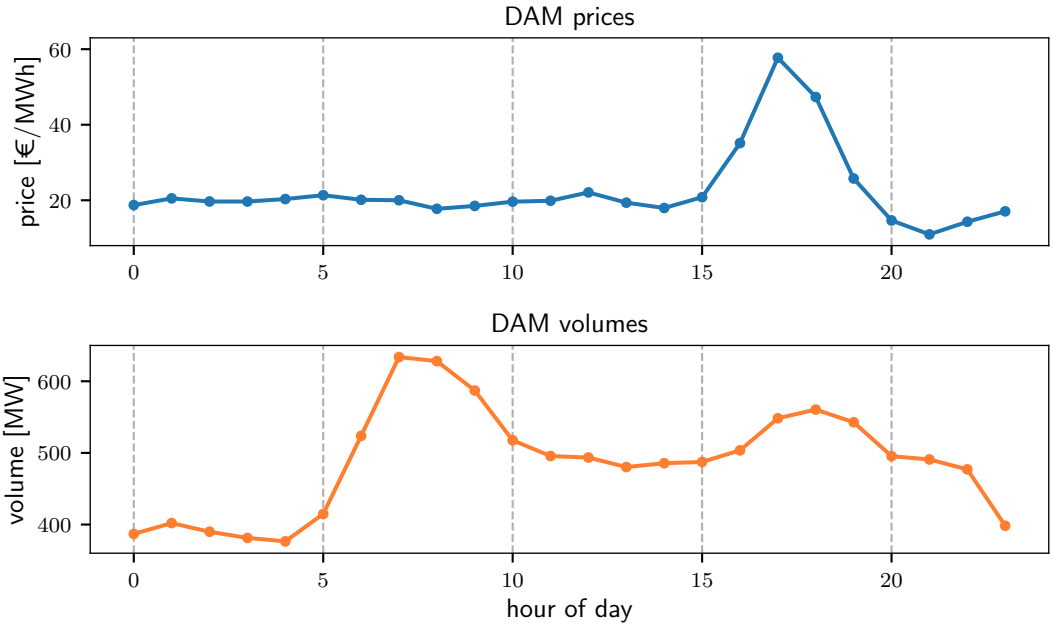


Figure 2.17: Standard deviation of day-ahead price (top) and volume (bottom) for each hour of the day.

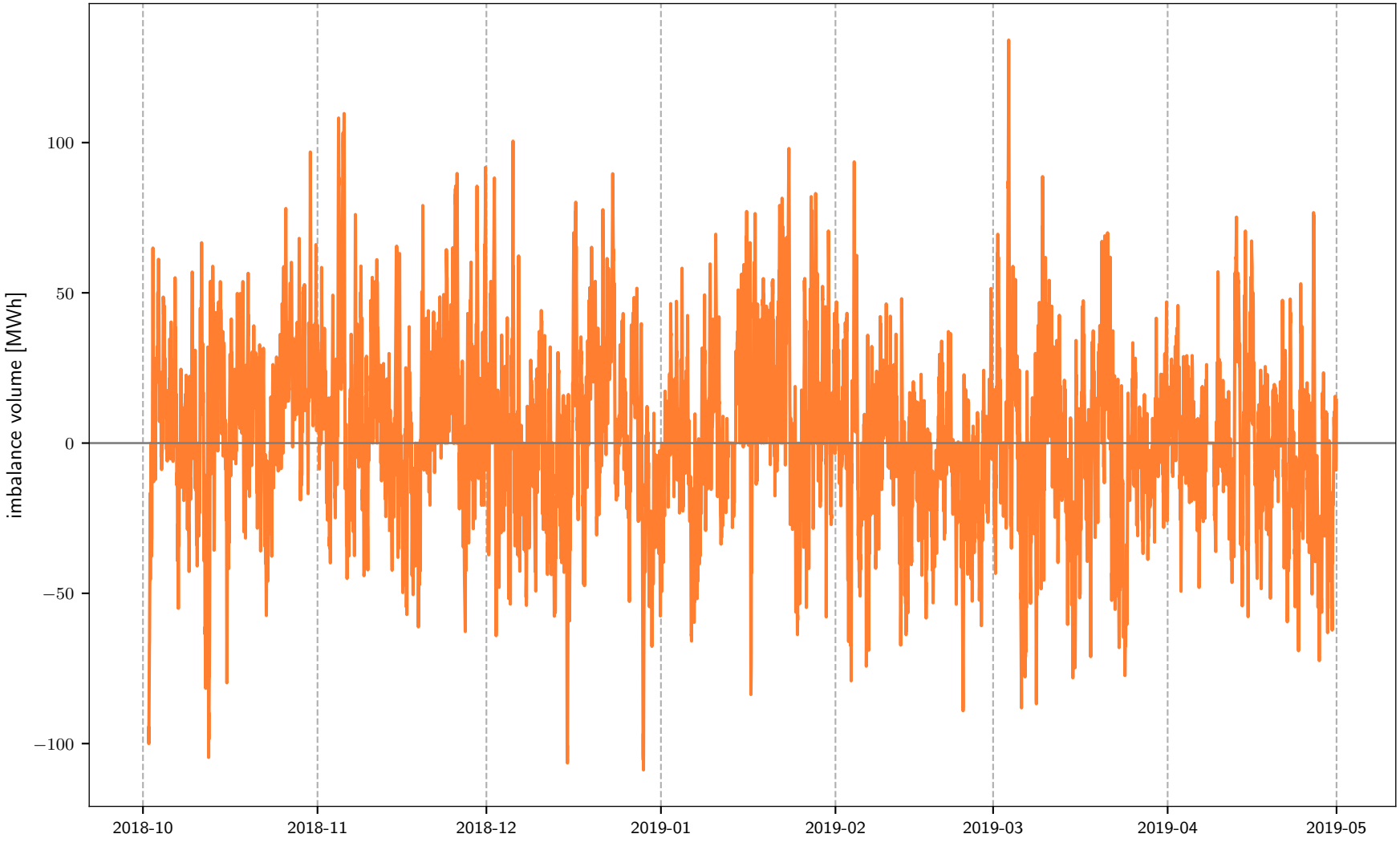


Figure 2.18: Time plot of imbalance volumes.

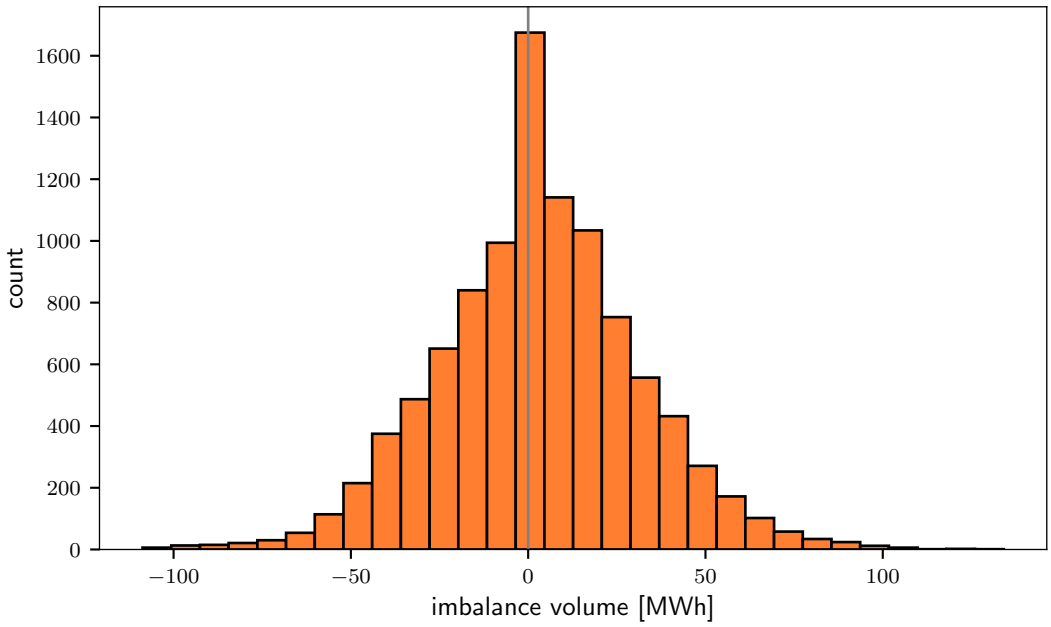


Figure 2.19: Distribution of imbalance volumes.

Analysing the distribution of imbalance volumes for each ISP, it was found that the tendency to leave the system short was in fact limited to the evening hours corresponding to peak demand. The boxplots in Figure 2.20 clearly show how distributions have their mean and median around zero for most trading periods except for those between ISP 34 and 43, i.e. from 16:30 to 21:00, where distributions are positively biased.

The cumulative and 30-day rolling frequency of up-regulation are plotted in Figure 2.21 to provide a more dynamic picture of how the variable has changed over time. The graph highlights how the tendency to leave the system systematically short has been very marked in the first months, but with a clear decreasing trend afterwards.

The imbalance price time series is plotted in Figure 2.22 (note the y-axis is broken in the graph). The graph shows the high volatility of the series and the occurrence of negative prices as well as very high prices. In particular, the strike price fixed at 500 €/MWh was exceeded in nine trading periods: twice on 2/10/2018 and 9/10/2018, once on 23/11/2018, and four times on 24/01/2019 where the record price of 3,773.69 €/MWh was registered. The causes leading to this specific event

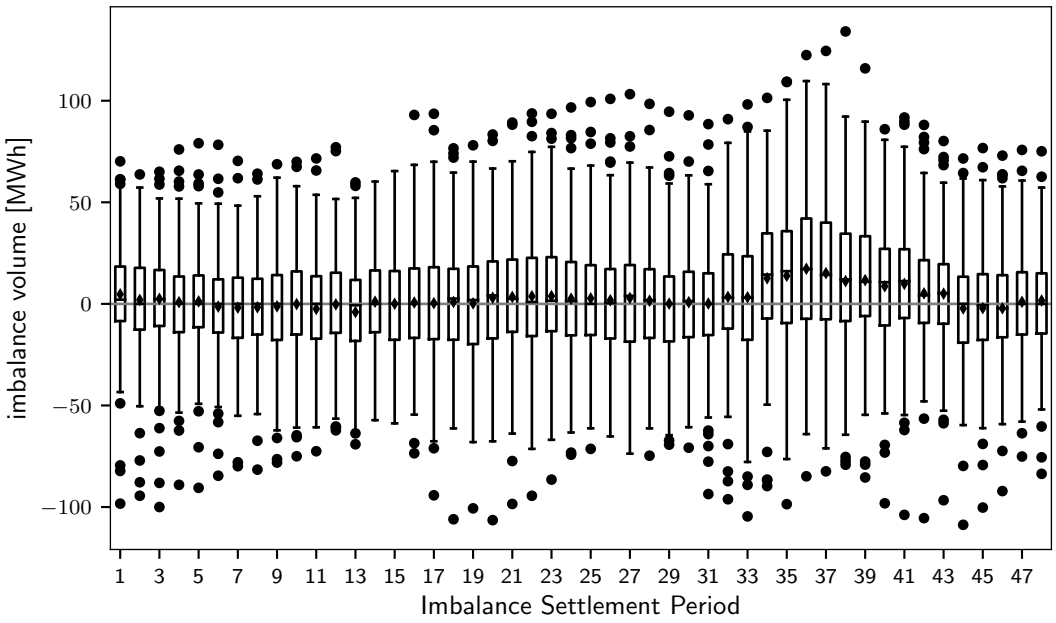


Figure 2.20: Boxplots of imbalance volumes for each Imbalance Settlement Period. The diamond marker indicates the mean of the distribution.

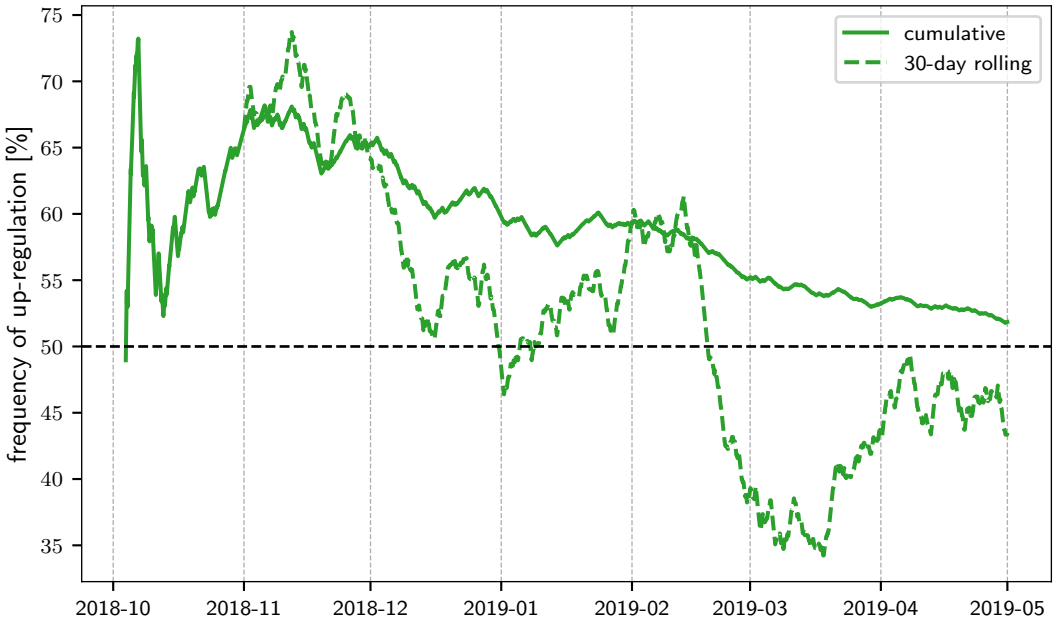


Figure 2.21: Cumulative (solid line) and 30-day rolling (dashed line) frequency of up-regulation.

are illustrated in detail in Appendix B. Whenever the imbalance price exceeds the strike price, the price is effectively capped at 500 €/MWh for all participants, with suppliers having to pay the system back the difference between the energy price

and the strike price. Therefore, imbalance prices above 500 €/MWh will be capped at this value for use in all the analysis performed in the remainder of the work, as producers are eventually remunerated at this price for their generation.

The scatter plot of imbalance prices versus volumes in Figure 2.23 shows the positive correlation between the two variables and that negative prices occurred mainly for negative imbalance volumes. It is also interesting to note that three of the nine instances where the strike price was exceeded (highlighted in orange) happened during down-regulation.

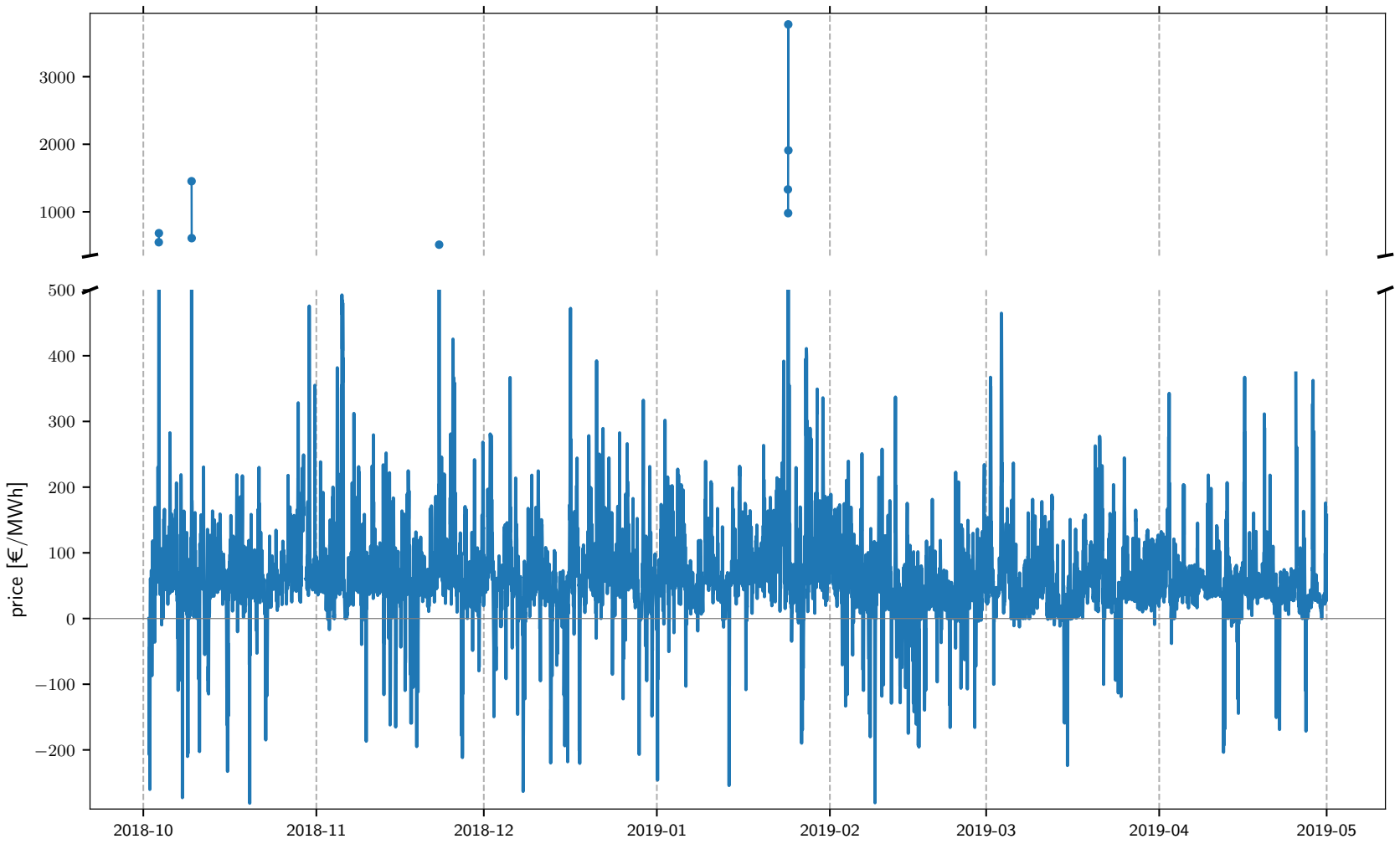


Figure 2.22: Time plot of imbalance prices. Note the y-axis is broken for a clearer visualisation of the time series while including very high prices.

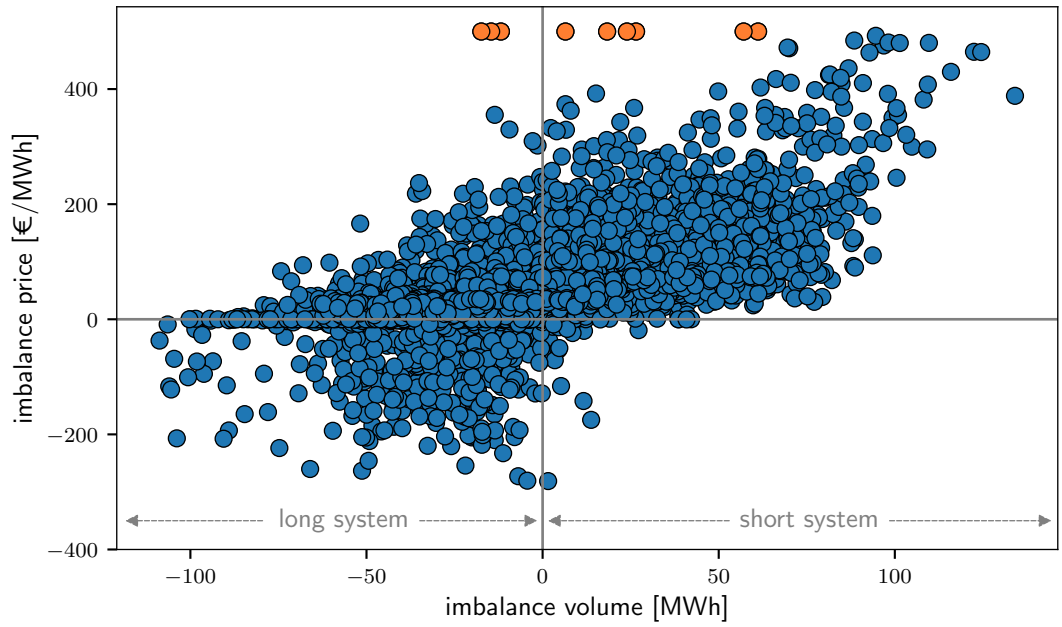


Figure 2.23: Scatter plot of imbalance prices vs imbalance volumes. Periods where the strike price was exceeded are highlighted in orange.

The distributions of imbalance prices during up- and down-regulation are shown in Figures 2.24 and 2.25, respectively. As expected, the histograms show how prices were generally higher when the system was short; as already noted from the previous scatter plot, negative prices occurred mainly when the system was long. During up-regulation, the average price was 92.78 €/MWh, with a minimum of –281.16 €/MWh and a maximum of 1,453 €/MWh (capped at 500). During down-regulation, the average price was 25.92 €/MWh, with a minimum of –280.45 €/MWh and a maximum of 3,773.69 €/MWh (capped at 500).

The weekly seasonal plot of BM prices and volumes are shown in Figures 2.26 and 2.27. As opposed to the case of DAM quantities, prices and volumes in the BM do not show any evident seasonal pattern. Nevertheless, Figure 2.26 highlights that during weekdays (i.e. Monday to Friday) negative prices occur more in late night/early morning hours.

Although BM quantities did not show any particular seasonality, imbalance prices on average followed a similar daily pattern to day-ahead prices, as shown in Figure 2.28: lower during the night, increasing during the day, and highest during

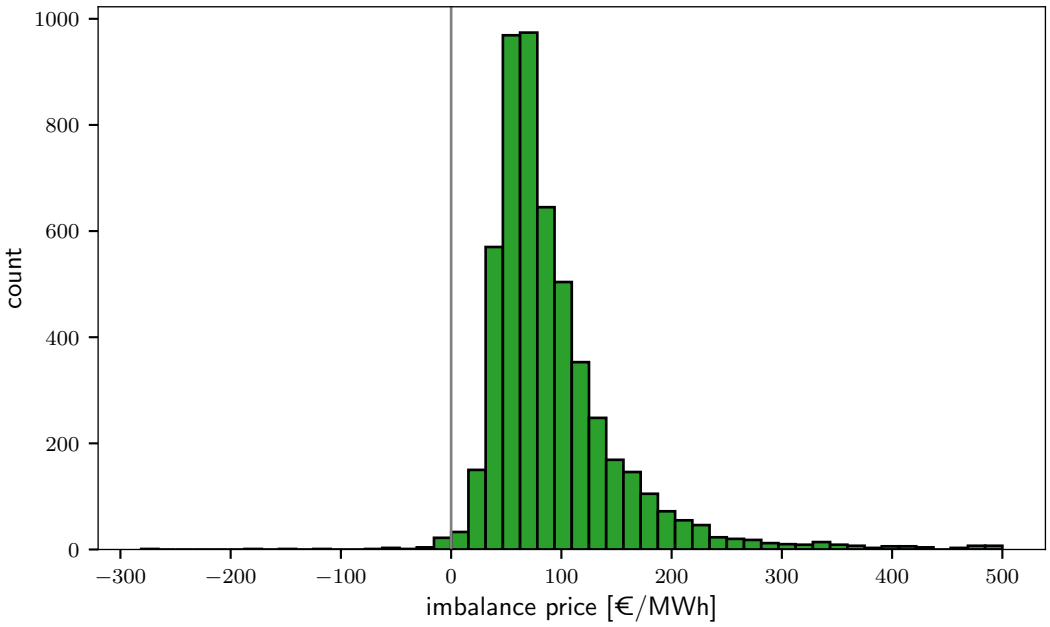


Figure 2.24: Distribution of imbalance prices during up-regulation.

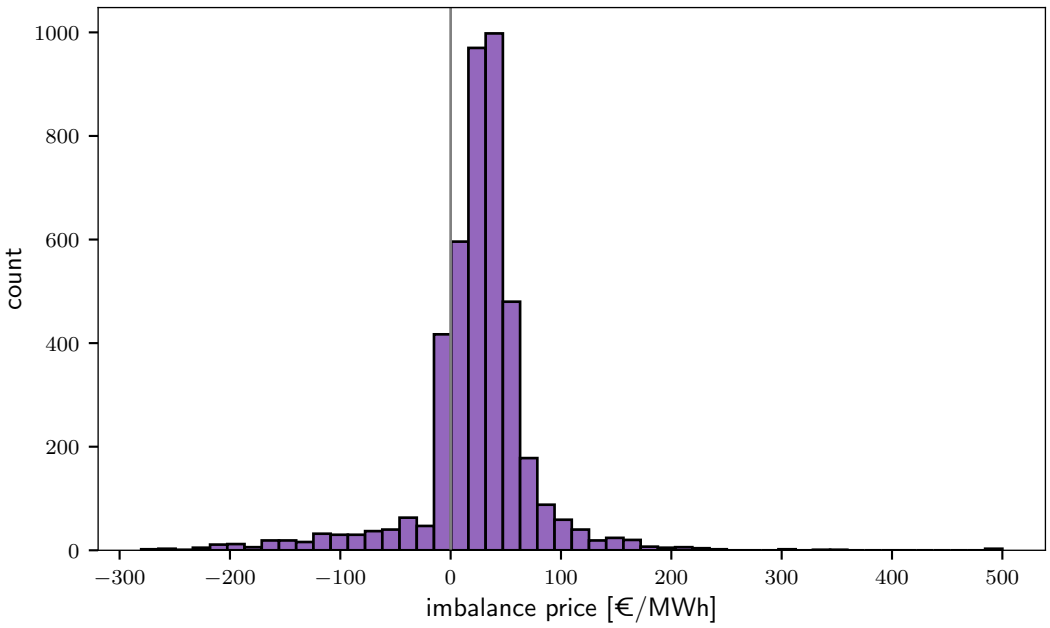


Figure 2.25: Distribution of imbalance prices during down-regulation.

peak demand hours. The bottom panel of the same Figure shows the average imbalance volumes for each ISP, from which the evening positive bias discussed earlier is evident. Evening hours are also those characterised by the highest variance of both prices and volumes, as displayed in Figure 2.29. The hourly behaviour of prices is further investigated by plotting prices during each 30-minute ISP separately, as

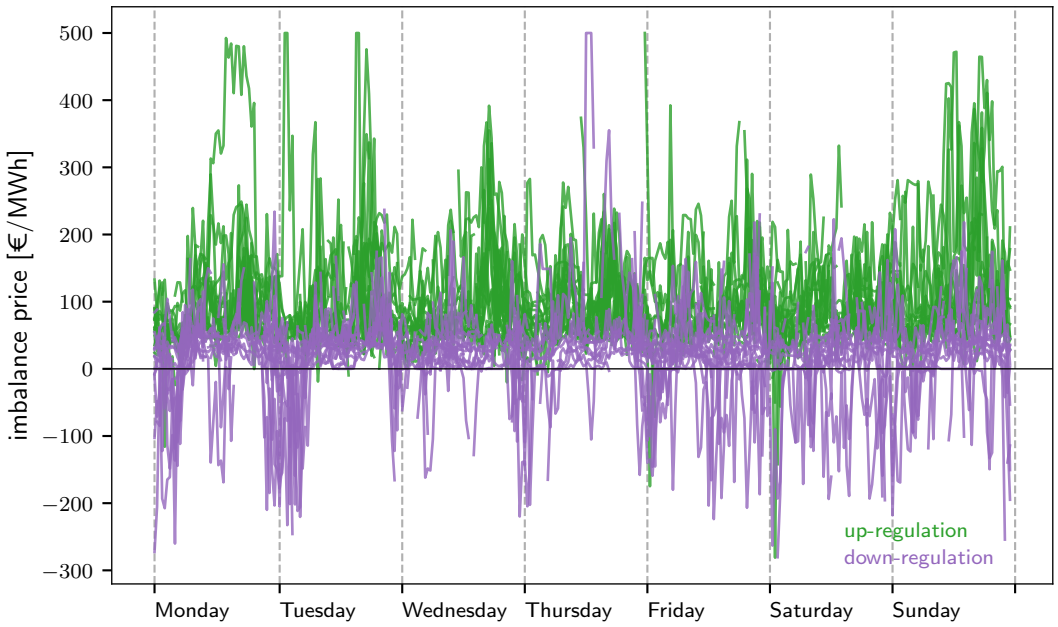


Figure 2.26: Weekly seasonal plot of imbalance prices.

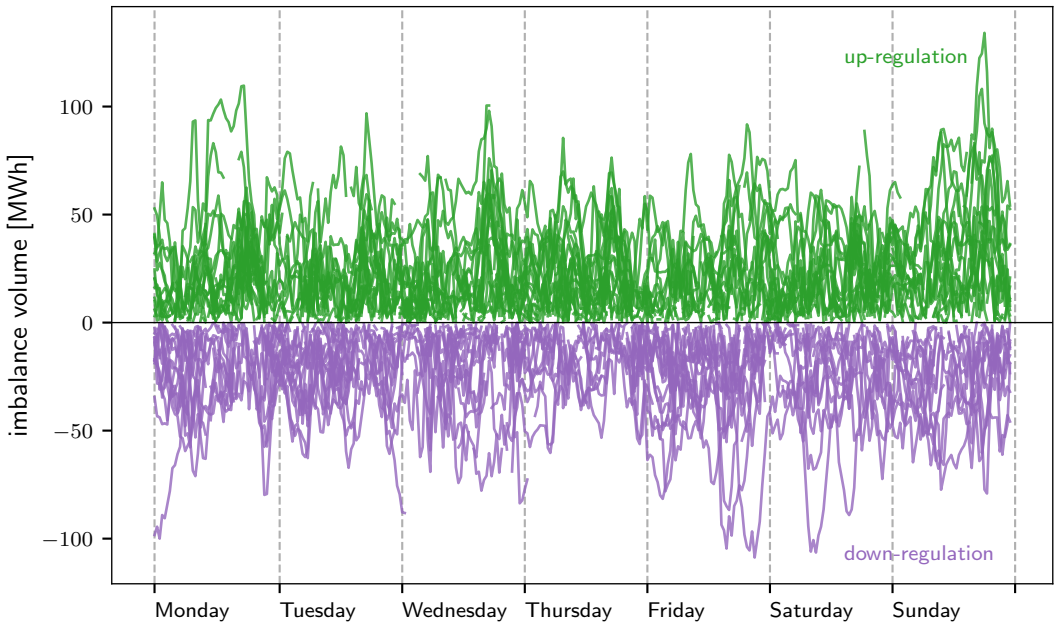


Figure 2.27: Weekly seasonal plot of imbalance volumes.

shown in Figures 2.30 through 2.33. The plots show that the variance of the series can change considerably depending on the time of the day. Moreover, the variance of imbalance prices over a single trading period is substantially higher than that of day-ahead prices (cf. Figure 2.14 and 2.15).

The distribution of prices during up- and down-regulation for each ISP is shown

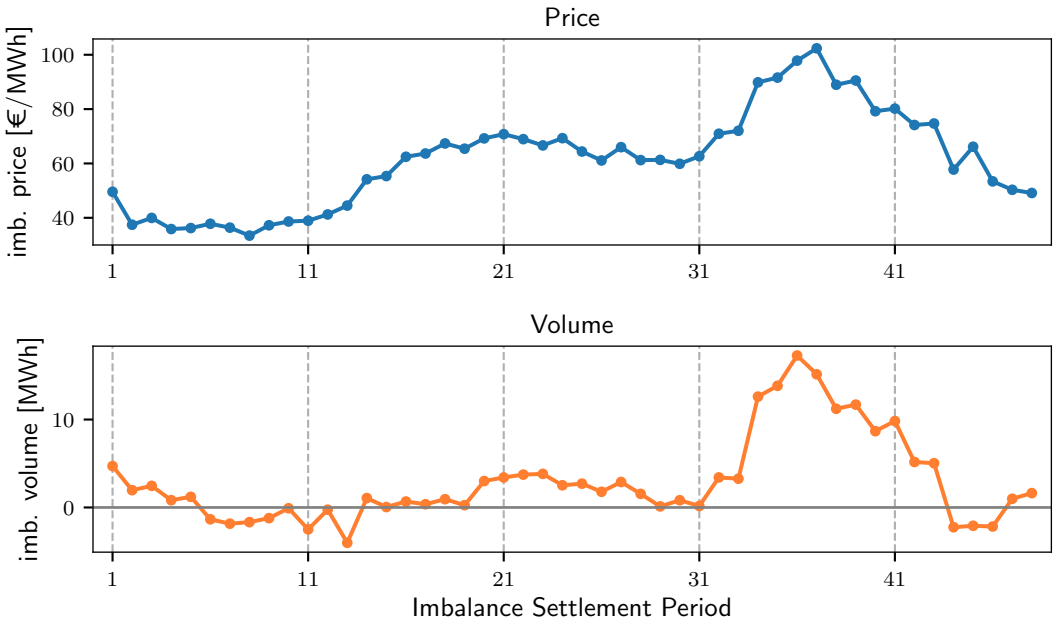


Figure 2.28: Average imbalance prices (top) and volumes (bottom) for each Imbalance Settlement Period.

Figure 2.29: Standard deviation of imbalance prices (top) and volumes (bottom) for each Imbalance Settlement Period.

via the boxplots in Figure 2.34, where it can be observed, as expected, that prices are generally higher during up-regulation across all trading periods.

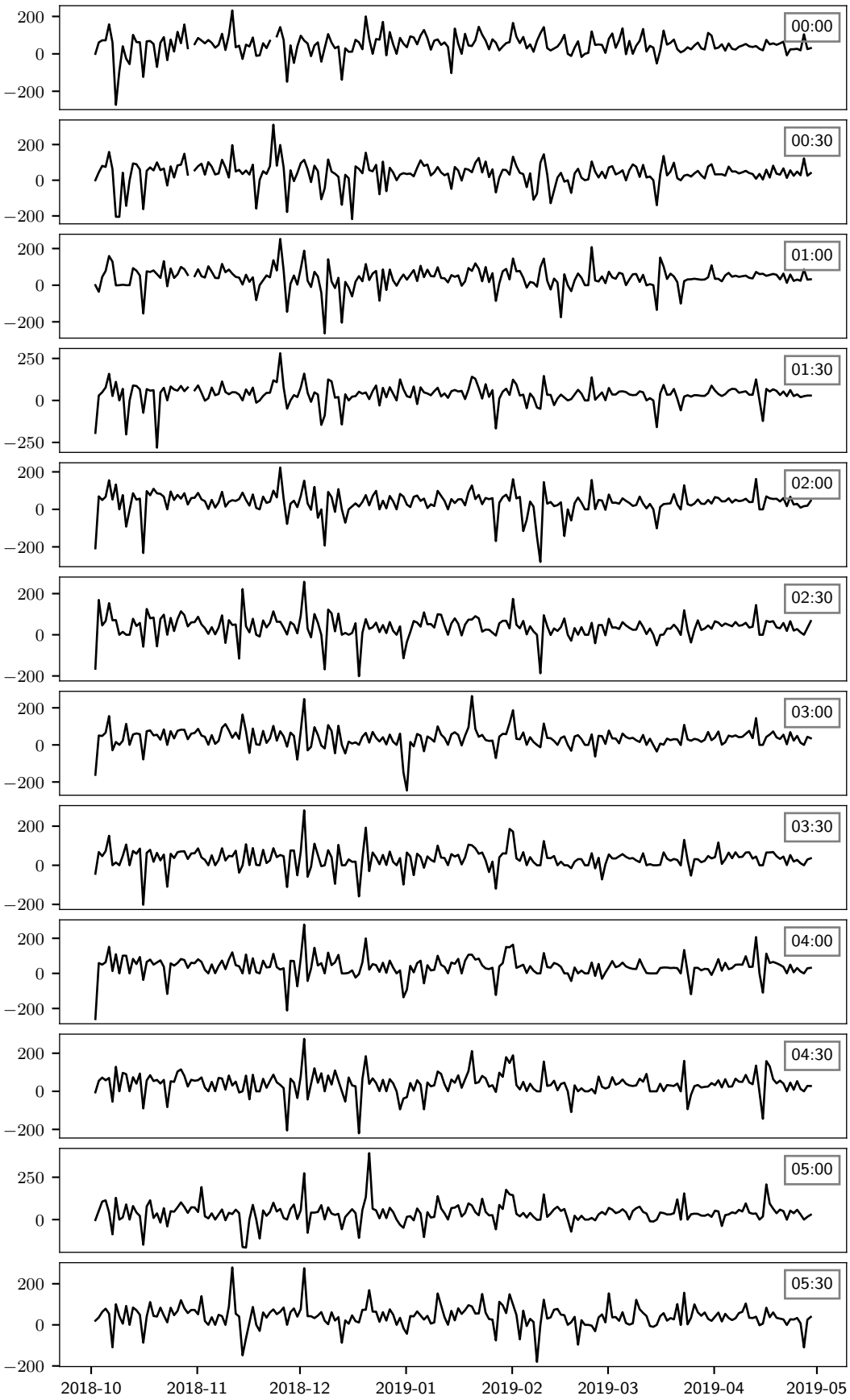


Figure 2.30: Time plot of BM prices (expressed in €/MWh) for each 30-minute trading period (part 1).

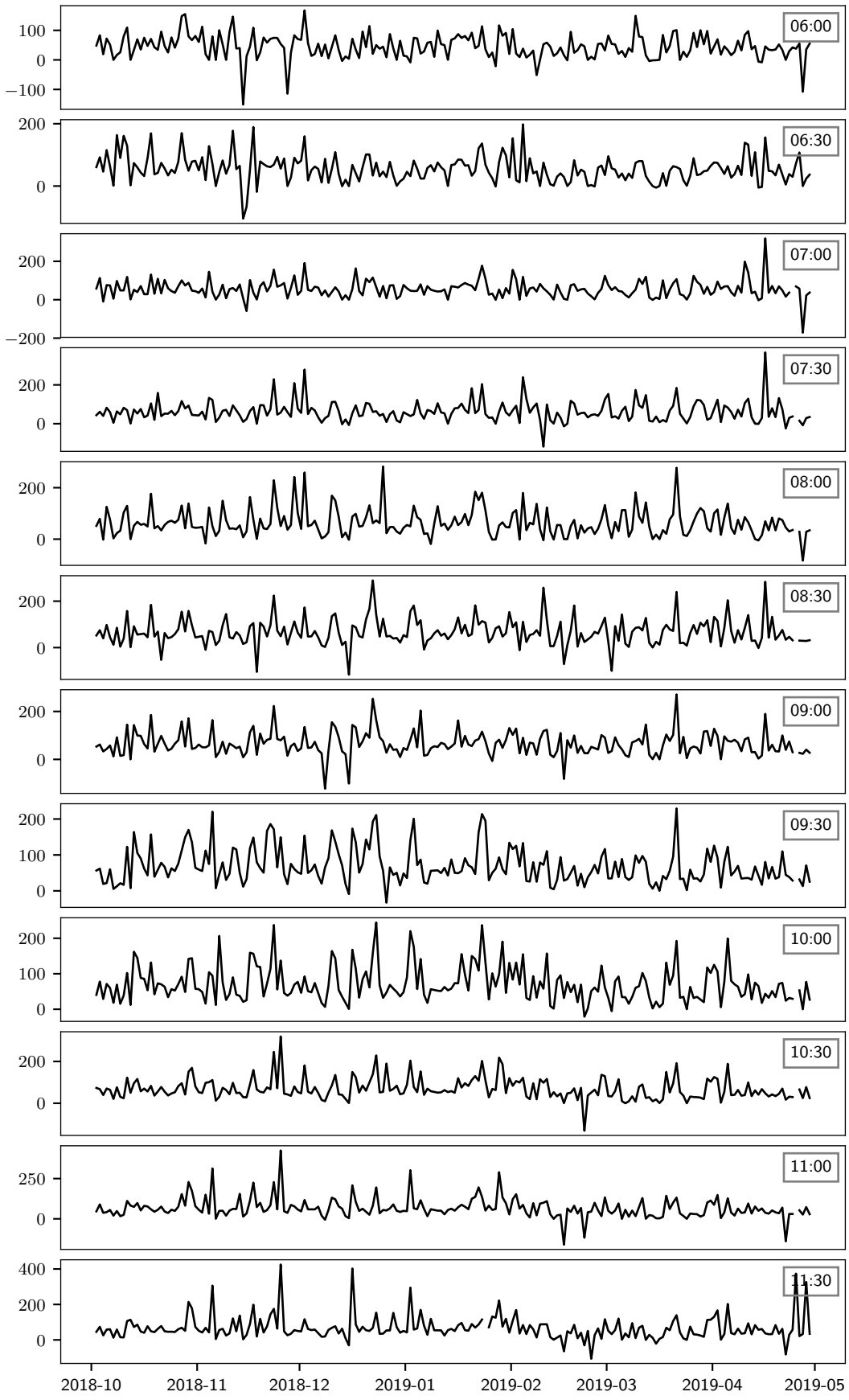


Figure 2.31: Time plot of BM prices (expressed in €/MWh) for each 30-minute trading period (part 2).

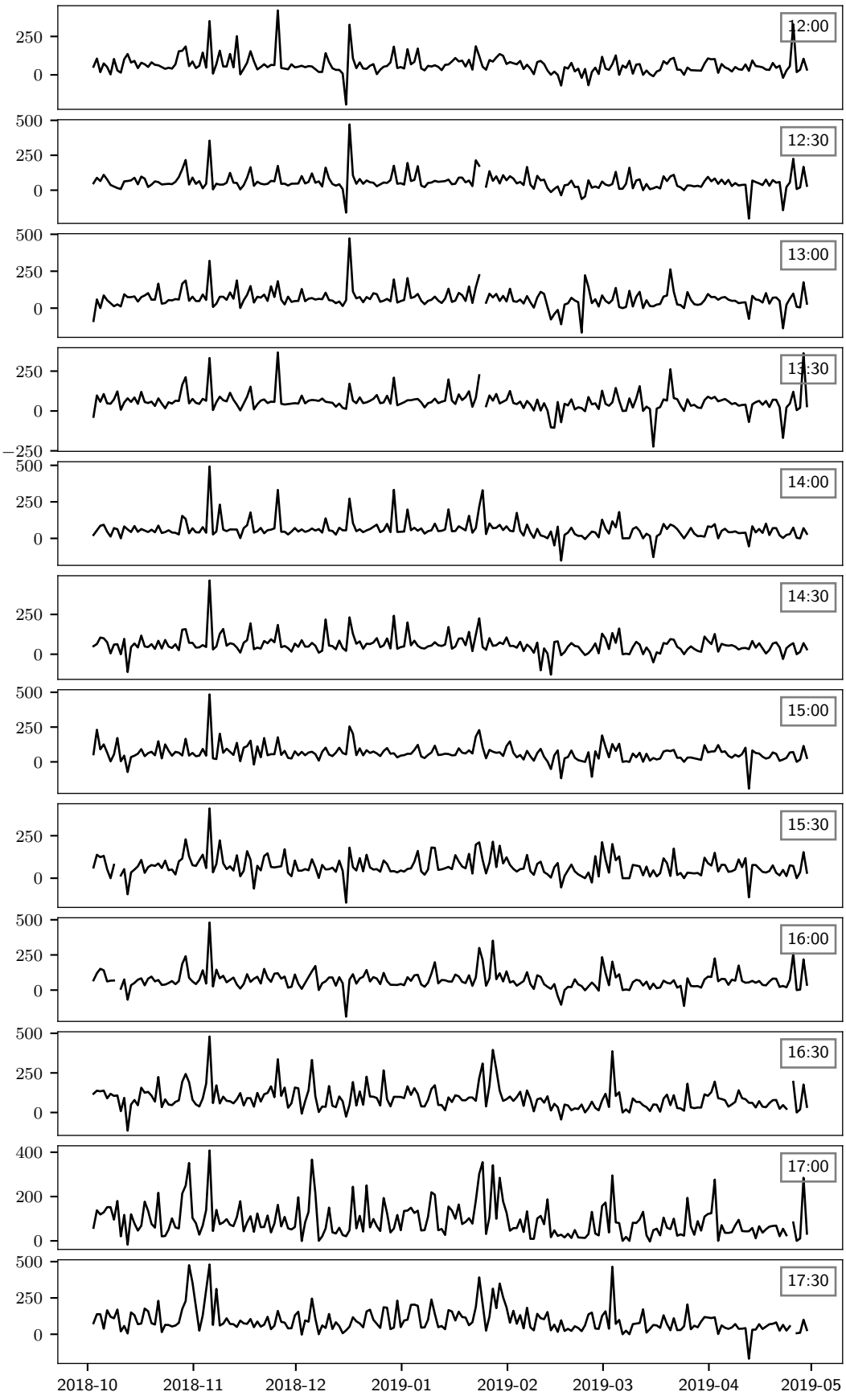


Figure 2.32: Time plot of BM prices (expressed in €/MWh) for each 30-minute trading period (part 3).

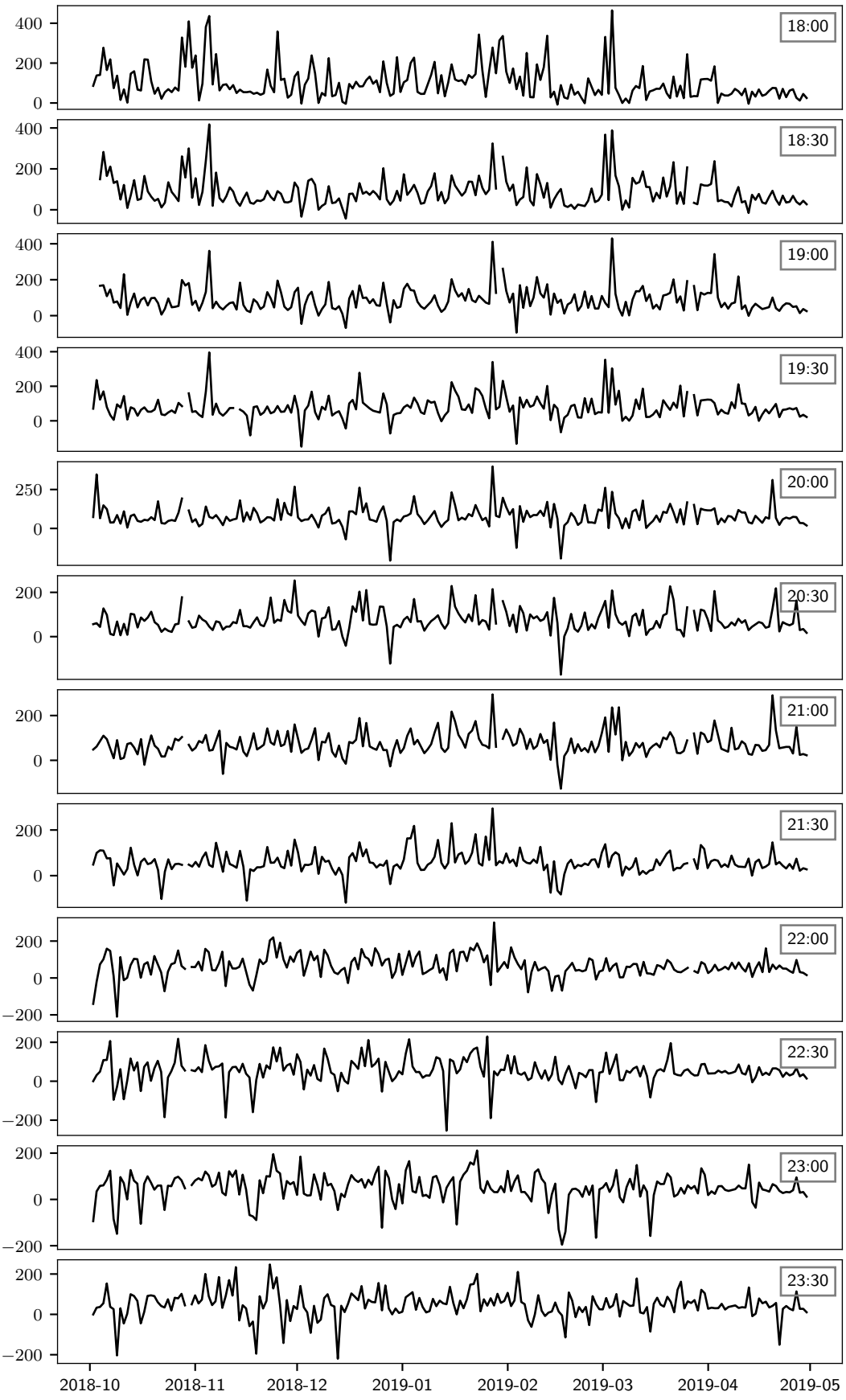


Figure 2.33: Time plot of BM prices (expressed in €/MWh) for each 30-minute trading period (part 4).

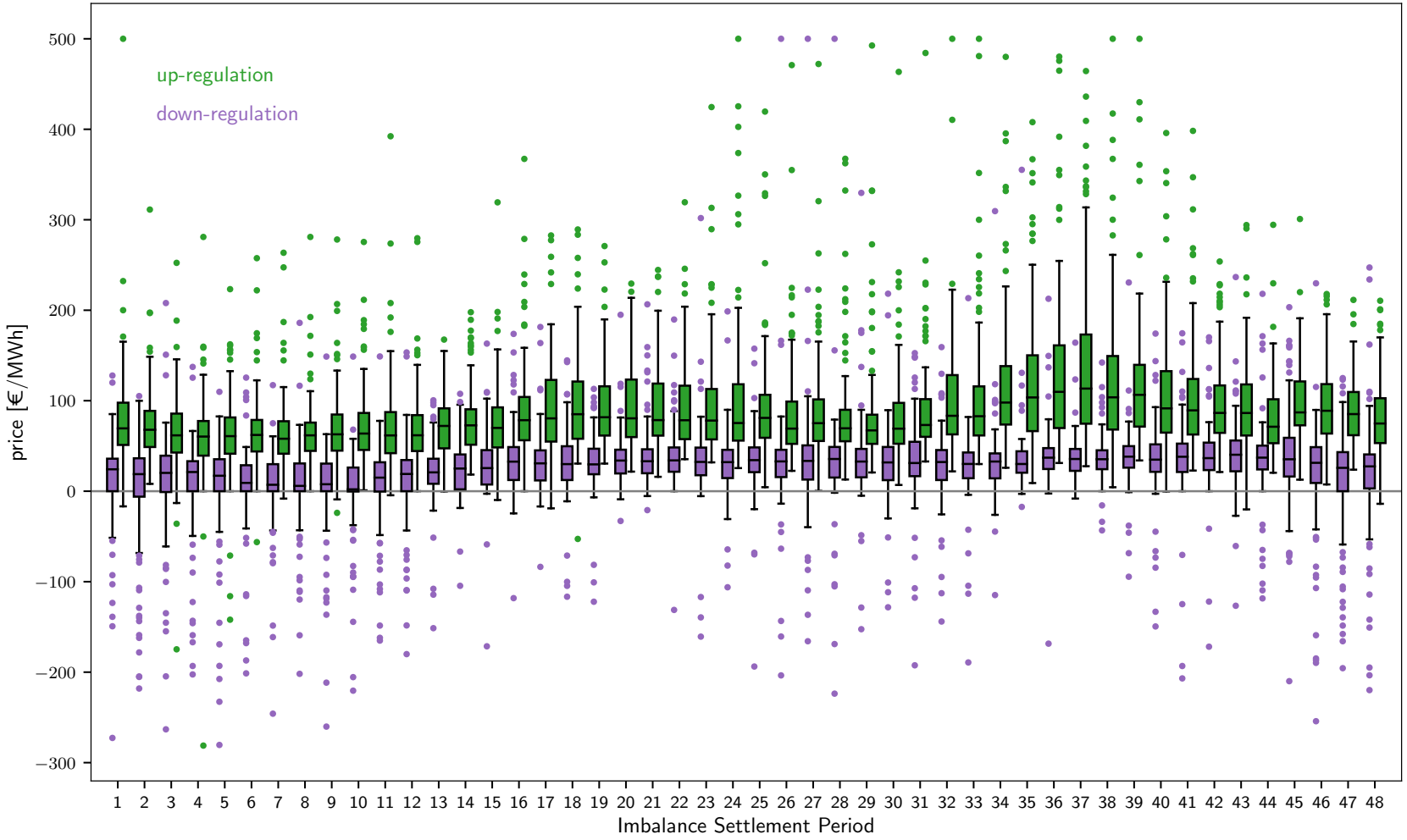


Figure 2.34: Boxplots of imbalance prices during up-regulation (green) and down-regulation (purple) for each Imbalance Settlement Period.

Zero and negative prices occurred more often at night and early morning, as shown in Figure 2.35, while prices above 150 €/MWh tended to occur more during the evening and especially around 18:00, as shown in Figure 2.36. Note that when negative imbalance prices occur, generators who have a day-ahead position are impacted depending on their own imbalance: if generators are decremented (due to a generation surplus), they are charged at a negative price, which translates to getting paid to reduce generation. If generators are incremented, on the contrary, they are paid the negative price, which results in a cost.

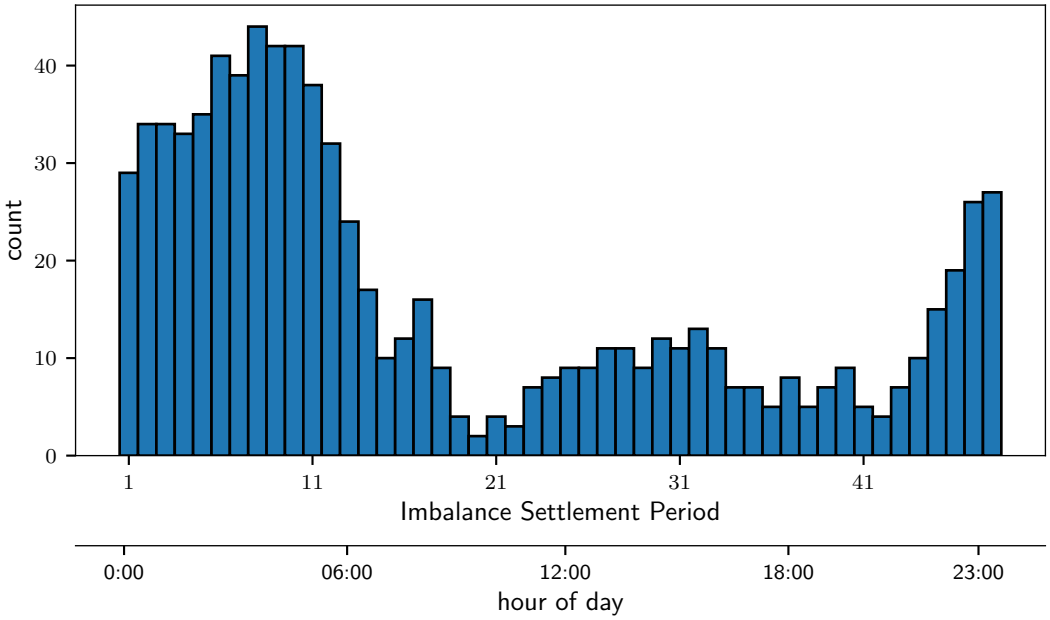


Figure 2.35: Number of instances with zero or negative prices for each Imbalance Settlement Period in the day.

As shown in Figure 2.37, the average prices and absolute value of imbalance volumes showed a decreasing trend through the last three months of the analysed period, with prices remaining more consistently at lower levels. The volatility of prices has also fallen over time, as shown by the decreasing trend in price variance from the second month of operation onwards in the top panel in Figure 2.38. The moving averages of imbalance volumes plotted in Figure 2.39 also show that over time, the tendency to have a positive imbalance (short system) in the balancing market has decreased, confirming the patterns observed in Figure 2.21. These

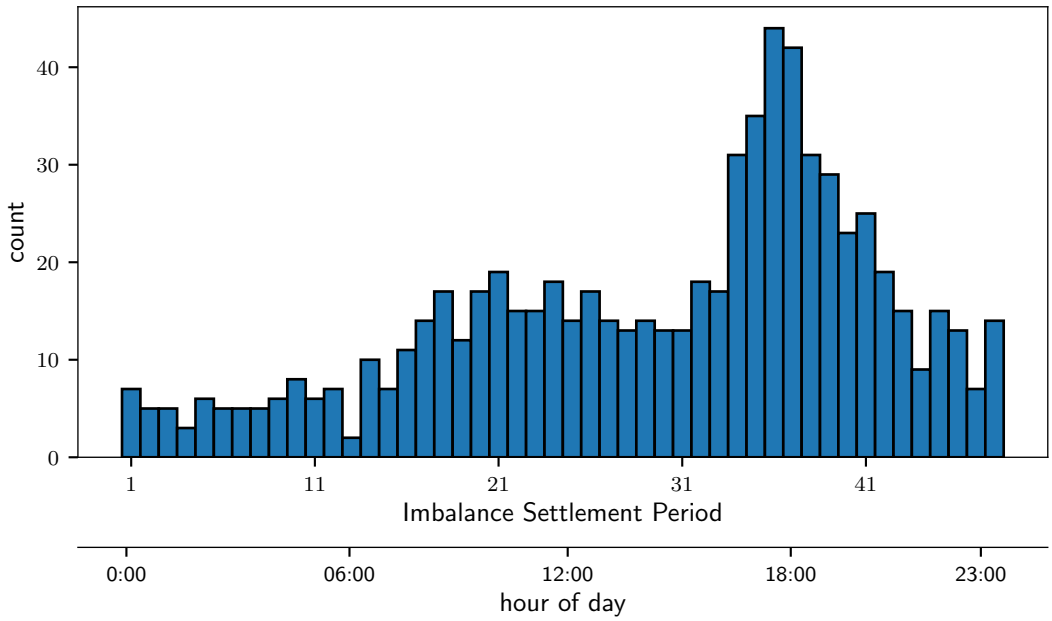


Figure 2.36: Number of instances with prices above 150 €/MWh for each Imbalance Settlement Period in the day.

trends suggest that the initially limited market experience of participants has been one of the main drivers for the high volatility of outcomes.

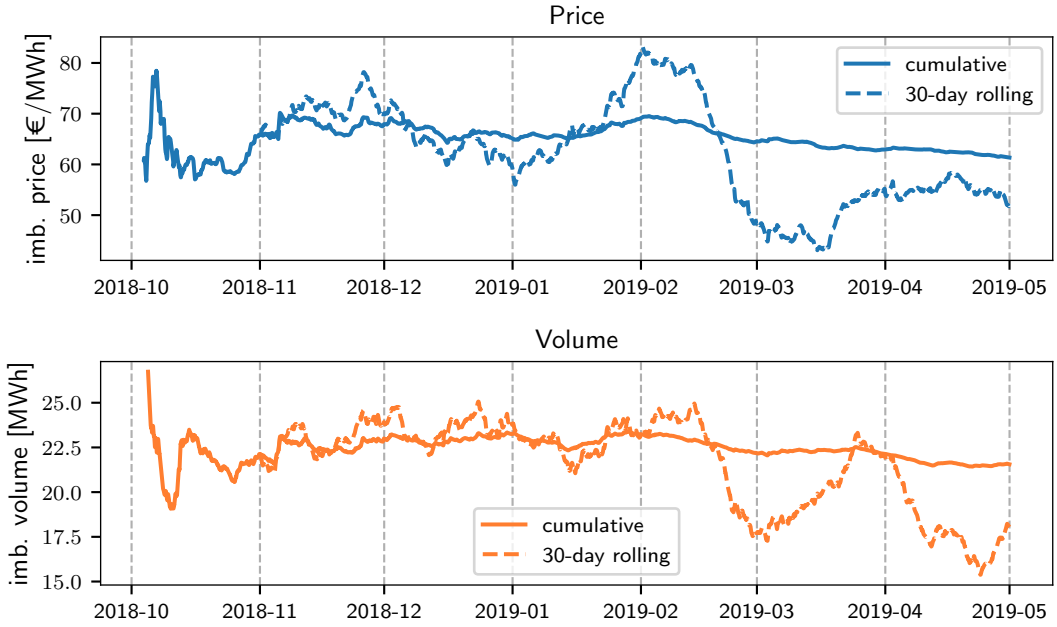


Figure 2.37: Cumulative (solid line) and 30-day rolling (dashed line) average of imbalance prices (top) and absolute imbalance volumes (bottom).

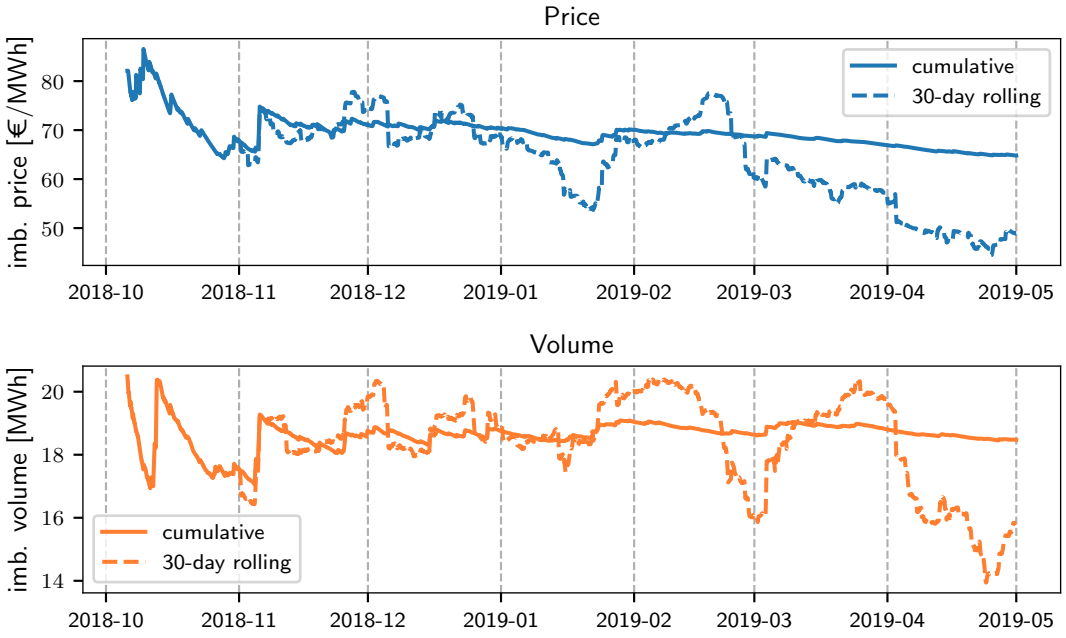


Figure 2.38: Cumulative (solid line) and 30-day rolling (dashed line) standard deviation of imbalance prices (top) and absolute imbalance volumes (bottom).

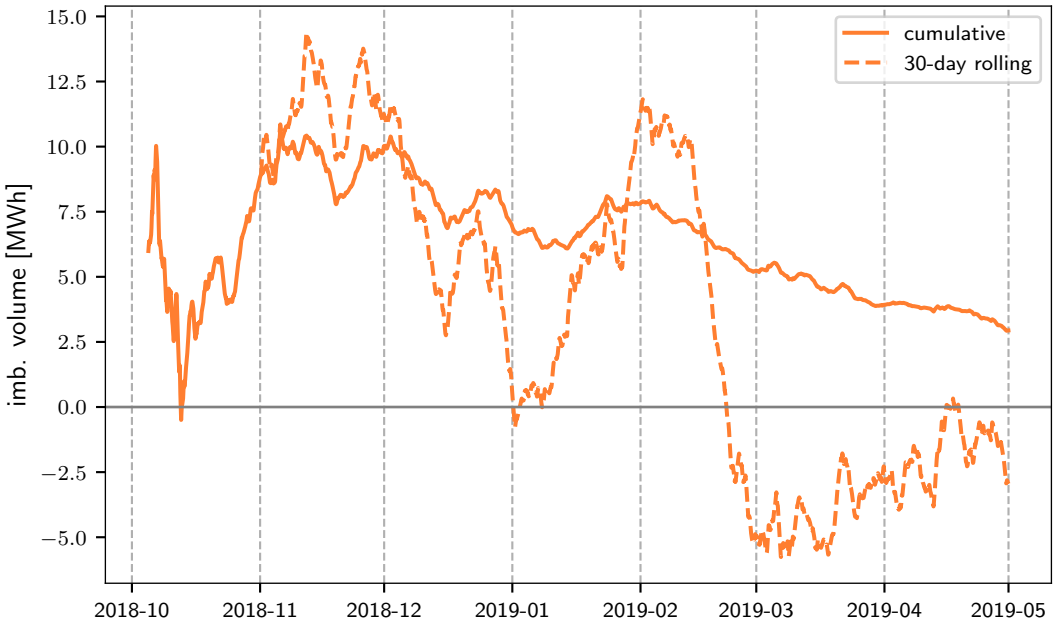


Figure 2.39: Cumulative (solid line) and 30-day rolling (dashed line) average of imbalance volumes.

2.2.2.3 Anomalous Imbalance Prices

With the implementation of the new market design, a number of issues have arisen which have generally been related to the operation of the Balancing Market. In

particular, the first months of operation of the I-SEM have been characterised by an unusually high number of counter-intuitive imbalance prices. More specifically, the imbalance price was frequently higher than the day-ahead price when the market was long, or vice versa, the imbalance price was lower than the day-ahead price when the market was short.

As explained in detail in Appendix A, the marginal energy action price (PMEA) is the most "expensive" price[1] that can contribute to setting the imbalance price, so that when the system is short, the final price can be set equal to or lower than PMEA, and when the system is long, the price can be set equal to or higher than PMEA. The existing pricing mechanism uses offers or bids to set the PMEA and thus determine the imbalance price, regardless of whether they are on the correct side of the energy imbalance of the system as a whole. During the market design phase, the scenario where the PMEA would be set by an action on the wrong side of the energy imbalance was envisaged to take place in a small number of periods, but in actual market operation this has not been the case. In a high number of 5-minute imbalance pricing periods (IPP), there have been no offers or bids available on the correct side of the energy imbalance to set the PMEA. As a result, those IPP (and often, after the averaging, the corresponding imbalance settlement periods) have been set at a counter-intuitive price, thus not providing an effective signal for participants to balance their positions.

To facilitate the analysis, the difference between the imbalance and day-ahead prices is defined as *price difference* (sometimes referred to as *price penalty* or *balancing premium*):

$$\pi^{diff} = \pi^{imb} - \pi^{DAM} \tag{2.1}$$

One would normally have that:

- price difference is zero ($\pi^{imb} = \pi^{DAM}$) when there is no regulation, i.e. when the system is balanced (imbalance volume is zero);

[1]The most expensive price is the highest priced offer during up-regulation and the lowest priced bid during down-regulation.

- price difference is positive ($\pi^{imb} > \pi^{DAM}$) during up-regulation, i.e. when the system is short (positive imbalance volume); and

- price difference is negative ($\pi^{imb} < \pi^{DAM}$) during down-regulation, i.e. when the system is long (negative imbalance volume).

Table 2.1 summarises the possible balancing market states and the corresponding imbalance volume, need for regulation, imbalance price and price difference.

system	**imb. volume**	**regulation**	**price difference**
balanced	zero	no	$\pi^{imb} = \pi^{DAM} \rightarrow \pi^{diff} = 0$
short	positive	up	$\pi^{imb} > \pi^{DAM} \rightarrow \pi^{diff} > 0$
long	negative	down	$\pi^{imb} < \pi^{DAM} \rightarrow \pi^{diff} < 0$

Table 2.1: Balancing Market states and corresponding market quantities.

In Figure 2.40, price difference is plotted against the corresponding imbalance volume for each trading period during the first seven months of I-SEM operation. The numerous instances of anomalous imbalance prices, corresponding to the points in the second and fourth quadrants, are highlighted in orange.

During up-regulation, price anomalies were very frequent, as shown in the histogram in Figure 2.41, with price difference being negative 37% of the times. The scatter plot in Figure 2.42 shows that there was no particular correlation between the occurrence of price anomalies (again, highlighted in orange), and the value of day-ahead prices.

During down-regulation, price anomalies were less frequent but still noticeably high, as shown in the histogram in Figure 2.43, with price difference being positive 14% of the time. Once more, no particular correlation was found between the occurrence of price anomalies (in orange), and the value of day-ahead prices, as visible from the scatter plot in Figure 2.44.

Figure 2.45 shows the cumulative and 30-day rolling frequency of counter-intuitive prices since the start of market operations: a significant drop around

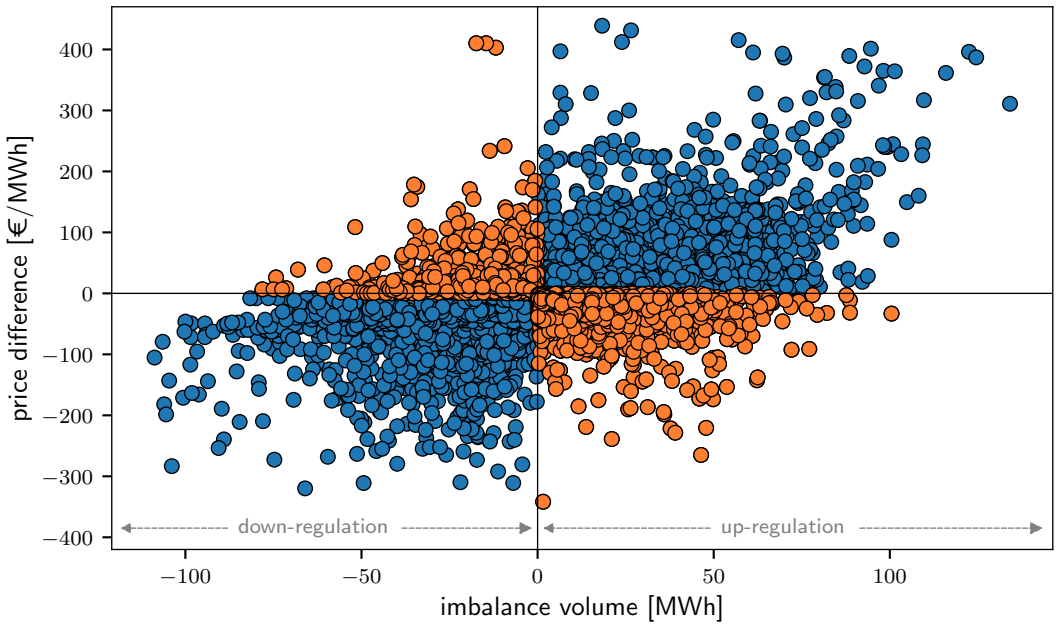


Figure 2.40: Scatter plot of price difference vs imbalance volume. The instances of anomalous imbalance prices are highlighted in orange.

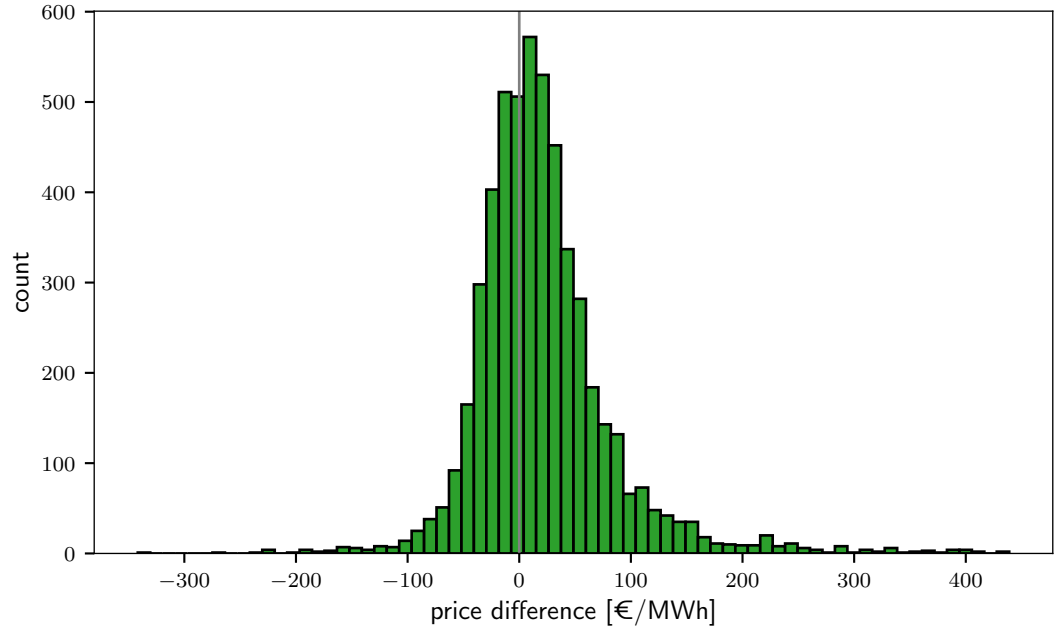


Figure 2.41: Histogram of price difference during up-regulation.

February and an overall decreasing trend are evident, with anomalous prices occurring on average 25% of the time after seven months. The trend is due to the falling number of anomalous prices when the system is short, as shown in Figure 2.46, where the cases of up- and down-regulation are visualised separately.

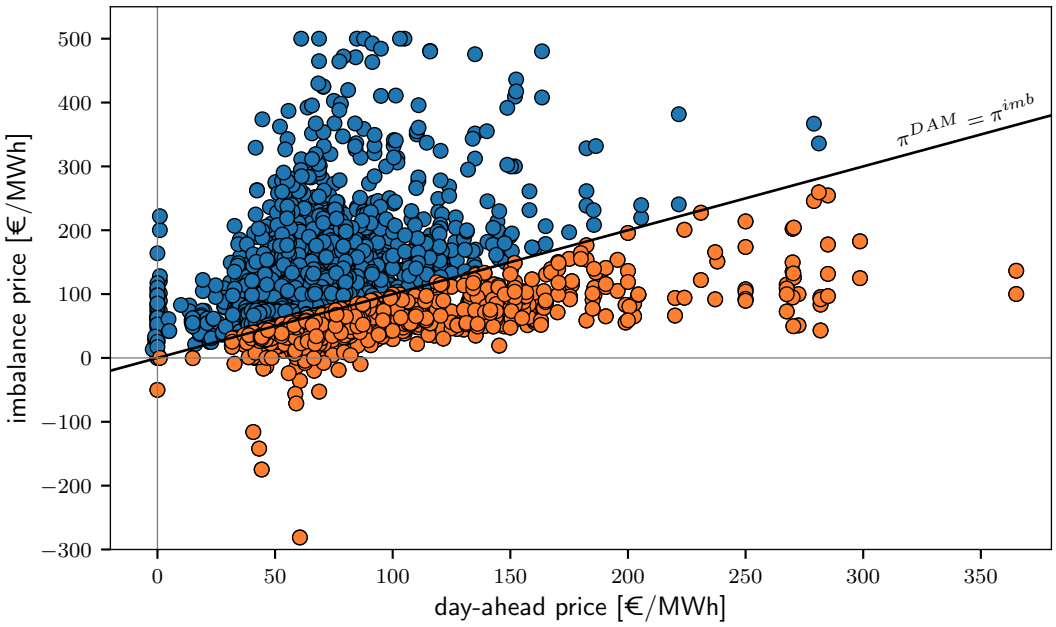


Figure 2.42: Scatter plot of imbalance prices vs day-ahead prices during up-regulation. Price anomalies are highlighted in orange, while the diagonal indicates the 45-degree line where $\pi^{DAM} = \pi^{imb}$.

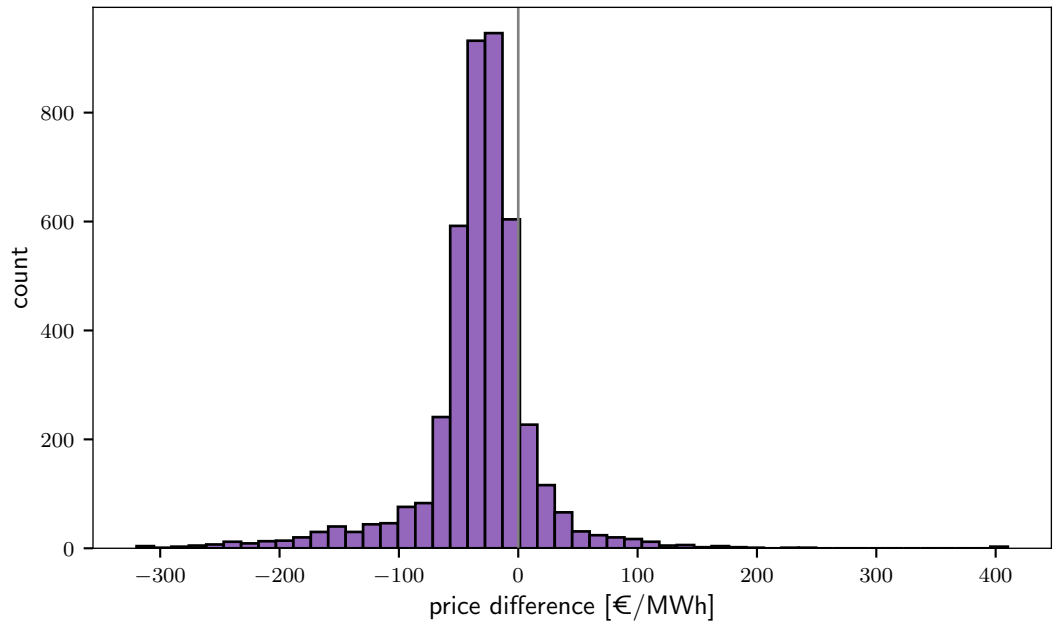


Figure 2.43: Histogram of price difference during down-regulation.

The graph also shows how price anomalies are consistently more frequent during up-regulation.

The possibility of a daily pattern in the occurrence of anomalous imbalance

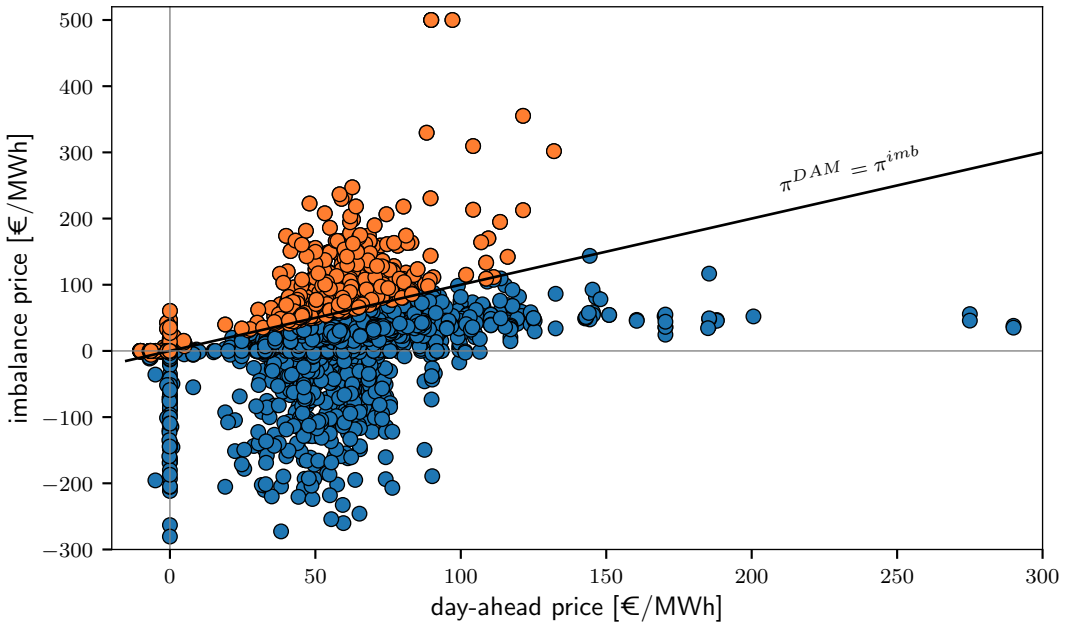


Figure 2.44: Scatter plot of imbalance prices vs day-ahead prices during down-regulation. Price anomalies are highlighted in orange, while the diagonal indicates the 45-degree line where $\pi^{DAM} = \pi^{imb}$.

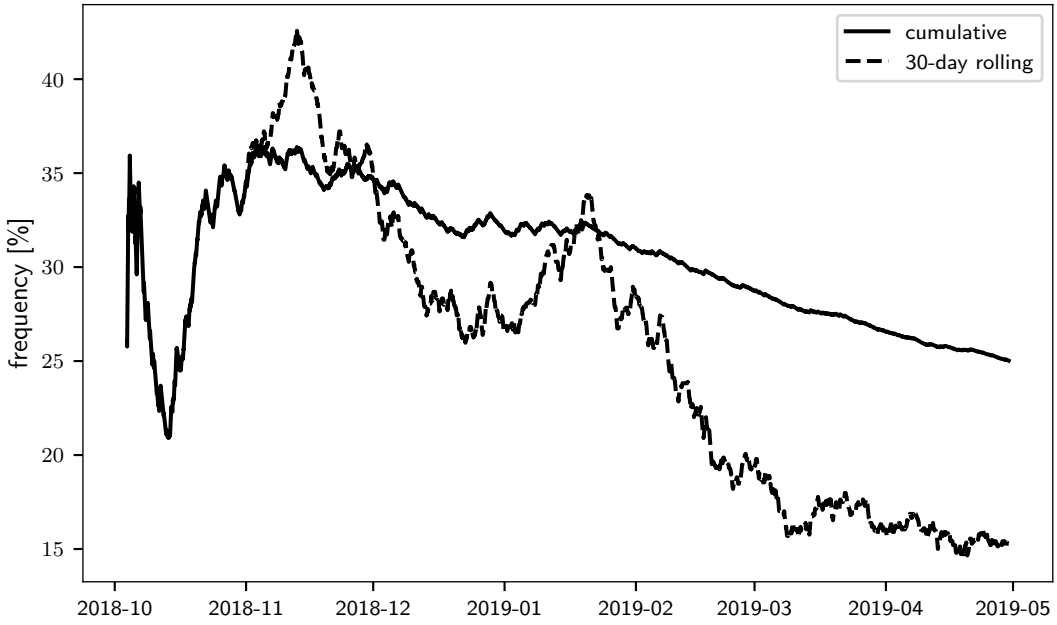


Figure 2.45: Cumulative (solid line) and 30-day rolling (dashed line) frequency of imbalance prices anomalies.

prices was also investigated. The frequency and count of instances with counter-intuitive prices for each imbalance settlement period are plotted in Figure 2.47, showing that anomalies tend to occur slightly more during the evening, but overall

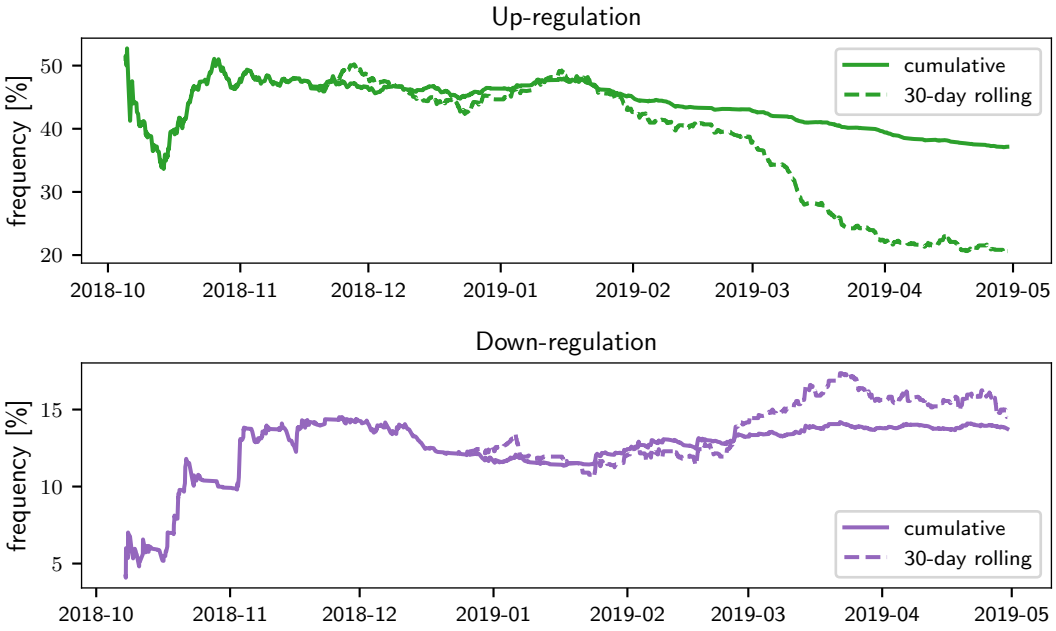


Figure 2.46: Cumulative (solid line) and 30-day rolling (dashed line) frequency of imbalance prices anomalies during up-regulation (top) and down-regulation (bottom).

there is no obvious temporal trend.

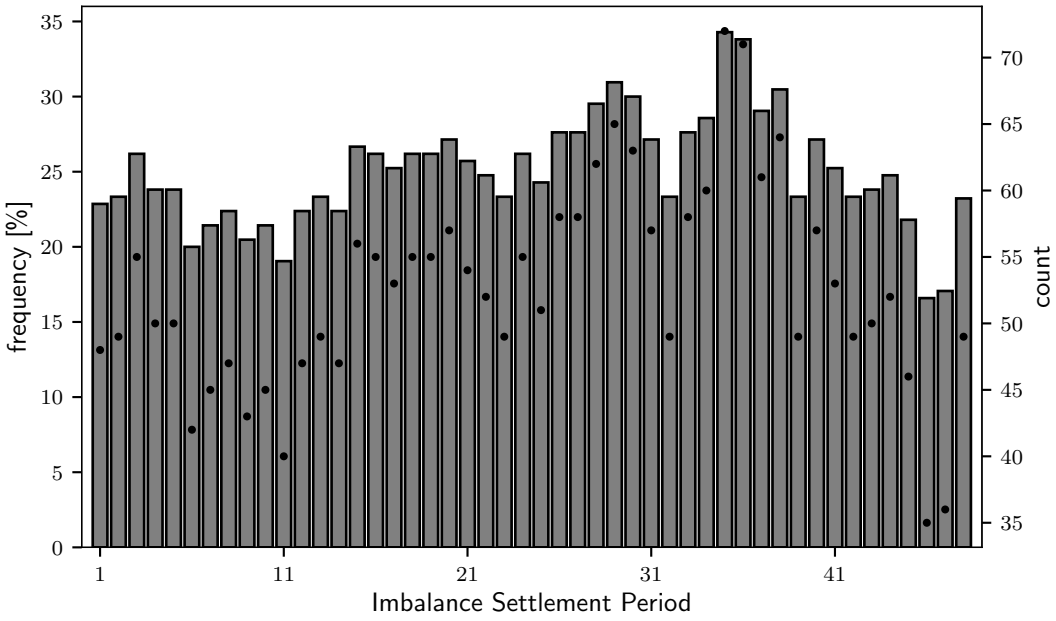


Figure 2.47: Frequency of anomalous imbalance prices for each Imbalance Settlement Period. The black dots indicate the count of instances, with values readable on the right y-axis.

Detailed studies carried out by the SEMO [38] found that the application of

locational constraints in the imbalance price calculation where these same restrictions are not present in the ex-ante markets was one of the major contributing factors to the numerous counter-intuitive prices registered since the start of the market operations under the new arrangements. Locational constraints are not considered in the ex-ante market design of the I-SEM. However, these same constraints were able to influence the calculation of the imbalance price. In fact, any units whose output could be identified as constrained for non-energy reasons — thus including the units subject to transmission system based constraints — were to be flagged out of the pricing process (see the Flagging step in Appendix A). This has impacted the formation of imbalance prices to a level beyond that expected during the market design, leading to extreme price events and counter-intuitive prices in the Balancing Market. Following the SEM Committee Decision for the Regulatory Authorities in relation to the "Modification Proposal Mod_09_19" [38], locational constraints have been removed from imbalance pricing calculation from the 2nd May 2019.

2.2.2.4 Impact of Wind Power Penetration on Market Quantities

The *penetration* of an energy source is the fraction of the total demand that the source covers. Due to the variable and uncertain nature of renewable energy sources, when their penetration in a power system is at significant levels, the operation and outcomes of the electricity markets can be impacted in various ways.

The variability of stochastic generation due to meteorological reasons can be very high within the time scales of relevance in both the day-ahead and balancing markets. For example, wind power generation can change substantially from one hour to the next, thus affecting the variability of system net demand and eventually, that of the wholesale electricity prices. On the other hand, unpredicted variations at shorter time scales (from 5 to 30 minutes) have to be compensated through the balancing market. Consequently, stochastic generation and variability thereof can impact the need for and direction of balancing actions, as well as the magnitude of

the net system imbalance.

The level of stochastic renewable energy penetration also has a direct impact on the clearing of day-ahead auctions via the so-called "merit-order effect", presented in detail by Sensfuß et al. in [39]. As discussed in Section 2.1.1, prices and volumes in the day-ahead market are determined through auction-based mechanisms. Each generation bid is defined by a quantity-price pair, where the price depends directly on the short-run marginal cost of the generating plant, i.e. the cost of producing another unit of energy. Bids are then ranked in order of increasing price following the merit-order principle and form the market supply curve. Renewable energy producers generally have no short-run marginal costs and thus bid at zero price. Consequently, renewable energy sources like wind and solar are prioritised in this competitive market framework because their offers are the first to be accepted (unless other generators bid at negative prices). When clearing the day-ahead market, the result is that increased levels of renewable energy penetration "shift" the supply curve to the right, yielding higher cleared volumes and lower system marginal prices.

In fact, what impacts the electricity markets is not the actual generation from stochastic sources, but rather the forecast generation and its associated uncertainty [40] together with the zero marginal cost of the renewable energy. Indeed, bids and offers placed in the day-ahead market are based, among other factors, on the knowledge at that time about future power generation. Approaching delivery time, then, the need for balancing stems from the deviation between day-ahead accepted bids and the actual generation. Therefore, it is the forecasting errors that affect the market quantities in the balancing market, rather than the actual generation from renewable energy sources. The variable of interest then becomes the *predicted wind power penetration*, which for a given trading period is defined as the ratio between the forecast wind power generation in the system and forecast demand.

During the period under analysis, wind power generation on the system was

1,629 MW on average, with a maximum of 3,950 MW and a minimum of 31 MW. Wind power penetration, i.e. the ratio of wind power generation to total demand, was equal to 38.0% on average and ranging from a minimum of 0.8% to a maximum of 85.2% for a single period, as shown in Figure 2.48. Note that this includes exports of wind energy to GB via the interconnectors, hence values can be higher than the 65% SNSP limit discussed in Section 1.2. The average values of predicted wind

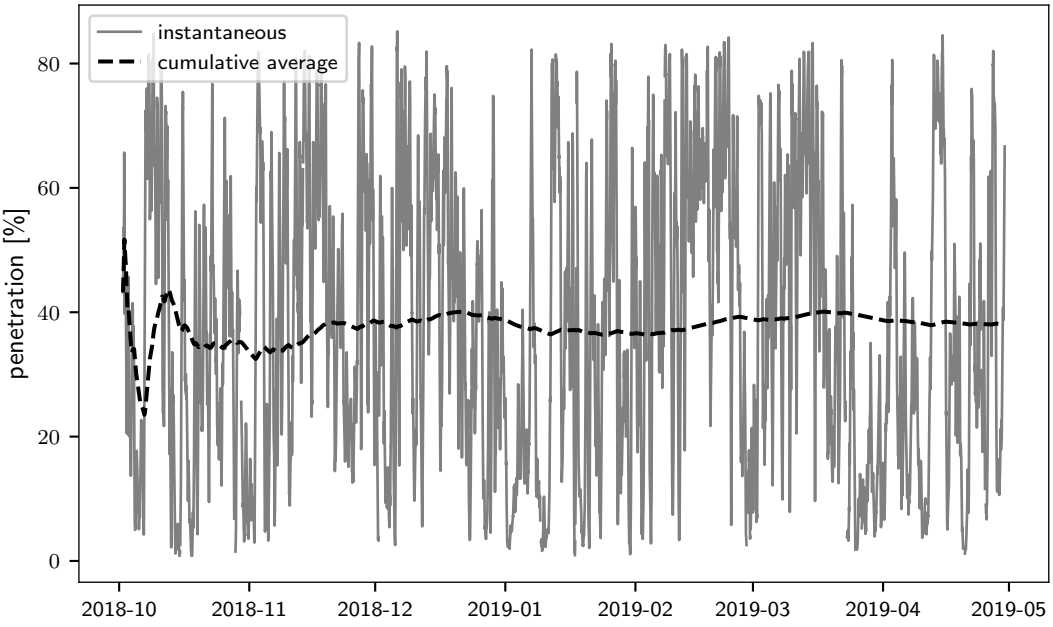


Figure 2.48: Instantaneous (solid) and cumulative average (dashed) observed penetration of wind power generation.

power penetration for each Imbalance Settlement Period (half-hour) are shown in Figure 2.49. It is not surprising to see how the trend is somewhat the inverse of average demand (cf. Figure 2.16): penetration is highest at night with peaks in the early morning, then decreases during the morning and reaches the lowest values in the evening.

The aim of this thesis is to develop forecast models for the value of wind energy. Therefore, it is important to understand the impact of system-level wind generation on the market quantities of relevance for wind energy trading. In the remainder of this Section, statistical analyses are carried out to investigate the impact of predicted wind power penetration, hereinafter also *predicted penetration*, on three

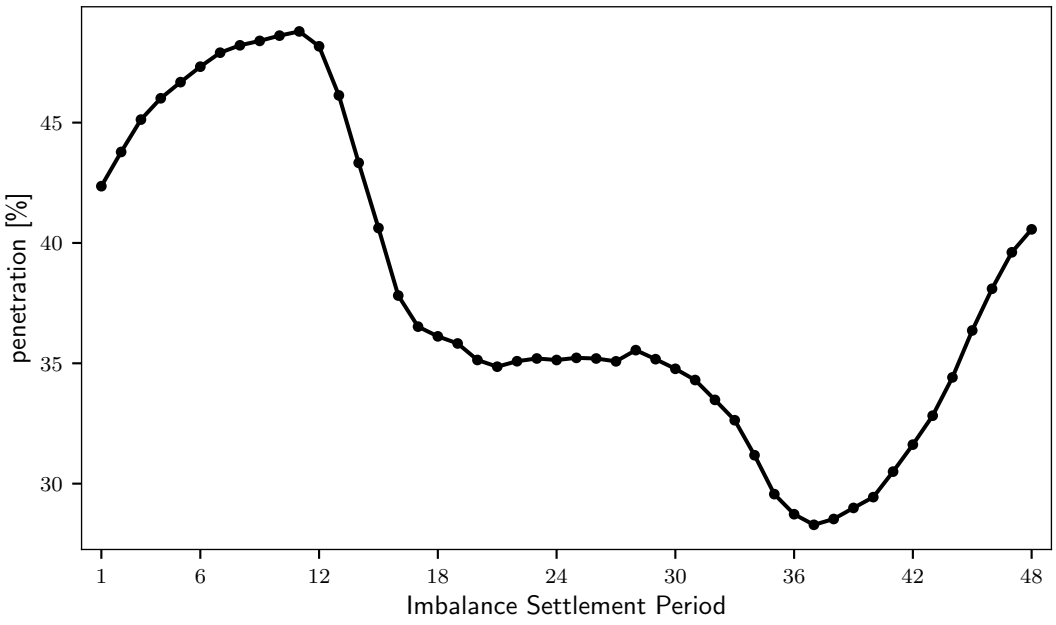


Figure 2.49: Average predicted wind power penetration for each ISP.

different quantities in the Irish electricity market:

- day-ahead prices,
- balancing direction (or imbalance sign), and
- imbalance price anomalies.

These ex-post analyses are carried out on the I-SEM market data described in the previous Sections of this Chapter, covering a seven-month period from the start of market operations in October 2018 to the end of April 2019. Day-ahead and balancing market quantities have hourly and half-hourly temporal resolution, respectively. Forecasts of wind power generation and demand for the all-island of Ireland are point forecasts issued before day-ahead gate closure for each trading day of interest.

Impact on Day-Ahead Prices

The first empirical study aims at investigating the impact of predicted wind power penetration on day-ahead prices. Because of the merit-order effect described earlier, prices are expected to decrease as penetration increases.

As highlighted in Section 2.2.2.1, the hour of the day has a strong influence on wholesale prices. Therefore, to establish the empirical relationships between day-ahead prices and predicted penetration, the price time series was segmented into intervals based on the hour of the day to obtain better fits. Data were split into three 8-hour groups as follows:

- from hour 7 to 14 (H07–H14),
- from hour 15 to 22 (H15–H22), and
- from hour 23 to 6 (H23–H06).

These ranges were chosen using a heuristic approach based on similar levels of penetration and price (depicted in Figure 2.50) and correlation between the two variables (shown in Figure 2.51).

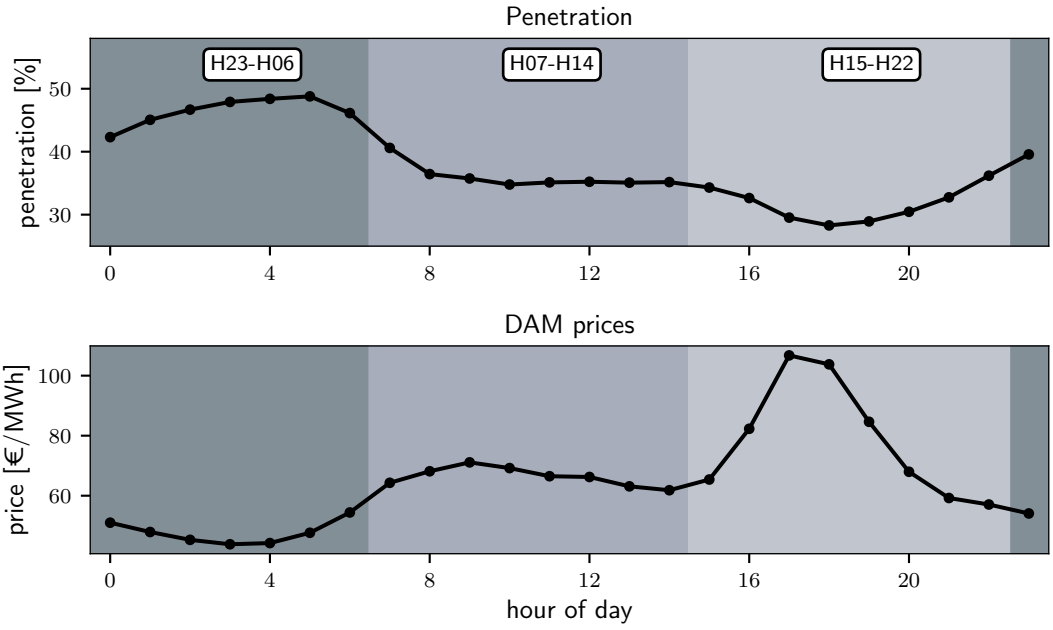


Figure 2.50: Average predicted wind power penetration (top) and day-ahead prices (bottom) for each hour of the day. Each time group is highlighted in a different colour.

For each time group, price values were binned by wind penetration level into intervals of width 0.10 (i.e. 10% wind penetration as a percentage of total demand). The empirical mean, standard deviation, and quantiles with nominal levels 0.05 and

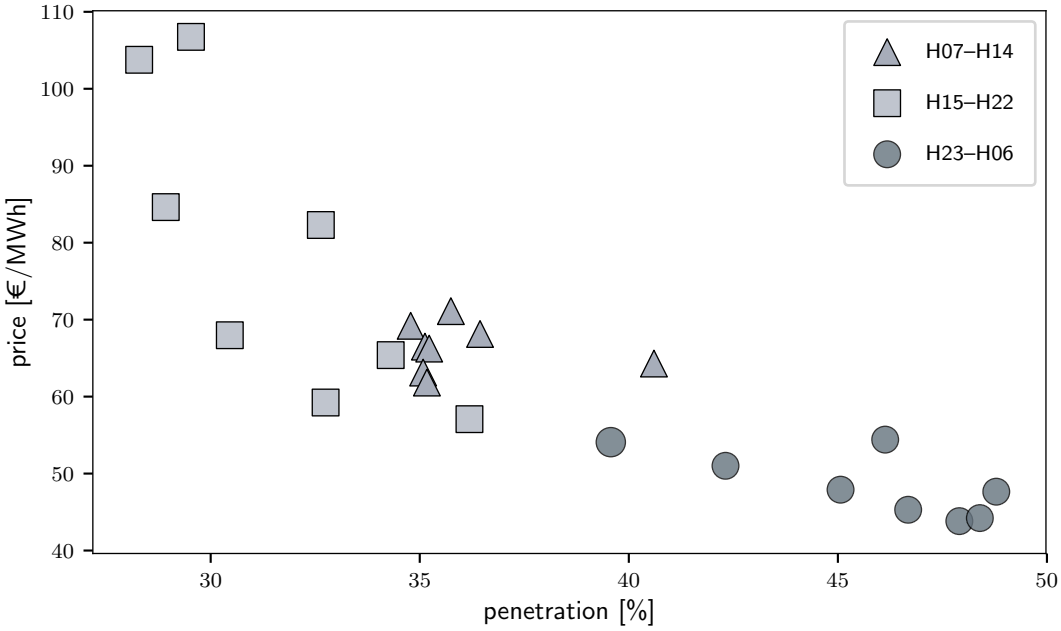


Figure 2.51: Scatter plot of the hourly averages of predicted wind power penetration against day-ahead price. Each time group is plotted in a different colour.

0.95 were then calculated for each interval and quadratic spline functions fitted to these values to find the trends presented below. Figures 2.52 to 2.54 illustrate the results of the analyses for the three groups, with the observed values plotted as grey dots and the fitted mean and quantile trends as black and orange curves, respectively. The mean and quantile values for various penetration levels are also reported in Table 2.2.

predicted penetration	**mean**			$\mathbf{q_{5\%}}$			$\mathbf{q_{95\%}}$		
	07-14	15-22	23-06	07-14	15-22	23-06	07-14	15-22	23-06
0%	86.00	110.91	59.71	54.13	53.70	44.05	141.31	265.75	76.00
20%	67.33	78.17	57.75	42.16	47.41	40.21	93.36	132.54	75.71
40%	62.09	67.00	54.61	44.65	43.92	37.35	76.87	95.05	68.43
60%	54.42	54.41	47.45	35.79	35.00	31.83	69.29	80.77	63.72
80%	31.09	33.65	19.03	–1.57	22.48	–5.28	48.60	41.87	47.27
100%	0.00	-	2.66	0.00	-	–0.06	0.00	-	8.00

Table 2.2: Mean, 0.05- and 0.95-quantile of day-ahead prices for each group of hours at various levels of predicted wind power penetration. Values are expressed in €/MWh.

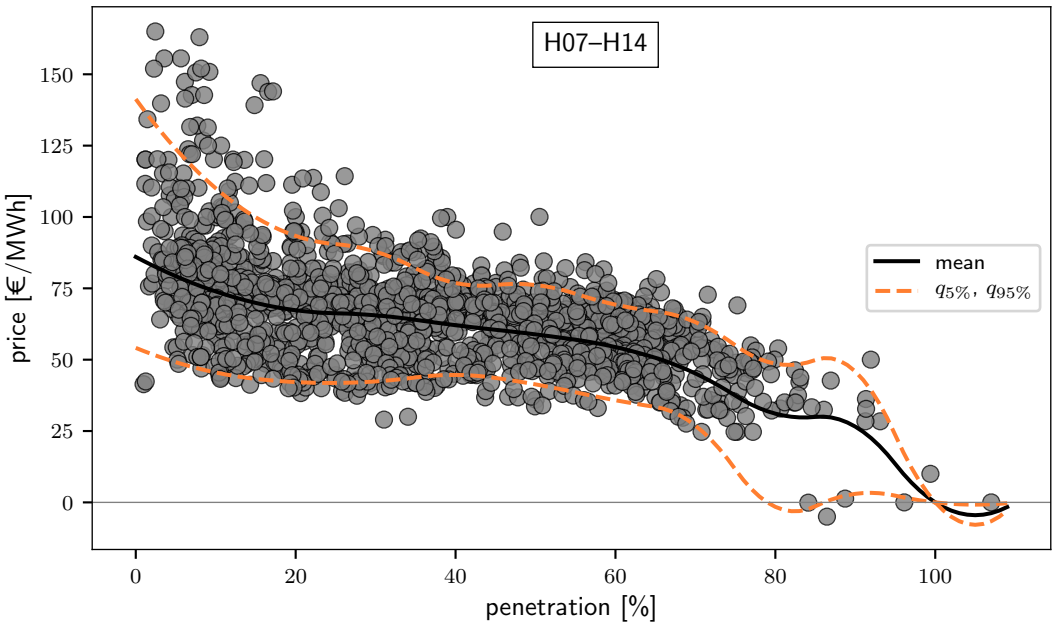


Figure 2.52: Impact of predicted wind power penetration on day-ahead prices between hours 7 and 14.

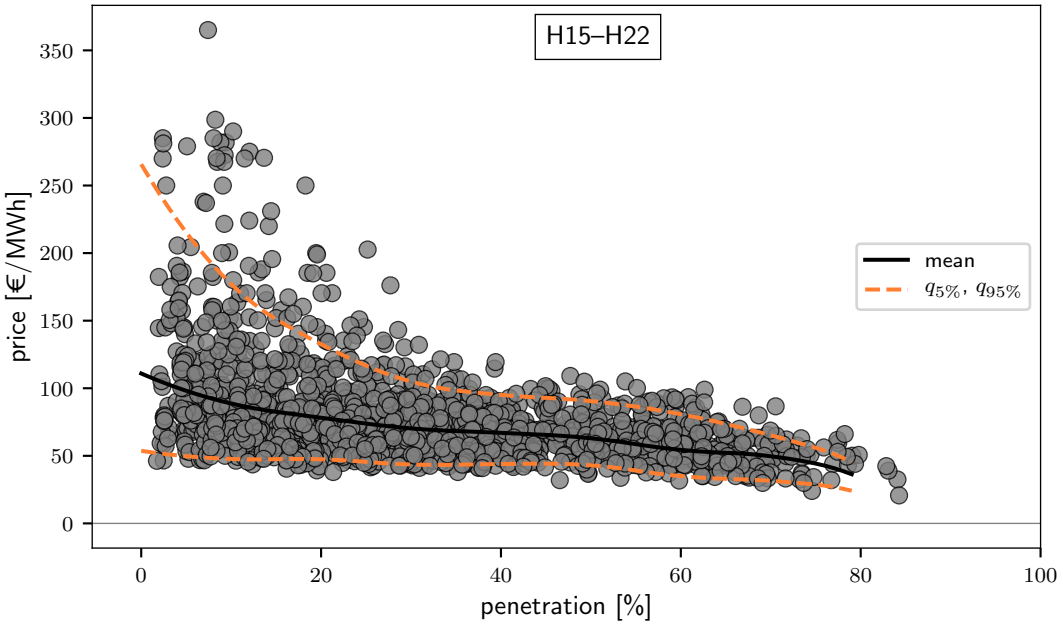


Figure 2.53: Impact of predicted wind power penetration on day-ahead prices between hours 15 and 22.

It is evident from Figures 2.52 and 2.53 that high prices occur less and less for larger values of penetration, as visible from the steep initial drop of the 0.95-quantile trends. At night, on the other hand, prices are often equal to zero for

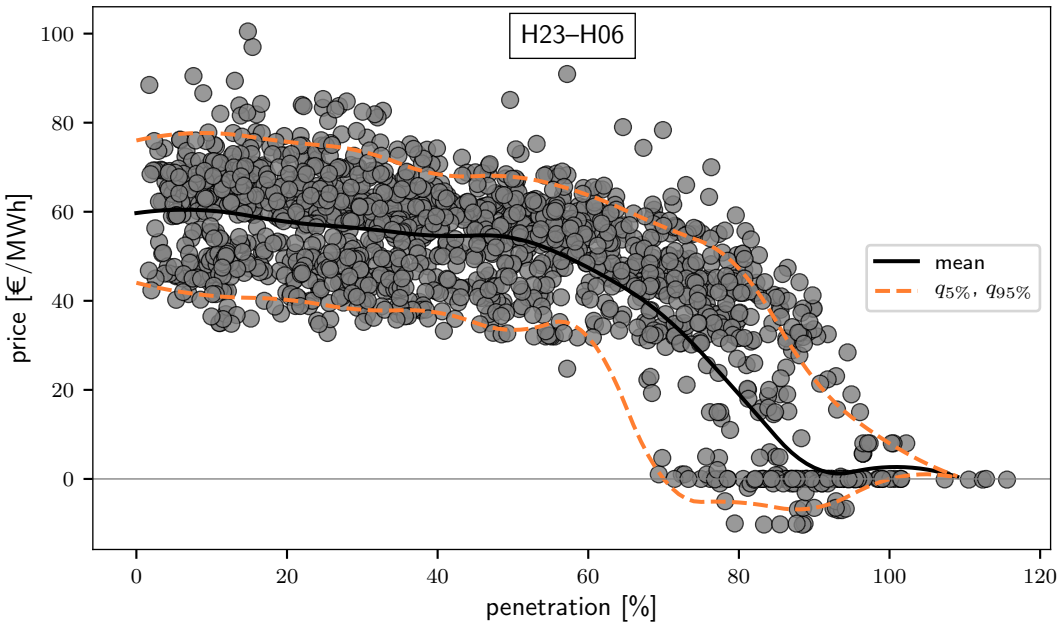


Figure 2.54: Impact of predicted wind power penetration on day-ahead prices between hours 23 and 6.

penetration levels above 70% and at times take even negative values, as shown in Figure 2.54.This occurs because thermal plants with long start-up times might want to stay on during such periods of low demand and high wind generation, and can be willing to pay to operate at their minimum output in these intervals. Note that due to the lower demand, night time is the period when predicted wind power penetration goes above 80% most often and reaches the highest values.

The mean trends in day-ahead prices reported in Table 2.2 are shown together in Figure 2.55 to allow a better comparison across the three time groups. As expected, average prices decrease for increasing levels of predicted wind power penetration for all hours. The extent of this reduction, however, does change with the time of the day as well as the level of penetration. During late afternoon and evening hours — corresponding to the H15–H22 group — average prices drop steeply for penetration up to 20% and keep decreasing at a lower rate. In the morning and early afternoon hours — that is, the H07–H14 group — average prices show a very similar behaviour. The observed behaviour for values of penetration above 80% should be interpreted with caution, since the number of observations used to

derive the trend was particularly low within those intervals. During night hours — corresponding to the H23–H06 group — the reduction in average price is moderate for penetration levels up to 50% but becomes more and more accentuated for larger values.

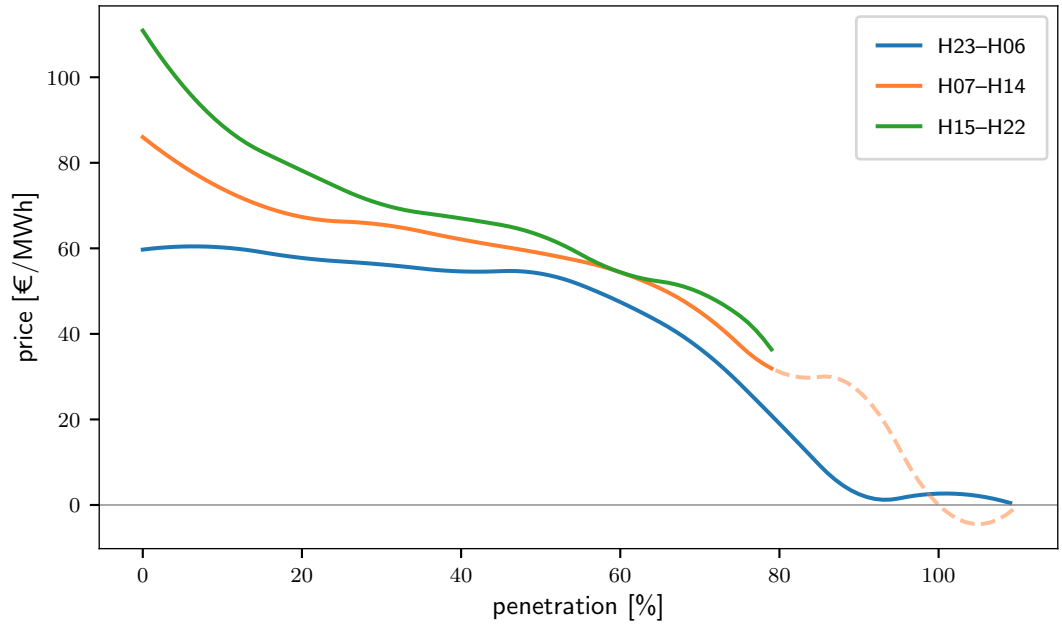


Figure 2.55: Average day-ahead price trends as a function of predicted wind power penetration for each group of hours. Transparent dashed lines indicate that the number of observations used to derive the trend was low and that the pattern should be interpreted with caution.

Another interesting effect is on the variance of prices, as shown in Figure 2.56, where the trend of the standard deviation of prices for each time period is plotted against the predicted penetration. During the morning, afternoon, and evening hours, the variability of prices decreases for increasing levels of predicted penetration, sharply at first and then more slowly. Again, the trend of the H07–H14 group for values above 80% should be dealt with caution. At night, on the contrary, the variance of prices is relatively constant for penetration levels up to 60%, while for larger values it increases noticeably as a consequence of the frequent instances of zero and negative prices.

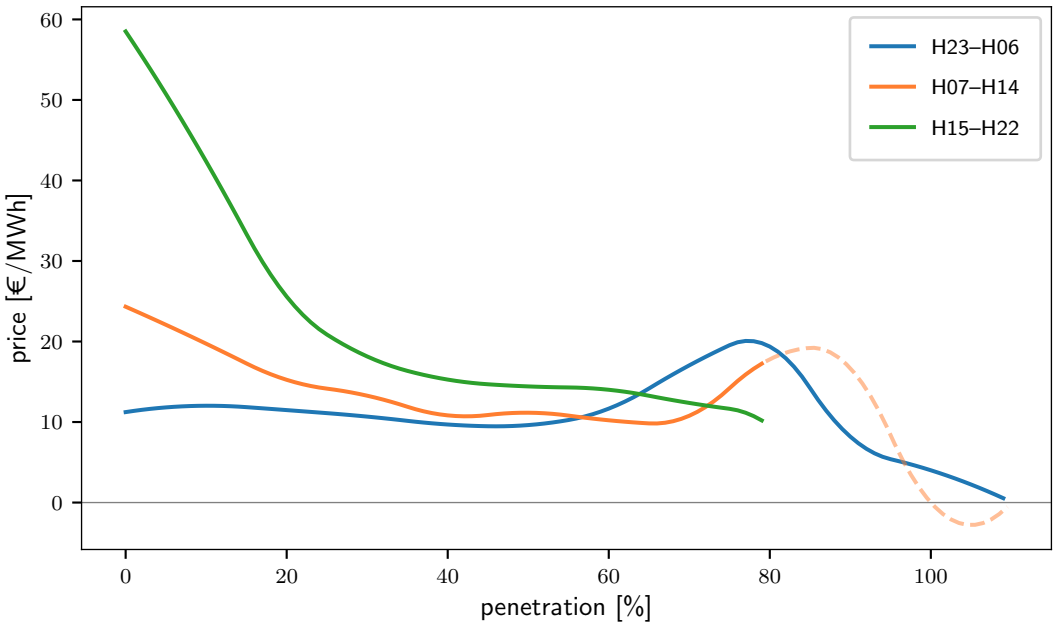


Figure 2.56: Standard deviation trends of day-ahead prices as a function of predicted wind power penetration for each group of hours. Transparent dashed lines indicate that the number of observations used to derive the trend was low and the pattern should be interpreted with caution.

predicted penetration [%]	**st.dev.**		
	07-14	15-22	23-06
0	24.32	58.51	11.23
20	15.24	25.58	11.48
40	10.83	15.27	9.69
60	10.22	14.01	11.66
80	17.79	9.55	19.37
100	-	-	4.01

Table 2.3: Standard deviation of day-ahead prices for each group of hours at various levels of predicted wind power penetration. Values are expressed in €/MWh.

Impact on Imbalance Direction

A second ex-post analysis was carried out to investigate the impact of predicted wind power penetration on the conditional probability of the imbalance direction, stet the system is not balanced. The study is performed following a similar approach to that applied on day-ahead prices. The analyses in Section 2.2.2.2 highlighted the tendency to have a positive imbalance during certain ISPs, i.e. trading periods (cf. Figure 2.20). Therefore, the half-hourly time series of imbalance vol-

umes was segmented based on the time of the day. For continuity with the previous study, the same ranges were used, that is:

- from hour 7 to 14 (H07–H14), thus including ISPs 15 through 30;
- from hour 15 to 22 (H15–H22), i.e. ISPs 31 through 46; and
- from hour 23 to 6 (H23–H06), i.e. ISPs 47, 48, and 1 through 14.

Having excluded the periods where the system was balanced and regulation was not needed, data in each group were binned by predicted penetration level into 0.10-wide intervals. For each interval, the number of trading periods with positive and negative imbalance volumes were counted and the results for the three time groups are presented in Figures 2.57, 2.59 and 2.61. Recall that a positive/negative imbalance volume means the system was short/long and, over the trading period, there was a net need for up/down-regulation. The observed frequency of up-regulation was then calculated for each penetration interval, and quadratic spline functions fitted to these values to find the trends plotted in Figures 2.58, 2.60 and 2.62. A frequency of 50% indicates that up- and down-regulation are equally frequent. Therefore, up-regulation is more frequent when the curve is above the dashed grey line in the graphs, and vice versa when the curve is below. Note that where the green curve is dashed and transparent, this indicates that the number of observations used in the calculations were low and that limited significance should be given to the resulting trends.

The results show that overall up-regulation is prevalent at low values of predicted penetration, and then become less frequent as penetration gets larger. Possibly, this could be caused by wind units bidding cautiously and selling less than their expected generation in the ex-ante markets. As penetration levels increase, this leads to the system being long and a growing need for down-regulation.

Again, the specific trends vary depending on the time of the day. Between hours 7 and 14 (Figure 2.58), the observed frequency of up-regulation decreases almost linearly for increasing levels of predicted penetration; up-regulation is prevalent for

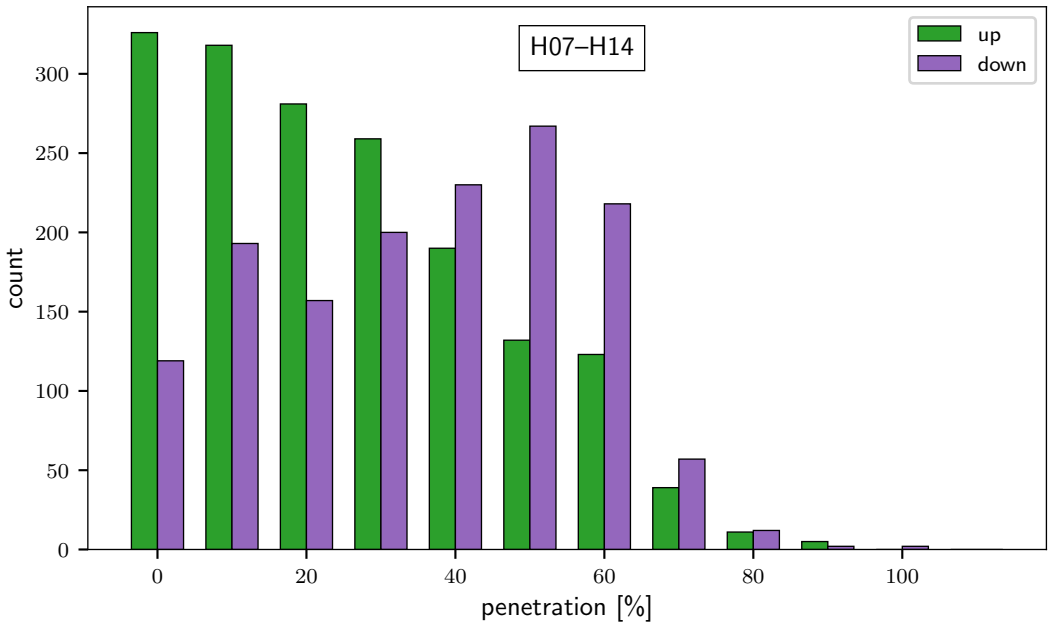


Figure 2.57: Number of trading periods with net up-regulation (green) and down-regulation (purple) as a function of predicted wind power penetration between hours 7 and 14.

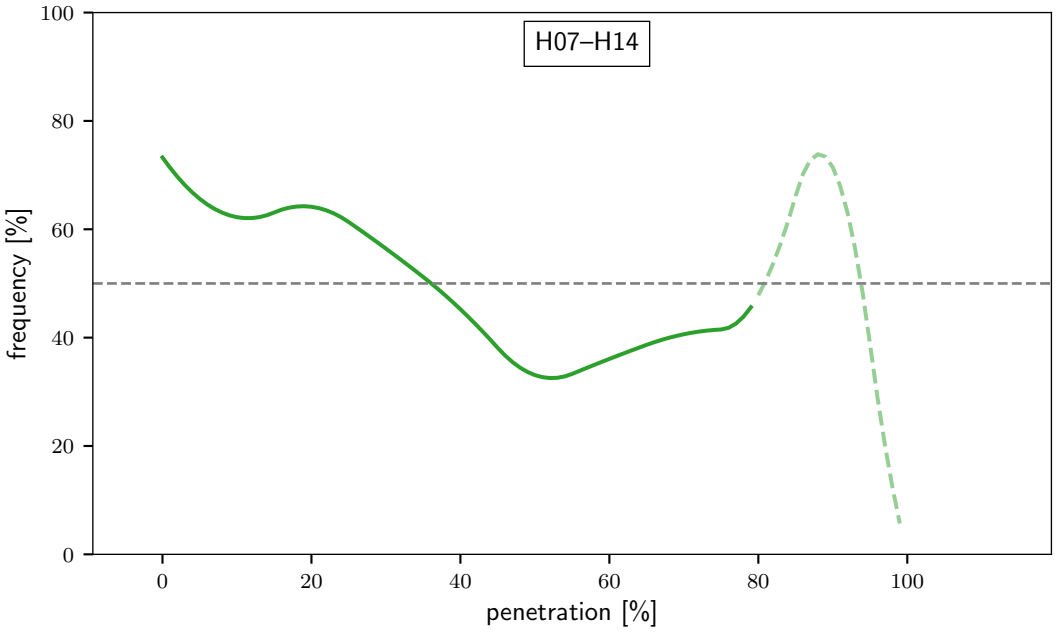


Figure 2.58: Frequency of up-regulation as a function of predicted wind power penetration between hours 7 and 14.

values of penetration up to 36%, with down-regulation then becoming more frequent. Between hours 15 and 22 (Figure 2.60), up-regulation is more frequent for values of penetration below 30%, while for higher values up- and down-regulation

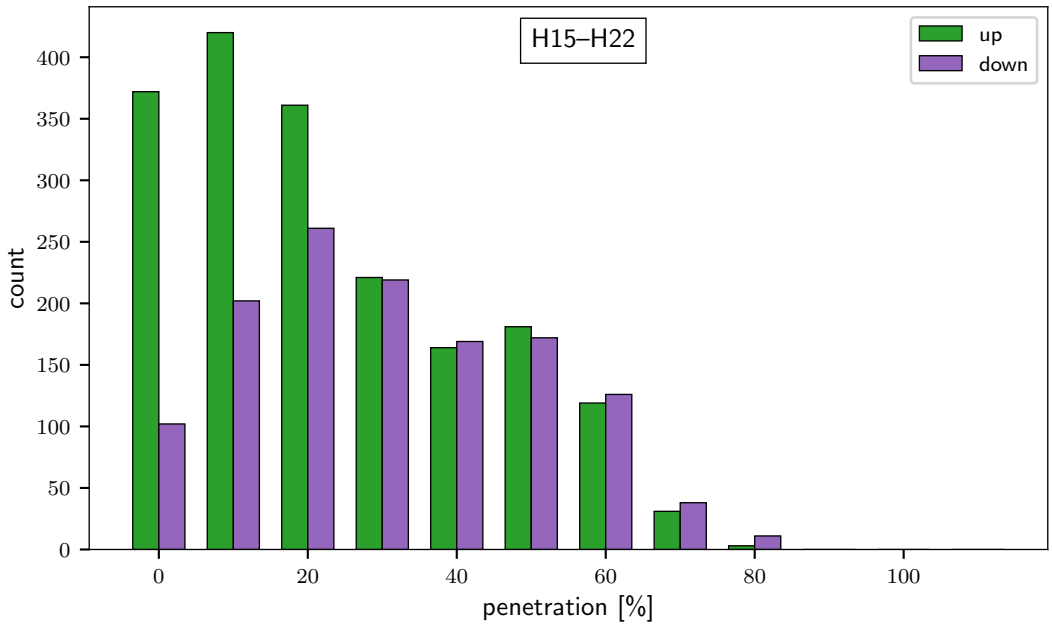


Figure 2.59: Number of trading periods with net up-regulation (green) and down-regulation (purple) as a function of predicted wind power penetration between hours 15 and 22.

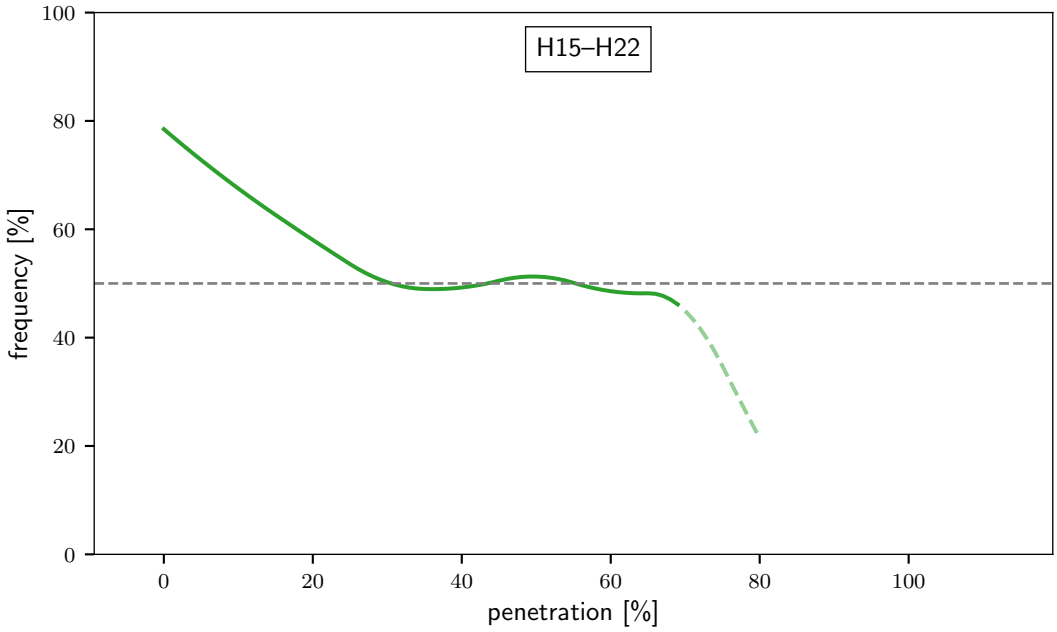


Figure 2.60: Frequency of up-regulation as a function of predicted wind power penetration between hours 15 and 22.

are equally frequent. The overall predominance of up-regulation during this time period can be explained by the tendency for the system to be left systematically short during these hours (cf. Figure 2.20). Between hours 23 and 6 (Figure 2.62),

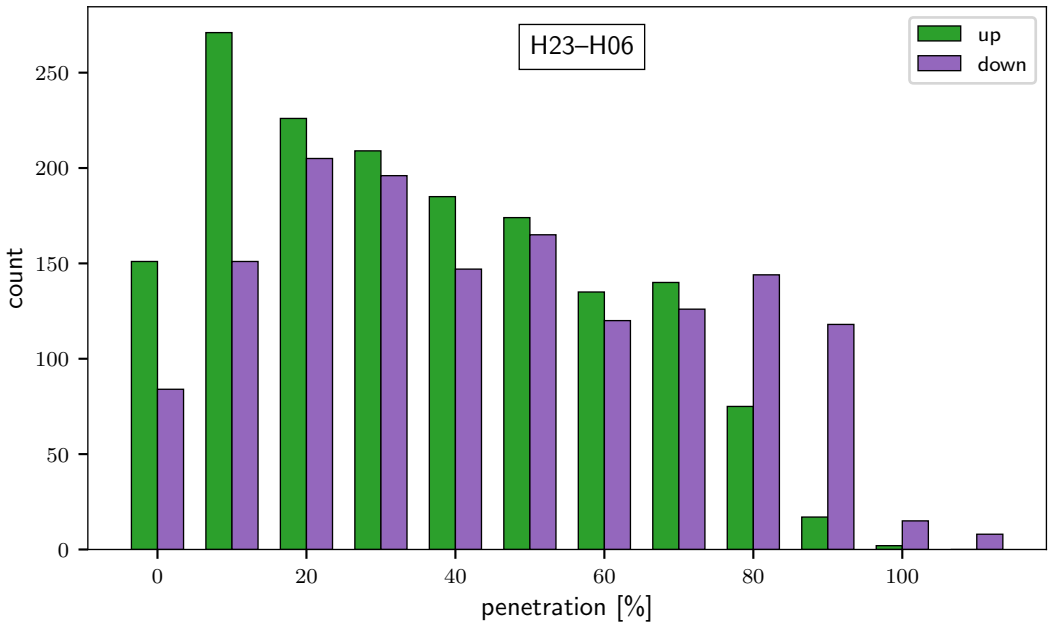


Figure 2.61: Number of trading periods with net up-regulation (green) and down-regulation (purple) as a function of predicted wind power penetration between hours 23 and 6.

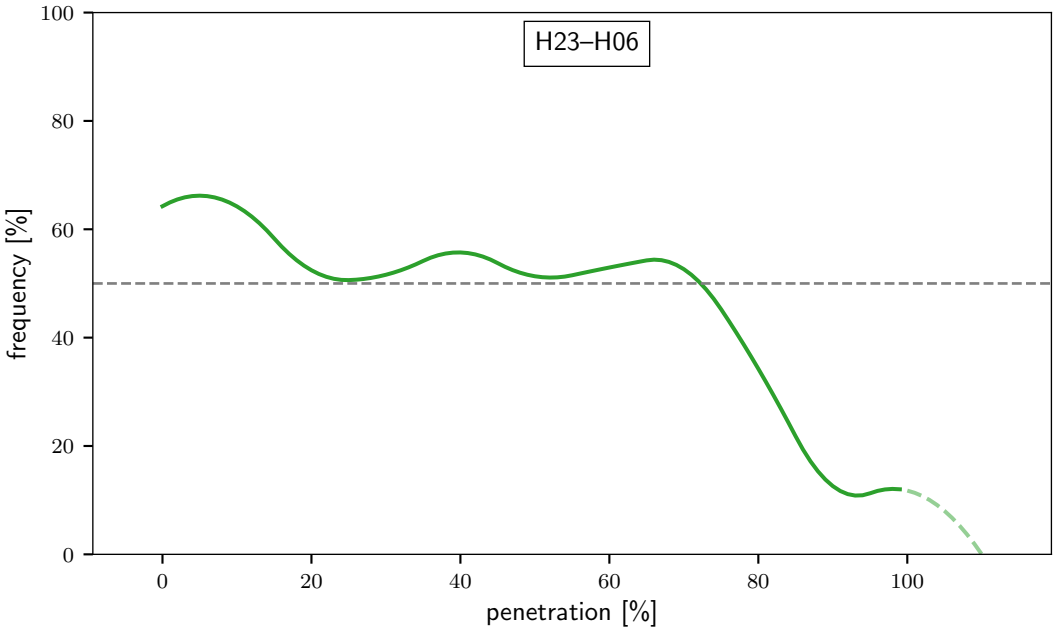


Figure 2.62: Frequency of up-regulation as a function of predicted wind power penetration between hours 23 and 6.

up-regulation is again more frequent for low values of predicted penetration. With penetration levels between 25% and 72%, up-regulation is only slightly more frequent than down-regulation, and for larger values the latter becomes dominant.

predicted penetration [%]	**up-regulation** [%]		
	07-14	15-22	23-06
0	73.3	78.5	64.3
20	64.2	58.0	52.4
40	45.2	49.2	55.7
60	36.1	48.6	52.9
80	-	-	34.2
100	-	-	11.8

Table 2.4: Frequency of up-regulation (expressed as %) for each group of hours at various levels of predicted wind power penetration.

Impact on Anomalous Imbalance Prices

In the third ex-post analysis, the impact of predicted wind power penetration on the occurrence of price anomalies in the Balancing Market is investigated. Recall that imbalance prices are deemed "anomalous" when they are lower than the day-ahead price when the system is short, and vice versa when the system is long. Otherwise, they are deemed "regular". The investigations carried out in Section 2.2.2.3 highlighted the overall high frequency of anomalous prices (see for example Figure 2.45) as well as the lack of any clear hourly pattern (cf. Figure 2.47). Therefore, while the approach followed in this study is similar to the previous two studies, here, data were not divided based on the time of the day. On the other hand, anomalies occurred significantly more when the system was short rather than long (see for example Figure 2.46). For this reason, the analyses were performed not only on the complete data set, but also considering the cases of up- and down-regulation separately. Trading periods were binned by predicted wind power penetration into intervals with width equal to 0.10. For each interval, the number of periods where anomalous prices occurred were counted, and results for the complete data set are shown in Figure 2.63. The frequency of anomalous prices was calculated over each interval for the complete data set and for the up- and down-regulation cases separately. Quadratic spline functions were then fitted to these values to find the trends illustrated in Figure 2.64. The overall trend is plotted as a solid orange line, while

the up- and down-regulation cases are plotted as dash-dotted green and purple lines, respectively. Note that for higher penetration levels, trends were calculated using a relatively low number of observations, and therefore are plotted as dashed transparent lines to remark their limited significance.

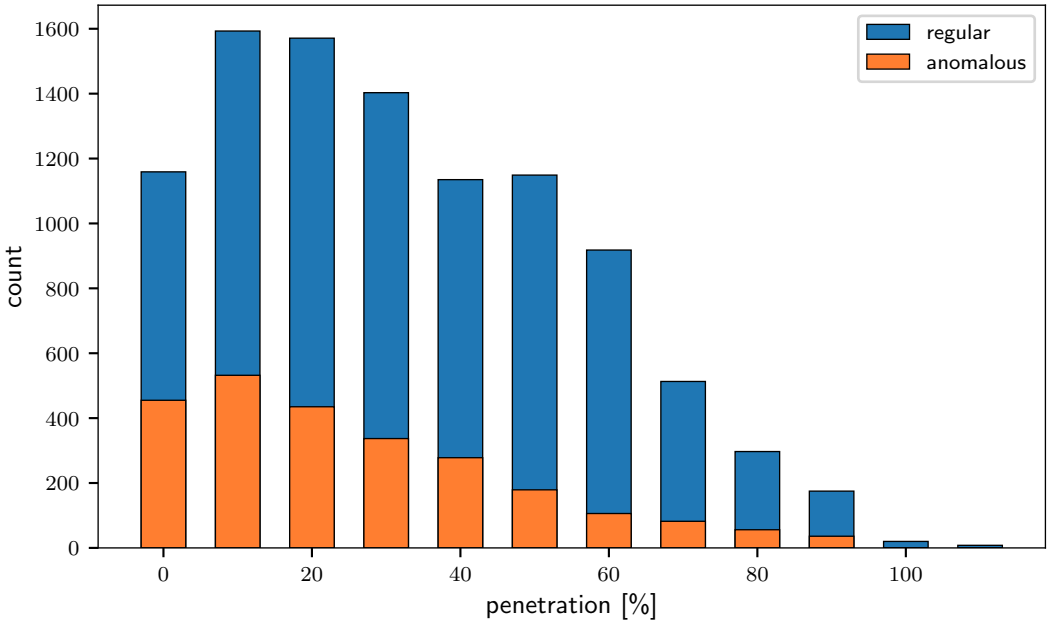


Figure 2.63: Number of periods with regular (blue) and anomalous (orange) imbalance prices at different levels of predicted wind power penetration.

predicted penetration [%]	**price anomalies**		
	overall	up-regulation	down-regulation
0	39.3	50.3	9.2
20	27.7	39.6	14.6
40	24.5	33.4	17.9
60	11.5	23.3	3.9
80	18.9	11.2	27.5
100	-	-	-

Table 2.5: Frequency of imbalance price anomalies (expressed as %) at various levels of predicted wind power penetration.

The results of the empirical analysis show that the overall frequency of price anomalies was highest for small values of predicted penetration, with anomalies occurring 39% of the time when no generation from wind was expected in the system. As penetration levels increase, the frequency decreases almost linearly

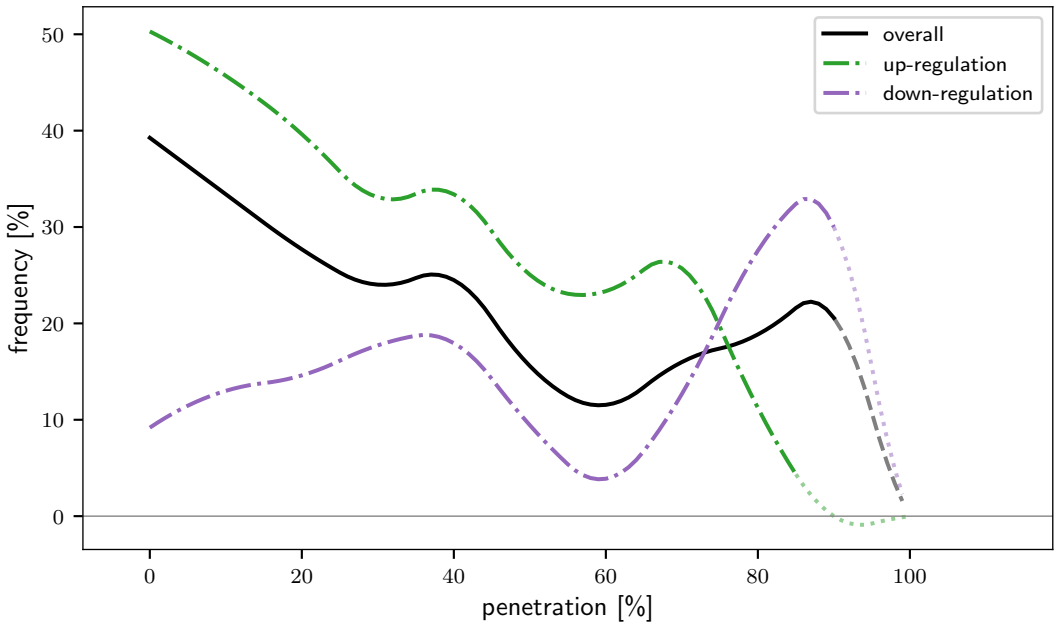


Figure 2.64: Overall frequency of imbalance price anomalies (solid black line) as a function of predicted wind power penetration. Trends are also shown separately for the up- and down-regulation cases (green and purple dash-dotted lines, respectively).

until it reaches a minimum of 12% when penetration is equal to 59%. When the forecast percentage of wind power in the system increases further, the overall trend reverses and price anomalies become more frequent again.

The analysis in Section 2.2.2.3 showed that price anomalies occur significantly more during up-regulation (cf. Figure 2.41). The results in the previous Section also show that up-regulation is prevalent for low levels of predicted penetration across all hours (Figures 2.58, 2.60 and 2.62). As a consequence of these two features, the frequency of anomalies is highest when there is no wind in the system and decreases as the penetration level increases (since, in parallel, the frequency of up-regulation decreases).

As the percentage of wind power becomes more significant, down-regulation becomes relatively more frequent. At the same time, price anomalies during down-regulation become more frequent, too, as shown by the dash-dotted purple line in Figure 2.64. In fact, with high penetration of wind power in the system, the network becomes more and more constrained. Most units are flagged out during the

"flagging-and-tagging" process (see Appendix A), leaving few and often very expensive units available for the imbalance price calculations. During down-regulation, this results in the marginal energy action price being set in the opposite direction of the system imbalance, resulting in an imbalance price above the day-ahead price. As a result, the overall frequency of price anomalies increases again for large levels of predicted penetration.

2.2.2.5 Summary of Results

The analyses of market data from the first seven months of operation of the I-SEM have highlighted a number of key features which are summarised here. The DAM was the dominant energy market in terms of liquidity, with more than 93% of the volumes being traded there. Prices have followed volumes closely, showing marked daily and weekly seasonality. Both variables followed the usual daily pattern, with values being low during the night, increasing in the morning with a first peak around 09:00 and then again in the evening with a second, larger peak around 18:00. Weekly effects were observed as well, with values for both variables being generally higher during week days and lower during the weekend. The DAM price time series has also been characterised by non-constant mean and variance. Volatility has been relatively high in general, with several instances of zero or negative prices at night (when demand is low) and price spikes in the evening (when demand is high).

In the BM, the imbalance volume time series was found to be positively biased with a tendency for the system to be left short (need for up-regulation), especially during the evening hours. This tendency has been very evident at the start of market operations, but with a clear decreasing trend over time. Neither the imbalance price nor the imbalance volume time series showed any marked seasonality. Values for both series have decreased on average over time. Imbalance prices have been extremely volatile during the period under analysis, although such volatility has been decreasing through the months, possibly reflecting the initially limited market experience of participants. A summary of some descriptive statistics for

the relevant variables from the DAM and BM is presented in Tables 2.6 and 2.7.

	mean	**min**	**max**	**st.dev.**
Day-Ahead Market				
Price [€/MWh]	64.35	–10.29	365.00	29.50
Volume [GW]	4.56	2.72	6.84	0.79
Balancing Market				
Imbalance volume [MWh]	2.94	-	-	28.24
Absolute imbalance volume [MWh]	21.56	-	-	-
Up-regulation volume [MWh]	23.62	-	134.11	18.75
Down-regulation volume [MWh]	–21.88	–108.75	-	17.51
Price [€/MWh]	62.07	-	-	79.24
Up-regulation price [€/MWh]	92.78	–281.16	1,453.00	61.27
Down-regulation price [€/MWh]	25.92	–280.45	3,773.69	51.89

Table 2.6: Mean, minimum and maximum values, and standard deviation of variables from the Day-Ahead and Balancing Markets.

up-regulation	**down-regulation**	**no regulation**
51.9%	42.5%	5.6%

Table 2.7: Empirical probability (frequency) of regulation direction.

The first months of operation of the I-SEM have also been characterised by an unusually high number of counter-intuitive imbalance prices, thus not providing an effective signal for participants to balance their positions. Price anomalies occurred on average 25% of the time and while no clear hourly pattern was found, anomalies were noticeably more frequent during up-regulation (37%) than down-regulation (14%). The overall frequency of anomalous imbalance prices has considerably decreased over time.

Finally, the impact of predicted wind power penetration on three market quantities of relevance for wind energy trading was investigated. Day-ahead prices decreased on average for increasing levels of predicted penetration across all hours during the day. Their variability also decreased for increasing amounts of wind in the system during morning and evening hours, but increased during night hours. Up-regulation was prevalent for low values of predicted penetration, becoming less

frequent as penetration increased. The frequency of imbalance price anomalies was highest for low levels of predicted penetration, decreasing for penetration levels up to 60% and then increasing again.

Chapter 3

Literature Review

Firstly, this Chapter provides a summary of the current methods used for time series forecasting and forecast evaluation in Section 3.1. Then, the academic literature of relevance for this work is reviewed. The state-of-the-art in short-term forecasting of wind power generation is presented in Section 3.2, followed by an analogous description for day-ahead forecasting of electricity prices in Section 3.3. A review of the academic literature on short-term value forecasting of wind power for energy trading is finally provided in Section 3.4.

3.1 Time Series Forecasting

In this Section, an overview of the general aspects of time series forecasting is given. The main typologies of forecast are presented in Section 3.1.1. The general framework for the evaluation of forecast quality is then discussed in Section 3.1.2, including the definition of benchmarks and error measures for point and probabilistic forecasts.

3.1.1 Forecast Types

Let us define Y as the random variable that needs to be forecast and y its realisation (or observation). A series of random variables each associated with a time point defines a *stochastic process* $\{Y_t : t = 1, \ldots, N\}$, and the realisation of the stochastic

process is a *time series* $\{y_t : t = 1, \ldots, N\}$. The *probability density function* (PDF) $f(y)$ defines the probability of observing a given value of the random variable Y and has to satisfy the following properties:

$$f(y) \geq 0 \qquad \forall y \in \mathbb{R}$$

$$\int_{-\infty}^{\infty} f(y)dy = 1$$

$$\mathbb{P}\left[a < Y < b\right] = \int_{a}^{b} f(y)dy$$

where $\mathbb{P}[\cdot]$ indicates the probability. The *cumulative distribution function* (CDF) $F(y)$ of the continuous random variable Y with density function $f(y)$ is given by:

$$F(y) = \mathbb{P}\left[Y \leq y\right] = \int_{-\infty}^{y} f(t)dt \qquad \text{for } -\infty < y < \infty \tag{3.1}$$

where

$$F(y) \geq 0 \quad \forall y \in \mathbb{R}$$

$$\lim_{y \to -\infty} F(y) = 0$$

$$\lim_{y \to \infty} F(y) = 1$$

Due to this definition, one can also write

$$\mathbb{P}\left[a < Y < b\right] = F(b) - F(a) \qquad \text{and} \qquad f(y) = \frac{dF(y)}{dy}$$

if the derivative exists [41]. The *expectation* (or mean value) $\mathbb{E}$ of the continuous random variable Y with probability density function f_Y is:

$$\mathbb{E}\left[Y\right] = \int_{-\infty}^{\infty} y f_Y(y)dy$$

A *forecast* $\hat{\cdot}_{t+h|t}$ is an estimate for time $t + h$ conditional on information up to time t. That means that being at time t, the goal is to predict what will happen

at time $t + h$, where h is referred to as the *forecast horizon* (or lead time, or look-ahead time). Forecasts are usually issued for regularly spaced horizons, that is $h = 1, 2, \ldots, H$, where H is the forecast *length* and the regular spacing is the forecast *temporal resolution* [40].

A *point forecast* informs on the conditional expectation of the variable:

$$\hat{y}_{t+h|t} = \mathbb{E}\left[Y_{t+h|t} \,|\, \Omega_t, M, \hat{\theta}\right] \tag{3.2}$$

where Ω_t is the information set available at time t, M is the model, and $\hat{\theta}$ is the vector of estimated parameters. The prediction is the average of the possible values that Y_{t+h} could take conditional on the knowledge at time t. Point forecasts are often referred to as "deterministic". This terminology, however diffused, is conceptually incorrect since it implies that there is no uncertainty associated with the realisation of the prediction.

A *probabilistic forecast*, on the other hand, provides more information about potential future outcomes through a prediction of the PDF (or CDF) of Y_{t+h} or of some summary features. A *density forecast* $\hat{f}_{t+h|t}$ (or $\hat{F}_{t+h|t}$ for the CDF) provides a full description of the PDF (or CDF) of Y_{t+h}:

$$Y_{t+h} \sim \hat{f}_{t+h|t} \qquad \text{or} \qquad Y_{t+h} \sim \hat{F}_{t+h|t} \tag{3.3}$$

where $\hat{f}_{t+h|t}$ and $\hat{F}_{t+h|t}$ are the predicted PDF and CDF for Y_{t+h}, respectively, given the information available at time t.

A *quantile forecast* $\hat{q}^{\alpha}_{t+h|t}$ is an estimate of the quantile $q^{(\alpha)}_{t+h}$ for the random variable Y_{t+h}:

$$\hat{q}^{\alpha}_{t+h|t} = \hat{F}^{-1}_{t+h|t}(\alpha) \tag{3.4}$$

where $\alpha \in [0, 1]$ is the *nominal level* and $\hat{F}$ is the predicted CDF for Y_{t+h}. It can be seen as a probabilistic threshold level for the variable, meaning that the predicted

probability of Y_{t+h} to be less than or equal to $\hat{q}^{\alpha}_{t+h|t}$ is equal to α:

$$F\left(\hat{q}^{\alpha}_{t+h|t}\right) = \mathbb{P}\left[Y_{t+h} \leq \hat{q}^{\alpha}_{t+h|t}\right] = \alpha \tag{3.5}$$

A *prediction interval* (or *interval forecast*) defines a range of values within which the variable may lie with a certain probability $\beta \in [0, 1]$, which is defined by its lower and upper bounds:

$$\hat{I}^{\beta}_{t+h|t} = \left[\hat{q}^{\underline{\alpha}}_{t+h|t}, \hat{q}^{\bar{\alpha}}_{t+h|t}\right] \tag{3.6}$$

where $\hat{q}^{\underline{\alpha}}_{t+h|t}$ and $\hat{q}^{\bar{\alpha}}_{t+h|t}$ are the quantile forecasts representing the bounds, and $\beta = \bar{\alpha} - \underline{\alpha}$ is called the *nominal coverage rate*, so that:

$$\mathbb{P}\left[Y_{t+h} \in \hat{I}^{\beta}_{t+h|t}\right] = \beta \tag{3.7}$$

The intervals are commonly centered on the median of the PDF, resulting in the so-called *central prediction intervals*, so that:

$$\underline{\alpha} = 1 - \bar{\alpha} = \frac{1 - \beta}{2} \tag{3.8}$$

Scenarios issued at time t for a set of forecast horizons $h \in \{1, 2, \ldots, H\}$ consist in a set of J time trajectories:

$$\hat{z}^{(j)}_t = \left\{\hat{y}^{(j)}_{t+1|t}, \hat{y}^{(j)}_{t+2|t}, \ldots, \hat{y}^{(j)}_{t+H|t}\right\} \quad \text{for } j = 1, \ldots, J \tag{3.9}$$

The time trajectories can be seen as equally likely samples of $\hat{F}_{Z_t}$, the predictive CDF of the multivariate random variable $Z_t = \{Y_{t+h} \,|\, h = 1, \ldots, H\}$. Note that since Z_t is a vector formed by the random variables that characterise the stochastic process for the following H lead times, it covers their marginal densities and their interdependence structure.

3.1.2 Forecast Evaluation

The performance of a forecast can be evaluated in multiple ways depending, for example, on the criteria and measures adopted in the evaluation, what the application of the forecast is, or the features of the forecast that one is interested in assessing. Murphy in [27] identified three types of "goodness" for the verification of weather forecasts, although the same principles apply for most kinds of forecasts:

- *consistency*: "*the correspondence between forecasters' judgements and their forecasts*";
- *quality*: "*the degree of agreement between forecasts and observations*"; and
- *value*: "*the benefits (economic or other) realised by end users through the use of the forecast as input for decision-making*".

It should be noted that while forecast quality is independent of the case under analysis, forecast value does depend on the operational problem at hand. Forecast quality will be discussed in the remainder of this Section, while forecast value will be addressed in Section 3.4.

The objective of forecast quality evaluation is to appraise the accuracy of the predictions; for this reason, the first fundamental recommendation is to perform the assessment on an independent evaluation period, that is, on data that were not used in any way to identify, train or estimate the model. The most common procedure is to split the available data set into a *training* (or *in-sample*) and a *test* (or *out-of-sample*) period. The model is fitted to the values in the training set and then evaluated on the test set. An alternative approach is *time series cross-validation* as defined by Hyndman and Athanasopoulos in [42] (also referred to as "evaluation on a rolling forecasting origin"). At every time step, the test set consists of a single period (the one being forecast) and the training set includes all previous observations. Forecast errors are then computed for each test set and their average determines the accuracy of the forecast. This strategy is particularly useful when dealing with a data set of limited length. However, since model parameters are

re-estimated at every step, it can be time consuming and computationally heavy, especially with non-linear models. Other strategies adopting a similar principle are used to perform model selection in regression models, such as leave-k-out and k-fold cross-validation [43], but their implementation is more difficult with time series since data are not independent.

In almost every study, data contain erroneous or abnormal values, missing entries or incomplete information [44], all of which can corrupt the quality of the data and hence the quality of the analysis [45]. Therefore, it is important to perform a set of preliminary actions on the data before any analysis is carried out to obtain a set of the highest possible quality and reflective of the functioning and conditions of the observed system [46]. This process is referred to as *data preparation* (also data cleaning, cleansing, or pre-processing). A detailed discussion on the preparation of data for forecast evaluation is provided in [47].

To contextualise evaluation results, the performance of advanced forecasting methods can be compared against that of some *benchmarks*. These are usually simple methods that can still provide competitive forecasts, which advanced methods are expected to outperform. In time series forecasting, the most relevant ones include:

- *persistence* (also "naïve method"): at time t, the forecast for all lead times is equal to the latest measurement available.

$$\hat{y}_{t+h|t} = y_t \qquad \forall h \tag{3.10}$$

- *moving-average predictor*: the forecast is the average of the last n measured values.

$$\hat{y}_{t+h|t} = \frac{1}{n}\sum_{i=0}^{n-1} y_{t-i} \qquad \forall h \tag{3.11}$$

- *climatology*: at time t, the forecast is the average of all available measurements.

$$\hat{y}_{t+h|t} = \bar{y} \qquad \forall h \tag{3.12}$$

- *seasonal naïve method*: a time series has a seasonal pattern when its values are affected by factors with a known and fixed period (for example, the day of the week or the hour of the day). When data exhibit a seasonality with period S, the forecast for time $t+h$ is:

$$\hat{y}_{t+h|t} = y_{t+h-S(k+1)} \qquad \forall h \tag{3.13}$$

  where S is the seasonal period and k is the integer part of $(h-1)/S$ [42]. For example, having daily data with weekly seasonality ($S = 7$), the forecast of future Monday values is equal to the last observed Monday value.

A first *qualitative* assessment of a forecast can be carried out through the visual inspection of predictions and associated errors. For example, time plots of forecast and observed values can help identify periods in the time series where the forecast performed particularly badly or well. A scatter plot of the same two series could show if the accuracy of the forecast is dependent on the variable level. Histograms and boxplots[1] can be used to analyse the distribution of prediction errors. These representations contain more information concerning the error dispersion and variation than a single statistic like standard deviation or root mean square error, and help to identify features like bias, symmetry, skewness, tails, quartiles and outliers. Moreover, histograms allow to quantify the frequency of occurrence of errors below or above a certain level [48].

Point forecasts

At the base of any *quantitative* analysis of point forecasts is the concept of *forecast error* [49], defined as the difference between the observation at time $t+h$ and the forecast for that same time issued at time t:

$$e_{t+h|t} = y_{t+h} - \hat{y}_{t+h|t} \tag{3.14}$$

[1] A boxplot normally shows the median, the first and third quartiles Q_1 and Q_3 (i.e. the 25^{th} and 75^{th} percentiles) and thus the inter-quartile range (IQR), and the outliers (i.e. values that exceed Q_1 and Q_3 by 1.5 times the IQR).

Scores are used to summarise aspects of the forecast accuracy and are usually expressed as a function of the look-ahead time h. Considering an evaluation period of length T, the most common scores include:

$$\text{bias: } bias(h) = \frac{1}{T}\sum_{t=1}^{T} e_{t+h|t} \tag{3.15}$$

$$\text{mean absolute error: } MAE(h) = \frac{1}{T}\sum_{t=1}^{T} \left|e_{t+h|t}\right| \tag{3.16}$$

$$\text{root mean square error: } RMSE(h) = \sqrt{\frac{1}{T}\sum_{t=1}^{T} e_{t+h|t}^2} \tag{3.17}$$

$$\text{standard deviation: } SDE(h) = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T} \left(e_{t+h|t} - \bar{e}_h\right)^2} \tag{3.18}$$

The bias indicates the systematic error in the forecast, while MAE, RMSE, and SDE are negatively oriented scores (i.e. the lower, the more accurate the forecast). Bias and MAE are directly related to the values of the variable since they are associated with the first moment of the forecast error. RMSE and SDE, on the other hand, reflect the variance of the forecast error since they are associated with its second moment, and are more impacted by large errors. These scores are normalised by the length of the evaluation period and therefore are independent of it. Nevertheless, if the length of the period is limited, the forecast quality assessment can be subject to significant uncertainty.

In order to have scale-independent errors and compare the performance of forecasts across different data sets, *percentage errors* can be calculated as:

$$p_{t+h|t} = \frac{e_{t+h|t}}{y_{t+h}}$$

Scores are calculated in the same way, with a commonly used score being the mean

absolute percentage error:

$$MAPE(h) = \frac{1}{T} \left| p_{t+h|t} \right|$$

These type of measures, however, have several disadvantages: they assume a meaningful zero, the error is undefined when $y_{t+h} = 0$, and distributions are heavily skewed when y_{t+h} is close to zero. For these reasons, their use is generally not recommended in the literature (see [50–52] for a discussion on this) and errors should be normalised by a fixed value (for example, the wind farm's max capacity). As a robust alternative to percentage errors, Hyndman and Koehler in [50] advocate the use of *scaled errors*, where the forecast error is divided by the in-sample MAE from the persistence forecast (cf. Equation (3.10)):

$$q_{t+h|t} = \frac{e_{t+h|t}}{\frac{1}{n-1} \sum_{i=2}^{n} |y_i - y_{i-1}|} \tag{3.19}$$

where $y_i, i = 1, \ldots, n$ are the in-sample values. Related measures such as the Mean Absolute Scaled Error (MASE) and the Root Mean Square Scaled Error (RMSSE) can be calculated directly from this definition. However, these errors have not been employed much in the literature to date, as pointed out by Weron in [34].

As discussed earlier, when evaluating an advanced model, its performance is usually compared to that of a benchmark or simpler model. The relative *improvement* is quantified through a skill score given by:

$$Imp(h) = 1 - \frac{Sc_{adv}(h)}{Sc_{ref}(h)} \tag{3.20}$$

where $Imp \leq 1$, Sc_{adv} is the score value for the advanced method and Sc_{ref} the score for the reference model. Note that Sc can be any of the statistics introduced earlier (e.g. MAE). Another skill score introduced by Madsen et al. [48] is the

coefficient of determination for each look-ahead time:

$$R^2(h) = \frac{MSE_{clim}(h) - MSE_{adv}(h)}{MSE_{clim}(h)} \tag{3.21}$$

where MSE_{clim} is the mean squared error for the climatology method (cf. Equation (3.12)). This coefficient represents the ability of the model to explain the variance of the data, ranging from 0 for useless predictions to 1 for perfect predictions. Therefore, its use should be limited to the training data set for model selection rather than on the test set for forecast evaluation.

When two competing forecasts are evaluated to find which one is more accurate, the *Diebold-Mariano* (DM) *test* [53, 54] should be used. Indeed, the scores introduced so far such as the MAE give an indication of the forecast accuracy on the test data set. However, the fact that forecast A had a lower MAE than forecast B on the test set under analysis does not imply that forecast A is more accurate than B in population. In the DM test, forecast errors are taken as primitives and a loss function is defined by assigning a penalty to the errors (typically, the loss function is quadratic or linear). The null hypothesis is that the two competing forecasts have equal predictive accuracy, i.e. the expected loss is equal. Then, testing for the null hypothesis corresponds to doing an asymptotic z-test of the hypothesis that the mean of the loss differential series is zero. The result of the DM test determines if the difference in accuracy between the two forecasts is statistically significant.

Probabilistic forecasts

For probabilistic forecasts, the quantitative evaluation is less straightforward since forecasts take the form of probability distributions while observations are real-valued series. The assessment of the predictive ability of probabilistic forecasts builds on two main properties that are required of them: calibration and sharpness [55]. *Calibration* (also referred to as *reliability*) refers to the consistency between forecast and observed probabilities by verifying if the forecast respects the

probabilistic contract. For a quantile forecast $\hat{q}^{\alpha}_{t+h|t}$ with nominal level $\alpha \in [0, 1]$, one expects that observations y_{t+h} are less than $\hat{q}^{\alpha}_{t+h|t}$ $(\alpha \times 100)$ percent of the time (see Equation (3.5)). For an interval forecast $\hat{I}^{\beta}_{t+h|t}$ with nominal coverage rate β, one expects observations to be covered by the interval $\beta \times 100$ percent of the time; since an interval forecast is in fact composed by two quantile forecasts, one has to evaluate the calibration of these two quantiles. Therefore, calibration is a joint property of forecasts and observations. In practice, it is assessed taking a frequentist approach by assessing the reliability of each of the quantile forecasts defining the prediction. For a given quantile forecast $\hat{q}^{\alpha}_{t+h|t}$ and the corresponding observation y_{t+h}, the *indicator variable* is defined as:

$$\xi^{\alpha}_{t,h} = \mathbf{1}\left\{y_{t+h} < \hat{q}^{\alpha}_{t+h|t}\right\} = \begin{cases} 1 & \text{if} \quad y_{t+h} < \hat{q}^{\alpha}_{t+h|t} \\ 0 & \text{if} \quad y_{t+h} \geq \hat{q}^{\alpha}_{t+h|t} \end{cases} \tag{3.22}$$

The *empirical level* of the quantile forecast is then calculated as the mean of the $\left\{\xi^{\alpha}_{t,h}\right\}$ time series over the set of T quantile forecasts in the evaluation period:

$$\bar{\xi}^{\alpha}_{h} = \frac{1}{T}\sum_{t=1}^{T} \xi^{\alpha}_{t,h} \tag{3.23}$$

The difference between nominal and empirical levels of the forecast can be interpreted as its probabilistic bias [56]. When multiple quantiles are evaluated for various nominal levels, the calibration assessment can be summarised in a Quantile-Quantile diagram [57] where empirical quantiles are plotted against the forecast ones, or in a reliability diagram [58] where the empirical level is plotted against the nominal level. In both diagrams, points are connected with a curve and the closer this is to the diagonal line the more reliable the forecast is.

Sharpness measures the concentration of probability by evaluating how tight the prediction densities are, and is thus a property of the predictive distribution only. Narrow predictive intervals are more informative than wide ones and enable end-users to make decisions more easily. The width of an interval forecast is given

by the distance between its two quantile bounds:

$$s_{t,h}^{\beta} = \hat{q}_{t+h|t}^{\bar{\alpha}} - \hat{q}_{t+h|t}^{\alpha} \tag{3.24}$$

The sharpness of the interval forecast is then calculated as the average width over the evaluation period:

$$\bar{s}^{\beta}(h) = \frac{1}{T}\sum_{t=1}^{T} s_{t,h}^{\beta} \tag{3.25}$$

Sharpness can also be visually evaluated via boxplots of the predictive intervals, as proposed by Bremnes in [59] (also called "sharpness diagram" by Gneiting et al. in [55]).

As mentioned earlier, the quality of a probabilistic forecast needs to be determined by calibration and sharpness [60]. However, when the forecast end-user wants a single evaluation criterion to describe the quality of the forecast method, *skill scores* can be used. In this context, a skill score needs to provide information on both reliability and sharpness simultaneously. Moreover, the score has to be proper [61], meaning that the perfect forecast obtains the best possible score, and a forecast with higher skills gets a better score [62]. Two of the most common skill scores currently employed are the Continuous Ranked Probability Score [55] and the pinball loss function [60]. Both scores are negatively oriented, with a lower score indicating a better forecast, and admit a minimum value of 0 for a perfect probabilistic forecast. The *Continuous Ranked Probability Score* (CRPS) is used to evaluate predictive densities and is calculated as:

$$CRPS_{t,h} = \int_{y} \left(\hat{F}_{t+h,t}(y) - \mathbf{1}\left\{ y_{t+h} \leq y \right\} \right)^2 dy \tag{3.26}$$

where $\hat{F}_{t+h,t}(y)$ is the predictive density for y, and y_{t+h} the corresponding observation. The value corresponds to the area between the predictive and observed CDFs. The CRPS score value is then calculated as the average for each of the

predictive densities and corresponding observations over the evaluation period:

$$CRPS(h) = \frac{1}{T}\sum_{t=1}^{T} CRPS_{t,h} \tag{3.27}$$

This skill score, which has the same unit as the variable of interest, can be seen as the generalisation of the MAE score (cf. Equation (3.16)) in a probabilistic framework.

The *pinball loss function* (also referred to as the *negative quantile-based score* (NQS)) is the other main proper scoring rule for probabilistic forecasting. This skill score was chosen as the single error measure at the Global Energy Forecasting Competition in 2014 [63]. For a quantile forecast $\hat{q}^{\alpha}_{t+h|t}$ with $\alpha \in [0, 1]$ as the target quantile, the score is defined as:

$$L\left(\hat{q}^{\alpha}_{t+h|t}, y_{t+h}\right) = \begin{cases} (1-\alpha)\left(\hat{q}^{\alpha}_{t+h|t} - y_{t+h}\right) & \text{if } y_{t+h} < \hat{q}^{\alpha}_{t+h|t} \\ \alpha\left(y_{t+h} - \hat{q}^{\alpha}_{t+h|t}\right) & \text{if } y_{t+h} \geq \hat{q}^{\alpha}_{t+h|t} \end{cases} \tag{3.28}$$

where y_{t+h} is the observation used for the evaluation. When evaluating a quantile forecast, the score is averaged over all periods in the evaluation set over the same forecast horizon:

$$L(h) = \frac{1}{T}\sum_{t=1}^{T} L\left(\hat{q}^{\alpha}_{t+h|t}, y_{t+h}\right) \tag{3.29}$$

To evaluate a full predictive density, the average of the scores for each of the quantiles defining the density is calculated.

Note that, as advocated for example by Gneiting et al. in [55] or Pinson et al. in [60], the evaluation of probabilistic forecasts should be based on the paradigm of "*maximising the sharpness of the predictive distribution subject to calibration*". In other words, one should start from evaluating the calibration. Once the model is proved to be calibrated, then its performance can be compared with that of other models in terms of sharpness and skill scores.

When probabilistic forecasts are employed to predict binary events, they can be

considered as "probabilistic classifiers" and evaluated by examining their *Receiver Operating Characteristics* (ROC) *curves* [64]. Given a classifier, the true-positive rate is the ratio between true positives and total positives, and the false-positive rate is the ratio between false positives and total negatives. A ROC graph is a two-dimensional graph with false-positive rate on the x-axis and true-positive rate on the y-axis, and provides a graphical illustration of the trade-off between hit rates and false alarm rates of classifiers [65]. It follows that point (0,1) represents perfect classification, while the diagonal $y = x$ represents the strategy of randomly guessing a class, for which false and true positive rates are equal. The graph can be used to visualise and compare the performance of different models, as the ROC curve of a superior forecast will lie above and to the left of that of an inferior forecast. Classifiers in the lower left-hand side of the graph are considered "conservative", as they require strong evidence to classify positives; conversely, those in the upper right-hand side are considered "liberal" because positive classifications are made with weak evidence. A single-valued measure of accuracy that can be extracted from the graph is the *area under the curve* (AUC), i.e. the portion of the area that lies between the curve and the x-axis. Therefore, AUC = 0.50 for a random guess and AUC = 1.00 for a perfect classifier, and the greater the area, the better the performance of the classifier. An example of ROC curves is presented in Figure 3.1, from which one can conclude that both models A and B perform better than random guessing, and that model B (AUC = 0.68) is more skilful than A (AUC = 0.57).

When evaluating the predictive accuracy of binary forecasts, performances can be compared using the *Brier Score* (BS) defined by Brier in [66]. For n predictive probabilities $p_i \in [0, 1]$ and corresponding realisations $o_i \in \{0, 1\}$, the score is defined as the mean squared forecast error and is given by:

$$BS = \frac{1}{n} \sum_{i=1}^{n} (p_i - o_i)^2 \tag{3.30}$$

where the variable o_i denoting the binary outcome takes the value 1 or 0 depending on whether the event occurred or not. As a proper scoring rule, the BS rewards

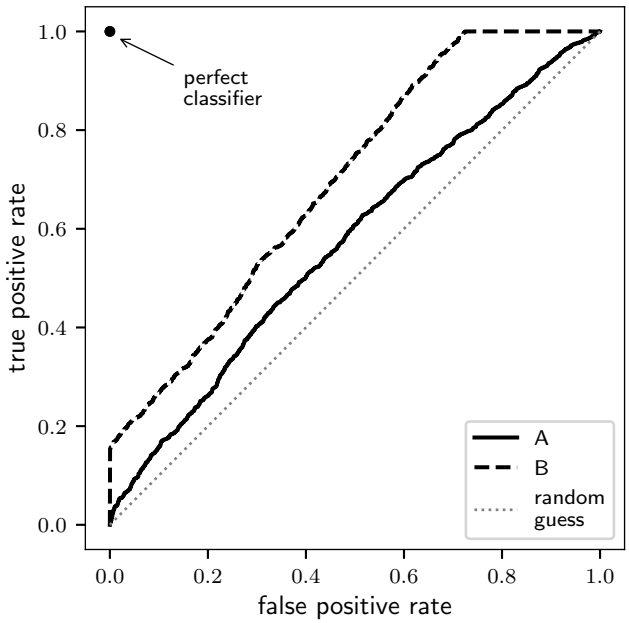


Figure 3.1: Example of ROC curves. The diagonal dotted line represents the performance of a random forecast.

simultaneously reliability and sharpness, since it depends on both the correlation between forecast and observed probabilities, and the frequency distribution of the binary predictions, as demonstrated by Murphy in [67]. The score is negatively oriented with a lower result indicating a superior model [68], and has a minimum value of 0 for a perfect forecast and a maximum of 2 for the worst possible forecast. If they are correct, confident forecasts (i.e. those close to 0 or 1) obtain a lower score than prudent forecasts (i.e. close to 0.5); on the other hand, they are penalised more severely if they are on the wrong side of 0.5 (the loss function is quadratic).

3.2 Short-term Wind Power Forecasting

One of the most significant challenges for the integration of wind power in the electrical grid is the variability of wind resources at time scales ranging from minutes to days, and wind power forecasting represents one of the key tools to address this challenge. For example, market operators and TSOs need system-level forecasts of wind power generation for unit commitment and economic dispatch; utilities and

energy traders use forecast information to participate in the day-ahead (or spot) and balancing (or real-time) markets.

At a high level, wind power forecasting models can be divided into two groups:

- time series models, which use purely statistical approaches to forecast wind speed and power from recent observations of the relevant variables; no actual weather forecasts are used.

- models using forecast values obtained from numerical weather prediction (NWP) models as an input to predict wind power through different approaches.

The former perform better for forecast horizons from a few minutes up to 3-6 hours — in what is referred to as *very short-term* forecasting — and for predictions of monthly, seasonal, or yearly averages. Models using numerical weather predictions work best for forecast horizons from a few hours up to two days, where the influence of atmospheric dynamics becomes more important, as shown by Landberg in [69]. Within this time frame, that is with forecast horizons up to 48 hours, we talk of *short-term* forecasting.

To maximise the use and value of electricity generated from wind, the trading platform of primary interest is the day-ahead market. The relevant forecast horizon generally ranges between 12 and 48 hours, where according to Giebel et al. in [70] the best performances are achieved by models including NWP inputs. Therefore, these will be the models reviewed in the remainder of this Section.

3.2.1 Short-term Forecasting using NWP

As explained by Manwell et al. in [5], short-term wind power forecasts generally need two broad categories of input data:

- forecasts of wind speeds and other relevant weather variables at the location of interest; and

- information on the wind farm, such as technical characteristics of the installed wind turbines, wind farm layout, description of the terrain (orography, roughness, obstacles, etc.), and historical observations from the site (if available).

The former are obtained from the NWP models. The latter entail the modelling (explicit or implicit) of the power system's behaviour in response to a number of external and internal factors, typically in the form of a *power curve*. Generally, the forecasting process consists of (at least) the following two steps:

- forecasts of the explanatory variables are obtained from a single or multiple NWP models; and
- forecast wind speed is converted into electric power.

Depending on the theoretical approach followed in the method, these steps might be formulated more or less explicitly within the process, or directly combined into a single passage. The work by Jensen et al. [71] shows that the use of wind speed predictions with subsequent conversion to wind power using autoregressive models is effective for horizons up to 8-12 hours. Beyond this, the use of separate wind speed forecasts does not provide any advantage over the direct estimation of wind power.

NWP model

An NWP model is usually characterised by three main components: the dynamical centre, which deals with the equations of the adiabatic flow; the physics pack, including the equations that describe the variability of meteorological processes; and the data assimilation code [72]. The choice of the NWP model is critical and depends on several criteria: the spatial and temporal resolution required, the timing of forecast updates, forecast horizon, geographical area, the accuracy required, and computational costs. The output is a comprehensive forecast of the overall state of the atmosphere. A list of currently available global and regional NWP models is presented by Foley et al. in [73].

Generally, however, the grid resolution of the model is larger than the area of the wind farm of interest. Kariniotakis et al. in [74] and Giebel et al. in [75] found that estimating the wind speed at the wind farm location can provide more accurate inputs to the wind power forecast and therefore improve its performance. For this reason, physical or statistical *downscaling* is often performed to get more precise local conditions from coarse NWP outputs [76]. The physical approach employs meso- and micro-scale models similar to the NWP but run at a higher resolution over a limited area, which return wind speed data at the hub height of the turbines. The statistical approach uses a transfer function generated with data from an actual wind farm in the region to predict wind power generation at other wind farms within the same area.

Power curve

Most forecasting tools involve the conversion of meteorological forecasts (in particular, of wind speed) into electrical power using a power curve. The quality of the power curve modelling can strongly affect the accuracy of the wind power forecast, contributing to 10 to 15% of the prediction error, as found by Paiva et al. in [77].

The power curve provided by the manufacturer of the wind turbine gives the relationship between wind speed and power at a particular air density; its typical shape is shown in Figure 3.2. The minimum speed at which the turbine starts generating useful electrical power is known as the *cut-in speed*; the wind speed at which the rated power, i.e. the maximum output power of the electrical generator, is generated is called *rated speed*; the maximum wind speed at which the turbine is allowed to generate power and is usually limited by engineering design and safety constraints is called *cut-out speed*. This power curve is usually calculated following the procedure designed in the International Standard IEC 61400-12-1 [78], where the "method of bins" is applied.

The use of the manufacturer's power curve represents the easiest approach to converting forecast wind speed to power. The curve, however, is neither site-specific [79] nor does it take into account the wear and tear of the turbine [80]

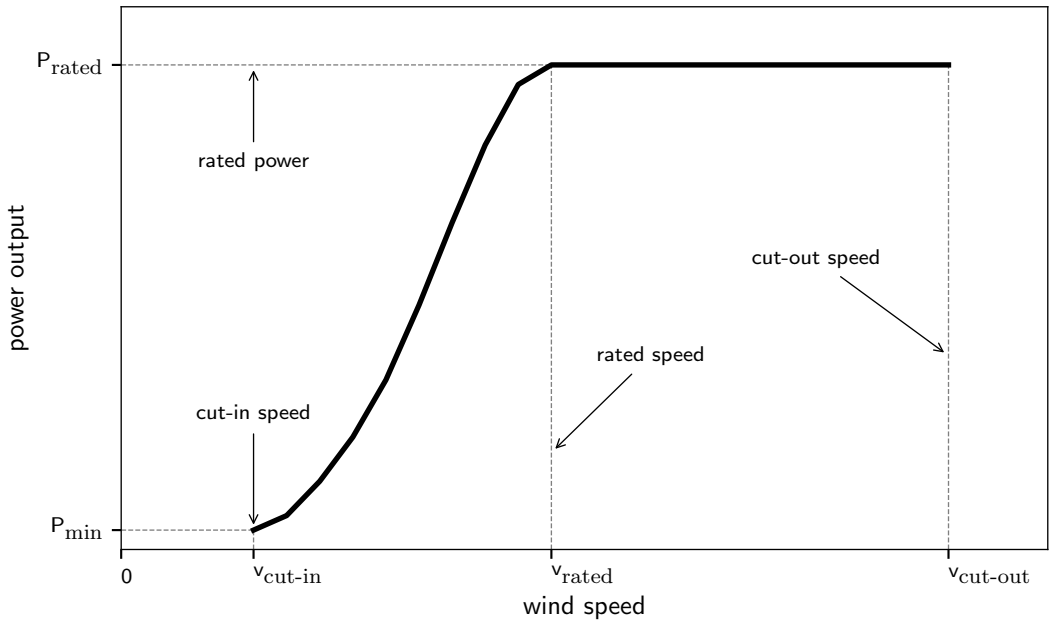


Figure 3.2: Typical power curve of a wind turbine.

and therefore its blind application to different sites cannot be completely correct. The probability density function of the power curve is in fact affected by various environmental parameters as discussed by Jafarian et al. in [81], and the modelling of a site-specific curve can be highly advantageous. This is particularly the case for large wind farms, where advanced methods that take into account heterogeneous flow fields and wake effects can significantly reduce the forecast error, as illustrated by Collins et al. in [82].

The various power curve modelling techniques presented in the literature can be classified into parametric and non-parametric. The former employ mathematical formulations to express the relationship between wind speed and power output. Available techniques include polynomial expressions [83, 84], the maximum principle method [85], dynamic and probabilistic approaches [86, 87], the "ideal power curve" approach [79] and logistic functions [88]. Whereas, non-parametric techniques aim at finding the empirical relationship without necessarily developing characteristic equations. Possible modelling approaches include the use of Copula distributions [89], spline interpolation [90], and artificial intelligence [91]. Broad reviews of power curve modelling methodologies are presented by Lydia et al. in [92]

and by Carrillo et al. in [93].

Upscaling

If the objective is to predict wind power generation for an area rather than for a single wind farm, *upscaling* is eventually performed. The forecast power output from a sample number of wind farms in the region of interest serves as input for an upscaling algorithm that estimates the generation in the whole region. This avoids the time-consuming and computationally costly task of predicting the output for each individual wind farm and then summing them together. As proven by Focken et al. in [94] and [95], the forecast error for distributed wind farms is reduced compared to that of a single site thanks to the so-called "spatial smoothing effects": the fluctuations of combined power output from distributed wind farms are damped, resulting in decreased fluctuations of the overall power output for the region. In the same vein, the studies by Hasche in [96] and by Miettinen and Holttinen in [97] find that both average and largest errors reduce when forecasts are aggregated over large geographical regions. The correlation between forecast errors is generally weak with distance, so the magnitude of the error reduction scales with the size of the region, i.e. the larger the region the larger the reduction [98]. Similarly, Miettinen et al. in [99] observe that while the forecasting skill for an aggregated wind farm fleet is heavily impacted by the accuracy of a single site, the site-specific accuracy becomes less significant as the number of sites aggregated together increases.

Forecasting Approaches

Adopting the classification used by Wang et al. in [100], forecasting models can be categorised into physical, statistical, and hybrid based on their approach. *Physical* methods use physical considerations as far into the process as possible to obtain a detailed description of the lower atmosphere and the best possible estimates of local wind speeds [101]. Global databases of meteorological measurements or atmospheric meso-scale models are used for this purpose. Wind speed values are then

converted into power based on a power law such as those described in the previous Section. Alternatively, computational fluid dynamics (CFD) is used to simulate the wind aerodynamics and calculate the power production of the wind turbines with the actuator disc model [102]. The physical approach is ideal when no historical data are available. However, large computational systems are required to run the simulations and the acquisition of physical data can be challenging. Often, statistical post-processing techniques are applied to remove any systematic errors in the forecast. Popular techniques include model output statistics (MOS) [103], Kalman filtering [104], artificial neural networks [105], and combination of individual forecasts [106, 107].

In the *statistical* approach, historical data are analysed to find the empirical relationship between a set of explanatory variables (e.g. wind speed, wind direction, temperature, etc.) and power output measurements. Forecasts of the explanatory variables are obtained from the NWP model and the fitted relationship is used to estimate future values of the power output. With this approach, meteorological processes are not represented explicitly, as input variables are mapped to power outputs in only one step. For this reason, these models are also referred to as "black box". On the other hand, historical data are needed to estimate the model's parameters.

The relationship between explanatory variables and power generation is often modelled using statistical techniques employed in time series analysis. Sometimes, this is referred to as the *conventional statistical approach* [108, 109]. Widely diffused techniques include: autoregressive integrated moving average (ARIMA) and its subsets (AR, MA, ARMA), Kalman filtering, and the Box-Jenkins methodology. If online data are available, recursive techniques can be highly advantageous to retune the model with more recent measurements. Alternatively, *learning approach* methods can be used. Sometimes, these are referred to as artificial intelligence (AI) or "grey box" methods. Rather than being expressed analytically, the relationship between the output and input variables is learned (hence, the name) from

the historical time series. For this reason, a very large amount of historical data is required to train the model, making this approach unsuitable for newly or recently installed wind farms. The main machine learning approaches are artificial neural networks (ANN), fuzzy systems, support vector machines (SVM), and more recently decision tree techniques and Bayesian methods.

Hybrid methods combine together different approaches to take advantage of the strengths of each modelling technique. Methods can result from the combination of physical and statistical approaches, or of different statistical approaches. The objective is to improve the forecast accuracy. Moreover, the combination approach often reduces the risk of large errors during extreme events, as found by Hibon and Evgeniou in [110].

An exhaustive review of the state-of-the-art approaches for wind power forecasting is provided by Jung and Broadwater in [109]. An overview of physical and statistical methods is also given by Lei et al. in [108], while recent publications employing learning approaches are discussed by Foley et al. in [73]. Finally, comprehensive reviews of commercially available models are given by Giebel and Kariniotakis in [8] and by Monteiro et al. in [72].

3.2.2 Uncertainty Forecasting

The wind power forecasting methods discussed so far produce a single value for each look-ahead time, corresponding to the conditional expectation of the wind power output. In the literature, these are referred to as *point* (or, less properly, deterministic) forecasts. Wind power *uncertainty forecasting*, on the other hand, provides additional probabilistic information on future wind power generation. Such information, for example, can bring advantages in decision-making processes [28, 111, 112] and power system operation [113, 114].

Following the classification used by Zhang et al. in [115], the three main representations of uncertainty in the context of wind power forecasting are: probabilistic forecasting, risk index, and scenario forecasting. *Probabilistic* forecasts are the most

common representation of forecast uncertainty, where the power output is treated as a random variable and the associated uncertainty can be expressed by:

- probability density functions (PDF) or cumulative distribution functions (CDF);
- quantiles and intervals;
- discrete probabilities; and
- moments of the probability distribution (for example, mean, median, variance, skewness).

PDF and CDF are the most general representation and all others can be derived from them. Interval forecasts give the range of values where the output can lie in with a specified probability. This is the representation that has attracted the most attention among end-users thanks to the intuitiveness of its visualization, as observed by Dobschinski et al. in [116].

The two main approaches for the construction of predictive distribution are the parametric and non-parametric approach. In the *parametric* approach, the forecasting error is assumed to follow a pre-defined distribution and very few parameters are involved, making the estimation process simple and the computational costs low. Gaussian and Beta distributions are common choices (see for example [117] and [118]). However, due to the non-linear shape of the power curve, the distribution of wind power forecast errors depends on the wind speed regime [119] and can be skewed and heavy-tailed [120]. For this reason, other distributions are also employed, like in the work by Hodge et al. in [121] who find that hyperbolic distributions are better suited to model errors from day-ahead wind power forecasts, as they represent more accurately the skewness, narrower peak, and fatter tails of the distribution. Otherwise, multiple distributions can be combined; this is the case, for example, of the work by Miettinen et al. in [99], where beta and Laplace distributions are used together to model the forecast uncertainty. On the other hand, the assumption of the distribution shape may not be reasonable in some cases, and the shape can change significantly with the forecast horizon. In *non-parametric*

(or distribution-free) approaches, no assumption on the distribution is made. The predictive PDF or CDF of wind power are viewed as a set of density or quantile forecasts, respectively, and each forecast is estimated separately. As a consequence, computational costs are considerably higher compared to the parametric case. Numerous non-parametric methods are proposed in the literature, including adaptive resampling [122], quantile regression [123], kernel density estimation [124], artificial intelligence [125], and ensemble forecasting [126, 127]. Beside these statistical and machine learning approaches, the employment of stochastic differential equations to generate probabilistic forecasts are starting to be investigated (for example, see [128]). Stochastic differential equations enable capturing the dynamics in the input information, allowing to describe the wind power dynamics better.

Probabilistic forecasts are generally more complicated than point predictions and their direct use in decision making is more difficult; for this reason, *risk indices* have been developed to give a more simplified form to the uncertainty information. A risk index is defined as a single value reflecting the expected level of error of the wind power forecast, with higher risk values corresponding to less reliable (more uncertain) forecast outputs. For example, Meteo-Risk Index (MRI) [129] and Normalised Prediction Risk Index (NPRI) [130] have been proposed in the literature to quantify the spread of members in an ensemble forecast.

Since the uncertainty associated with a forecast is generated for each horizon independently, the correlation between points at different times is not considered. It follows that probabilistic forecasts do not provide any information on how uncertainty varies over the forecasting horizon. In many time-dependent and multi-stage decision-making processes, however, spatio-temporal dependencies have to be taken into account. For this reason, *scenario forecasting* has been developed as an alternative approach [131]. A scenario is formed by a series of point forecasts over a period of look-ahead times and a set of scenarios (also referred to as "time trajectories") produces the desired time-dependent information among different prediction horizons [132]. Therefore, scenarios constitute the most suitable input for dynamic

stochastic optimisation problems, while also allowing the visualisation of the multivariate probabilistic distribution of forecast wind power.

A complete review of state-of-the-art wind power uncertainty forecasting can be found in the work by Zhang et al. in [115]. The results of the wind power forecasting track in the 2014 Global Energy Forecasting Competition[2] also provide an overview of probabilistic methodologies. The main features of the winning methods are summarised by Hong et al. in [63]. In this article, it is also highlighted that within the energy forecasting field (which includes wind power, solar, price, and demand), wind power is the domain where probabilistic forecasting is more mature.

As part of its activities, the IEA Task 36 on Forecasting for Wind Energy has been working to investigate the current use of uncertainty forecasts in the electric power industry. As discussed by Bessa et al. in [9], point forecasts are still predominant in utility practice, even though truly optimal decisions and risk hedging are only possible with the adoption of uncertainty forecasts. In fact, uncertainty in weather forecast and power production is one of the primary causes of volatility in electricity markets and a driver for competition, as well as playing a role in security constraints and risk management. Nevertheless, despite uncertainty forecasts being widely available, current business practice is at best focused on the so-called "multiple-supplier" approach, where different point forecasts from various providers are used; this approach, however, cannot be considered as probabilistic, since the spread of a set of point forecasts seldom reflects a realistic weather or wind power production uncertainty.

In an earlier study conducted by Möhrlen et al. in [25], the authors carried out a combination of interviews and questionnaires with different players in electricity markets of various countries. The outcomes of this study showed that a gap exists in the understanding of probabilistic forecasts and their use in the business practice, and it is mainly this lack of understanding that leads to mistrust towards

[2]Website: `http://www.drhongtao.com/gefcom`

these solutions. For example, one common misconception diffused among end users is that if a decision has to be a Boolean yes/no or a single value, then the forecast information should come as a single value, too. There is also a lack of trust towards uncertainty information, often connected with the erroneous perception of probabilistic forecasts being associated with speculation. On the contrary, there is little to no concern among end users for the overwhelming amount of information that the use of probabilistic forecasts might introduce; but again, they are concerned with how to make use of such additional information. Another relevant finding is that a large penetration of renewable sources in the electricity market is significantly correlated with the use of uncertainty forecasts by market players, as for example in the case of Denmark or Portugal. Since renewables are increasing their share in the electricity generation mix of most countries in the world, it is thus fair to assume that the demand for uncertainty forecasts will rise in the near future.

3.2.3 Performance of Forecasts

Errors in forecasts of renewable energy generation (but also, for example, of load or electricity prices) are most often driven by weather forecast errors. In wind power forecasting, the largest part of the prediction error comes from the NWP models, as explained by Giebel et al. in [8]. The uncertainty in the meteorological forecast is amplified or dampened by the conversion into power, i.e. the power curve, which also determines the shape of the forecast error distribution. Due to the shape of the power curve (cf. Figure 3.2), wind power forecast errors do not follow a Gaussian distribution and prediction intervals are usually not symmetric around the point prediction, as found by Lange in [117]. Moreover, the level of predicted wind speed introduces some non-linearities to the estimation of the intervals. For example, near the cut-out speed, where power can vary dramatically for small variations in the wind speed. Typical error patterns are: amplitude errors, where the magnitude of the power output is misjudged; and phase errors, where the changes in generation

are not timed correctly. The latter have a much larger influence on error scores such as those presented in Section 3.1.2.

It is not possible to draw any conclusion as to which one forecasting approach or model performs best, since site dependencies are always significant. Indeed, a model performing well for one wind farm might not perform as accurately on another. However, there are some features that can be observed in the performance of any model. The first, more obvious one, is that forecast accuracy decreases for increasing forecast horizons [70]. This is evident for example from Figure 3.3, where the root mean square error of the forecasts increases for all models as the forecast length increases; the rate of increase varies significantly depending on the model and forecast length. Persistence (cf. Equation (3.10)) and the so-called "new reference model" (NewRef in the Figure) introduced by Nielsen et al. in [133] are very simple approaches which outperform more complex models for very short prediction horizons (less than three hours). However, their errors increase steeply with forecast length. Climatology (cf. Equation (3.12)), the other basic approach consisting of predicting the mean value for all times (Mean in the Figure), outperforms persistence for lead times higher than 15 hours. The two NWP-based models (HWP and HWP/MOS in the Figure) outperform all other models for forecast horizons larger than four hours, and their accuracy degrades much less rapidly as forecast length increases. Moreover, it can be noted that the use of simple postprocessing techniques (MOS) in the HWP/MOS model (known commercially as Prediktor [134]) results in marked improvements.

The performance of models also changes with the type of terrain the wind farm is located in. Generally, smaller error values are observed in a flat terrain wind farm compared to a complex terrain farm. Martì et al. found in [135] that there is a significant increase of the magnitude and dispersion of prediction errors when the terrain complexity increases. Furthermore, offshore wind farms have slightly higher error values than flat terrain wind farms. Similar conclusions are drawn by Kariniotakis et al. in [136], where it is also highlighted that the spatial resolution of

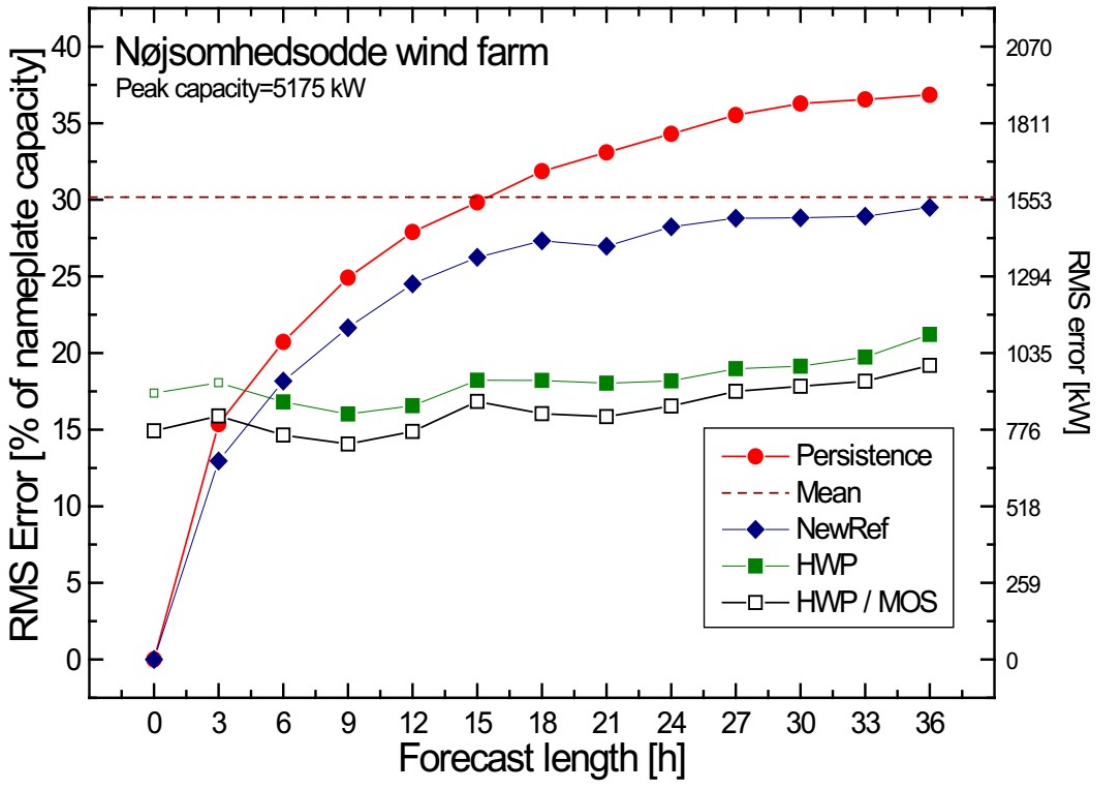


Figure 3.3: Root mean square error at different forecast lengths for various prediction models. Source: Giebel et al. [70].

the NWP models is critical for complex terrains. The impact of seasonal variability in Germany on forecast error was investigated by Lange and Focken in [137], who found that errors are larger in winter due to the higher levels of wind speed and larger uncertainty associated with storms, connected in particular with low pressure systems and their frontal zones. In summer, the main source of error is thermal stratification and especially at sunset in high pressure situations, when the large increase in wind speed is hard to predict. Errors also increase for more unstable atmospheric situations, as found by Pinson and Kariniotakis in [129]. On the other hand, the work by Lange and Heinemann [138] shows that forecasting uncertainty is not correlated with the magnitude of the wind speed forecast but does depend on the pressure, with errors being larger in low pressure situations.

The widely cited article by Madsen et al. [48] proposes a standardised protocol for evaluating wind power forecasts. The authors recommend that forecasts are assessed on a *test set* which was not used to train the model and that the evaluation

framework is defined clearly before any analysis is carried out. A minimum set of error measures are then suggested as well as some strategies for the visual inspection of errors. The persistence, moving-average, and climatology benchmark methods are also presented (cf. Equations (3.10), (3.11), and (3.12)). In the context of wind power forecasting, persistence is difficult to outperform for short lead times (0–6 hours) due to the quasi-stationarity of the atmosphere [69], while climatology is hard to beat after 24 hours. To take advantage of these aspects, another benchmark model was proposed by Nielsen et al. in [133] (sometimes referred to as the "new reference model"), which is in fact a weighting between the persistence and the mean methods.

Forecast errors and derived scores can be *normalised* by the installed capacity of the wind farm P_{nom} so that results are independent of its size, hence allowing for comparison among different sites. Alternatively, errors can be normalised by the average power production measured over the test period. This normalisation allows better assessment of the monetary consequences of the model errors as a function of the capacity factor of the wind farm, that is the ratio of the energy actually produced by the farm to the energy that could have been produced if the turbines ran at nominal power over a given time period [5].

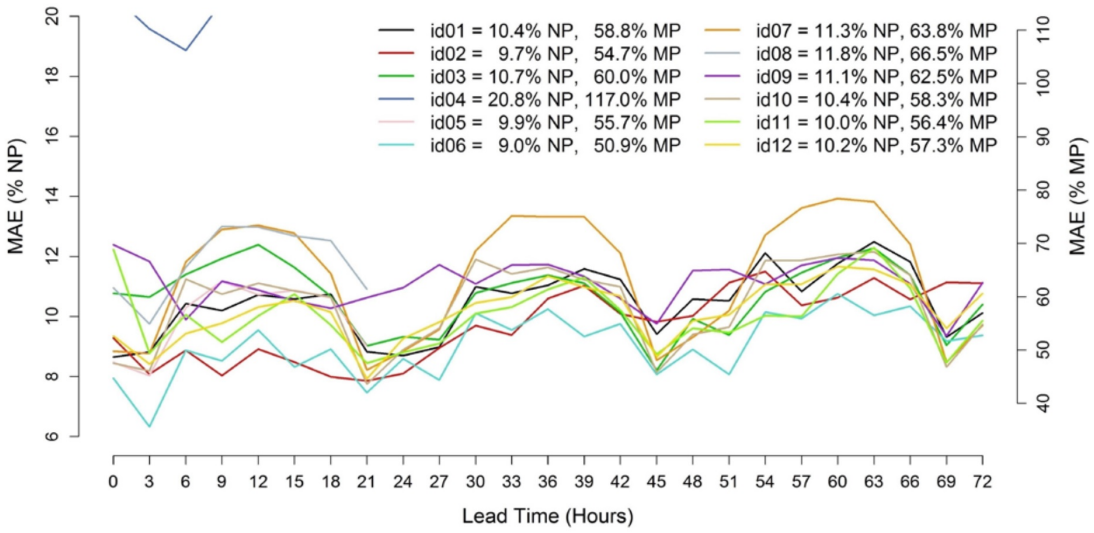
The review article [8] by Giebel and Kariniotakis reports that typical forecast accuracies as measured by RMSE are in the range of 9-14% of the installed capacity for a 24 hour horizon. When the forecast upscales to regional or national areas, RMSE is in the order of 3-5% of installed capacity. The article [139] by Sperati et al. also provides a good picture of the accuracies of state-of-the-art deterministic forecasts. The work reports the outcomes of the benchmarking exercise conducted on two different wind farms. The first one is located in Abruzzo (Italy) on complex terrain and has a nominal power of around 100 MW; the second one is located in Klim (Denmark) on flat terrain and has a nominal power of 21 MW. Participants were required to issue forecasts from 0 to 72 hours ahead with three-hourly resolution and were evaluated over a test period of one year. The values of the

mean absolute error for each look-ahead time for the various forecasts are shown in Figure 3.4 for the two wind farms. For the complex terrain test case (Abruzzo, Figure 3.4a), errors have quite similar trends with nearly all models showing a noticeable daily cycle, and errors being larger during evening and night hours. Note how for some models, the dispersion of error values can be significant. For the flat terrain test case (Klim, Figure 3.4b), errors again have similar trends, but the dispersion of error values is lower. Moreover, the tendency of forecast errors to increase for larger forecast horizons is very evident.

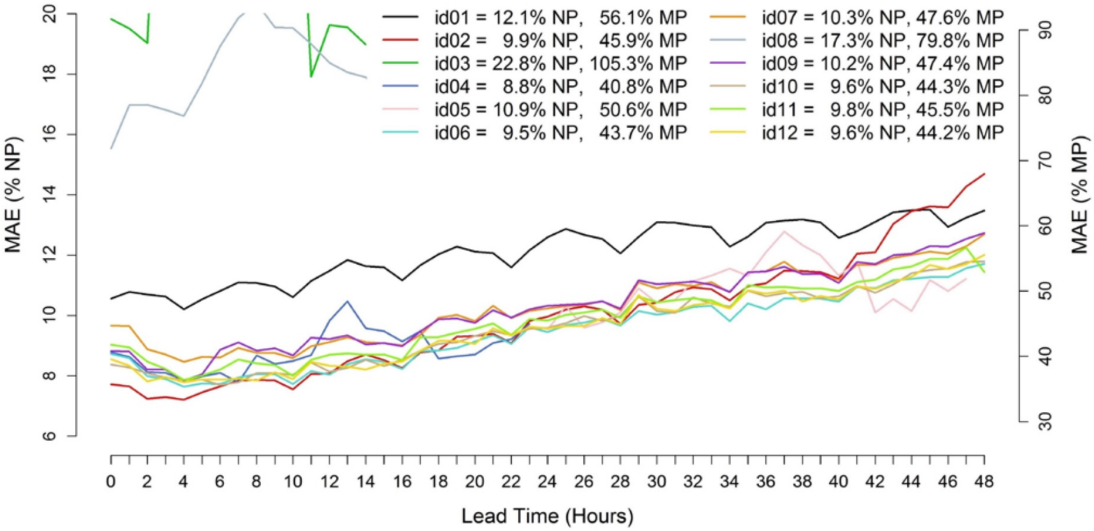
3.3 Short-term Electricity Price Forecasting

Prices in both day-ahead and balancing markets determine the financial obligations of market participants, as will be discussed in detail in Chapter 4. Therefore, price forecasting is of critical importance for power producers and energy traders, especially in the case of stochastic units like wind farms. Short-term electricity price forecasting (EPF) generally involves look-ahead times between a few hours and a few days. There is a significant body of literature on EPF, with extensive and detailed reviews to be found in the articles by Weron [34] and by Aggarwal et al. [140]. However, a large majority of the published papers are concerned with point forecasts only. Moreover, the focus has been primarily on day-ahead (or spot) market prices, with the forecasting of prices from the balancing market having received very little attention so far.

An overview of the literature on day-ahead market price forecasting is presented in Section 3.3.1, including a detailed review of statistical approaches, a discussion on their accuracy and on the approaches to forecasting series with multiple seasonalities. The literature on balancing market price forecasting is then reviewed in detail in Section 3.3.2. Note that although none of the papers cited in the next Sections investigate prices from the Irish electricity market, the framework considered is the same as that of interest in this work, that is, forecasting day-ahead and balancing market prices for the next day in wholesale electricity markets.



(a) Abruzzo (complex terrain) test case.



(b) Klim (flat terrain) test case.

Figure 3.4: Mean absolute errors for the Abruzzo (top) and Klim (bottom) wind farms. Values on the left y-axis are normalised by the nominal power (NP), while those on the right y-axis are normalised by the mean power (MP). Source: Sperati et al. [139].

3.3.1 Forecasting Day-Ahead Market Prices

Following the same classification criteria used by Weron in [34], EPF models can be categorised as:

- multi-agent models;
- fundamental models;
- reduced-form models;

- computational intelligence techniques; and
- statistical models.

A brief description of each approach is provided below, although the last two approaches are by far the most popular, with their share in the literature being somewhat equal.

Multi-agent models, which include production-cost, static equilibrium and agent-based models, simulate the behaviour of the electricity market system as a result of the interactions between its participants — referred to as "heterogeneous agents". They are flexible tools used to assess strategic behaviours in electricity markets because the competition between participants can be represented without limitations. In some cases (see the static equilibrium models developed by Ruibal and Mazumdar in [141]), these models can be employed to predict prices when no historical values are available but supply costs are known. On the other hand, they require the identification of the relevant players and of their bidding strategies, and how they interact with each other, which can be a difficult task and introduces significant modelling risk. While pricing processes are generated by matching demand and supply in the market, the main objective of these models is to simulate the mechanics and operations of the power system. Hence, their application is more focused on qualitative rather than quantitative results.

Fundamental models include the two subclasses: parameter rich; and parsimonious structural models [142]. They attempt to capture the functional associations between fundamental physical and economical drivers (for example, load, weather, etc.) and the price of electricity to describe the price dynamics. Predictions of the prices are then obtained by modelling and forecasting independently these fundamental inputs. The construction of these models is highly dependent on the availability of data from the market. Since some of the fundamental inputs may be collected with monthly or weekly frequency, fundamental models are more suitable for medium-term forecasts and are mainly employed in risk-management and derivative pricing [143]. Moreover, models are constructed under specific as-

sumptions about physical and economical relationships in the system. Violations of these assumptions significantly affect the price predictions, therefore making it very challenging to incorporate stochastic input fluctuations in the model.

Reduced-form models are finance-inspired models that aim to describe price dynamics in a simplified but realistic way. They are generally unable to provide accurate price forecasts at the hourly time scale, but can reproduce the main features of electricity prices (such as marginal distributions, price variations) at the daily level [144]. Mean-reverting jump-diffusion and Markov regime-switching models are two of the more popular alternatives since they provide a balance between the capacity to capture the main features of prices and model parsimony, so that the computational burden does not impede their online use in trading. A review and extended discussion on these models can be found in the book [144] by Benth et al.. Their common use is in derivatives pricing and risk analysis, although their application in volatility or price spike forecasting is reportedly good (see for example References [145] and [146]).

Computational intelligence (CI), or artificial intelligence (AI), *techniques* use a combination of learning, evolution and fuzziness to develop models able to adapt to complex dynamic systems. The most popular classes of CI techniques are artificial neural networks (ANN), fuzzy systems, support vector machines and evolutionary computation. Within the EPF literature, ANN have received the most attention [140, 147]. The main strength of CI models is their flexibility and ability to model complex and non-linear processes [34]. These models have an advantage over statistical ones when modelling the usually highly volatile, spiky, and non-linear electricity price processes. Often, however, this advantage in modelling ability does not translate into better forecasting skills, as found by Aggarwal et al. in [148]. Another significant limitation of CI techniques is that they require large amounts of historical data for training and calibration, which makes them unsuitable when historical data are unavailable or their size is limited.

Statistical models forecast the price through a mathematical combination of

historical prices and/or historical or present values of exogenous factors. These factors are typically system-level consumption (i.e. demand) and generation (for example, wind power), weather variables (for example, temperature, wind speed, precipitation, solar radiation), fuel costs and reserve margin (or surplus generation, i.e. the difference between available generation and predicted demand). Statistical EPF models have originated directly from the load forecasting literature, often simply by substituting prices for loads and loads for temperatures in the mathematical expressions of the model. Models can be categorised as additive or multiplicative, depending on whether the forecast variable is the sum or the product of a number of components. Within the EPF field, the former are significantly more popular [34]. The model's components can often have some physical meaning (for example, if load or temperature are used as predictors), a desirable feature which allows practitioners to better understand their behaviour. The main limitation of statistical models is the ability to model the usually non-linear and spiky behaviour of electricity prices; nevertheless, their performance in practical applications is comparable to that of computational intelligence models, which represent their non-linear alternative [148]. Given the data driven approach, the forecasting accuracy does not depend just on the algorithms employed and their numerical efficiency, but also on the quality of the processed data and the capacity to include fundamental factors (e.g. historical and forecast demand, weather forecasts, fuel prices) into the model [149].

The implementation of the trading strategies described in Chapter 4 requires day-ahead forecasts of electricity prices (specifically, prices from the day-ahead and balancing markets) with half-hourly resolution. The data available for the training and testing of the potential model include historical values of DAM and BM prices (since the start of the I-SEM) as well as forecast and observed values of some fundamental price drivers, namely demand and system-level wind power generation. However, due to the recent formation of the new market, the length of the data set is limited (seven months, overall). For the aforementioned reasons,

a statistical approach is chosen to forecast electricity prices in this work. Hence, the existing literature on this class of models is reviewed in more detail in the next Section.

3.3.1.1 State-of-the-art of statistical approaches

The *similar-day* method is one of the simplest statistical approaches in EPF. The rationale behind the method is to look into historical data for days which have similar characteristics to the day that needs to be forecast (for example, same day of the week, day of the year, holiday type), and then use those historical values as forecasts. A popular implementation of this approach is that proposed by Conejo, Contreras et al. in [150] (called the "naïve method" in the paper) where:

- a Monday is similar to the Monday of the previous week, and the same rule applies for Saturdays and Sundays; and
- a Tuesday is similar to the previous day, and the same rule applies for Wednesdays, Thursdays, and Fridays.

This method is commonly used as a benchmark when developing a new model and, as argued by the authors, advanced forecasting procedures that are not well calibrated fail to outperform this simple method with surprising frequency.

Another basic method which is very often used as a benchmark in load forecasting but less in EPF is *exponential smoothing* [151]. Here, the prediction is constructed from an exponentially weighted average of past observation:

$$\hat{y}_t = \nu y_t + (1-\nu)s_{t-1} = s_t$$

Each smoothed value s_t is the weighted average of the previous observations, where the weights decrease exponentially depending on the value of parameter $\nu \in [0, 1]$.

AutoRegressive Moving Average (ARMA) models are at the heart of all advanced time series methods in EPF. An $ARMA(p, q)$ model, where p is the order of the autoregressive part and q that of the moving average part, accounts for

both the random nature and the time correlations in the electricity price time series. This approach assumes the (weak) stationarity of the time series under analysis, which otherwise needs to be transformed first. The most simple and common transformation is differentiation, which was introduced by Box et al. in [152] and results in the so-called *AutoRegressive Integrated Moving Average* (ARIMA) or Box-Jenkins model. When the time series exhibits one or more seasonality (e.g. daily, weekly) and it is necessary to differentiate at lags longer than one, seasonal ARIMA (SARIMA) models can be used. The model parameters in all these AR-type models are generally estimated with a two-step process:

- model identification using information criteria (e.g. Akaike's Information Criterion, Schwarz's Bayesian Information Criterion) to find the optimal trade off between improvement in fit and model complexity/overfitting; and
- estimation of the coefficients, for example by least squares regression, recursive least squares, maximum likelihood, or prediction error method.

While there are some applications of ARIMA models or their variations[3] in EPF, they are most commonly used in combination with other approaches which will be presented next.

Electricity prices are heavily influenced by various *exogenous* factors, in particular load profiles, generation capacity, and weather conditions (for instance, temperature, wind). AR-type models relate the forecast variable (i.e. the price) to its own past, but are not able to capture the relationship with these fundamental variables. For this purpose, AR models can be combined with linear regression to form the so-called *AutoRegressive Moving Average* models with *eXogenous variables* (ARMAX). These are among the most popular approaches in EPF and, in the literature, they are also referred to as dynamic regression, transfer function, Box-Tiao, intervention or interrupted time series models.

Recall that the general purpose of regression is to determine the mathematical relationship between a number of independent (or predictor) variables and a depen-

[3]Here, variations refers to AR, ARMA, and SARIMA models.

dent (or forecast) variable. In its classical form, the relationship between variables is assumed to be linear and the coefficient estimation is based on least squares or maximum likelihood methods. An ARMAX model can in fact be interpreted as a multiple regression model where the error series is assumed to follow an ARMA process. The inclusion of exogenous variables follows the same mechanism for all ARMA-type models: the present value of the price is expressed as a linear combination of its past values, past values of the noise, and present and past values of the exogenous variable(s). The model's coefficients are then estimated using least squares, maximum likelihood, instrumental variable techniques or prediction error methods.

AR processes can also be coupled with regime-switching models, where the price series is represented by a set of separate states (or regimes), each modelled with a different process, and the regime is determined by an unobservable (latent) or observable variable. The former case corresponds to the Markov regime-switching models introduced earlier. Within the latter case, *Threshold AutoRegressive* (TAR) models and extensions thereof are most commonly used. In the basic case, the regime is determined by an observable variable in relation to a threshold value [153]. If the threshold variable is chosen as a lagged value of the price series itself, this results in the so-called Self Exciting TAR (SETAR) model [154]. Exogenous variables can also be included by simply replacing the AR process with an ARX one, leading for example to TARX [155] and SETARX [156] models.

The linear AR(X)-type models discussed so far assume a constant variance and covariance function (i.e. homoskedasticity). Price time series, however, exhibit non-linearities which in some cases are related to a non-constant conditional variance, or heteroskedasticity. The first model developed to address this issue was the *AutoRegression Conditional Heteroskedastic* (ARCH) by Engle in [157], where an autoregressive process is used to represent the conditional variance of the series. This approach was then extended by Bollerslev in [158] into the *Generalised AutoRegressive Conditional Heteroskedastic* (GARCH) model, where the conditional

variance additionally depends on a moving average of past conditional variances. GARCH models have been employed in short-term EPF in combination with AR-type processes with contrasting results, as will be discussed next.

Comparison of statistical forecasting techniques

As stated earlier, the cases where forecast models are based exclusively on an AR-type process are relatively rare in the literature. One example is in the work conducted by Cuaresma et al. in [159], where various AR and ARMA processes are used to forecast hourly spot prices from the German Leipzig Power Exchange. Conejo, Plazas et al. propose a model in [160] where the historical prices are first decomposed in a set of constitutive series using the wavelet transform, then values are forecast using ARIMA models fitted on the constitutive series, and finally the series of forecast prices is reconstructed through the inverse wavelet transform. A mixed model is proposed by García-Martos et al. in [161] to forecast market-clearing prices in Spain, where each of the 24 hourly time series of price is modelled separately using ARIMA models. The models are identified and estimated on training sets formed by either the whole hourly time series or by weekday prices only (i.e. from Monday to Friday). Prices on weekdays are then forecast with the models calibrated on weekdays only, while prices on weekends are forecast with the models calibrated on the whole time series (weekdays and weekends together).

Contreras et al. use ARIMA and ARIMAX models in [162] to predict day-ahead electricity prices in California and Spain. Both models are estimated on the full time series, i.e. on all hours, and explanatory variables are demand for the Californian market, and demand and available daily production of hydro units for the Spanish market. The two models give similar errors for the Spanish case, while the ARIMAX performs better in the Californian market. The same case study is considered by Nogales et al. in [163]. Here, the authors use ARMAX (called "transfer function" in the paper) and ARX (called "dynamic regression" in the paper) models to predict prices and consider only demand as the explanatory variable. Both models perform considerably better than the ARIMA used

by Contreras et al. in [162]. The same approach is used by Nogales and Conejo in [164] on a one-year data from the Pennsylvania-New Jersey-Maryland (PJM) Interconnection market in the USA. Again, the results indicate that the ARMAX model with demand as the explanatory variable (called "transfer function" in the paper) outperforms the standard ARIMA process. In a related work by Conejo, Contreras et al. [150], prices from the PJM Interconnection market are forecast using several models, including ARIMA, dynamic regression (ARX), transfer function (ARMAX), a multilayer perceptron (CI model) and a wavelet regression technique. In the study, the best performance is achieved by the AR-based models including exogenous variables.

Several models are considered by Knittel and Roberts in [165] to forecast spot prices in the Californian market: mean-reverting diffusions, jump diffusions, a seasonal ARMA model, an AR-EGARCH process (with demand as explanatory variable), and a seasonal ARMA model incorporating a polynomial of the temperature of order 3 as explanatory variables (i.e. a SARMAX model). This last model has the lowest errors as measured by RMSE, although the difference from the SARMA model is small. The authors also observe that the GARCH-based specification outperforms the other time series models during the California market crisis period (i.e. from May to August 2000). The same model, however, yields the worst of all forecasts in the two years before the crisis. A similar conclusion is reached by Garcia et al. in [166], who find that models including GARCH components outperform other statistical approaches only during highly volatile and spiky periods.

Weron and Misiorek in [167] compare a number of different time series models including TAR and TARX (with the exogenous variable being demand), and find their forecasting performance to be quite poor. In a similar study two years later, Weron and Misiorek [168] investigate the accuracy of 12 time series models for day-ahead EPF in the Californian and Nord Pool markets. The approaches include AR(X) models and their extensions, and mean-reverting jump diffusions. One of the two main findings of this work is that models using system demand or

air temperature as the exogenous variable perform better than pure time series models. The other finding is that semi-parametric models provide better point and interval forecasts and can perform well under various conditions. Interestingly, TARX models have, on average, the largest errors during the regular and less spiky test periods despite often being one of the best forecasts on single weeks, indicating that when the forecast is off, the associated error is large. Kristiansen proposes in [169] an ARX model based on that by Weron and Misiorek in [168], but where prices are modelled across the whole time series rather than on each single hour separately. Moreover, system wind power is considered together with demand as exogenous variables. However, the data period used is larger and more recent, so results are not directly comparable.

Extending the work by Weron and Misiorek in [170], Misiorek et al. develop in [171] a simple ARX model with lags of 24, 48 and 168 hours (corresponding to one, two, and seven days, respectively) and demand forecast as exogenous variable. Evaluated together with AR(X)-GARCH, TAR(X), and Markov regime-switching models, the model has the best or second best accuracy over different test periods. Jonsson et al. adopt a similar approach in [172] using an ARX model with the same lags as in [171] to account for residual autocorrelation and seasonal dynamics. System demand and wind power generation are included as explanatory variables based on a non-parametric and time-varying regression model to account for their non-linear and non-stationary influence on spot prices.

Forecasting multiple seasonal patterns

When forecasting hourly (or half-hourly) prices, both the daily and weekly seasonality of the series have to be captured (see Section 2.2.2.1). This can be achieved by using seasonal models (i.e. SARIMA(X)) or combining the autoregressive structure of the model with dummy variables, as for example in the works by Weron and Misiorek in [168] and by Kristiansen in [169]. Since each *load period* displays quite a distinct price profile due to the daily variation of demand, costs and operational constraints, another approach is to perform the forecasts across each period

separately. This results in 24 (or 48) separate models and consequently sets of parameters to be estimated, as done for example in References [147], [173], and [174]. The inspiration for this solution comes from the extensive literature on demand forecasting, where this so-called *multi-model* specification is usually favoured for short-term predictions (see, for example, References [175–177]).

Weron and Misiorek in [170] consider simple time series models, namely ARMA and ARMAX processes, to forecast market clearing prices in the Californian Power Exchange. The modelling was implemented separately across the hours, resulting in better forecasts compared to the single ARIMA specification for all hours proposed by Contreras et al. in [162]. A similar conclusion is reached in the study described earlier by Misiorek et al. in [171]. The mixed model by García-Martos et al. in [161] described above, where each hour is modelled separately, has a better accuracy than those by Contreras et al. in [162] and Conejo, Plazas et al. in [160], where all hours are modelled together. The results in the above-mentioned work [159] by Cuaresma et al. also showed that modelling each hour of the day separately brings a significant improvement compared to modelling the whole time series.

3.3.2 Forecasting Balancing Market Prices

As mentioned at the start of Section 3.3, the literature on balancing market price forecasting is very limited. Moreover, many of the works touching upon this topic concentrate only on forecast horizons of a few hours and after the clearing of the day-ahead market. In many cases, time-series-based models are borrowed from the literature on day-ahead price forecasting and adapted to the balancing market framework. Following the classification used by Klæboe et al. in [178], balancing price forecasting methods can generally be divided into two types: those that model the future balancing *state* (that is, whether the system will be long, short, or balanced) explicitly and those that do not include state information or do so implicitly. Many authors have pointed out the difficulty to forecast balancing market prices, especially when issuing predictions for the next day. In their review

of price forecasting statistical models [140], Aggarwal et al. observed that models for spot price forecasting generally achieve higher accuracy levels compared to those for balancing price forecasting.

A neural-network-based method is developed by Ma et al. in [179] to forecast day-ahead and real-time (i.e. balancing market) LMPs in the PJM and New England markets in the USA. The model consists of multilayer perceptron networks using the Decomposed Extend Kalman Filter. Since real-time prices are heavily affected by market congestions, a heuristic method is also developed to capture system transmission outages using published information on future outages. Real-time prices are forecast before and after the day-ahead market is cleared using two different neural networks. Results from the second network are more accurate, indicating that information from the cleared day-ahead market can improve balancing market forecasting. The authors observe that accurate predictions of these prices are hard due to their strong dependence on many unforeseeable factors, such as real-time demands, system operations, errors in demand or weather forecasting and behaviours of market participants.

Olsson and Söder present in [180] a model for the generation of balancing market price scenarios in the Nordic market. The simulated scenarios are used to construct scenario trees for trading tools which consider uncertainties based on stochastic optimisation. The price forecasting model is based on SARIMA and discrete non-time homogeneous Markov processes. Prices for upward and downward regulation are treated separately, and the assumption is made that these prices are constrained by the spot (i.e. day-ahead) price. Although not explicitly stated in the paper, this seems to imply that there is knowledge of day-ahead prices when forecasting balancing market prices; in other words, balancing prices are forecast after the day-ahead market closure. The authors observe that factors affecting the prices are the level of water in hydro systems, the market structure, and the demand for real-time balancing power (i.e. the imbalance volume). This demand, in turn, depends on unforeseen events in the system (e.g. production unit outages) and forecast errors

(e.g. wind power production, load, and temperature).

In a related work [181], Brolin and Söder extend the applicability of the previous model and improve the performance of the price simulations by including exogenous variables in the model. The probabilistic price model proposed in this paper is based on non-linear price series processes and uses day-ahead prices and real-time balancing demands as input variables. Again, the forecasts are issued after day-ahead market prices are published.

Jaehnert et al. develop a method in [182] to forecast balancing prices in southern Norway in order to investigate the use of interconnections for the trading of regulating power. The model takes into account the regulating volume and is split into a long-term and a short-term part. The long-term model describes the price-volume dependence in the balancing market and is based on the statistical model developed by Skytte in [183], although price difference is modelled here (instead of the balancing price). It includes a deterministic and a stochastic part and the three different states of upward, downward and no regulation are modelled separately. The deterministic part is based on linear regression, while the stochastic part is modelled by a normal distribution in the case of no regulation, and by the Extreme Value distribution in the upward and downward regulation cases. A short-term model based on the SARIMA process is then used to forecast the regulating states and generate volume scenarios in the next 48 hours. Given the forecast regulating states and volume scenarios, the price difference is finally estimated based on the statistical relationship found in the long-term model.

Ji et al. analyse the PJM 5-bus system in [184] and observe that the main challenge in forecasting real-time LMPs stems from the electricity prices being location dependent. Determining the LMPs then requires the optimisation of the power flow conditional on several transmission and generation constraints. A probabilistic technique is proposed to forecast real-time prices for forecast horizons of 6-8 hours. Prices are modelled as a time non-homogeneous Markov chain and a Monte Carlo technique is used to estimate transition probabilities, based on which

the future price distribution is computed. The technique is compared against two benchmarks: a deterministic baseline using the mean value of load to compute price trajectories and an ANN with two hidden layers. The proposed technique is the most accurate on average over the five buses, although the neural network performs better in the bus with the least spiky prices.

A stochastic programming model is developed by Boomsma et al. in [185] to analyse the potential of coordinated bidding for market participation trading in sequential markets, namely the Nordic day-ahead and balancing markets. Within this work, a method for the generation of market price scenarios is formulated. First, autoregressive models are fitted to day-ahead and balancing prices, which are described by a SARIMA process and an ARMAX process with spot prices as exogenous variables, respectively. Then, scenario tree sampling and reduction is used to obtain spot price scenario paths and balancing price scenario values. Balancing price forecasting, however, is developed only for 1-hour ahead horizons, and cleared day-ahead prices are used as an input to the forecasting model.

To the best of our knowledge, the survey by Klæboe et al. in [178] constitutes the only review of time-series-based methods for forecasting of balancing market prices. Five previously published (and in some cases extended) models are benchmarked for day-ahead point and interval forecasts. In fact, rather than the balancing market price, the models are developed here to predict the price difference (called "balancing penalty" in the paper). The analysed models are:

- an ARMA(1,1) model based on the approach used by Jaehnert et al. in [182];
- an ARX model based on that proposed by Boomsma et al. in [185] using current and past values of the day-ahead price as input variable;
- a regime-switching autoregressive model (ARM) based on that developed by Olsson and Söder in [180], with the Markov model that determines the switch being hour-specific rather than duration-dependent like in the original work;
- a regression model (EXO) based on that presented by Jaehnert et al. in [182]

using the balancing volume, the day-ahead price, and the overall power production in the zone as regression variables; and

- a similar-day (naïve) forecast based on that defined by Conejo et al. in [150].

The ARMA and ARX models forecast the price directly, with the balancing market state being defined implicitly. The ARM and EXO models, on the other hand, use explicit information on the balancing state. The study focuses on the Nord Pool NO2 price zone in Norway and a 13-week period is used for out-of-sample verification. The results for the point forecasts show that all four advanced models had a similar performance and provided better forecasts than the naïve model in all weeks except two. More interestingly, however, the authors observe that the average errors from these models are of the same magnitude as those that would be obtained from a constant forecast of the price difference equal to zero. This suggests that the information available before the day-ahead market closure is already reflected in the day-ahead prices, thus negating the value of generating forecasts of balancing prices at all. Indeed, these are issued based on the same a priori knowledge. The results for the probabilistic forecasts show that models without an explicit formulation of the balancing market state tended to overestimate the variance and therefore produced interval forecasts that are too wide. On the other hand, models including the balancing market state information explicitly reflected the distribution more closely and produced better interval forecasts. The authors conclude by stating that "*it is hard to predict the balancing market before the closure of the day-ahead market*" and recommend using models with balancing state information for probabilistic forecasting.

The forecasting of imbalance prices in the framework of day-ahead energy trading is considered by Jónsson et al. in [186], who propose an approach to predict the price difference (called "imbalance penalties" in the paper) in the Nord Pool market. Due to the dual pricing scheme adopted in the market for regulating power, the net balance of the system (i.e. the imbalance direction) induces a regime-switching behaviour in the prices depending on whether the system is in

up- or down-regulation. Consequently, the proposed forecasting approach consists of two parts: the estimation of the probability of the imbalance direction and the forecasting of the penalties conditional on the imbalance direction. Holt-Winters (exponential smoothing) models with a daily seasonal cycle are used to model both variables, and the two predictions are then combined with the law of total expectation to yield the expected imbalance costs. The exponential smoothing models are employed either in their standard formulation or conditioned on a number of exogenous variables, namely forecasts of day-ahead price, system load, and wind power penetration. The models are fitted on the whole time series, resulting in the estimation of one set of parameters for each model. Their performance is evaluated on a two-year period and they are found to outperform the climatology benchmark by a substantial margin.

A method based on hidden Markov models (HMM) is presented by Dimoulkas et al. in [187] to forecast balancing market volumes and prices 12-36 hours ahead. The focus again is on the Nordic market and the forecast variable is the price difference (called "BM premium" in the paper). Two separate models are trained for up- and down-regulation on the one-year data with a moving window approach and a forward-backward algorithms based on that by Kouemou and Dymarski in [188] is employed to find similar patterns in historical data and use them as forecasts. This model is compared against the Markov-SARIMA one developed by Olsson and Söder in [180], with results showing a slightly better accuracy than the former. Nevertheless, the authors observe that the scores obtained are still quite poor and the models are unable to capture well the balancing market movements, and conclude that factors influencing the BM prices need to be identified in further research.

To summarise, the review of the literature on day-ahead price forecasting has highlighted that CI and statistical models are the two most popular approaches. Due to the limited amount of historical data in the present study, CI techniques were discarded and the focus was restricted to statistical models. Within this class,

ARMAX models tend to achieve the best performance, and modelling each load period separately improves the forecast accuracy in comparison to having a single specification for all hours. On the other hand, the literature on imbalance price forecasting for the next day is very limited and mostly concerned with Scandinavian markets, and the evidence for one class of models outperforming the others is scarce.

As will be discussed extensively in Chapter 4, the interest is primarily in forecasting the sign of the difference between day-ahead and balancing market prices, rather than their value. In single-price balancing markets, this information can normally be deduced directly from the balancing state, as up-regulation (down-regulation) corresponds to a positive (negative) price difference (see Section 2.1.3). However, one of the highlights from the analyses performed on the I-SEM market data in Section 2.2.2 was the significant occurrence of price anomalies in the balancing market. During these instances, the sign of price difference would not be consistent with the direction of the system imbalance.

In light of this, two alternative approaches are developed within this work to forecast the sign of price difference. In one, the system imbalance is modelled as a binomial distribution and forecast with logistic regression models; the sign of price difference is then deduced directly from this forecast. In the other approach, a new method is proposed, where day-ahead and balancing market prices are forecast individually and these predictions are then combined to estimate the sign of price difference. ARMAX processes are chosen to forecast day-ahead and imbalance prices, with the multi-model approach being implemented to model the time series. The methodology used in the two approaches will be presented and explained in detail in Section 4.3.

3.4 Short-term Value Forecasting of Wind Power

The structure and operation of wholesale electricity markets were described in Chapter 2, explaining that participants are required to establish their position in the day-ahead market to sell their production. As discussed in Section 3.2, short-

term wind power forecasting represents one of the fundamental tools to enable the participation of wind power producers in electricity markets where the penetration of stochastic renewable sources is significant. In addition, knowledge of future market conditions can be beneficial to participants and enables them to trade strategically.

The most basic approach to trading consists of directly bidding the point forecasts of wind generation, and will be referred to as the *baseline* strategy. Currently, this is the common practice among most utilities in Europe and USA, as found by Möhrlen et al. in [25]. However, a more accurate forecast does not necessarily translate to a higher profit, as demonstrated by Pinson in [189], particularly in single-price imbalance markets where energy deviations can receive a more profitable price if they help bringing the system back to balance. Referring to the types of goodness defined by Murphy in [27], rather than a forecast of higher quality, wind energy traders are mostly interested in prediction methods that can maximise their benefits; that is, forecasts of higher value. For example, wind farm operators might want to apply a trading strategy that reflects their target trade-off between return and risk. However, advanced offering strategies tailored to the end-user's needs cannot be designed using wind power point forecasts only. For this purpose, it is necessary to move to a probabilistic framework and/or to integrate information on future market conditions (such as the direction of the system imbalance or electricity prices) in the decision-making process. Moreover, the participant's risk attitude should be modelled appropriately and included in the strategy. The definition of (acceptable and unacceptable) risks is provided by Artzner et al. in [190], where a framework is presented for the analysis and implementation of measures of risk. Risk attitudes for individuals are presented by Clemen in [191]; the author also discusses various approaches to representing preferences in decision-analysis that incorporate such attitudes. The work by Conejo et al. [192] provides an insight into risk management and the modelling of the uncertainties which affect trading decisions; risk measures such as variance, expected shortage, value-at-risk and con-

ditional value-at-risk are defined, and risk-neutral and risk-averse decision making are discussed in the context of stochastic problems. The market participants' attitudes towards risk and the measures relevant to this work will be presented in detail in Section 4.1.3.

The participation of wind energy in electricity markets and strategic bidding in the day-ahead market has been investigated with growing interest in recent years, with studies considering various power systems and prediction methodologies. However, the majority of the published works concentrates on markets adopting a dual-price imbalance settlement. Strategies developed for this type of markets account for the fact that energy deviations cannot lead to an extra profit; indeed, as discussed in Section 2.1.3, the participant's imbalances are paid the day-ahead price if they help to balance the system, otherwise they are penalised. A selection of the most cited works on this subject is reviewed next.

3.4.1 Dual-price Markets

The work by Pinson et al. in [28] is one of the first to show that the integration of uncertainty information in the decision-making process can increase the market value of wind power forecasts. The authors propose a revenue-maximising strategy in a stochastic optimisation framework where the optimal bid is derived based on probabilistic forecasts of wind power, quarterly or annual average values of prices, and a loss function designed heuristically to model the sensitivity of the market participant to penalties. In the paper, the participation of a wind farm in the Dutch market is simulated over one year. The results show that the revenue obtained with the proposed strategy is higher than that obtained by only using point forecasts of wind generation, and that further improvements can be obtained by increasing the resolution of the price forecasts. Interestingly, the authors observe that reducing the amount of energy in imbalance does not increase the participant's revenue, and in fact, the opposite is true.

The approach used by Pinson et al. in [28] is generalised by Zugno et al.

in [193] by including probabilistic rather than point forecasts of market prices. First, the revenue maximising strategy referred to as "Expected Utility Maximization" (EUM) is formulated by revisiting the "optimal quantile" strategy derived by Bremnes in [123]. The performance of the EUM strategy is tested by simulating the participation of a wind farm portfolio in the Nord Pool market over a 10-month period. The proposed strategy performs better than the baseline in terms of revenue but is quite volatile, with the participant incurring large isolated losses when price forecasts are inaccurate. Hence, the constraining of bids is proposed, whereby the bid is limited within a certain interval around the point forecast of wind power. With this approach, extreme bid values and consequently large imbalances — which could be penalised by the system operator or violate the price-taker assumption — are avoided. Moreover, the reduced level of expected imbalance results in a decreased exposure to the risk of significant losses for the market participant. It should be noted, however, that while the EUM bid is derived using the probabilistic forecasts of wind power and market prices, the constrained strategies do not use the uncertainty information associated with those forecasts, with the allowed interval being a function of the point forecast of wind power only. Tested over the same 10 months, the constrained strategies outperform both the point forecast and the EUM bids, showing that setting a constraint on the expected deviations can effectively control the risk of severe losses and reduce the impact of forecasting errors on long-term revenues. Nevertheless, the benefits introduced by these advanced strategies are relatively small: the revenue obtained with the best of the proposed strategies is 0.37% higher than that obtained employing the baseline strategy; put into context, a perfect wind power forecast would have increased the revenue by 4.32%.

A bidding strategy aimed at minimising the imbalance costs is formulated by Matevosyan and Söder in [194] as a stochastic optimisation problem. A large number of scenarios are generated for wind power generation and spot and imbalance prices, and mixed integer programming is employed to find the optimal bid. The

proposed strategy is tested in the Nordic power market over six and 29 days and in both cases, the profit yielded is higher than or equal to that obtained from the baseline strategy.

A computationally simpler alternative to stochastic optimisation is proposed by Dent et al. in [195], where the analytical expressions to determine the optimal day-ahead market bids for risk-neutral generators are derived. For the general case of unknown prices and wind generation, a sloped bid curve that reflects the opportunity cost of trading in the day-ahead rather than real-time market is optimal; in other words, the optimal quantity is a function of the ratio of the expected short and long penalties (i.e. the price differences). The closed form methods outlined here require only the specification of the expected values of the random variables, rather than their full joint distributions or a scenario tree, and numerical methods are sufficient to solve the optimisation problem. The proposed strategy is applied to the British market but results from the example are very limited and a comparison with alternative approaches is not offered, impeding a critical appraisal of the outcomes. The risk-averse case is also investigated in the same work but only under the assumption of known spot prices, which is an unrealistic condition in day-ahead trading.

Deterministic and stochastic offering strategies are empirically assessed by Rahimiyan et al. in [196] by simulating the participation of a 100-MW wind farm in the Iberian Peninsula electricity market on an out-of-sample data set of four months (one month per season). The stochastic processes are described using seasonal ARIMA models and the strategies include the possibility of trading energy in the intraday ("adjustment" in the paper) market and consider risk management. The results show that in terms of expected profit and profit volatility, strategies based on a stochastic approach generally outperform the deterministic cases. The latter, however, are found to be more conservative overall, leading to higher values of the conditional value at risk (CVaR) (its definition and formulation are given in Section 4.1.3). Moreover, the performance of the stochastic programming ap-

proaches is significantly impacted by how accurately the market-related processes are modelled. Finally, the authors point out that the possibility of trading in the adjustment market would have beneficial effects not only for wind producers, who would increase their profit and the total energy traded, but also for the system as a whole, thanks to the reduction of wind energy deviation.

3.4.2 Single-price Markets

As mentioned earlier, strategic bidding in markets with a single-price imbalance settlement has received very limited attention. The optimal offering strategy for a risk-neutral producer in this type of markets is derived by Mazzi and Pinson in [197]. The solution — sometimes referred to as the *Zero/Max* strategy — is trivial and completely decoupled from the forecast wind power production. In fact, the bid maximising the expected revenue is *price-inelastic* and equal either to zero or to the whole wind farm capacity, depending only on the sign of the expected difference between imbalance and day-ahead prices. The same solution for a risk-averse participant is found by Morales et al. in [40], where the authors present also a risk-averse strategy that aims at maximising the CVaR rather than the expected profit but only for the dual-price case. Although mathematically correct, the profit-maximising Zero/Max strategy is hardly employable in practice for a number of reasons. Firstly, for a wind farm or a portfolio of wind farms of significant size, this strategy would cause large imbalances on the system which would inevitably impact the formation of prices, in particular those for the settlement of balancing actions. This way, prices would become endogenous variables, thus violating the price-taker assumption whereby the participant is too small to affect prices, and making the model of the market inconsistent. Secondly, to achieve the optimal scheduling of all units, bid quantities are expected to match the actual delivery as closely as possible; frequent and systematic large deviations caused by a market player would be unwanted by the system operator, and could lead to warnings and penalties from the market authorities. Finally, the potentially very large losses on

a single contract deriving from a wrong forecast of the sign of the price difference could cause financial problems to the participant.

Bidding strategies for participants with different levels of risk aversion are developed by Botterud et al. in [198], where the trading of wind energy in U.S. electricity markets with locational marginal prices is investigated. Uncertainties in wind power generation and prices are considered. Kernel density estimation is used to produce the probabilistic wind power forecast, while prices are forecast with rather basic models where weekdays and weekend days are modelled separately and a 4-week rolling window is used to estimate the necessary parameters. Expected profit, conditional value at risk, or utility function are then chosen as the decision criterion to reflect the risk attitude of the trader. The resulting optimization problem is solved using a generalised reduced gradient algorithm. The results from the four-month case study show that bids are lower under risk aversion, a tendency amplified with increasing volatility in imbalance prices. This is explained by the day-ahead prices being on average higher than the imbalance (here called "real time") prices and due to self-curtailment being allowed when imbalance prices become negative. Without a deviation penalty, the bids for profit-maximising and risk-taking strategies are directly defined by the difference between forecast imbalance and day-ahead prices and are independent of the wind power forecast, in accordance with Mazzi and Pinson in [197]. When a deviation penalty is applied, on the other hand, the optimal bids tend towards the average forecast wind power. The results, however, show that the baseline strategy not only yields similar results in terms of total profit (1.1% less than the profit maximising strategy), but also gives significantly lower risk exposure compared to any of the advanced strategies. The authors attribute the poor outcome from the latter to the inaccuracies in the price forecasts, although information on their performance are limited to a few summary statistics.

To address the issues related with the practical implementability of the Zero/Max strategy, Browell in [199] proposes a number of risk-constrained strategies based

on a probabilistic assessment of the sign of the net system imbalance ("system length" in the paper). Here, the forecast wind power generation is adjusted by a certain amount to exploit favourable imbalance prices while limiting the size of the expected imbalance. The adjustment factor is fixed a priori based on the participant's aversion to risk and the direction of the adjustment depends on whether the forecast probability of the system length is higher or lower than a critical value. This critical value is in turn a function of the point forecasts of day-ahead and imbalance prices obtained with ARMAX models. The proposed strategies were evaluated over a six-month period using data from the GB electricity market. The results show that when the allowed adjustments are modest, some strategies are able to increase the revenue while simultaneously reducing the participant's risk. Moreover, all strategies perform better when coupled with advanced rather than simple price forecasting models and can generate more revenue than the baseline strategy.

The attention on forecasting the value of wind energy for operators who trade in the day-ahead market has grown over the recent years. Nevertheless, the works that address electricity markets adopting a single-price imbalance settlement are few, and it is the objective of this research to advance the discussion on this subject.

The profit-maximising strategy proposed by Mazzi and Pinson in [197] does not account for risk aversion and is unfeasible in practice. The benefits obtained through the advanced strategies developed by Botterud et al. in [198] are relatively small compared to the baseline strategy, making it difficult to justify the adoption of such a complex approach. In some cases, the strategies proposed by Browell in [199] achieve results valuable to wind farm operators in terms of revenue and risk exposure. However, while the methodology is developed in a probabilistic framework, uncertainties in future market conditions are not fully integrated, as only point predictions of prices are used and adjustment factors are fixed beforehand.

In the present work, novel probabilistic value forecasting models are developed

for wind farms operating in electricity markets with single-imbalance price settlement. Value is measured in terms of revenue and exposure to risk, and the bidding strategies are designed for wind farms whose size is small relative to the power system so that their practical implementability is guaranteed. The proposed models combine probabilistic short-term predictions of wind power and of relevant electricity market quantities. Offers are then formulated to reflect the participant's risk profile, conditional on the uncertainty in future wind power generation and electricity market states. The objective is to enable wind farm operators to improve the value of their energy generation by increasing the revenue while simultaneously controlling their exposure to risk.

Chapter 4

Methodology

This Chapter concentrates on the methodology used to develop the value forecasting models for day-ahead trading of wind energy in markets adopting a single-imbalance pricing scheme.

The proposed models employ forecast information on wind power generation and electricity market conditions under a fully probabilistic framework to increase the producer's revenue while controlling its risk exposure. The flowchart in Figure 4.1 illustrates the steps involved in the methodology, which are discussed in detail in the following Sections. First, the problem of trading wind energy in the day-ahead and balancing markets is outlined in Section 4.1. Section 4.2 describes the main characteristics of the wind power forecasts available for the wind farm under study. Then, the two alternative approaches developed to forecast the sign of price difference are explained in Section 4.3. Finally, the different bidding strategies are formulated in Section 4.4.

4.1 Problem Formulation

The aim of this Section is to present the problem of trading wind energy in markets with a single imbalance pricing scheme. The assumptions made in this study are outlined and justified in Section 4.1.1. The revenue for a wind power producer is derived in Section 4.1.2 and the rationale behind strategic bidding explained. The

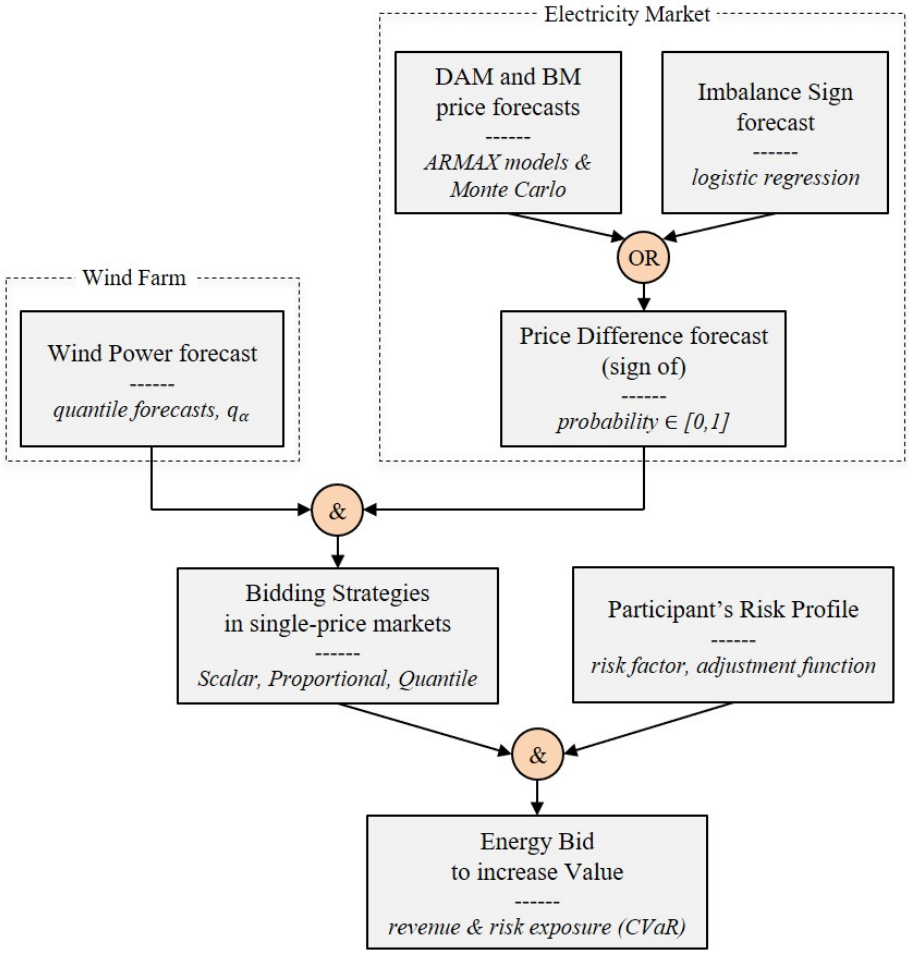


Figure 4.1: Flowchart of the methodology.

risk attitudes of market participants and measures of risk are then discussed in Section 4.1.3.

4.1.1 Assumptions

In this study, the following assumptions are made:

a. The participant trades only in the day-ahead and balancing markets. Intra-day markets give stochastic producers the opportunity to refine their contracted position taking advantage of forecast information issued closer to delivery time and thus generally more accurate. However, these markets are often characterised by very low liquidity (see the work by Weber in [35] as well as the analysis of markets' liquidity in Ireland in Section 2.2.2.1) and therefore the possibility of taking any corrective action in these trading platforms

is discarded here.

b. The participant is considered as a *price-taker*, meaning that market prices are not impacted by its bidding strategies. For this assumption to hold, the strategies developed in this work are aimed at wind farms whose size is small in comparison to the average system imbalance volume.

c. Traders cannot offer above the wind farm's installed capacity E_{max}. This is a common constraint in most markets.

d. Wind energy is bid at zero marginal cost. This is common practice for wind farm operators since they don't have short-run marginal costs.

e. Support mechanisms for renewable electricity such as feed-in tariffs, power purchase agreement (PPA)[1], or price premia are not considered in the revenue calculation, so that wind energy is treated like any other type of energy.

f. Wind power producers do not apply any control strategy on their own generation (for example, self-curtailment, storage). This assumption is made to focus on the value of forecasts in the market, where value here is intended as the benefits resulting from the use of predictions.

4.1.2 Revenue

For each trading period $t+h$, a wind power producer contracts at time t an amount of energy E_{t+h}^{DAM} in the day-ahead market, which is paid at the day-ahead (or spot) price π_{t+h}^{DAM}. At the time of delivery, the energy actually generated is E_{t+h} and if this quantity differs from the contracted one, this results in an *energy imbalance*:

$$d_{t+h} = E_{t+h}^{DAM} - E_{t+h} \tag{4.1}$$

[1]In Ireland, this is a contract under the Renewable Energy Feed-in Tariff (REFIT) under which the supplier agrees to purchase all of the output from the generator [200].

In a one-price balancing market, as explained in Section 2.1.3, this deviation is priced at the imbalance price π^{imb}_{t+h} independently of its direction. The revenue for the producer for trading period $t+h$ then is:

$$R_{t+h} = E^{DAM}_{t+h} \cdot \pi^{DAM}_{t+h} - \left(E^{DAM}_{t+h} - E_{t+h}\right) \cdot \pi^{imb}_{t+h} \tag{4.2}$$

where the first term on the right-hand side is the *day-ahead market revenue* and the second term is the *balancing market revenue*. This equation can be reformulated to isolate the revenue from the actual energy generated and the costs due to energy imbalances, yielding:

$$R_{t+h} = E_{t+h} \cdot \pi^{DAM}_{t+h} - \left(E^{DAM}_{t+h} - E_{t+h}\right)\left(\pi^{imb}_{t+h} - \pi^{DAM}_{t+h}\right) \tag{4.3}$$

The first term on the right-hand side is called the *perfect forecast revenue*, since this would be the revenue obtained having perfect predictions of the wind energy generation. The difference between imbalance and day-ahead prices was defined in Section 2.2.2.3 as *price difference*, although terms like "price penalty" or "balancing premium" are also used in the literature (see Section 3.3.2). Let us then define the last term in Equation (4.3) as the *balancing costs*:

$$C_{t+h} = \left(E^{DAM}_{t+h} - E_{t+h}\right)\left(\pi^{imb}_{t+h} - \pi^{DAM}_{t+h}\right) \tag{4.4}$$

so that the revenue now writes:

$$R_{t+h} = E_{t+h} \cdot \pi^{DAM}_{t+h} - C_{t+h} \tag{4.5}$$

From this last equation, it is clear that to maximise the revenue, balancing costs have to be minimised.

It is important at this point to underline that in a single-price market, balancing costs can be positive or negative. Indeed, looking at Equation (4.4), both factors in the product can have either a positive or negative sign. As a result, the producer's

revenue can be higher than the perfect forecast one if the energy imbalance has the opposite sign of the price difference. This is represented visually in Figure 4.2, showing that a producer's revenue decreases when the signs of price difference and energy imbalance are the same, but increases when the two signs are opposite.

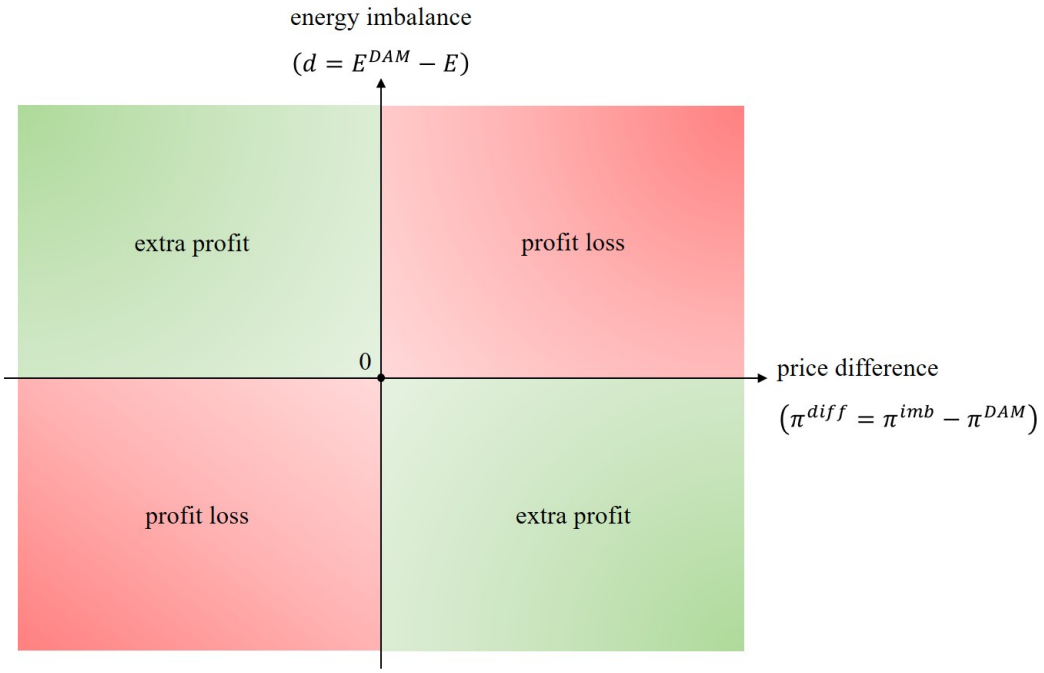


Figure 4.2: Balancing costs can be positive, hence resulting in a profit loss (red), when the signs of price difference and energy imbalance are the same, or negative, resulting in an extra profit (green), when the signs are opposite.

The expected revenue, also referred to as *expected monetary value* (EMV) in decision analysis [40], is found by taking the expectation $\mathbb{E}[\cdot]$ of Equation (4.3). Under the assumptions made in the previous Section 4.1.1, it results:

$$\mathbb{E}\left[\tilde{R}_{t+h}\right] = \mathbb{E}\left[\tilde{E}_{t+h}\right] \cdot \mathbb{E}\left[\tilde{\pi}^{DAM}_{t+h}\right] - \left(E^{DAM}_{t+h} - \mathbb{E}\left[\tilde{E}_{t+h}\right]\right) \cdot \mathbb{E}\left[\tilde{\pi}^{imb}_{t+h} - \tilde{\pi}^{DAM}_{t+h}\right] \quad (4.6)$$

where the $\tilde{\cdot}$ accent indicates a random variable. Note that the first term on the right-hand side is outside the control of the producer and can be regarded as a "fatal" component. In fact, the only decision variable in Equation (4.6) is the energy contracted at the day-ahead market E^{DAM}_{t+h} since prices are exogenous variables under the price-taker assumption.

If the objective of the producer then is to increase the expected revenue, the

expected balancing costs

$$\begin{aligned}\mathbb{E}\left[\tilde{C}_{t+h}\right] &= \left(E_{t+h}^{DAM} - \mathbb{E}\left[\tilde{E}_{t+h}\right]\right) \cdot \mathbb{E}\left[\tilde{\pi}_{t+h}^{imb} - \tilde{\pi}_{t+h}^{DAM}\right] = \\ &= \mathbb{E}\left[\tilde{d}_{t+h}\right] \cdot \mathbb{E}\left[\tilde{\pi}_{t+h}^{imb} - \tilde{\pi}_{t+h}^{DAM}\right]\end{aligned} \tag{4.7}$$

ought to be minimised. In particular:

- when $\mathbb{E}\left[\tilde{\pi}_{t+h}^{imb} - \tilde{\pi}_{t+h}^{DAM}\right] > 0$, the producer should bid less than the expected power generation, so that $\mathbb{E}\left[\tilde{d}_{t+h}\right] < 0$;
- when $\mathbb{E}\left[\tilde{\pi}_{t+h}^{imb} - \tilde{\pi}_{t+h}^{DAM}\right] < 0$, the producer should bid more than the expected power generation, so that $\mathbb{E}\left[\tilde{d}_{t+h}\right] > 0$; and
- when $\mathbb{E}\left[\tilde{\pi}_{t+h}^{imb} - \tilde{\pi}_{t+h}^{DAM}\right] = 0$, any bid yields the same expected revenue.

This highlights that the two variables of primary relevance for a producer who wants to bid strategically are the expected wind power generation and the sign of the expected price difference.

Two alternative approaches to estimate the sign of the price difference will be investigated in Section 4.3, and the term *market quantity forecast* will be used to refer to either of these approaches. In the first one, the probability distribution of the difference between imbalance and day-ahead prices is forecast. The methodology is presented in detail in Section 4.3.1. Writing the probability at time t of the price difference being positive on trading period $t + h$ as:

$$\hat{\delta}_{t+h|t} = \mathbb{P}_t\left[\pi_{t+h}^{imb} - \pi_{t+h}^{DAM} > 0\right] \tag{4.8}$$

where $\hat{\delta}_{t+h|t} \in [0, 1]$, it follows that:

- if $\hat{\delta}_{t+h|t} > 0.5$, the price difference is forecast positive, i.e. $\mathbb{E}\left[\tilde{\pi}_{t+h}^{imb} - \tilde{\pi}_{t+h}^{DAM}\right] > 0$;
- if $\hat{\delta}_{t+h|t} < 0.5$, the price difference is forecast negative, i.e. $\mathbb{E}\left[\tilde{\pi}_{t+h}^{imb} - \tilde{\pi}_{t+h}^{DAM}\right] < 0$; and

- if $\hat{\delta}_{t+h|t} = 0.5$, the price difference is forecast null, i.e. $\mathbb{E}\left[\tilde{\pi}^{imb}_{t+h} - \tilde{\pi}^{DAM}_{t+h}\right] = 0$.

In the second approach, the sign of the net imbalance volume of the system NIV_{t+h} is forecast. Hereinafter, this will also be referred to as the *imbalance sign*. The methodology is presented in detail in Section 4.3.2. As discussed in Section 2.1.3, the imbalance price is expected to be higher than the day-ahead price when the system is short, and lower when the system is long. Writing the probability at time t of the system being short on trading period $t + h$ as:

$$\hat{\gamma}_{t+h|t} = \mathbb{P}_t \left[NIV_{t+h} > 0\right] \tag{4.9}$$

where $\hat{\gamma}_{t+h|t} \in [0, 1]$, it follows that:

- if $\hat{\gamma}_{t+h|t} > 0.5$, the system is forecast short, hence $\mathbb{E}\left[\tilde{\pi}^{imb}_{t+h} - \tilde{\pi}^{DAM}_{t+h}\right] > 0$;
- if $\hat{\gamma}_{t+h|t} < 0.5$, the system is forecast long, hence $\mathbb{E}\left[\tilde{\pi}^{imb}_{t+h} - \tilde{\pi}^{DAM}_{t+h}\right] < 0$; and
- if $\hat{\gamma}_{t+h|t} = 0.5$, the system is forecast in balance, hence $\mathbb{E}\left[\tilde{\pi}^{imb}_{t+h} - \tilde{\pi}^{DAM}_{t+h}\right] = 0$.

4.1.3 Modelling Risk Attitudes

The attitudes of individuals towards risk are generally categorised in [191]:

- *risk-averse* if the individual is afraid of risk;
- *risk-taking* if he/she is seeking it; and
- *risk-neutral* if he/she is indifferent to it.

If expected values are used to make decisions, only the average outcome is considered while ignoring the range of possible values. The decision maker ignores the risk aspects associated with these possible alternatives and consequently is risk neutral, as explained by Clemen in [191]. Therefore, a producer who wishes to maximise their long-term expected revenue (i.e. the EMV in Equation (4.6)) while disregarding possible large balancing costs in the short term is considered to be risk-neutral.

Risk-averse producers, on the other hand, consider the variance of revenues or how far the revenue can decrease below the expected value. The objective sought is different from the maximisation of the expected revenue. Commonly, the objective function in this case penalises the lowest revenues, that is the left tail of the revenue distribution. The measures most commonly used to quantify risk in trading are the *Value at Risk* (VaR) and the *Conditional Value at Risk* (CVaR), also known as *expected shortfall* [40]. For a confidence level $\alpha \in [0, 1]$, the $VaR_{1-\alpha}$ is the is the $(1-\alpha)$-quantile of the empirical distribution of revenue over the period:

$$VaR_{1-\alpha}(\tilde{R}) = \max\left\{R \in S : \mathbb{P}\left[\tilde{R} < R\right] \leq 1-\alpha\right\} \tag{4.10}$$

where $\mathbb{P}[\cdot]$ indicates the probability and S is the support of the distribution. In other words, the $VaR_{1-\alpha}$ is a threshold value such that there is a $(1-\alpha)$ probability that the revenue will be lower than that threshold.

The $CVaR_{1-\alpha}$ is the expected value of the revenues that are lower than or equal to $VaR_{1-\alpha}$:

$$\begin{aligned} CVaR_{1-\alpha}(\tilde{R}) &= \mathbb{E}\left[\tilde{R} | \tilde{R} \leq VaR_{1-\alpha}(\tilde{R})\right] = \\ &= \int_0^{1-\alpha} VaR_x \mathrm{d}x \end{aligned} \tag{4.11}$$

In other words, this metric gives the expected revenue in the worst $(1-\alpha) \times 100$ percent of cases.

When evaluating risk-averse trading strategies, CVaR is generally preferred to VaR because it accounts for heavy tails in the probability distribution and is a "coherent" risk measure, as demonstrated by Artzner et al. in [190]. A possible objective function for a risk-averse producer is the maximisation of the $\text{CVaR}_{1-\alpha}$. Note that if $\alpha = 0$, this coincides with the risk-neutral case, while increasing values of α correspond to higher aversion to risk.

4.2 Wind Power Forecasts

The wind farm under study is located in Ireland on complex terrain. Day-ahead forecasts of wind power generation were provided by an anonymous wind farm operator. Four distinct sets of predictions are available: three consist of point forecasts only, and one of point and 80% interval forecasts (i.e. the 10^{th} and 90^{th} quantiles). All forecasts are at the wind farm level with 30-minute resolution, and were issued at 09:00 on the day before realisation. Owing to confidentiality restrictions, further information regarding the models cannot be discussed. The accuracy of the point and interval forecasts will be evaluated in Section 5.1.

4.3 Market Quantity Forecasting

As discussed in Section 4.1.2, the financial obligations of market participants depend on prices in both the day-ahead and balancing markets (cf. Equation (4.3)). In fact, it is primarily the sign of the price difference that is relevant to the producer in order to develop any advanced bidding strategy (cf. Equation (4.7)). In this Section, the two alternative approaches developed to forecast the sign of the price difference are explained. The first approach, where the probability distribution of price difference is derived from the individual forecasts of DAM and BM prices, is presented in Section 4.3.1. The second approach, where the sign of the system imbalance is directly forecast, is presented in Section 4.3.2.

The predictions generated with the methodologies explained in this Section will serve as inputs for the bidding strategies presented in Section 4.4, where risk-constrained trading of wind energy in the Irish market is investigated. In the forecasting procedures developed next, the practical constraints of participating in the day-ahead market (such as data availability and timing of bid submissions) are considered. Recalling that in the I-SEM the day-ahead gate closure for trading day D is at 11:00 on the previous day D–1, all models were fit in order to issue the forecasts at 10:00 on the day before realisation. The information available for

the estimation process includes historical data up to 10:00 of D–1 and day-ahead predictions of relevant variables for trading day D.

The data available for the study cover approximately seven months of operation of the Irish electricity market, from the 1st October 2018 to the 25th April 2019, this period being the maximum available at the time. The data set includes:

- historical DAM prices and volumes;
- historical BM prices and volumes;
- observed and forecast demand in Ireland; and
- observed and forecast wind power generation in Ireland.

All the above time series have half-hourly resolution except the DAM prices, whose values are hourly. To take advantage of the higher resolution of the other information and because market settlement occurs on a half-hourly basis, the DAM price series is discretized to this resolution, with the same value repeated for the two 30-minute periods in each hour. The demand and wind power forecasts are predictions issued at 09:00 for the next day and were obtained from an anonymous industry partner. They consist of point forecasts for the demand, and point forecast and 10th and 90th quantiles for the wind power generation.

Data were split into training and test sets, with model estimation performed on the former and out-of-sample evaluation on the latter. The initial training set covers a four-month period from the 1st October 2018 to the 28th January 2019 included (120 days, 58% of the total sample). The independent test set is almost three months long, ranging from the 29th January to the 25th April 2019 (87 days, 42% of the total sample). The partitions into training and test sets for the electricity price and net imbalance volume data are shown in Figures 4.3 and 4.4, respectively. We acknowledge that this partition is somewhat arbitrary and does not follow the more commonly employed 80/20 ratio. However, no strict guidelines exist in this regard and the minimum length of training and test sets can vary depending on

the total sample size and the forecast horizon of interest [42]. In the report on recommended practices for the evaluation of forecasts [201] by the IEA Task 36, the authors suggest a sample size of at least 90 days for an adequate evaluation. Given this recommendation and the total size of the data available for this study, the ratio discussed above was selected.

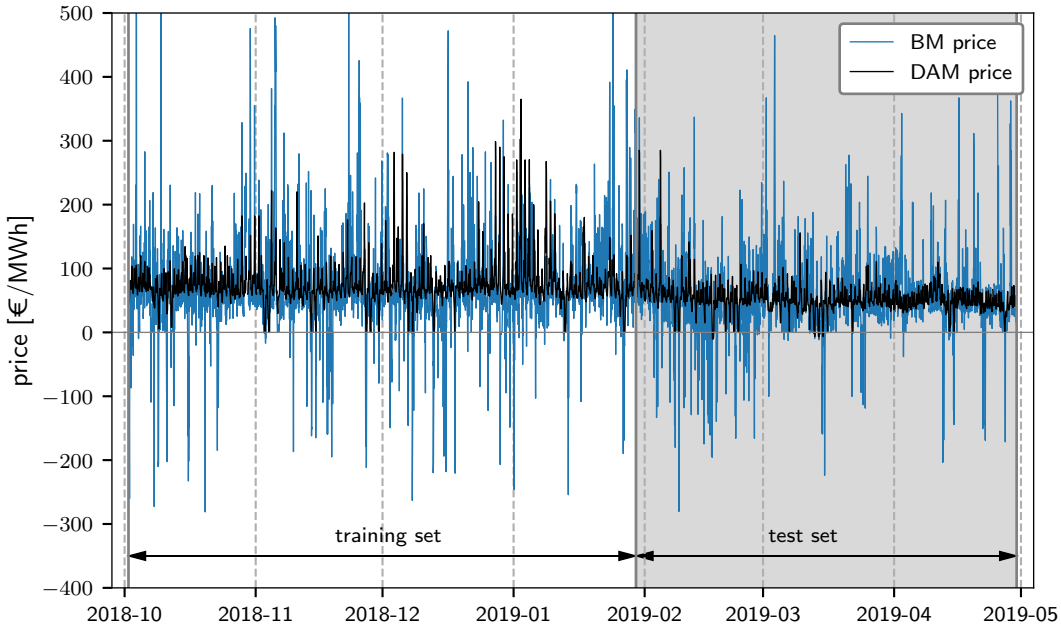


Figure 4.3: Training and test sets used for electricity price forecasting.

4.3.1 Electricity Prices Forecasting

The purpose of this Section is to explain the methodology developed to forecast the sign of price difference, where day-ahead and balancing market prices are forecast individually. First, the theoretical formulation of ARIMAX models is presented in Section 4.3.1.1. Then, in Section 4.3.1.2, the implementation of these models to forecast day-ahead and balancing market prices is described. Finally, the methodology used to combine the resulting quantile forecasts using the Monte Carlo method to obtain a probabilistic forecast of the price difference is explained in Section 4.3.1.3.

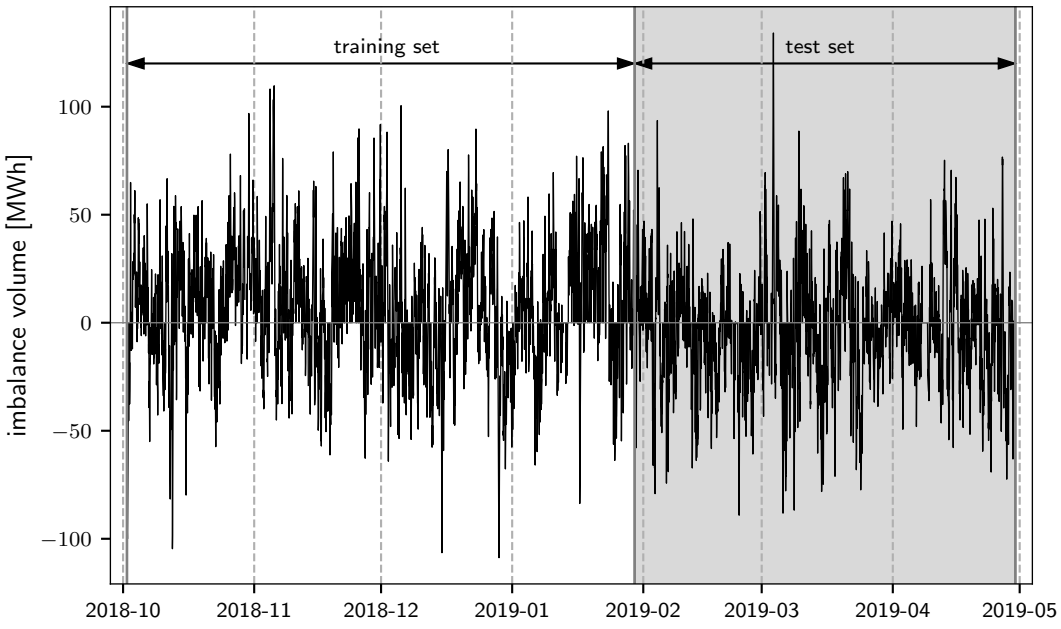


Figure 4.4: Training and test sets used for imbalance sign forecasting.

4.3.1.1 ARIMAX models

Auto-Regressive Moving Average (ARMA) models are, together with exponential smoothing, the most popular approaches to time series forecasting [42]. Moreover, as highlighted in the literature review in Section 3.3, models based on ARMA processes are very well established in electricity price forecasting. As the name suggests, the distinguishing feature of AR-type models is their ability to describe the autocorrelation in the data.

ARMA models are usually restricted to stationary data. Here, *stationarity* means that the properties of the time series do not change over time. Then, if y_t is a stationary time series, the distribution of $\{y_t, \dots, y_{t+h}\}$ does not depend on t for all h. Such a time series will show a stable mean and constant variance. It follows that time series with seasonality or trends are non-stationary. In these cases, *differencing* — i.e. computing the difference between consecutive observations — can remove the changes in the level of the variable values and stabilise the mean, thus making the time series stationary. The differenced series can be written as:

$$y'_t = y_t - y_{t-1}.$$

Sometimes, it can be useful to perform seasonal differencing, whereby the difference is calculated between an observation and the previous observation from the same season:

$$y'_t = y_t - y_{t-S}$$

where S is the number of seasons, i.e. the number of time periods until the seasonal pattern repeats again. For example, this value is equal to 12 for a monthly series with yearly periodicity, or 7 for a daily series with weekly periodicity. To stabilise the variance of the series, mathematical *transformations* can be helpful, with common choices being logarithmic and Box-Cox transformations [202].

The full Auto-Regressive Integrated Moving Average (ARIMA) model (where integration is the reverse of differencing in this context) can be written as:

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - B)^d y_t = c + (1 + \theta_1 B + \cdots + \theta_q B^q) e_t \tag{4.12}$$

and indicated as $ARIMA(p, d, q)$, where p is the order of the autoregressive part, d the number of first differencing, and q the order of the moving average part. In the Equation:

- B is the *backshift operator*, defined so that $By_t = y_{t-1}$;
- $(1-\phi_1 B-\cdots-\phi_p B^p)$ is the autoregressive part $AR(p)$, consisting of a linear combination of past values of the variable defined by parameters $\phi_1, \ldots, \phi_p$ [2];
- $(1+\theta_1 B+\cdots+\theta_q B^q)e_t$ is the moving average part $MA(q)$, using a combination of past forecast errors e_t defined by parameters $\theta_1, \ldots, \theta_q$; and
- $c = \mu(1 - \phi_1 - \cdots - \phi_p)$, where μ is the mean of the differenced series $y'_t = y_t - y_{t-1}$.

It is also possible to model seasonal data through seasonal differencing with

[2] In fact, an $AR(p)$ model can be seen as a multiple regression where lagged values of y_t are the predictors.

the so-called *Seasonal ARIMA* (or SARIMA) models, which are defined by simply adding seasonal terms to an ARIMA model. The general notation in this case is $ARIMA(p,d,q) \times (P,D,Q)_S$, where:

- the term (p,d,q) represents the order of the non-seasonal part;
- the term $(P,D,Q)_S$ represents the order of the seasonal part; and
- S is the number of seasons.

Therefore, the model equation will have the same form as Equation (4.12) but with each term multiplied by its seasonal counterpart, namely:

- seasonal $AR(P)$: $1 - \Phi_1 B^S - \cdots - \Phi_P B^{PS}$;
- seasonal $MA(Q)$: $1 - \Theta_1 B^S + \cdots + \Theta_Q B^{QS}$;
- seasonal differencing: $(1 - B^S)^D$.

Model selection and estimation

In this work, the model order for each time series will be selected using the `forecast` package [203] in `R` with a procedure based on the Hyndman-Khandakar algorithm described in [203]. To determine the value of d (that is, whether differencing is needed or not) objectively, the *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS) *unit root test* [204] is used. In this test, evidence is sought to reject the null hypothesis of the data being stationary. The values of p and q are then selected by minimising a chosen Information Criterion. The most widely used is the *Akaike's Information Criterion* (AIC) [205], which can be written as:

$$AIC = -2\log(L) + 2(p + q + k + 1) \tag{4.13}$$

where $\log(L)$ is the *log likelihood* of the data — i.e. the logarithm of the probability that the observed data come from the estimated model — and k is 0 if $c = 0$ (cf. Equation (4.12)) and 1 otherwise. The general idea is that the fit of the model (L)

is penalised with the number of parameters that need to be estimated to prevent over-fitting. However, the AIC can be misleading if the sample size is small and the use of the bias-corrected version of the AIC is recommended (see for example the book [41] by Walpole et al.). This is defined as:

$$AICc = AIC + \frac{2(k+2)(k+3)}{(n-k-3)} \tag{4.14}$$

where n is the number of observations used for estimation. In this work, given the limited length of data available, the model order is selected by minimising the AICc. The Hyndman-Khandakar algorithm then considers different possible combinations of p and q by varying their value in a stepwise search until the AICc can no longer be reduced. Once the model order is identified, the model parameters are estimated via the *maximum likelihood estimation* (MLE) technique. In practice, `R` finds the parameters that maximise the log likelihood function, given the selected values of p, d and q.

Forecasting

To calculate an h-step ahead point forecast from an ARIMA model, the model equation is expanded, t is replaced with $t+h$, future observations are replaced with their forecasts, past errors with the observed residuals, and future errors with zero. For example, let us consider an $ARIMA(1,1,1)$ model with no constant ($c=0$):

$$(1-\hat{\phi}_1 B)(1-B)y_t = (1+\hat{\theta}_1 B)e_t$$

where $\hat{\phi}_1$ and $\hat{\theta}_1$ are the estimated parameters. The one-step ahead forecast then is:

$$\hat{y}_{t+1|t} = (1+\hat{\phi}_1)y_t - \hat{\phi}_1 y_{t-1} + \hat{\theta}_1 \varepsilon_t$$

where ε_t is the *residual* at time t, i.e. the difference between the observed value and the corresponding fitted value. The two-step forecast is:

$$\hat{y}_{t+2|t} = (1 + \hat{\phi}_1)\hat{y}_{t+1|t} - \hat{\phi}_1 y_t$$

and so on for any forecast horizon.

Under the assumption of uncorrelated and normally distributed residuals, one-step prediction intervals are simply calculated as:

$$\hat{y}_{t+1|t} \pm c\hat{\sigma}$$

where $\hat{\sigma}$ is the standard deviation of the residuals, and the coefficient c depends on the coverage percentage of the interval (see Table 4.1 for some indicative values). The calculation of multi-step prediction intervals is complicated and goes beyond the scope of this work, and we refer to the book [206] from Brockwell and Davis for the mathematical details. If either of the assumptions on the residuals does not hold, then bootstrapping can be used to calculate the intervals [207].

Percentage	**Coefficients**
99	2.58
95	1.96
90	1.64
80	1.26
70	1.04
50	0.67

Table 4.1: Coverage percentages of prediction intervals and associated coefficients c for one-step forecasts.

The ARIMA models discussed above are time series models which allow the inclusion of information from past values of the series. Regression models, on the other hand, allow the inclusion of information from other relevant variables which may affect the system that has to be predicted. By combining the features of these two approaches, ARIMAX (or dynamic regression) models extend the purely

time series models with information from *exogenous* variables. To facilitate their mathematical formulation, a brief overview of regression models is presented first.

In a multiple linear regression model, the *forecast variable* y (also referred to as regressand or dependent variable) is assumed to have a linear relationship with a number of *predictor variables* $x_1, \ldots, x_m$ (also referred to as regressors, independent, exogenous or explanatory variables):

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_m x_{m,t} + e_t \tag{4.15}$$

where $x_1, \ldots, x_m$ are the m predictor variables, β_0 is the intercept, $\beta_1, \ldots, \beta_m$ are the coefficients (obtained via least square estimation) measuring the marginal effects of the predictors, and e_t is the error. In these models, errors are assumed to be *white noise*, that is: have zero mean, not be autocorrelated and not have any correlation with the predictor variables.

In ARIMAX models, errors from the regression are allowed to be autocorrelated. The error term in Equation (4.15) is assumed to follow an $ARIMA(p, d, q)$ process and is denoted as η_t instead of e_t, so that the final model is given by:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_m x_{m,t} + \eta_t \tag{4.16}$$

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - B)^d \eta_t = c + (1 + \theta_1 B + \cdots + \theta_q B^q) e_t \tag{4.17}$$

where only the error term e_t from the ARIMA model is assumed to be white noise.

As discussed earlier, the model order for the ARIMA part of the model is selected by minimising the AICc. This principle is also followed for the selection of the exogenous variables. Model parameters are then estimated using the MLE technique.

Recalling that forecasts from a regression model are obtained as:

$$\hat{y}_{t+h|t} = \hat{\beta}_0 + \hat{\beta}_1 x_{1,t+h} + \cdots + \hat{\beta}_m x_{m,t+h} \tag{4.18}$$

future values of the regressors need to be estimated. In order to forecast using an ARIMAX model, it is necessary to predict both the ARIMA and the regression parts of the model and combine the outcomes. Therefore, predictions of the exogenous variables are needed to forecast the regression part. It should be noted that the uncertainty associated with the forecasts of the predictors is not accounted for in the final prediction intervals; in other words, the forecasts of the predictors are treated as deterministic and consequently, prediction intervals should be considered as being conditional on those estimated values.

4.3.1.2 Forecasting DAM and BM prices

The analyses of historic market data presented in Section 2.2.2 have highlighted a number of aspects which should be considered when modelling the price processes with the purpose of forecasting their future values. Relevant features include: the impact of exogenous variables, seasonality and patterns at the weekly and daily timescales, and the volatility and heteroskedasticity of the time series. In this Section, the implementation of the ARIMAX models is presented while explaining how these features are accounted for in the construction of the models.

Two separate sets of models are developed to forecast the day-ahead and balancing market price time series, although the same approach and available information are used. The small number (nine in total) of extreme values observed in the imbalance price time series (see Figure 2.22) would have a large undesired influence on the estimated parameters of the models. Therefore, such outliers are capped at the strike price (i.e. 500 €/MWh) as this is the price effectively used in the market to pay and charge participants (see Section 2.2.2.2). The imbalance prices are forecast directly by fitting a single set of models to the time series, thus defining the balancing market state implicitly. The alternative approach would be to fit two sets of models, one for up-regulation and one for down-regulation prices, and selecting which one to use based on the expected regulating state (see the discussion on this in Section 3.3.2). However, here the aim is to develop an approach to forecasting

the sign of price difference which is different to that developed in Section 4.3.2, where the forecast sign of price difference is conditional on — and in fact, directly determined by — the estimated imbalance sign.

The variables included as predictors in the regression part of the ARIMAX models (that is, the exogenous variables) are:

- forecast demand;
- point and 80% interval forecast of the wind power generation, where the interval is calculated as the difference between the 90^{th} and 10^{th} quantiles (cf. Equation (3.6)); and
- predicted wind power penetration, calculated as the ratio between the (point) forecasts of demand and wind power generation (see Section 2.2.2.4 for the study on the impact of this variable on various market quantities).

In the analyses in Section 2.2.2, electricity prices show rather distinct profiles and features depending on the day of the week (see for example Figure 2.7 for DAM prices and Figure 2.26 for BM prices). Therefore, prices are modelled separately based on the day type, distinguishing between:

- *weekday*: from Monday to Friday; and
- *weekend*: Saturday and Sunday.

Furthermore, the price processes display characteristic daily patterns in response to the hourly variations of demand, spot prices, wind penetration, and behaviour of market participants. In particular, day-ahead prices exhibit a clear daily seasonality (see Figures 2.9 and 2.11) correlated with the changes in demand at the same timescale (see Figure 2.16). Despite a considerably higher volatility, an analogous pattern is followed on average by BM prices (see Figure 2.28). These prices tend to be negative more often at night (see Figure 2.35) and higher in the evening (see Figure 2.36). The influence of predicted wind power penetration

also varies with the time of the day (see the results of the empirical analyses in Section 2.2.2.4).

In light of these features, the modelling is implemented separately across the imbalance settlement periods. That is, instead of calibrating a single ARIMAX process for all hours, the time series for each half-hour are modelled individually, leading to 48 ARIMAX models per day type. This *multi-model approach* generally yields better performances for day-ahead predictions, as highlighted in the review of the literature in Section 3.3.1.1, and has multiple advantages which are described below.

- The separate time series have larger homogeneity compared to the complete one, resulting in simpler models (and generally more accurate forecasts).
- The daily seasonality is removed, since each element of the series now refers to the same hour of a different day; this way, the difficulty of dealing with multiple seasonal patterns is avoided.
- The impact of exogenous variables on the forecast variable can be modelled differently depending on the hour of the day.
- Instead of estimating next-day values with a single model with forecast horizons between 12 and 36 hours, predictions are computed as one or two step-ahead forecasts from the 48 individual models. This generally allows an improvement of the forecast accuracy. Recall that forecasts for trading day D are estimated using information available up to 10:00 on D–1. Hence, prices for settlement periods before 10:00 are calculated as one-step forecasts from the corresponding model. On the other hand, two-step forecasts are used for prices at 10:00 and later because the corresponding prices on D–1 are not yet available, the most recent information being from D–2. A visual explanation of this concept is given in Figure 4.5.

The drawbacks of the multi-model approach are due to the considerably higher number of models that need to be identified and the associated large number of

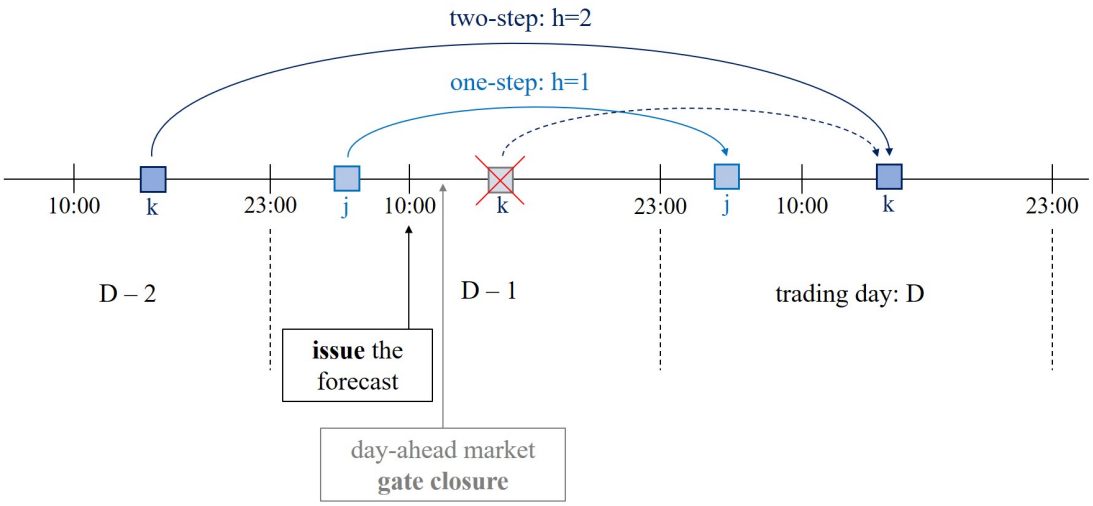


Figure 4.5: Use of one and two step-ahead forecasts depending on the hour of day.

parameters estimated. Indeed, for each day type, 48 models are selected, each requiring a set of parameters to be calculated. The main disadvantages are that:

- the modelling task can be computationally intensive; and
- it is impractical to investigate in depth the performance of each model separately.

To recap, the overall number of price sequences to be modelled is equal to 192: one for each of the 48 (half-hourly) settlement periods in either day-type group (weekday and weekend) for each market (day-ahead and balancing). This is depicted schematically in the tree chart in Figure 4.6.

Another relevant feature of the market data highlighted in Section 2.2.2 is the high volatility of the price series (see the time plots in Figures 2.4 and 2.22) with evidence of heteroskedasticity (see the variations of the moving standard deviation in Figures 2.6 and 2.38). In an attempt to accomodate the time-varying nature of these processes when estimating model parameters, two alternative dynamic approaches are implemented to train the forecasts:

- *recursively expanding* (also called "time series cross-validation" by Hyndman and Athanasopoulos in [42]), and

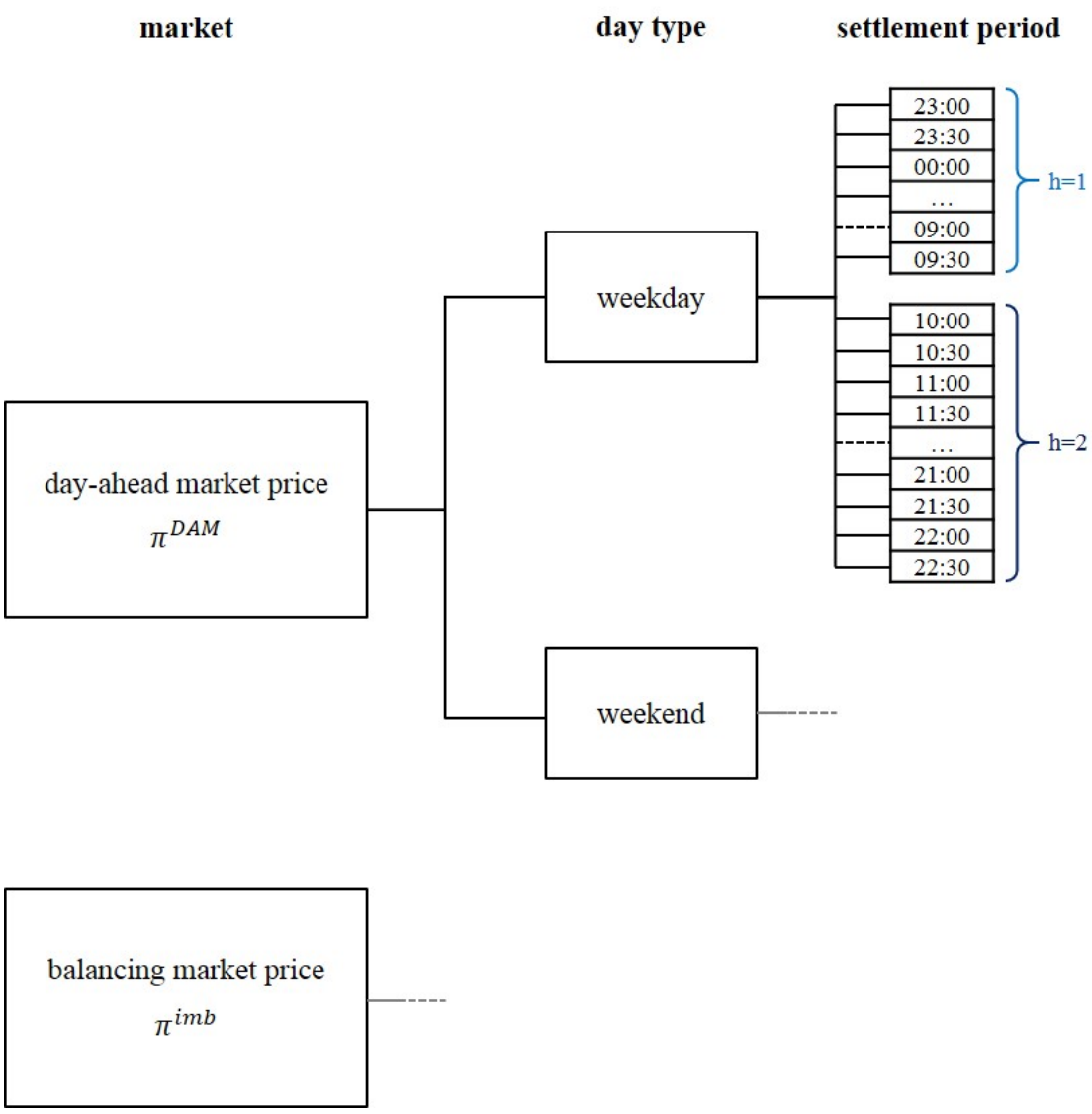


Figure 4.6: Overview of the multi-model approach adopted for electricity price forecasting, showing the segmentation of models based on market, day type, and imbalance settlement period (half-hour).

- *rolling* (also referred to as "moving window").

In the first, the length of the training set increases with each iteration, as illustrated schematically in Figure 4.7, and the ARIMAX model is re-selected and re-estimated at each step before the forecast is computed. This approach allows the use of more observations as they become available, a desirable feature given the limited size of the data set. In the second approach, the length of the training set is fixed and as new observations become available, they replace older ones, as illustrated in Figure 4.8. Again, the ARIMAX model is re-selected and re-estimated before

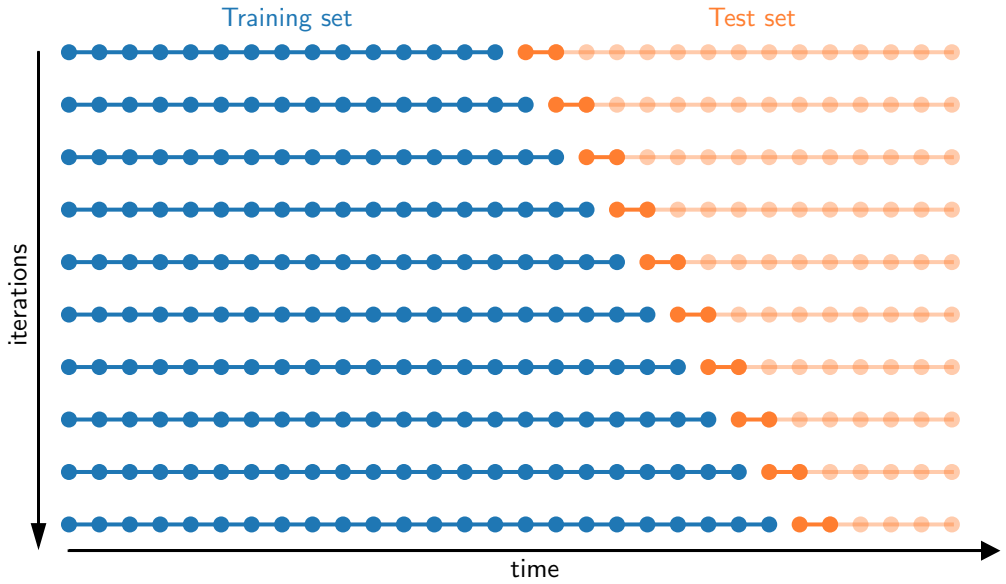


Figure 4.7: *Recursively expanding* approach: training and test sets at each iteration, with one- and two-step forecasts highlighted in dark orange.

computing the forecast at each step. This way, for a given window size N, only the N-latest observations are used, preventing previous observations from influencing the model selection and estimation. In this study, the window length is set equal to 120 days, resulting in a test set of 87 days.

The output from each model consists of half-hourly one and two step-ahead point forecasts and quantiles of the forecast distribution with 5, 10, 25, 75, 90 and 95% nominal levels. From these quantiles, having assumed normally distributed forecast errors, it is possible to calculate the central prediction intervals with 50, 80, and 90% nominal coverage rates (cf. Equation (3.6)). An example of the point and interval forecasts generated by a model is shown in Figure 4.9.

The source code developed to generate the forecasts of day-ahead and balancing market prices can be found in Appendix C.1.

The various forecasts are then evaluated. First, the point forecasts obtained from the two approaches for both DAM and BM prices are assessed against a variation of the benchmark method proposed by Conejo et al. in [150]. Even though balancing market prices are generally less seasonal then day-ahead ones,

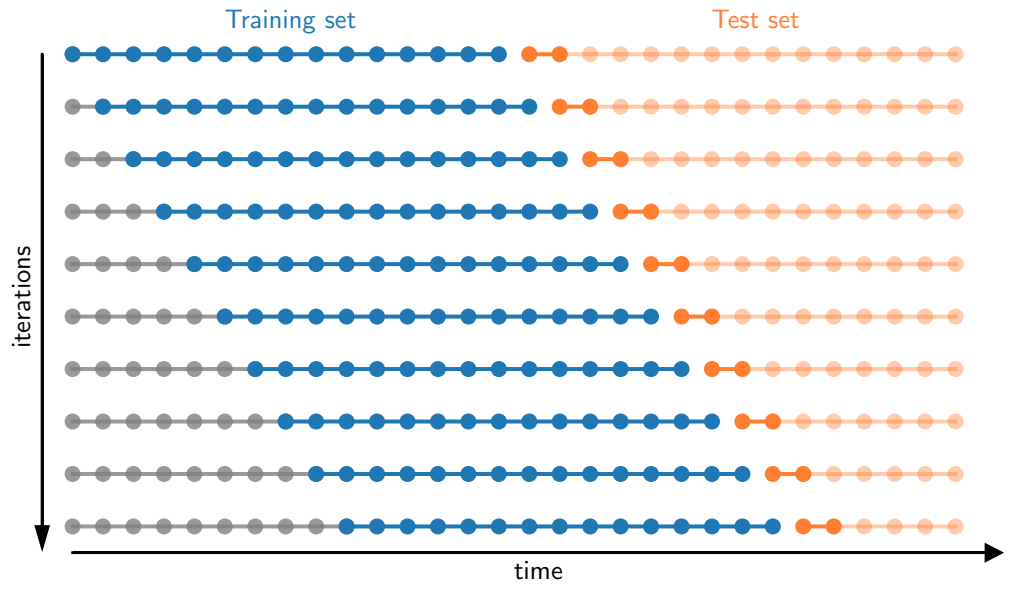


Figure 4.8: *Rolling* approach: training and test sets at each iteration, with one- and two-step forecasts highlighted in dark orange.

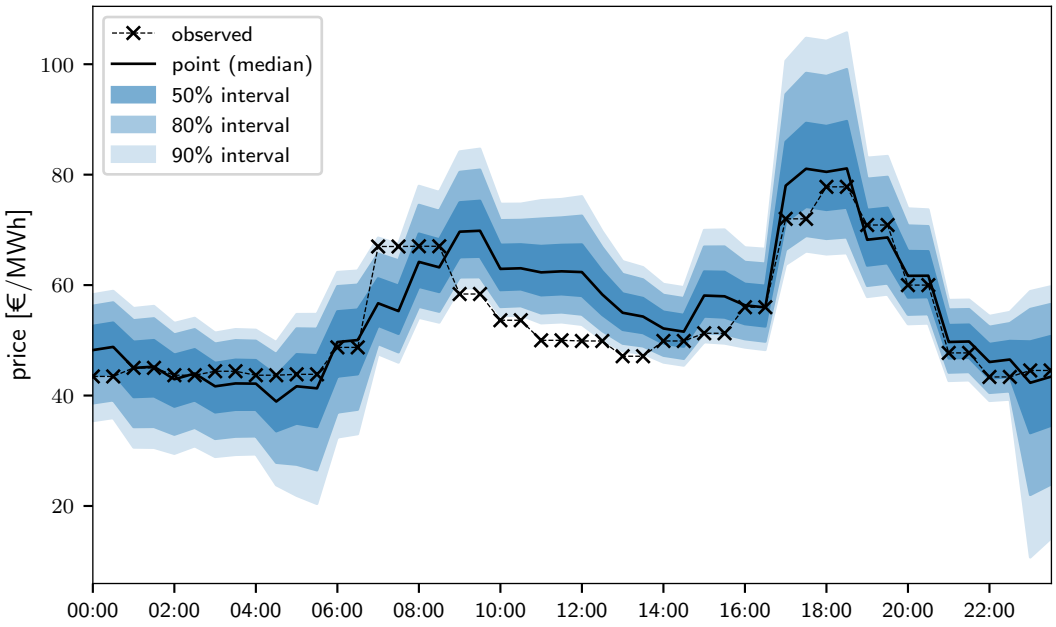


Figure 4.9: Example of point and interval forecasts and observed prices for a sample day.

this method is commonly used as a benchmark for next-day forecasts of both types of electricity price (see for example References [34] and [178]).

Recall that in the original version of the "similar-day" method already outlined

in Section 3.3.1.1:

- forecasts for Monday are equal to the values observed on Monday of the week before, and the same applies for Saturday and Sunday;
- forecasts for Tuesday are equal to the values observed on the previous day, and the same applies for Wednesday, Thursday and Friday.

However, in the context of our work, values for Tuesday are forecast at 10:00 on Monday and not all the necessary data are available for the previous day (Monday) so data from the previous weekday (i.e. Friday) are used as required. Similarly for Wednesday forecasts, values for periods after 10:00 are taken from the previous Monday, and the same applies for Thursdays and Fridays. Denoting the price at time T on day D as $\pi_{T,D}$, the method can be expressed in mathematical terms as:

$$\hat{\pi}_{T,D} = \begin{cases} \pi_{T,D-1} & \text{if} \quad T < 10{:}00, D \in \{\text{Tuesday, Wednesday, Thursday, Friday}\} \\ \pi_{T,D-4} & \text{if} \quad T \geq 10{:}00, D = \text{Tuesday} \\ \pi_{T,D-2} & \text{if} \quad T \geq 10{:}00, D \in \{\text{Wednesday, Thursday, Friday}\} \\ \pi_{T,D-7} & \text{if} \quad \forall T, D \in \{\text{Monday, Saturday, Sunday}\} \end{cases} \tag{4.19}$$

This is also represented visually in the chart in Figure 4.10, where the darker blocks represent the group of settlement periods before 10:00. The source code developed to implement the benchmark model can be found in Appendix C.1.

Once it is verified that both forecasting approaches outperform the benchmark, DM tests [53] are performed to determine which approach is more accurate for each market price.

Quantile forecasts are then examined. For each half hour, climatology is used as the benchmark, as suggested by Winkler in [208]; that is, the historical quantiles of the price distributions calculated on the training data set. The quantile forecasts from the two approaches are evaluated against the benchmark in terms of

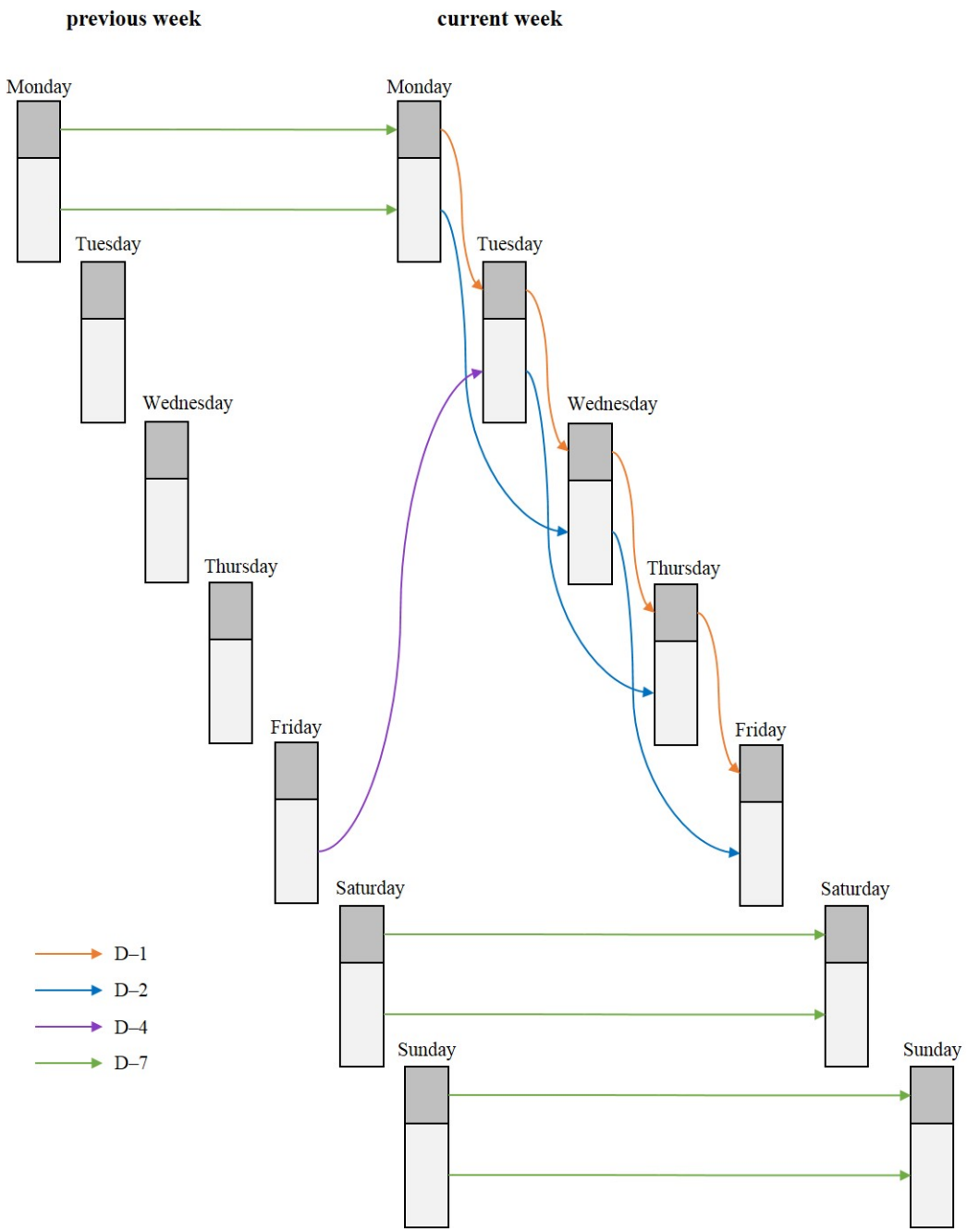


Figure 4.10: Variation of the "similar-day" benchmark method for short-term electricity price forecasting proposed by Conejo et al. in [150]. For each day, the darker block represents the group of periods before 10:00.

reliability and sharpness, and the pinball loss score is calculated to determine the best approach.

4.3.1.3 Monte Carlo simulations

In a second stage, Monte Carlo simulations are executed to estimate the probability of the price difference being positive from the quantile forecasts of DAM and BM prices. The use of quantile rather than point predictions as input to the simulations allows for the estimation of the sign of price differences in a probabilistic framework. Monte Carlo methods are computational algorithms that use repeated random sampling to obtain numerical results [209]. The accuracy of the results is inversely proportional to the number of samples. Indeed, these methods rely heavily on the "law of large numbers" and the "central limit theorem", by which the distribution of large samples converges towards the underlying population's distribution [210].

For any trading period $t + h$, the proposed method consists of the steps below.

1. Two quadratic spline functions are interpolated to the forecast quantiles of DAM and BM prices, respectively, to represent their probability pseudo-distribution (a similar procedure, for example, is employed by Zugno et al. in [193]).

2. Two random numbers are generated to assign values to the DAM and BM prices through their interpolated distributions.

3. The difference between the two prices is calculated and stored.

4. Steps 2 and 3 are repeated 10,000 times and the sequence of results generated is transformed into a frequency distribution, from which the frequency of positive price difference can be calculated.

Through this algorithm, an estimate of the probability of positive price difference (cf. Equation (4.8)) is obtained for each trading period $t + h$, that is:

$$\hat{\delta}_{t+h|t} = \mathbb{E}\left[\delta_{t+h|t}\right] = \mathbb{E}\left[\mathbb{P}_t\left[\pi^{imb}_{t+h} - \pi^{DAM}_{t+h} > 0\right]\right] \tag{4.20}$$

The simulations are run using the `random` module [211] and the `pandas` library [212] in `python`. The source code developed to perform this task can be

found in Appendix C.1.

The outcomes and the evaluation of the price forecasting methods developed in this Section will be presented in Section 5.2.

4.3.2 Forecasting the Imbalance Sign

In this Section, the methodology used to forecast the system imbalance sign is explained. Recall that a positive imbalance ($NIV > 0$) corresponds to a need for up-regulation, meaning that the system is short. Consequently, balancing price is expected to be higher than the day-ahead price, yielding a positive price difference. Vice versa, a negative imbalance ($NIV < 0$) means that the system is long and down-regulation is needed, yielding a negative price difference. These conditions were summarised earlier in Table 2.1, which for convenience is re-proposed here in Table 4.2. Therefore, the forecast probability of the system being short will be used as a proxy for the probability that the price difference will be positive.

system	**imb. volume**	**regulation**	**price difference**
balanced	zero	no	$\pi^{imb} = \pi^{DAM} \rightarrow \pi^{diff} = 0$
short	positive	up	$\pi^{imb} > \pi^{DAM} \rightarrow \pi^{diff} > 0$
long	negative	down	$\pi^{imb} < \pi^{DAM} \rightarrow \pi^{diff} < 0$

Table 4.2: Balancing Market states and corresponding market quantities.

Logistic regression models

Let us assign the values 0 and 1 to the (half-hourly) trading periods where the system is long and short, respectively, so that the imbalance sign time series follows a binomial distribution. The process can then be modelled with a *logistic regression model*, a case of the Generalized Linear Model which fits responses that follow a binomial distribution [213]. The objective is to estimate the probability γ_{t+h} that the imbalance will be positive (i.e. the system will be short) at time $t + h$ (cf. Equation (4.9)) conditional on a set of m predictors (or explanatory variables)

$x_{1,t+h}, \ldots, x_{m,t+h}$ which are related to the responses, that is:

$$\gamma_{t+h} = \mathbb{P}\left[NIV_{t+h} > 0 | x_{i,t+h}, i = 1, \ldots, m\right] \tag{4.21}$$

The model to be fitted is a multiple linear regression of the binomial probability on the m predictors:

$$\gamma_{t+h} = \beta_0 + \beta_1 x_{1,t+h} + \cdots + \beta_m x_{m,t+h}$$

To relate the predictors to the mean of the responses, the *logit* link function is used so the logistic model is given by:

$$\log\left(\frac{\gamma_{t+h}}{1+\gamma_{t+h}}\right) = \vec{\beta} \times \vec{x}_{t+h} \tag{4.22}$$

where $\vec{\beta}$ is the vector of parameters β_i to be estimated and $\vec{x}_{t+h}$ is the vector of predictors $x_{i,t+h}$.

The predictors can be defined a priori or be selected with a *backward selection* procedure where the model with the smaller AIC is chosen in a stepwise search, as described by Venables and Ripley in [214]. The backward selection algorithm proceeds as follows:

1. The AIC is calculated for the model including all possible predictors (full model).

2. Each predictor is considered for removal from the model one at the time.

3. The predictor removed from the model is the one whose removal results in the largest decrease in AIC. Note that once a predictor is removed it is not considered for re-entry.

4. Steps 2 and 3 are repeated until the removal of any of the remaining predictors fails to reduce the AIC.

Once the model is selected, the parameters $\vec{\beta}$ are estimated with the method of

MLE. The probabilities can then be forecast using the parameter estimates $\hat{\vec{\beta}}$ and predictions of the explanatory variables $\vec{x}_{t+h}$ by solving Equation (4.22) for $\hat{\gamma}_{t+h}$ (inverse logit function):

$$\hat{\gamma}_{t+h} = \frac{\exp\left(\hat{\vec{\beta}} \times \vec{x}_{t+h}\right)}{1 + \exp\left(\hat{\vec{\beta}} \times \vec{x}_{t+h}\right)} \tag{4.23}$$

From the analyses in Section 2.2.2.2, imbalance volumes show a distinct profile depending on the hour of the day, with a marked tendency to have positive imbalances in the evening (see the boxplots in Figure 2.20). A multi-model approach is adopted analogous to that implemented when forecasting prices in Section 4.3.1: a separate logistic regression model is fit for each (half-hourly) settlement period. Again, one-step predictions will be used for imbalance signs before 10:00 and two-step predictions for imbalance signs after 10:00 (cf. Figure 4.5).

Predictors are chosen from the following range of variables:

- forecast demand;
- point and 80% interval forecasts of the wind power generation;
- predicted wind power penetration; and
- day of the week.

Note that the day of the week is included as a categorical predictor to capture any possible weekly seasonality. Models are then selected with two alternative strategies:

1. the full model (i.e. including all explanatory variables) is fit for all settlement periods;
2. the model with the smaller AIC is chosen via backward selection.

The first strategy is straightforward but more exposed to the risk of overfitting, a risk avoided with the second strategy, which is computationally more intensive.

Another relevant characteristic of the imbalance volumes time series is its non-stationarity (see for example Figures 2.21 and 2.39). Similarly to the price forecasting case, two alternative approaches are implemented to train the logistic regression models in order to capture these non-stationary effects:

- recursively expanding (see Figure 4.7); and
- rolling (see Figure 4.8).

In both cases, models are re-estimated at each iteration. This results in the development of four forecasting methods:

- recursively expanding with model re-estimation;
- recursively expanding with stepwise model re-selection and re-estimation;
- rolling with model re-estimation; and
- rolling with stepwise model re-selection and re-estimation.

In this work, models are selected and estimated using the `glm` and `step` functions from the package `stats` [215] in `R`. The source code developed to generate the forecasts of imbalance sign can be found in Appendix C.2.

Firstly, each of these models is tested against a benchmark and then their accuracies are compared. For each settlement period, the historic proportion of the system being short (calculated on the training set) constitutes the benchmark forecast. This corresponds to what Brier calls the "climatological probabilities" in [66] and is employed as a benchmark, for example, by Browell in [199] (where is referred to as "empirical proportions") and by Jónsson et al. in [186]. The performance of the forecasts is evaluated by examining their ROC curves and calculating the Brier scores, as explained in Section 3.1.2. The outcomes and the evaluation of the imbalance sign forecasting methods developed in this Section will be presented in Section 5.3.

4.4 Bidding Strategies

In this Section, two benchmark and three novel bidding strategies for a market participant trading wind energy in the day-ahead market are presented. First, the two limit cases of seeking to minimise and maximise energy imbalances are presented. Then, the three risk-constrained strategies specifically developed in this work are described. The strategies that require a prediction of the price difference sign, $\mathbb{E}\left[\tilde{\pi}_{t+h}^{imb} - \tilde{\pi}_{t+h}^{DAM}\right]$, can be implemented using either of the two probabilistic market quantity forecasts developed in Section 4.3, namely the probability of the price difference being positive $\hat{\delta}_{t+h|t}$ and the probability of the system imbalance being positive $\hat{\gamma}_{t+h|t}$.

4.4.1 Imbalance Minimisation

The *Imbalance Minimisation* (ImbMin) strategy aims to minimise the exposure to balancing costs regardless of their sign. The participant concedes the opportunity of making extra profits when their energy imbalance has the opposite sign of that of the system; on the other hand, losses are limited when the two imbalances go in the same direction. The objective is to minimise the energy imbalances in absolute terms; in other words, to have imbalances as close to zero as possible. Since the loss function is symmetric piecewise linear, the optimal bid is the median of the predictive distribution of wind power generation (as demonstrated by Gneiting in [216]):

$$E_{t+h}^{DAM} = \hat{E}_{t+h|t} \tag{4.24}$$

Note that this strategy does not require any market quantity forecast but only the forecast of wind power generation. In fact, a point forecast is sufficient for this purpose as long as the model produces the median, rather than the mean, of the distribution. As mentioned in Section 3.4, this strategy constitutes the current industry practice among most utilities in Europe and the USA. In the case study in Chapter 5, it will therefore establish a benchmark to examine possible

improvements of the advanced strategies.

4.4.2 Imbalance Maximisation

The *Imbalance Maximisation* (ImbMax) strategy aims to maximise the exposure to balancing costs to take advantage of the arbitrage opportunities arising when the producer's energy imbalance goes in the opposite direction of the system imbalance. Given that the producer cannot bid above its installed capacity, the optimal bid is:

$$E_{t+h}^{DAM} = \begin{cases} 0 & \text{if} \quad \mathbb{E}\left[\tilde{\pi}_{t+h}^{imb} - \tilde{\pi}_{t+h}^{DAM}\right] > 0 \\ E_{max} & \text{if} \quad \mathbb{E}\left[\tilde{\pi}_{t+h}^{imb} - \tilde{\pi}_{t+h}^{DAM}\right] < 0 \end{cases} \tag{4.25}$$

Note that this strategy requires a market quantity forecast to estimate the sign of the price difference but no forecast of the wind power generation.

As previously discussed in Section 3.4, this strategy — which in the literature is sometimes referred to as the "Zero/Max" strategy — can only be implemented if the size of the producer's portfolio is small compared to the average system imbalance volume, otherwise the price-taker assumption is violated. Moreover, the producer is exposed to potentially large losses whenever the market quantity forecast is incorrect. Finally, the system operator might not accept the frequent and systematic large deviations caused by the market player, leading to warnings and penalties from the market authorities. Nevertheless, as this strategy represents the state-of-the-art in academic research, it will establish a second benchmark for comparative purposes with the novel strategies.

In the following Sections, the three novel advanced bidding strategies are presented, where both wind power and market quantity forecasts are used to define the bid in order to increase the revenue while controlling the exposure to imbalance charges.

4.4.3 Scalar Adjustment

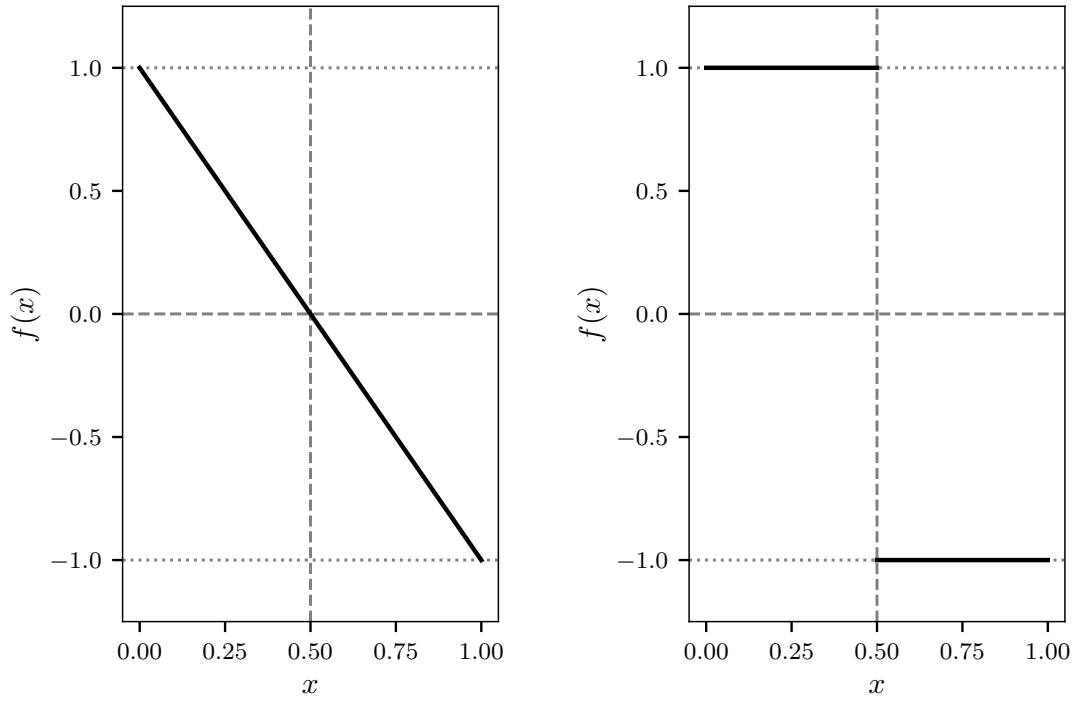
In the *Scalar Adjustment* (Scal) strategy, the bid quantity is equal to the forecast wind power increased or decreased by a fraction of the installed capacity:

$$E_{t+h}^{DAM} = \hat{E}_{t+h|t} + \rho \cdot f\left(x\right) \cdot E_{max} \tag{4.26}$$

where $\rho \in [0, 1]$ is the *risk factor*, $x \in [0, 1]$ is the market quantity forecast, $f : x \in [0, 1] \rightarrow f\left(x\right) \in [-1, 1]$ is the *adjustment function*, and $E_{t+h}^{DAM} \in [0, E_{max}]$. Note that the bid quantity has to be bounded between 0 and E_{max}, so negative values of E_{t+h}^{DAM} are set equal to zero and values exceeding the installed capacity are set equal to E_{max}. The maximum possible adjustment is defined a priori based on the risk appetite of the trader via the parameter ρ. The direction and magnitude of the adjustment are determined by the adjustment function $f(\cdot)$, which in turn depends on the market quantity forecast x. This function gives great flexibility of use since it can be defined to reflect the sensitivity of the producer to the uncertainty associated with the market quantity forecast. Figure 4.11 shows two examples of possible functions. On the left, f is a linear curve where the size of the adjustment increases linearly in value when the prediction is more confident — that is, closer to 0 or 1 — and tends to zero when the prediction is uncertain — that is, around the value of 0.5. On the right, f is a simple step (or piecewise constant) function, resulting in the size of the adjustment being fixed and equal to ρE_{max}. Both the linear and the step functions give a positive adjustment when $x < 0.5$ and a negative one when $x > 0.5$. It should be noted that the case where $\rho = 0$ corresponds to the ImbMin strategy; on the other hand, if the step function in Figure 4.11b is chosen and $\rho = 1$, this corresponds to the ImbMax strategy.

4.4.4 Proportional Adjustment

In the *Proportional Adjustment* (Prop) strategy, the bid quantity is equal to the forecast wind power increased or decreased by a factor proportional to the forecast



(a) Linear. **(b)** Step.

Figure 4.11: Examples of adjustment functions $f(x)$.

generation itself:

$$E_{t+h}^{DAM} = \hat{E}_{t+h|t} \cdot (1 + \rho \cdot f(x)) \tag{4.27}$$

where $\rho \in [0, 1], f : x \in [0, 1] \rightarrow f(x) \in [-1, 1], E_{t+h}^{DAM} \in [0, E_{max}]$. Again, the bid quantity has to be bounded between 0 and E_{max}, so values exceeding the installed capacity are set equal to E_{max}. Similarly to the Scal case, the maximum adjustment is defined a priori based on the appetite for risk of the producer via the parameter ρ, while the direction and size are determined through the adjustment function $f(x)$. The same principles discussed earlier for f apply here. Furthermore, it should be noted that, regardless of how this function is defined, the size of the expected energy imbalance is proportional to the expected generation. This is a desirable feature of the strategy, as observed by Browell in [199] (where an homologous strategy is called "Multiplicative Adjustment"), as it implies that the exposure to balancing costs and therefore to risk is large only when the expected revenue is large as well, and vice versa.

4.4.5 Quantile Bid

Recall from Section 3.1.1 that the forecast quantile $\hat{q}^{\alpha}$, where $\alpha \in [0, 1]$ is the nominal level, is the value of the distribution for which:

$$\mathbb{P}\left[E \leq \hat{q}^{\alpha}\right] = \alpha \tag{4.28}$$

so that α is the probability that the actual energy will be below $\hat{q}^{\alpha}$ (cf. Equation (3.5))[3]. It also follows that:

$$\mathbb{P}\left[E > \hat{q}^{\alpha}\right] = 1 - \alpha = \mathbb{P}\left[E < \hat{q}^{1-\alpha}\right] \tag{4.29}$$

Writing Equation (4.28) in terms of the energy imbalance $d = E^{DAM} - E$:

$$\alpha = \mathbb{P}\left[E - E^{DAM} \leq \hat{q}^{\alpha} - E^{DAM}\right] = \mathbb{P}\left[-d \leq \hat{q}^{\alpha} - E^{DAM}\right]$$

If the quantity $\hat{q}^{\alpha}$ is offered, it results that:

$$\mathbb{P}\left[d \geq 0 \,|\, E^{DAM} = \hat{q}^{\alpha}\right] = \alpha$$

Therefore, bidding the quantile $\hat{q}^{\alpha}$, the probability that the expected energy imbalance d will be positive is equal to α. Similarly from Equation (4.29), it can be shown that:

$$\mathbb{P}\left[d < 0 \,|\, E^{DAM} = \hat{q}^{1-\alpha}\right] = \alpha$$

That is, bidding the quantile $\hat{q}^{1-\alpha}$ results in a probability α that the energy imbalance d will be negative.

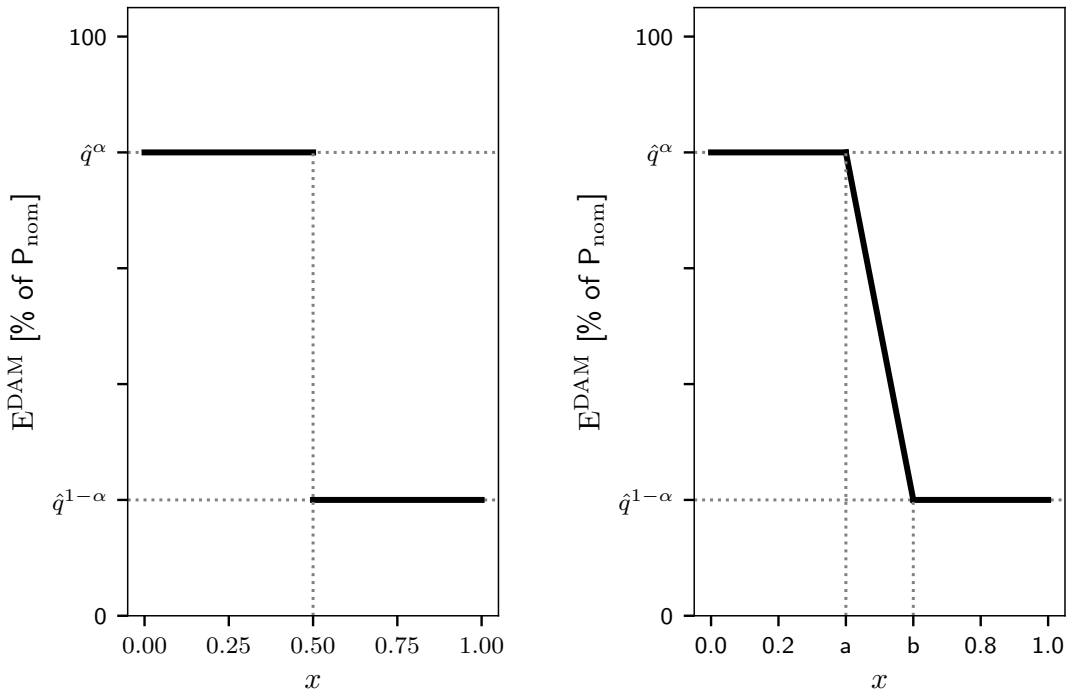
In the *Quantile Bid* (Quant) strategy, the offered quantity is a function of the market quantity forecast and of the forecast quantiles $\hat{q}^{\alpha}_{t+h|t}$ and $\hat{q}^{1-\alpha}_{t+h|t}$ of wind power generation, where α is the probability that the realisation of $d_{t+h|t}$ has positive or

[3]In this paragraph, time indices are discarded in order to lighten the notation.

negative sign, as shown above. The strategy can be written as:

$$E_{t+h}^{DAM} = \begin{cases} \hat{q}_{t+h|t}^{\alpha} & \text{if } x < a \\ g\left(\hat{q}_{t+h|t}^{\alpha}, \hat{q}_{t+h|t}^{1-\alpha}, x\right) & \text{if } x \in [a, b] \\ \hat{q}_{t+h|t}^{1-\alpha} & \text{if } x > b \end{cases} \tag{4.30}$$

The value of α is set based on the desired trade-off between revenue and risk, giving this strategy the valuable property of modelling explicitly the uncertainty associated with the wind power forecast. The function $g(\cdot)$ and parameters a and b can be defined heuristically based on the sensitivity of the participant to the uncertainty of the market quantity forecast x. Figure 4.12 shows two examples of possible implementations of the strategy. On the left panel, the bid quantity



(a) Step (a=b=0.5).

(b) Piecewise linear (a=0.4; b=0.6).

Figure 4.12: Example of two implementations of the Quant strategy. On the left, $a = b = 0.5$, hence no function $g(\cdot)$ is actually defined. On the right, $a = 0.4$ and $b = 0.6$ and $g = \hat{q}^{\alpha} - \frac{\hat{q}^{\alpha} - \hat{q}^{1-\alpha}}{b-a} \cdot (x - a)$

.

is simply $\hat{q}^{\alpha}$ or $\hat{q}^{1-\alpha}$ depending on whether the price difference is expected to be

positive ($x \leq 0.5$) or negative ($x > 0.5$), respectively. On the right panel, the bid is a piecewise linear function with the two quantiles being interpolated linearly in the central region of the market quantity forecast x delimited by parameters a and b to hedge against price uncertainty when the forecast is more cautious (i.e. around 0.5).

Chapter 5

Results and Discussion

This Chapter evaluates the results of the methodology developed in this Thesis. In the case study, the day-ahead trading of wind energy in the Irish electricity market is simulated over a test period. Energy is traded employing the bidding strategies and associated market quantity forecasts developed in this work. First, the wind power forecasts for the wind farm under study are assessed in Section 5.1. Next, the two market quantity forecasts are evaluated. Results for the price difference forecasts are presented in Section 5.2, and for the imbalance sign forecasts in Section 5.3. Finally, the performance of the various bidding strategies is analysed in Section 5.4.

5.1 Wind Power Forecasts

This Section evaluates the accuracy of the wind power forecasts introduced in Section 4.2. The forecasts are evaluated over the seven-month period from the 1st October 2018 to the 25th April 2019 and all the analyses that follow are based on the framework and error measures defined in Section 3.1.2. Values of observed and forecast wind power are normalised by the installed capacity of the wind farm P_{nom} and thus lie between 0 and 1; forecast errors and derived scores are then expressed as a percentage of the installed capacity. The four prediction models are indicated as A, B, C and D, with the first three consisting of point forecasts only

and the last including point and 80% interval forecasts. Owing to confidentiality restrictions, further information regarding the models and the forecast providers cannot be discussed.

First, the point forecasts are assessed. The histograms and boxplots in the next three Figures allow the visual inspection of the errors from each model. Figure 5.1, where the distribution of forecast errors as well as the maximum likelihood Gaussian distribution fit are displayed, shows that errors from all four models do not follow a normal distribution and are positively skewed (i.e. the right tail is longer). The four models are all positively biased (i.e. on average they forecast lower energy output than is realised), and the largest errors (both positive and negative) are observed for models B and C, as seen in Figure 5.2. The boxplots in Figure 5.3 show the distribution of forecast errors by hour of the day. There is no visible pattern in any of the four subplots, suggesting that there is no clear dependence of errors on the hour of the day and hence, on the forecast horizon. This is not surprising, as the wind farm is located on complex terrain.

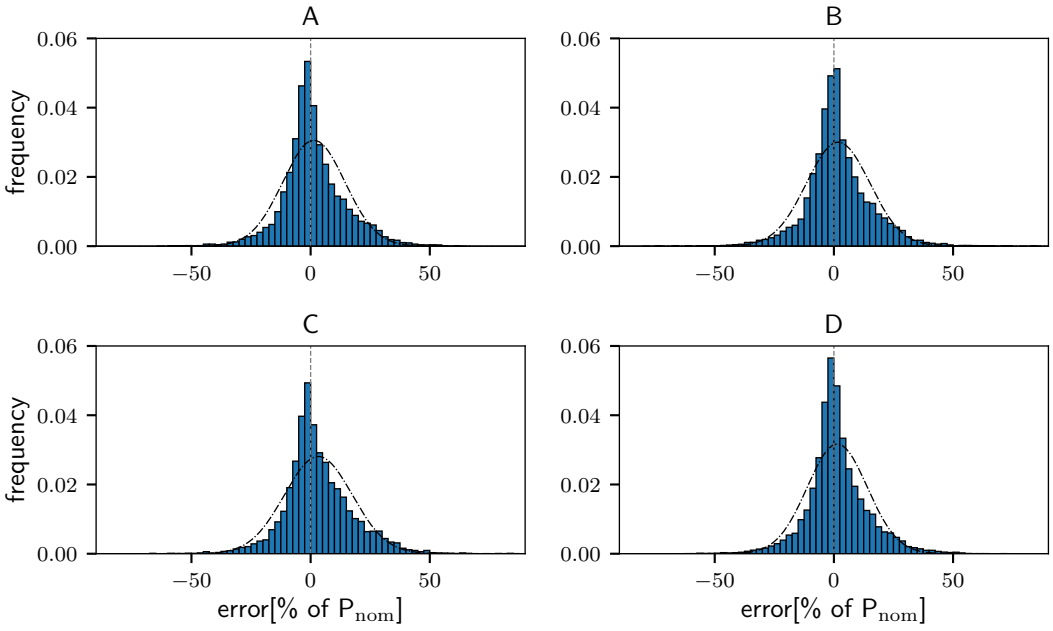


Figure 5.1: Distribution of wind power forecast errors for each model. The black dash-dotted line corresponds to the maximum likelihood Gaussian distribution fit.

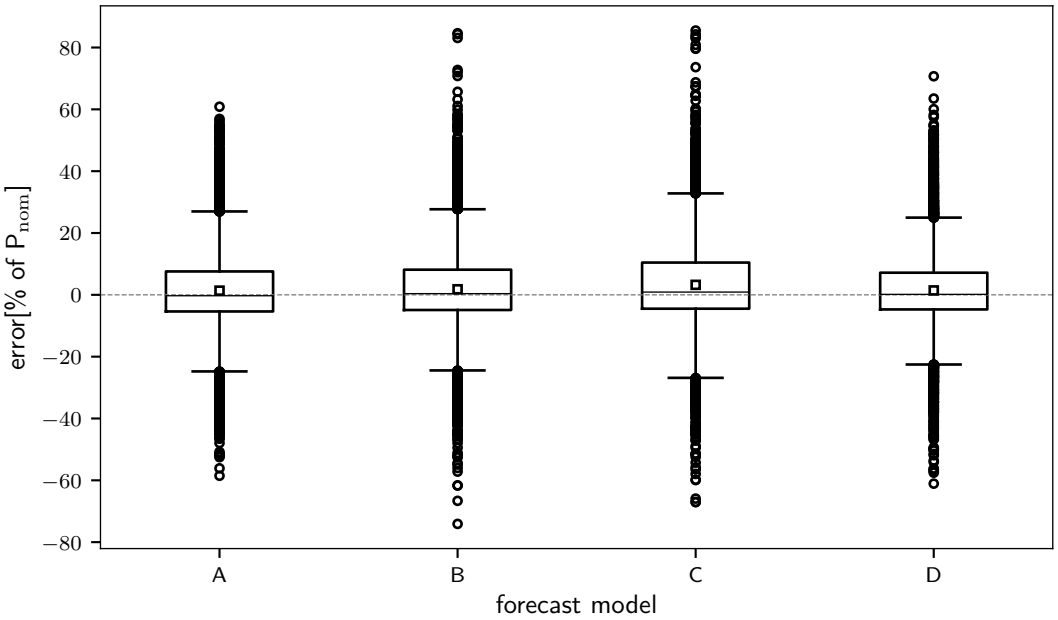


Figure 5.2: Boxplots of wind power forecast errors for each model. The squares indicate the mean of the distribution (i.e. the forecast bias).

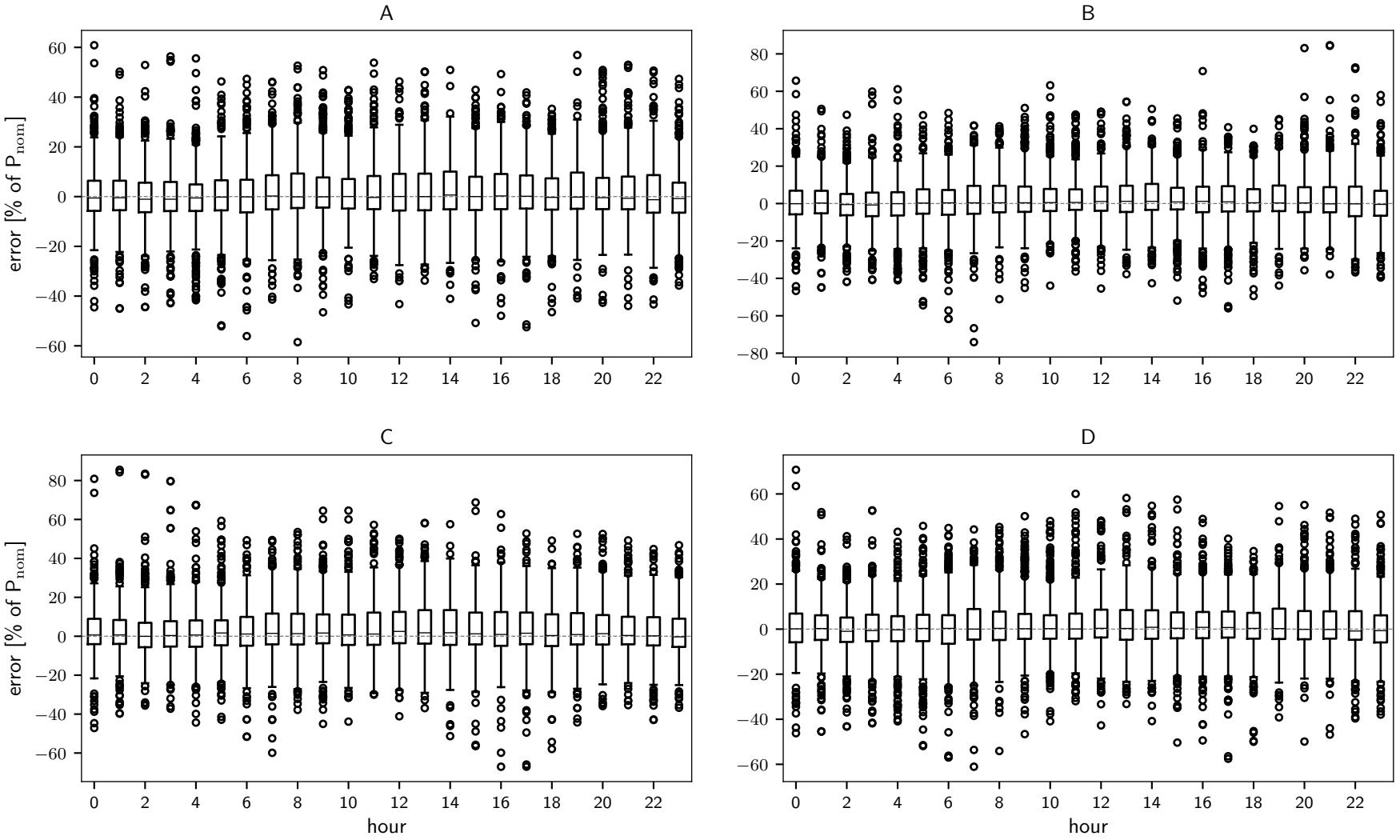


Figure 5.3: Boxplots of wind power forecast errors by hour of the day for each model.

Model	bias	MAE	RMSE
A	1.35	9.36	13.12
B	1.79	9.42	13.41
C	3.23	10.35	14.57
D	1.42	8.85	12.68

Table 5.1: Evaluation scores of the wind power point forecasts. All scores are expressed as a percentage of the installed capacity P_{nom}.

Table 5.1 presents the bias, MAE and RMSE of the forecasts calculated over the whole evaluation period, while Figures 5.4 and 5.5 show the MAE and RMSE, respectively, calculated for each hour separately. These results indicate that D is the most accurate model as measured by MAE and RMSE and across all hours. To confirm the observation made earlier in Figure 5.3 on the errors not depending on the hour of the day, it is worth noting from Figures 5.4 and 5.5 that the accuracy of the four models does not degrade with a clear trend for larger look-ahead times.

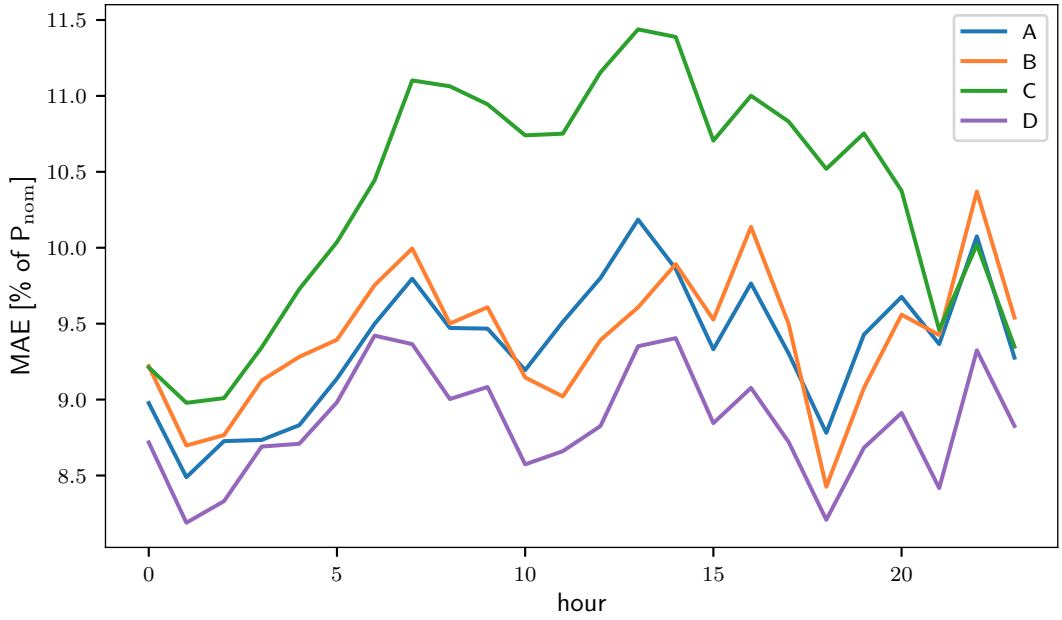


Figure 5.4: Mean Absolute Error of wind power forecasts as a function of the hour of day.

The cumulative and 30-day rolling MAE is also calculated for the four models and results are shown in Figures 5.6 and 5.7, respectively. In both Figures, model D emerges as the most consistently accurate of the four throughout the entire period.

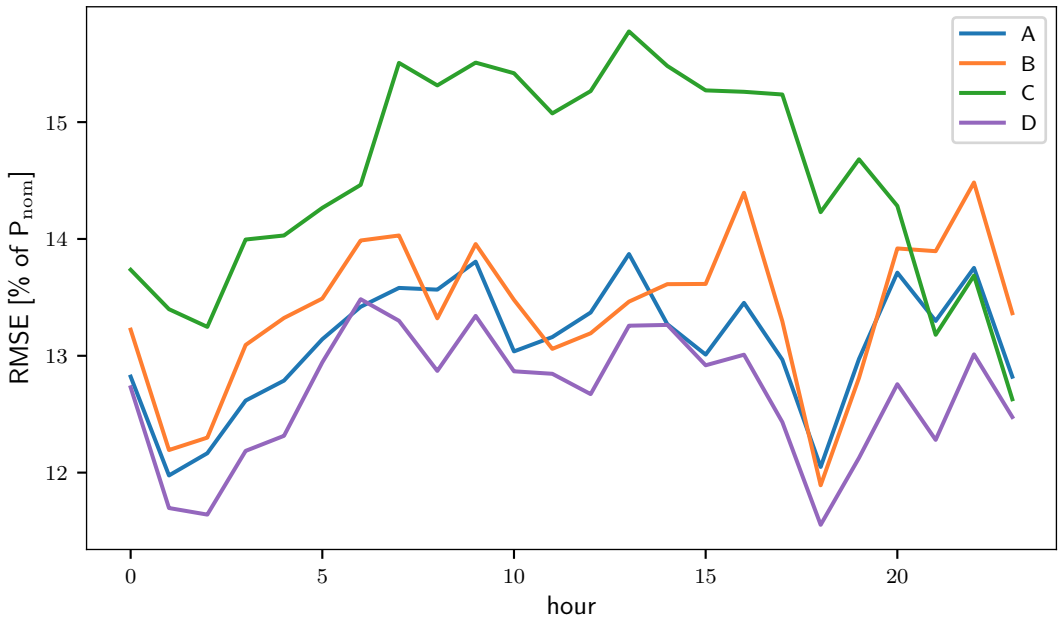


Figure 5.5: Root Mean Square Error of wind power forecasts as a function of the hour of day.

It is worth noting that the accuracies of the models follow similar patterns, possibly indicating the use of the same NWP model by the various forecast providers.

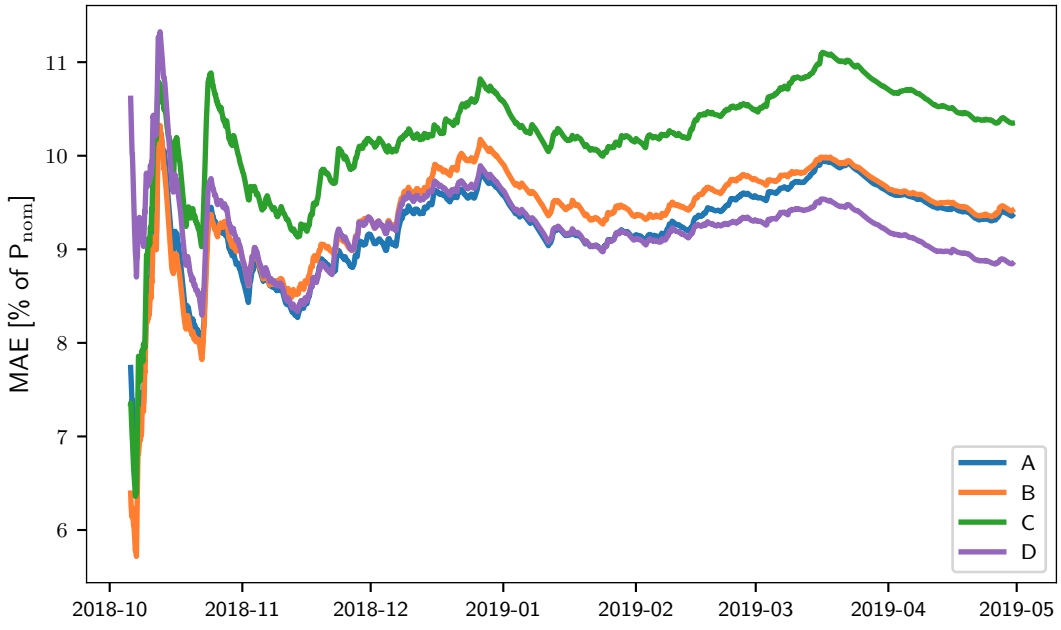


Figure 5.6: Cumulative MAE of wind power forecast models.

Finally, the DM test is applied to model D versus the three others to verify whether the methods have significantly different accuracy over each of the 48 fore-

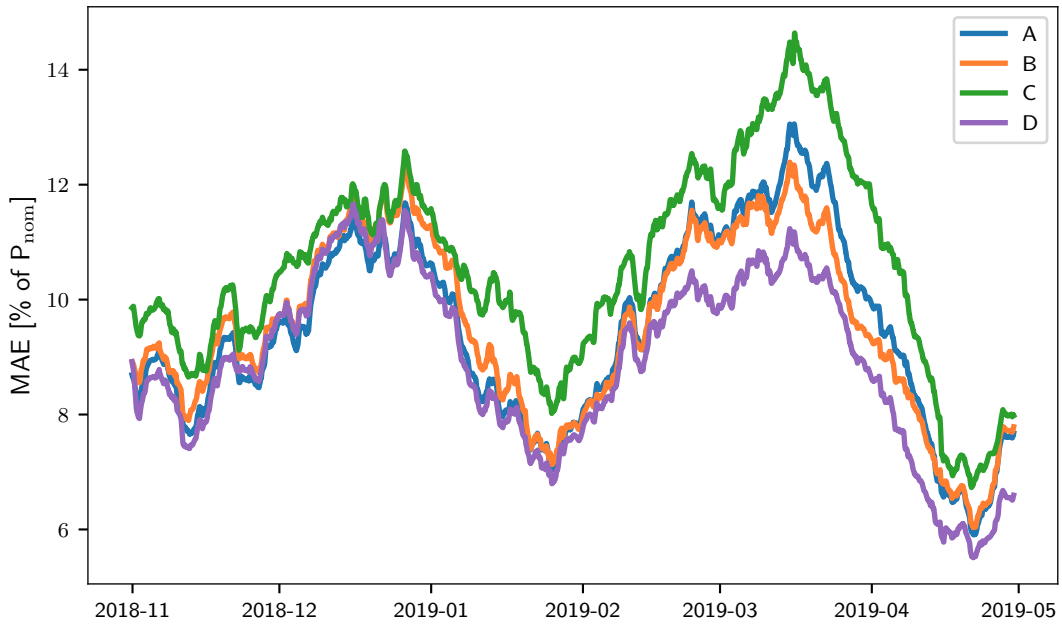


Figure 5.7: 30-day rolling MAE of wind power forecast models.

cast horizons[1]. When testing D versus A, B and C, the null hypothesis that the models have the same accuracy is rejected in 12, 14, and 31 cases, respectively. A second DM test is applied on these periods to verify the null hypothesis that the other model is significantly better than D. The results of this second DM test are tabulated in Table 5.2. In all cases, the p-values returned by the tests are lower than the 5% significance level, meaning that the hypothesis of the alternative model being more accurate than D can be rejected. Hence, it can be concluded that model D is the more accurate of the four point forecasts.

Next, the 80% prediction interval from D is evaluated. The calibration is assessed by calculating the empirical levels of the 10^{th} and 90^{th} quantile forecasts defining the interval. The results are presented in the reliability diagram in Figure 5.8, showing that the reliability curve is flatter than the 45-degree line, with the 10^{th} quantile being systematically overestimated (probabilistic bias equal to –7%) and the 90^{th} quantile systematically underestimated (probabilistic bias equal to 13%). This means that the quantiles are not perfectly calibrated and intervals

[1]Recall that the forecast resolution is 30 minutes, and therefore forecasts are issued for 48 distinct look-ahead times.

hour	A vs D	B vs D	C vs D	hour	A vs D	B vs D	C vs D
00:00	-	-	-	12:00	0.01908	-	0.01392
00:30	-	0.01008	-	12:30	0.02128	-	0.00956
01:00	-	-	-	13:00	-	-	-
01:30	-	0.00439	-	13:30	-	-	0.02300
02:00	-	-	-	14:00	-	0.02263	0.01031
02:30	-	-	-	14:30	-	-	0.01942
03:00	-	-	-	15:00	-	-	-
03:30	-	-	-	15:30	0.01920	0.00933	0.00701
04:00	-	-	0.02166	16:00	0.01533	0.02380	0.00131
04:30	-	-	-	16:30	0.00967	0.00535	0.00164
05:00	-	-	-	17:00	0.00055	0.00786	0.00264
05:30	-	-	-	17:30	-	5.92E-05	0.00907
06:00	-	-	0.00800	18:00	-	-	0.00016
06:30	-	-	-	18:30	-	-	1.89E-08
07:00	-	0.00265	0.00047	19:00	-	0.02424	2.59E-06
07:30	0.01395	0.00389	0.00023	19:30	0.00741	-	7.59E-09
08:00	0.00744	-	4.07E-07	20:00	-	-	2.20E-14
08:30	-	-	2.72E-05	20:30	-	-	0.00047
09:00	0.01388	-	9.88E-16	21:00	-	-	0.00043
09:30	-	-	3.05E-06	21:30	-	-	0.01188
10:00	-	-	1.72E-09	22:00	-	-	0.01669
10:30	-	-	1.33E-08	22:30	-	0.02013	-
11:00	0.02471	-	0.00107	23:00	-	0.00303	-
11:30	0.00694	-	0.00800	23:30	-	0.01166	-

Table 5.2: p-values returned by the DM tests applied to models A, B, and C versus D. The null hypothesis is that the first model is significantly more accurate than D.

are on average too narrow.

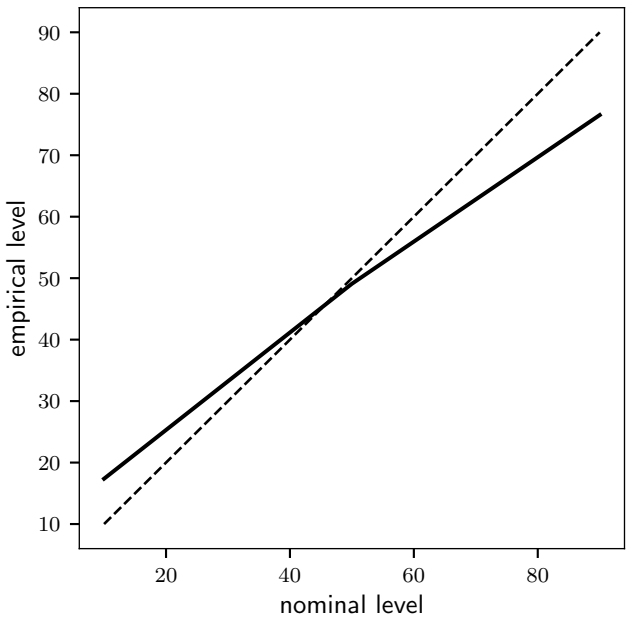


Figure 5.8: Reliability diagram of the 80% interval forecast. The dashed line is the diagonal indicating a perfectly calibrated forecast.

To examine the sharpness, the boxplots of the prediction intervals for each hour of the day are shown in Figure 5.9, and the average widths (cf. Equation (3.25)) are tabulated in Table 5.3. It can be observed that while there is no clear pattern in the variation of the medians, the mean and the range of the distributions increase with the forecast horizon, resulting in an increasing trend for the sharpness. This indicates that the uncertainty associated with the predictions increases with larger look-ahead times.

Point and quantile forecasts from model D will be employed in Section 5.4 to implement the bidding strategies described in Section 4.4.

5.2 Electricity Price Forecasts

This Section evaluates the electricity price forecasting methods developed in Section 4.3.1. The ARIMAX models for DAM and BM prices were trained on a 4-month period from the 1st October 2018 to the 28th January 2019 and are eval-

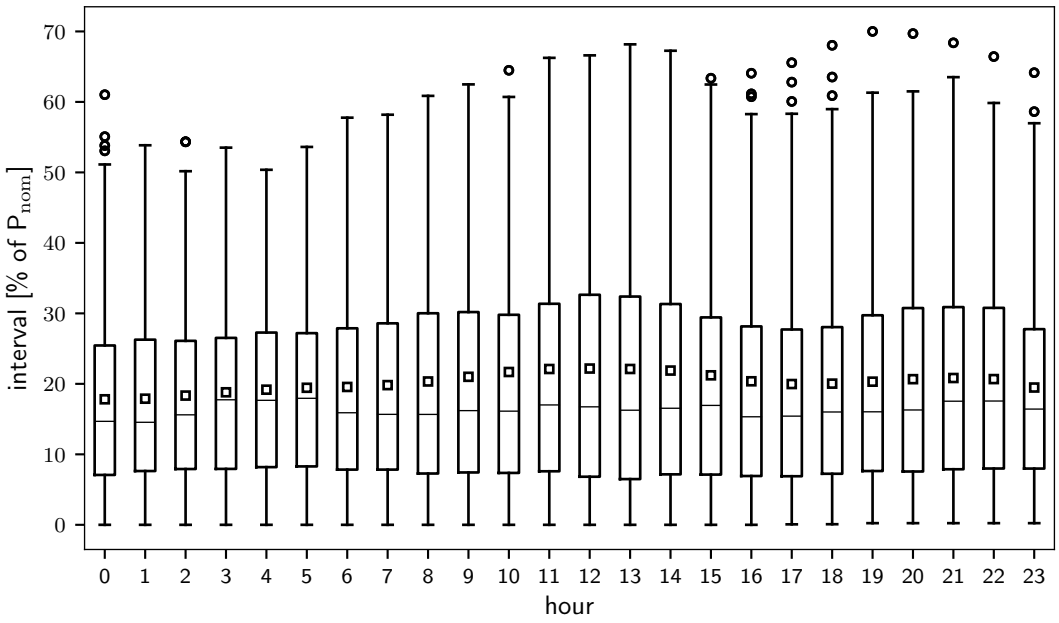


Figure 5.9: Boxplots of the 80% predictive intervals by hour of day. The squares indicate the mean of the distribution (i.e. the sharpness).

uated on the 3-month test period from the 29th January 2019 to the 25th April 2019. The forecasting processes were implemented using `R` and `python` and run on an Intel Core i7 processor with 3.4 GHz CPU and 16 GB memory. The total processing time to produce the DAM and BM price forecasts was under three and a half hours, while running the Monte Carlo simulations took under one and a half hours. As for the evaluation of wind power forecasts, the analyses are based on the framework and error measures defined in Section 3.1.2. Results are presented for DAM prices first, and then BM prices. In both cases, an overview of the ARIMAX models' fit is given first, followed by the evaluation of the point forecasts and of the quantile forecasts. Finally, the results of the Monte Carlo simulations are presented and evaluated.

5.2.1 Day-ahead Market Prices

As explained in Section 4.3.1.2, in both the "Rec. expanding" and "Rolling" approaches, the ARIMAX model is re-selected and re-estimated at each iteration. Figures 5.10 and 5.11 display the occurrence of models with the same order for the

hour	sharpness	hour	sharpness
0	17.8	12	22.2
1	17.9	13	22.1
2	18.3	14	21.9
3	18.8	15	21.2
4	19.2	16	20.4
5	19.4	17	20.0
6	19.6	18	20.0
7	19.8	19	20.3
8	20.3	20	20.7
9	21.0	21	20.8
10	21.7	22	20.7
11	22.1	23	19.5

Table 5.3: Sharpness of the 80% predictive intervals by hour of day. Values are expressed as a percentage of P_{nom}.

"Rec. expanding" and "Rolling" approaches, respectively. For both cases, models of order (0,1,1) are the most common and account for more than 30% of the fitted models. Interestingly, the second most common models are those of order (0,0,0). These are in fact multiple linear regression models (cf. Equation (4.18)), where future prices are estimated exclusively through the predictions of the exogenous variables. Recall that the variables included as predictors in the regression part of the ARIMAX models are: forecast demand, point and 80% interval forecasts of wind power generation at system level, and predicted wind power penetration.

Point forecasts

The forecast errors from the two advanced methods are displayed in Figures 5.12 and 5.13. The histograms and the boxplots show that errors have very similar distributions, although large negative errors are more frequent in the "Rolling" approach. Both approaches are also negatively biased, meaning that they tend to overpredict prices.

The performance of the two advanced methods is compared to that of the benchmark — the "similar-day" method presented in Section 4.3.1.2 (cf. Equation (4.19)) — in terms of model fit and accuracy scores. First, the three methods

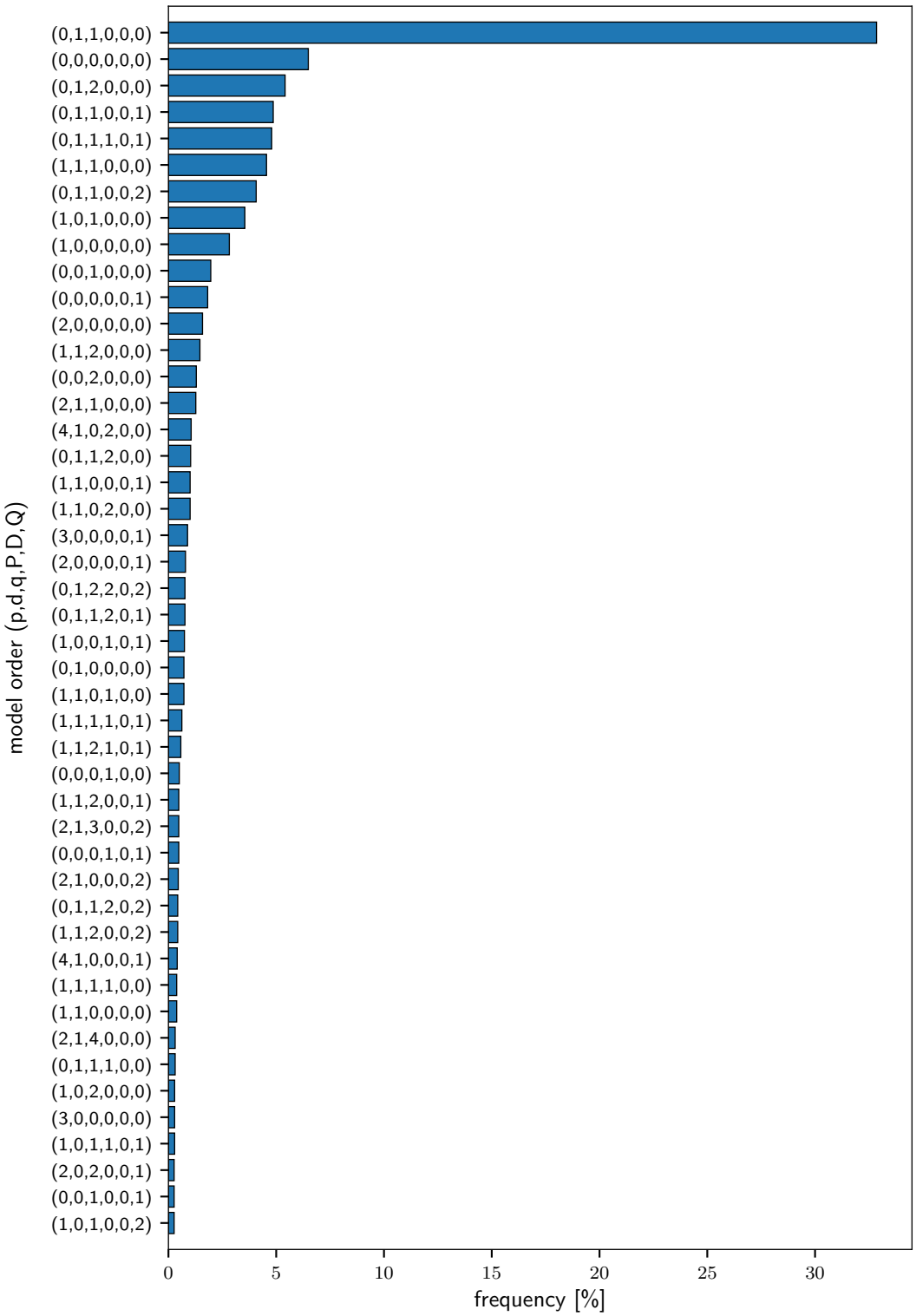


Figure 5.10: Frequency of DAM models with the same ARIMA order for the "Rec. expanding" approach.

are analysed through scatter plots, as illustrated in Figure 5.14. Here, forecast prices are plotted against the corresponding observed values; the red diagonal is

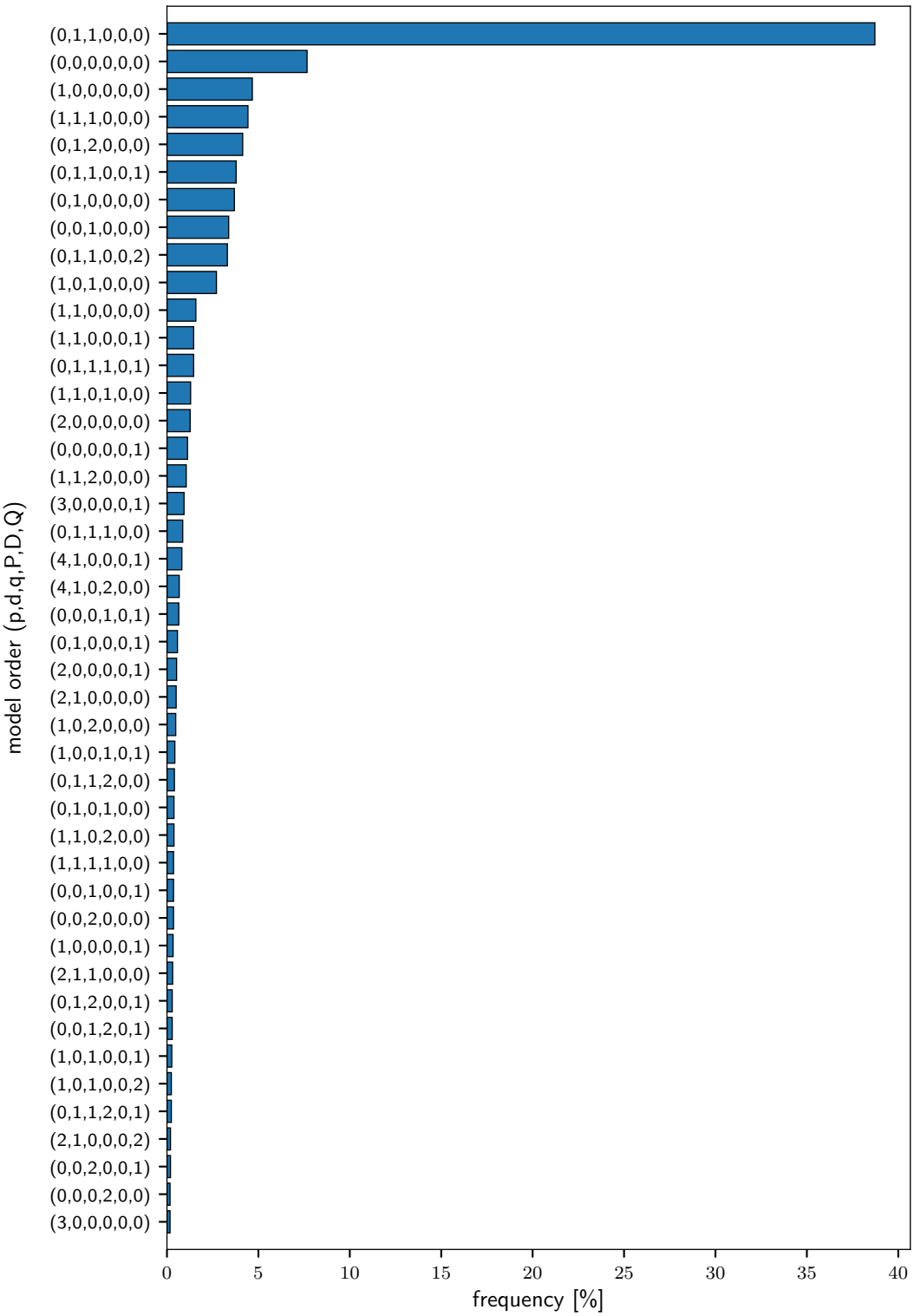


Figure 5.11: Frequency of DAM models with the same ARIMA order for the "Rolling" approach.

the $y = x$ line representing the perfect forecast, and the coefficient of determination (R^2) is displayed in blue for each model. From the Figure, it can be observed that

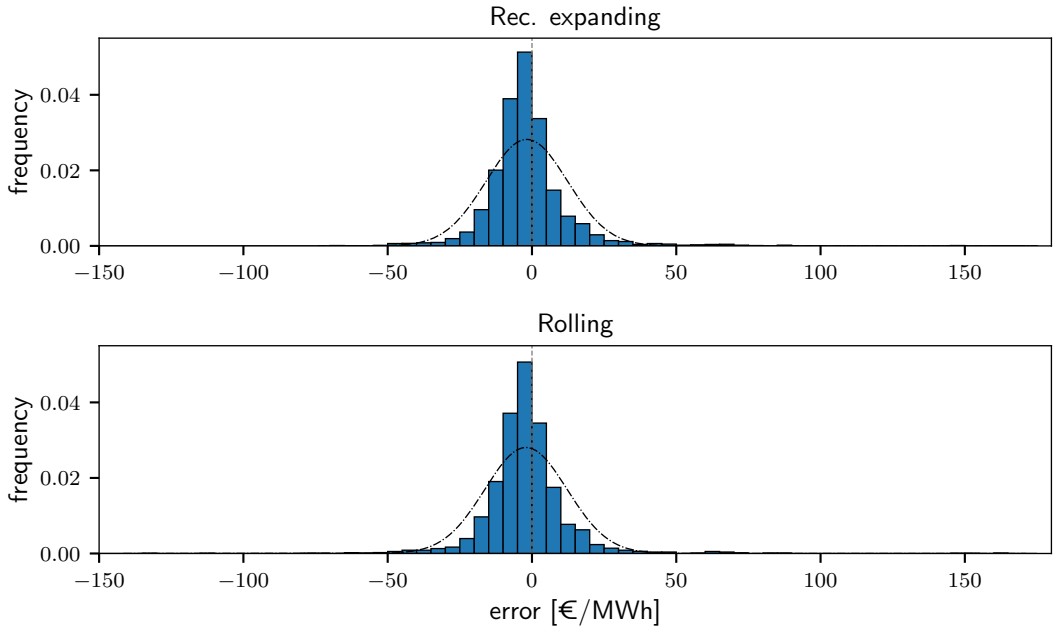


Figure 5.12: Distribution of DAM price forecast errors for the "Rec. expanding" (top) and "Rolling" (bottom) approaches.

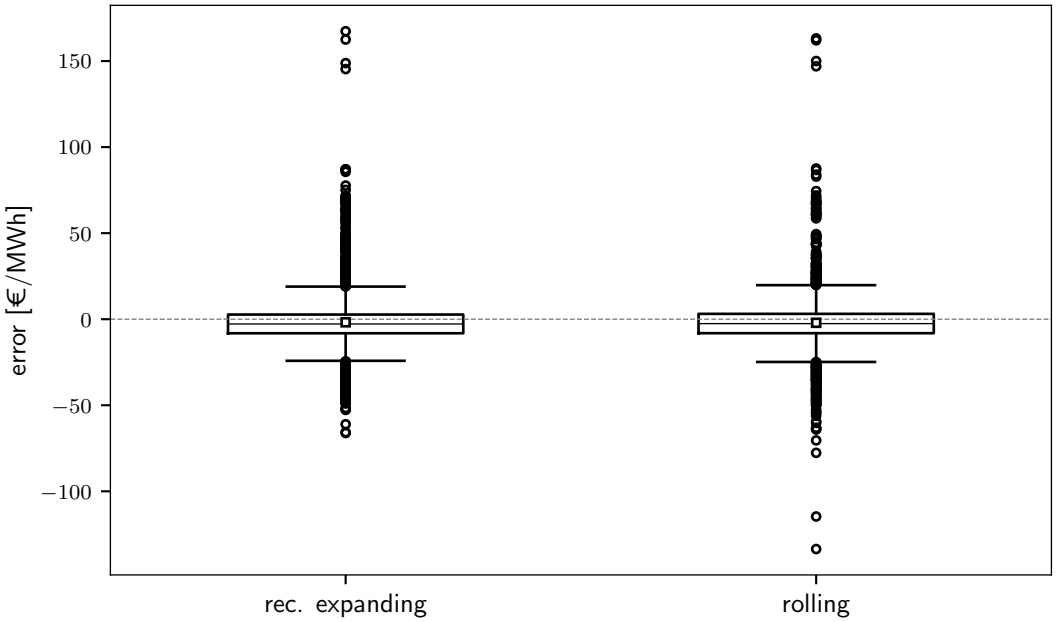


Figure 5.13: Boxplots of DAM price forecast errors for each approach. The squares indicate the mean of the distribution (i.e. the forecast bias).

the two advanced methods are better fits than the benchmark, as points lie closer to the diagonal and the value of R^2 is higher. Moreover, the two advanced methods tend to underpredict when prices are high (i.e. larger than 100 €/MWh).

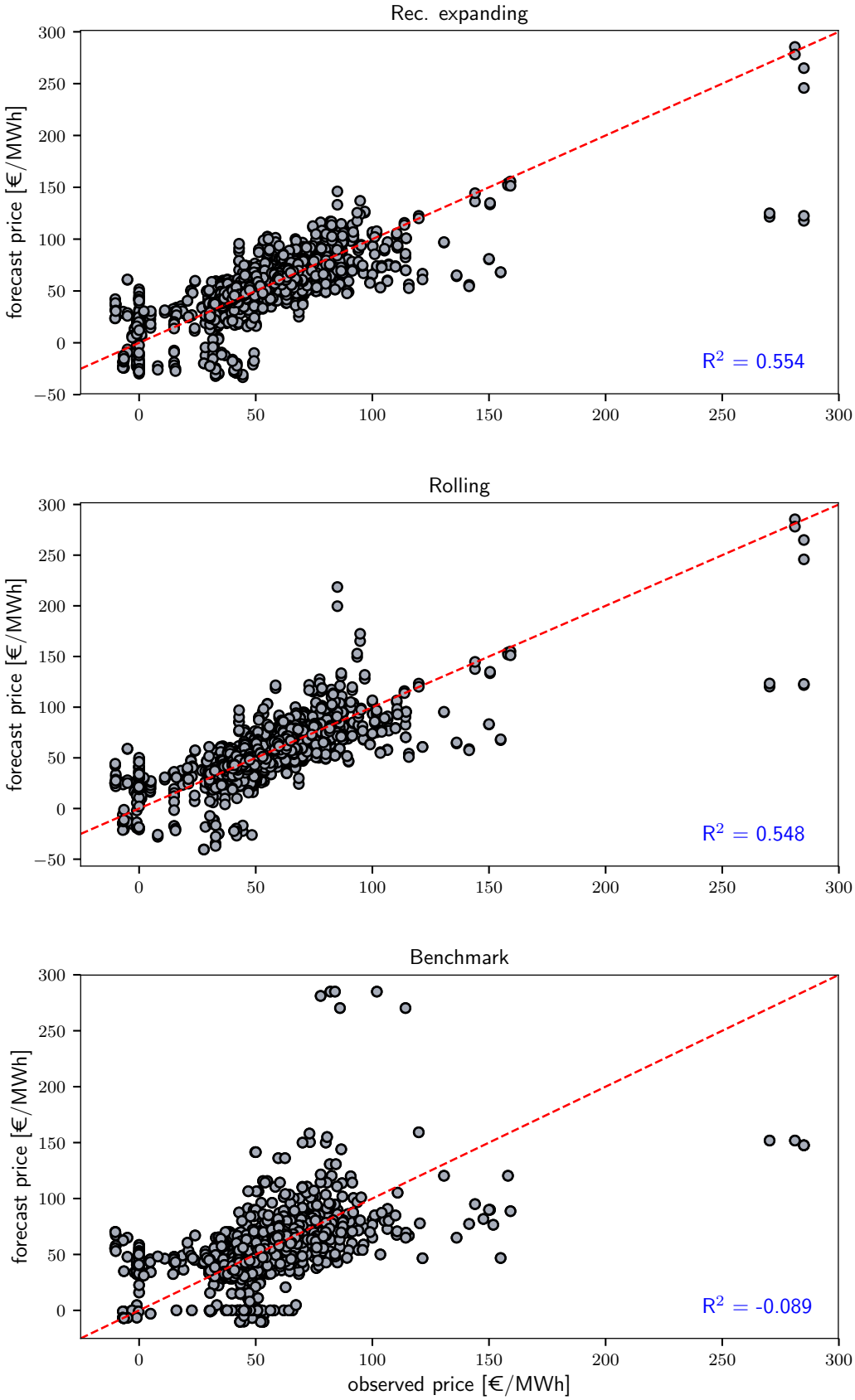


Figure 5.14: Scatter plot of observed vs forecast day-ahead prices for each forecasting method. The coefficient of determination of each model is reported in blue. The red line represents the perfect forecast.

The bias, MAE, RMSE, MASE and RMSSE of the three methods calculated over the test period are given in Table 5.4. The accuracy of the two advanced methods is almost identical and markedly better than that of the benchmark, as measured by MAE and RMSE and their scaled homologous. The MASE and RMSSE (cf. Equation (3.19) are also calculated separately for each half-hour period — namely, the Imbalance Settlement Period[2] — and the results are shown in Figures 5.15 and 5.16, respectively. The graphs confirm that the two advanced methods perform similarly and both better than the benchmark over all periods. Interestingly, scaled errors from all three methods follow similar trends both in terms of magnitude and dispersion: lower in the morning and evening, and highest at night. Finally, since separate sets of models have been developed for weekdays and weekend days, the MAE and RMSE are calculated for each day of the week separately and results are shown in Figures 5.17 and 5.18. Again, the "Rec. expanding" and "Rolling" methods consistently perform better than the benchmark. Even more interestingly, the errors are low and very similar between Tuesday and Friday, but almost twice as large on Monday and on the weekend. The reason for this is that forecasts for these days are generated using less recent information: Monday prices are predicted by the weekday models using three-day old information (i.e. from the previous Friday), and Saturday and Sunday prices are predicted using information from six and seven days before (i.e. the previous weekend). In future work, this initial condition problem should be addressed by including more recent information in the training data set. For example, the models estimating Monday prices could include Sunday historical data; similarly, models estimating weekend prices could include Friday historical data.

The MAE of the two methods is calculated on a moving-window basis to investigate how the accuracy has changed over the test period. Figure 5.19 shows the cumulative MAE of the two advanced forecasts, and Figure 5.20 the two-week rolling MAE. These graphs show that even though by the end of the test period

[2]Note that while DAM prices are hourly, the forecasts are half-hourly to match the resolution of BM prices.

Model	bias	MAE	RMSE	MASE	RMSSE
Rec. expanding	–1.79	9.11	14.26	0.61	0.91
Rolling	–2.08	9.08	14.36	0.60	0.89
Benchmark	–1.86	13.28	22.29	0.86	1.29

Table 5.4: Evaluation scores of the DAM price point forecasts. All scores are expressed in €/MWh, except for MASE and RMSSE in per unit.

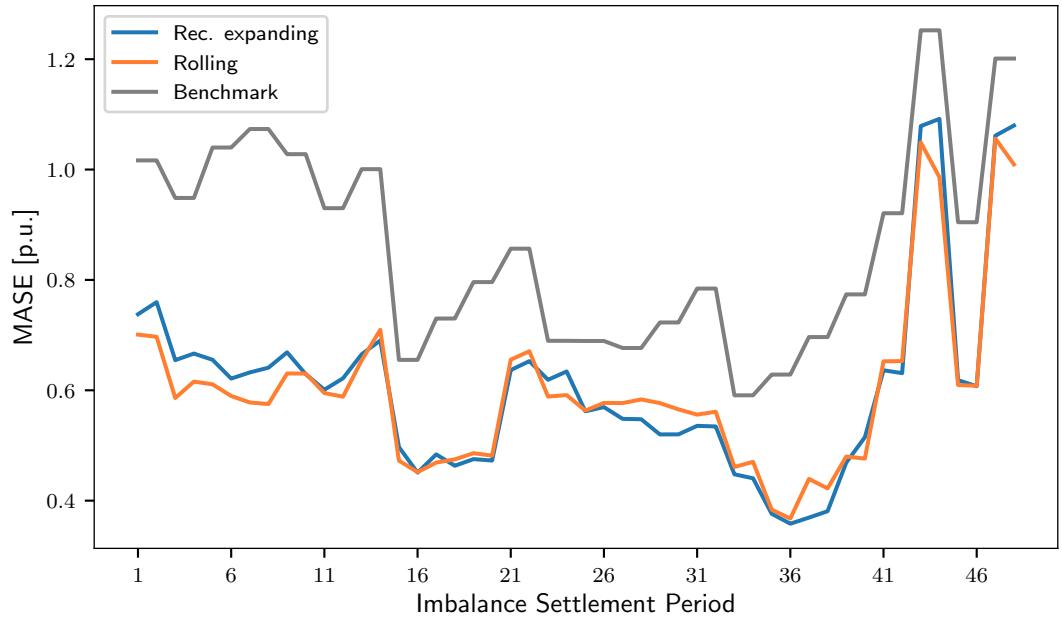


Figure 5.15: Mean Absolute Scaled Error of DAM price forecasts for each half-hour.

the two approaches have almost identical accuracies, the "Rec. expanding" approach is more accurate during the first half (until mid-March) and the "Rolling" approach during the second. An explanation for this can be found by looking back at the study of day-ahead market data in Section 2.2.2.1. A generally decreasing trend was observed in average prices and their volatility from February onwards (cf. Figures 2.5 and 2.6). While the larger and expanding training data set gives an initial advantage to the "Rec. expanding" approach, the use of older historical data for model calibration penalises its accuracy as the price level changes. The "Rolling" approach, on the other hand, is more dynamic as models are trained using only recent historical data, and is able to better capture the non-stationarity of the price series.

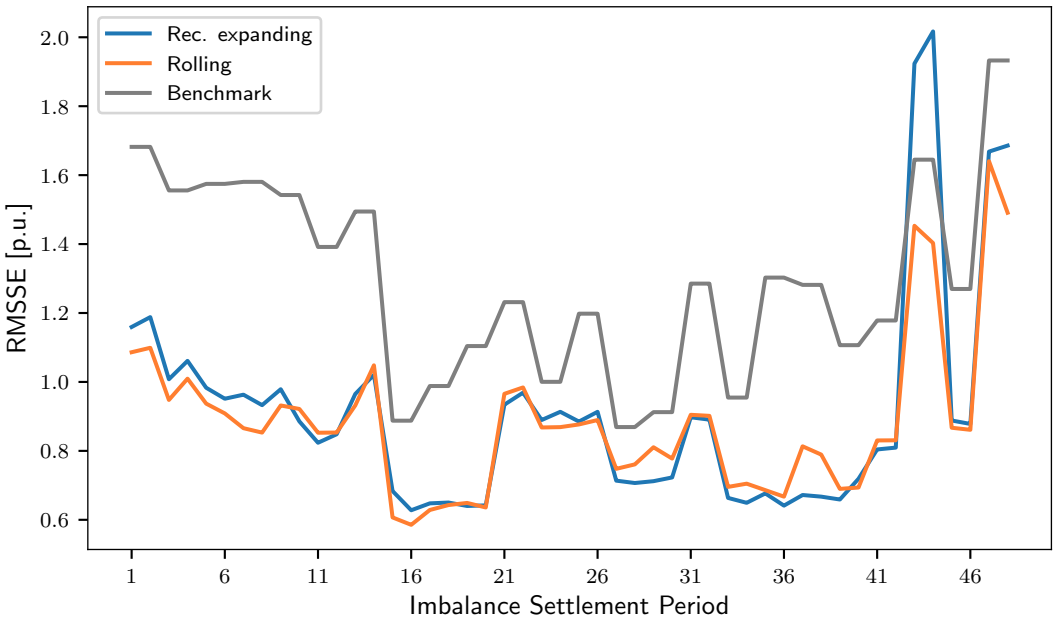


Figure 5.16: Root Mean Square Scaled Error of DAM price forecasts for each half-hour.

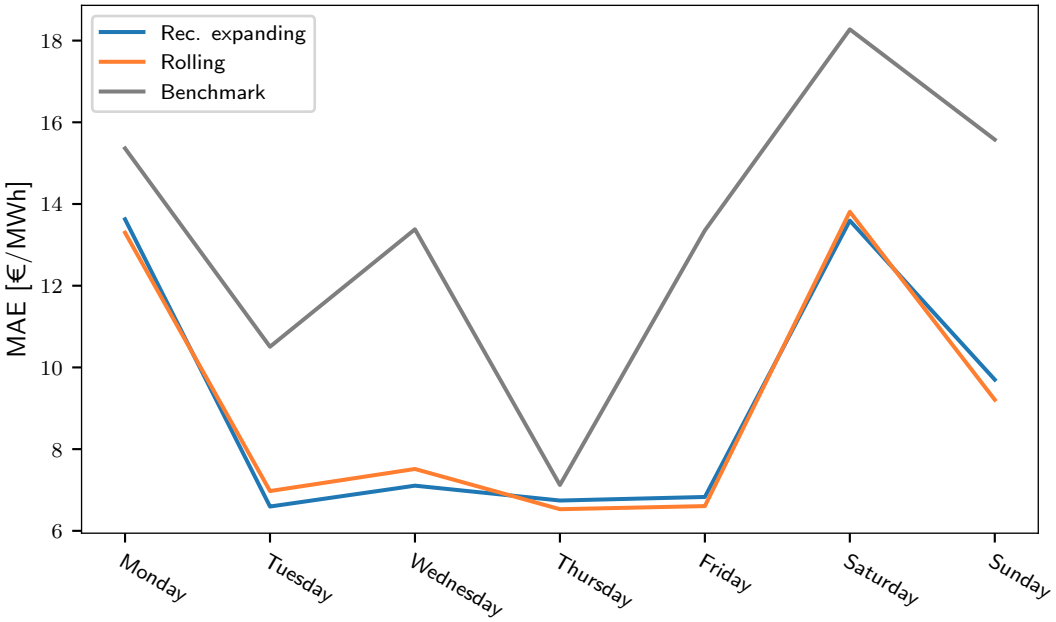


Figure 5.17: Mean Absolute Error of DAM price forecasts for each day of the week.

To conclude the analysis on point forecasts of DAM prices, having shown that both advanced methods outperform the benchmark and have very similar scores, DM tests are applied to determine if one of the two methods is more accurate than

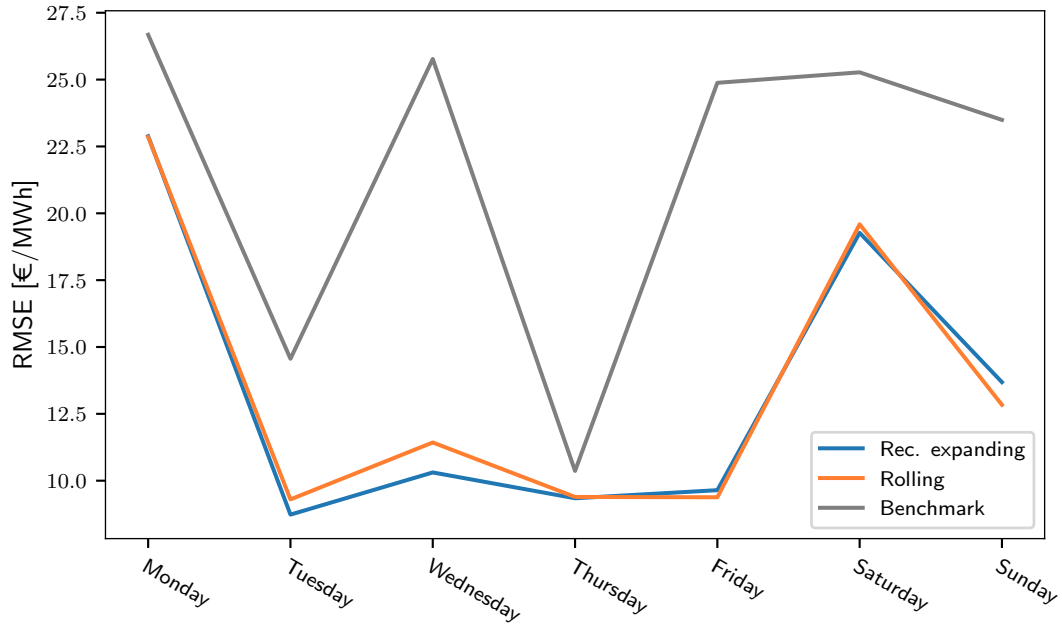


Figure 5.18: Root Mean Square Error of DAM price forecasts for each day of the week.

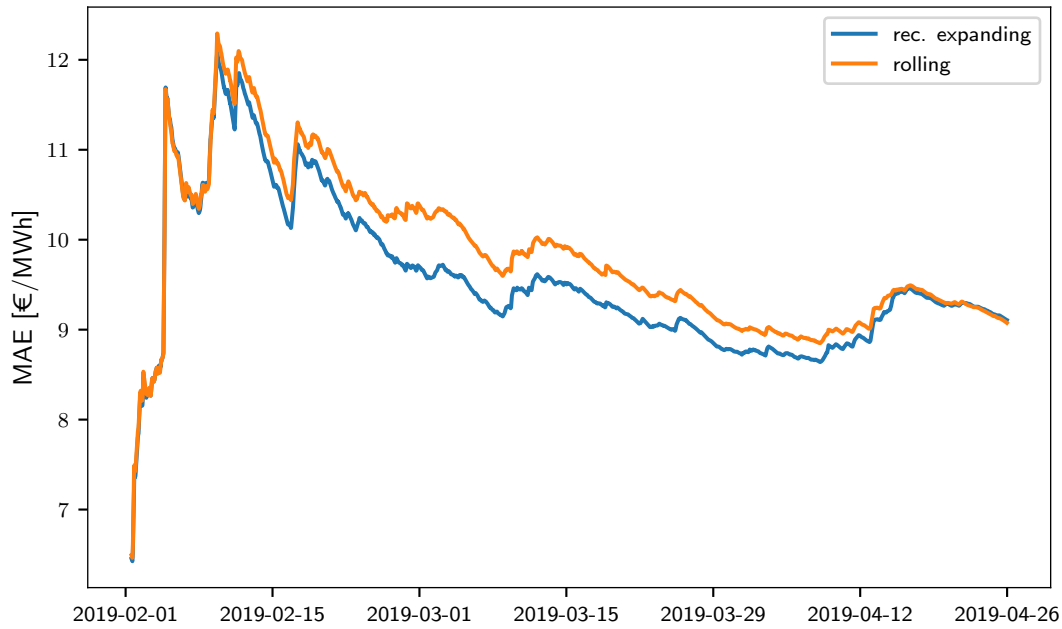


Figure 5.19: Cumulative Mean Absolute Error of DAM price forecasts.

the other. First, a DM test is applied to verify whether or not the two methods have the same accuracy. A second DM test is then applied on the periods where the hypothesis of same accuracy is rejected. The null hypothesis in this case is that the "Rec. expanding" method is significantly more accurate than the "Rolling"

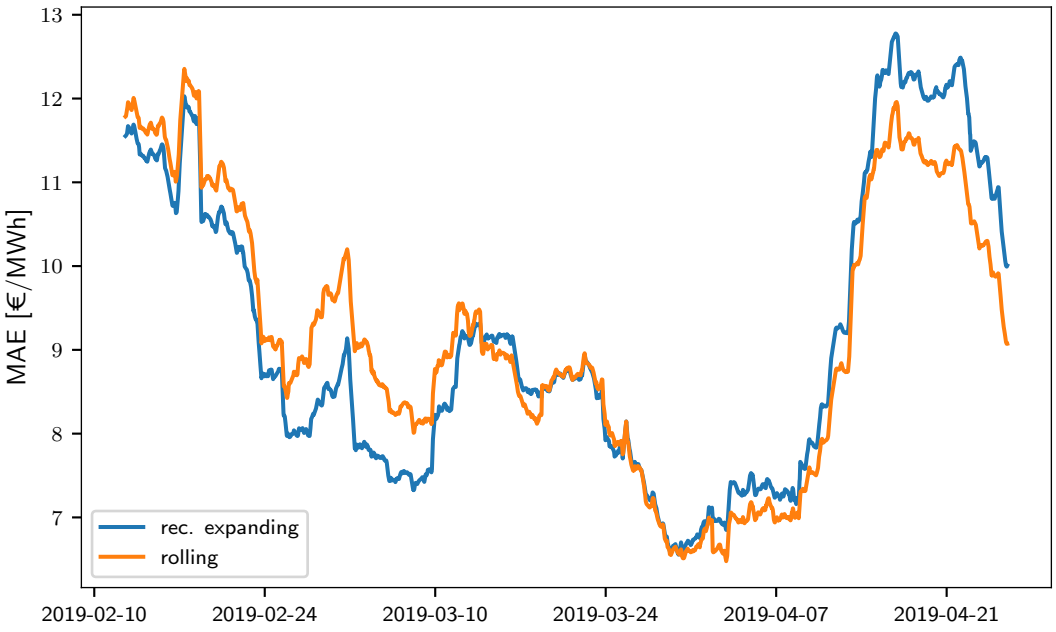


Figure 5.20: Two-week rolling Mean Absolute Error of DAM price forecasts.

and the results are presented in Table 5.5. The hypothesis of same accuracy is rejected on only four out of 48 periods, and the p-values returned by the second test are lower than the 5% significance level on three periods and higher on one. Given the results, it is difficult to distinguish between the two models, leading to the conclusion that they have the same accuracy.

Probabilistic forecasts

The verification of the quantile forecasts is now presented. Recalling that the quantile forecasts generated have nominal levels equal to 5, 10, 25, 75, 90 and 95%, the prediction intervals have 50, 80, and 90% nominal coverage rates. Figure 5.21 depicts the reliability diagram for the quantile forecasts from each method and empirical levels are tabulated in Table 5.6 , while the pinball loss scores are tabulated in Table 5.7. Both "Rec. expanding" and "Rolling" models are significantly more reliable than the benchmark and overall, the "Rolling" quantiles are slightly better calibrated than the "Rec. expanding" ones. Nevertheless, both advanced methods tend to overestimate the quantiles, especially at lower nominal levels.

In future work, statistical post-processing techniques could be applied to im-

hour	Rec. exp vs Roll	hour	Rec. exp vs Roll	hour	Rec. exp vs Roll
00:00	-	08:00	-	16:00	-
00:30	-	08:30	-	16:30	-
01:00	-	09:00	-	17:00	-
01:30	-	09:30	-	17:30	-
02:00	-	10:00	-	18:00	-
02:30	-	10:30	-	18:30	-
03:00	-	11:00	0.001776	19:00	-
03:30	-	11:30	0.000082	19:30	0.024718
04:00	-	12:00	-	20:00	-
04:30	-	12:30	-	20:30	-
05:00	-	13:00	-	21:00	-
05:30	-	13:30	-	21:30	-
06:00	-	14:00	-	22:00	-
06:30	-	14:30	-	22:30	-
07:00	-	15:00	-	23:00	-
07:30	-	15:30	0.984557	23:30	-

Table 5.5: p-values returned by the DM tests applied to "Rec. expanding" models versus "Rolling" ones for DAM price forecasting. The null hypothesis is that the first model is significantly more accurate than the second.

	nominal level					
Model	5	10	25	75	90	95
Rec. expanding	15.3	22.3	40.5	80.8	89.3	92.3
Rolling	13.7	21.8	38.8	80.3	90.2	92.3
Benchmark	42.6	54.0	81.4	96.9	99.0	99.5

Table 5.6: Empirical levels of the quantile forecasts of DAM prices for each model.

	nominal level					
Model	5	10	25	75	90	95
Rec. expanding	1.35	2.12	3.63	3.91	2.50	1.73
Rolling	1.37	2.15	3.68	3.84	2.45	1.70
Benchmark	4.98	7.38	10.85	7.88	4.79	3.09

Table 5.7: Pinball loss scores of the quantile forecasts of DAM prices for each model.

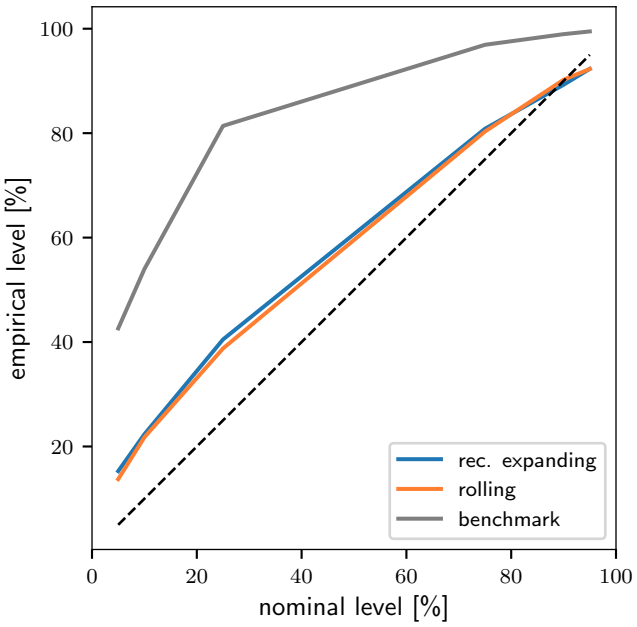


Figure 5.21: Reliability diagram of the DAM price quantile forecasts. The dashed line is the diagonal indicating a perfectly calibrated forecast.

prove the reliability of the forecasts. Adaptive short-term bias-correction [107] can be implemented to reduce the systematic overestimation of the quantiles. The probabilistic bias for each quantile is calculated over the previous N days, and the bias-corrected forecast is obtained by shifting up/down the forecast quantile depending on the bias sign and size. Typically, spot prices evolve smoothly within a trading day. However, as prices are forecast separately for each trading period in the multi-model approach, resulting in quantile forecasts being rather uneven across hours. In this case, smoothing procedures can be applied to the quantile estimates, for example, using a moving average over a N-hour window [217]. Finally, as the record of forecast prices increases, Model Output Statistics (MOS) can be applied [103]. The basic principle behind these techniques is that by using a forecast system long enough, additional information on its performance become available. Assuming that the current forecasts are affected by the same miscalibration, the past performance information can be incorporated in the model to produce improved forecasts. For example, in the case of Gaussian distributions, the mean and standard deviation of the forecast and observed prices is calculated

over the past N days. The distribution of current forecasts can then be corrected so that their mean and variance correspond to those of the observed prices [218].

The two advanced methods are also very similar in terms of sharpness, as shown in Figure 5.22, with intervals always sharper than those from the benchmark. The average interval width is investigated in more detail by calculating the sharpness for each ISP (half-hour) separately. The results are shown in Figure 5.23, where it emerges that the "Rolling" approach produces tighter intervals in the morning when one-step ahead predictions are used, while the "Rec. expanding" is tighter during the afternoon and evening when two-step predictions are used (cf. Figure 4.5). As expected, the plots show that intervals tend to be larger (hence, the prediction is more uncertain) during the evening peak-demand hours, where prices are higher and more volatile.

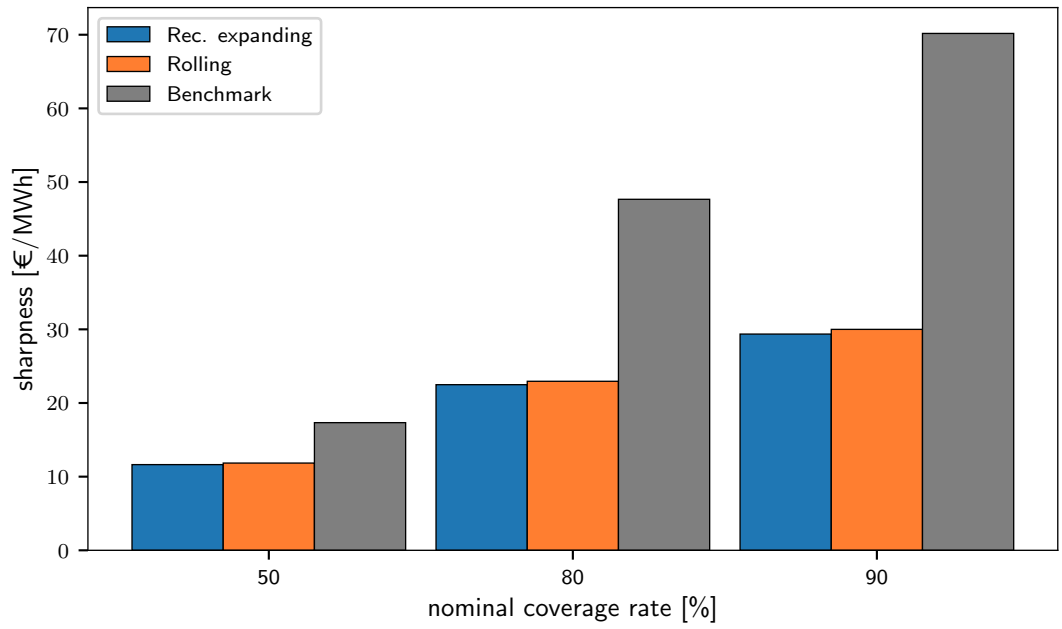


Figure 5.22: Sharpness of the DAM price quantile forecasts.

Complying with the paradigm of "*maximising the sharpness of the predictive distribution subject to calibration*" discussed in Section 3.1.2, the benchmark model is discarded due to its poor calibration and the overall pinball loss score is calculated for the two advanced methods. The score is equal to 2.540 for the "Rec. expanding" method and to 2.533 for the "Rolling" method. Again, the difference between the

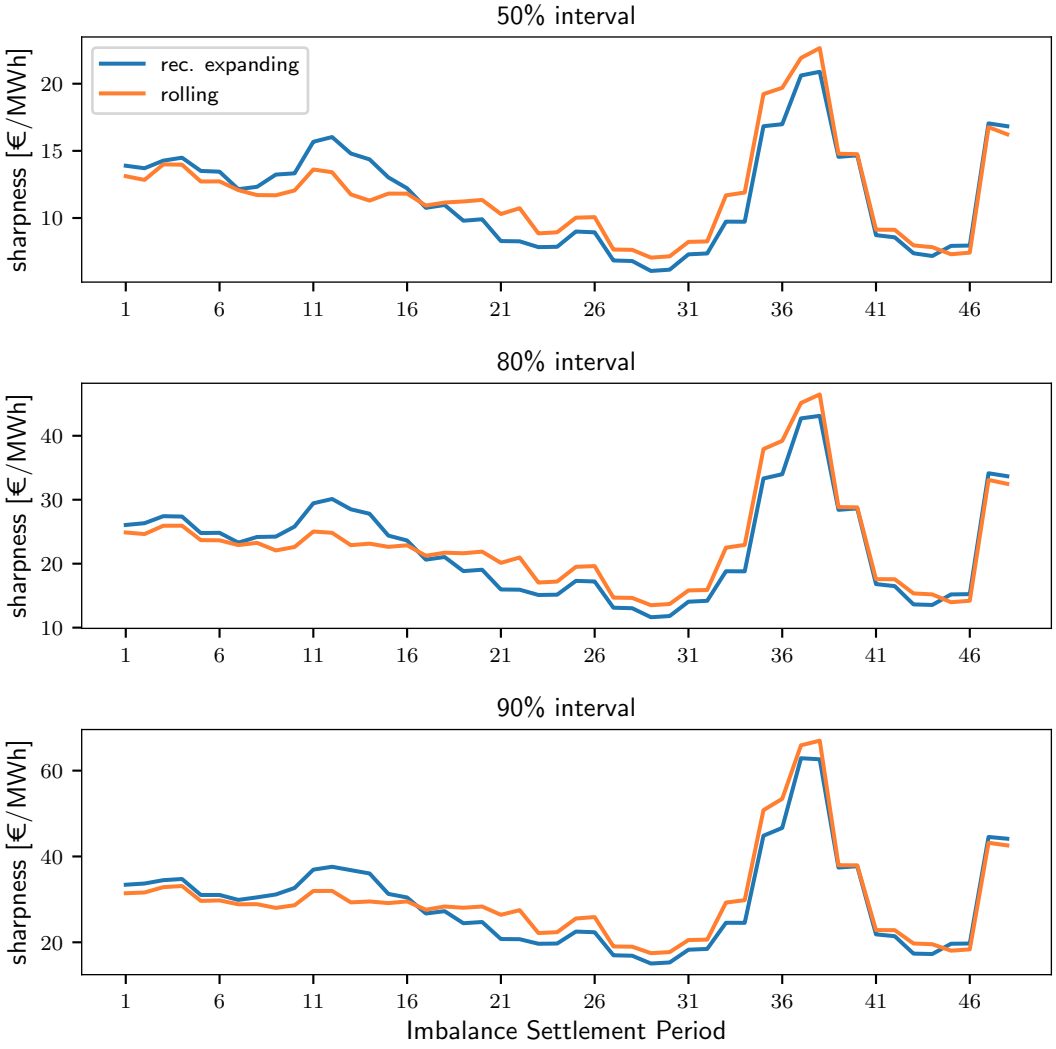


Figure 5.23: Half-hourly sharpness of the DAM price quantile forecasts from the "Rec. expanding" (blue) and "Rolling" (orange) approaches.

methods is minimal and it is not possible to conclude which of the two advanced approaches provides better quantile forecasts.

The difference in performance between the two advanced methods is very small and it is not possible to conclude that either is superior. Nevertheless, the "Rolling" approach is found to generate slightly more accurate point predictions as well as better quantile forecasts over the testing period. This finding suggests that the use of historical data on a moving-window basis to train the models was more beneficial. In fact, the spot price process is characterised by very marked seasonalities and a strong dependence on the level of demand — which in turn

varies with the time of the year — and training the models using only the more recent observations helped to better capture these seasonal variations. Therefore, the "Rolling" method is selected as the forecasting method for DAM prices and its predictions will be used as inputs to the Monte Carlo simulations discussed in Section 4.3.1.3.

5.2.2 Balancing Market Prices

Following the same framework for BM prices, an overview is given of the occurrence of models with the same order for the "Rec. expanding" and "Rolling" approaches in Figures 5.24 and 5.25, respectively. For both cases, models of order (0,0,0) — i.e. multiple linear regression models — are by far the most common and account for more than 80% of the models' fit. This result confirms the autoregressive and seasonal components in the imbalance price process are very weak, and the prediction of future values has to rely primarily on (forecasts of) the exogenous variables.

Point forecasts

The forecast errors from the two advanced methods are displayed in Figures 5.26 and 5.27. The histograms and the boxplots show that errors have very similar distributions, with both approaches being negatively biased and with large positive and negative errors.

The performance of the two advanced methods is compared to that of the benchmark, that is the "similar-day" method, in terms of model fit and accuracy scores. First, the three methods are analysed through scatter plots, as illustrated in Figure 5.28. From the Figure, it can be observed that all three models give a poor fit, with values of R^2 being close to zero or negative. The fit from the benchmark is noticeably poorer than the advanced methods. All three forecasts fail to predict most instances of negative prices and tend to underpredict as values get larger than 100 €/MWh. The bias, MAE, RMSE, MASE and RMSSE of

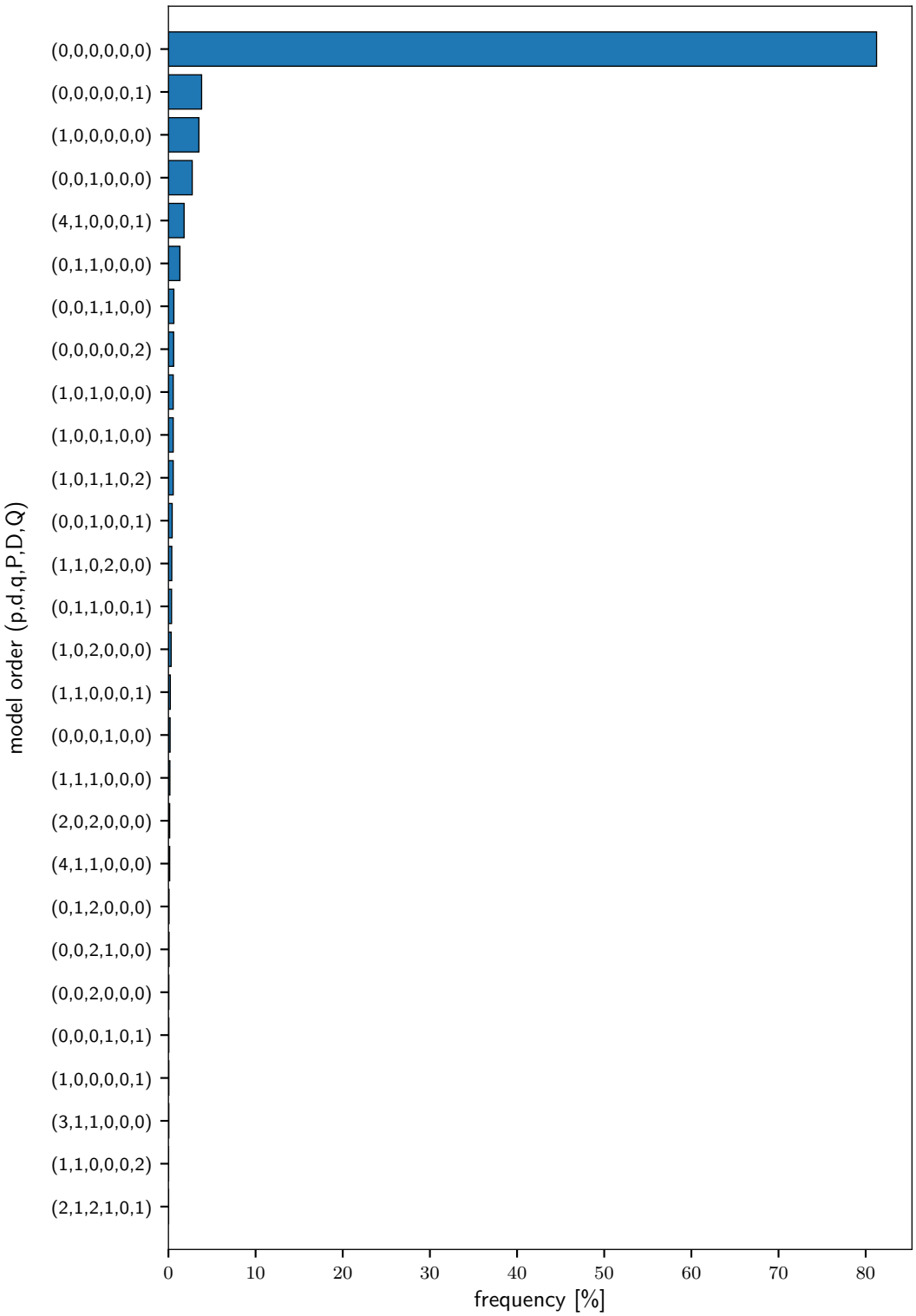


Figure 5.24: Frequency of BM models with the same ARIMA order for the "Rec. expanding" approach.

the three methods calculated over the test period are given in Table 5.8. The accuracy of the two advanced methods as measured by MAE and RMSE and their

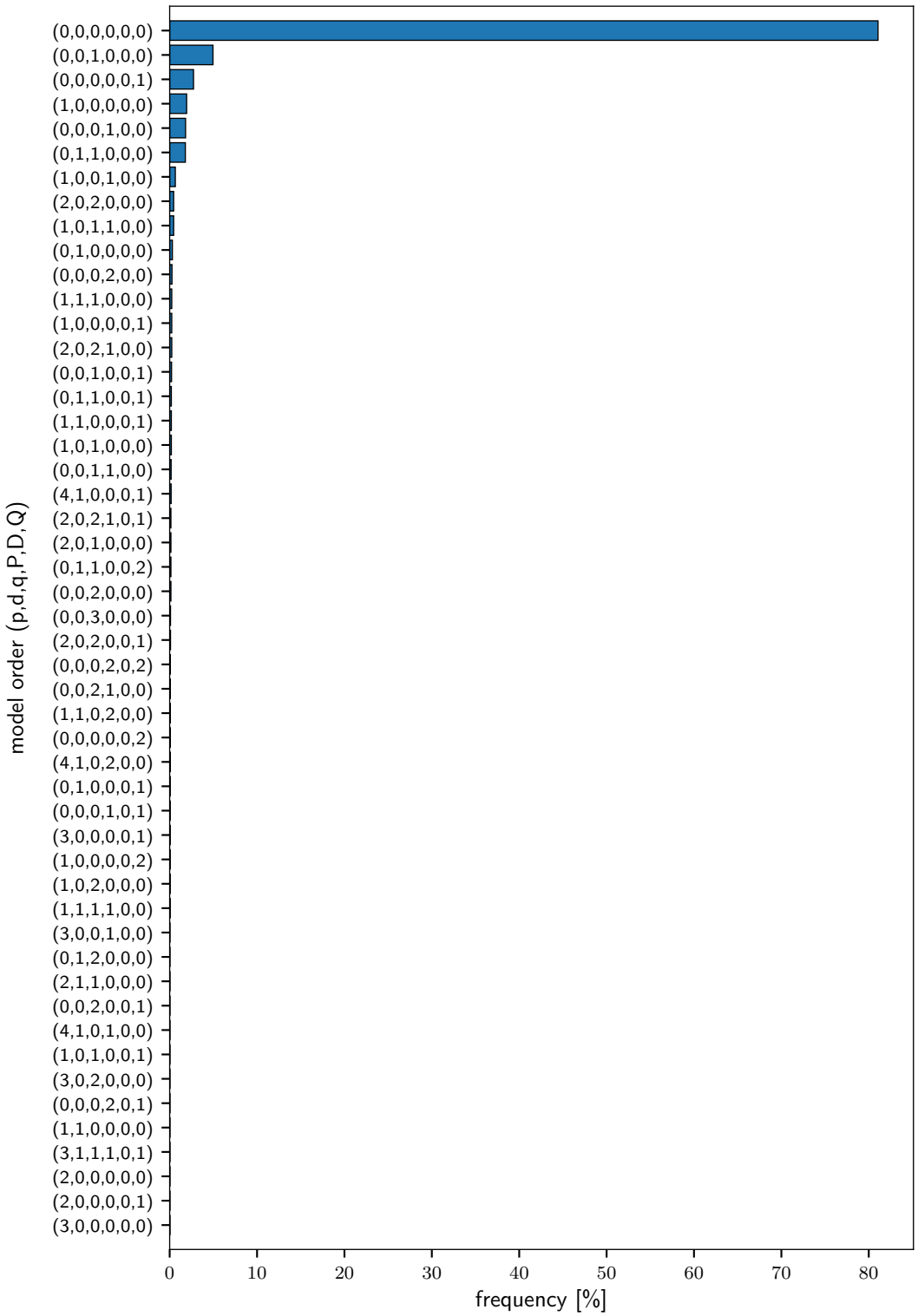


Figure 5.25: Frequency of BM models with the same ARIMA order for the "Rolling" approach.

scaled homologous is markedly better than that of the benchmark, with the "Rec. expanding" approach yielding slightly lower scores. Comparing these scores to

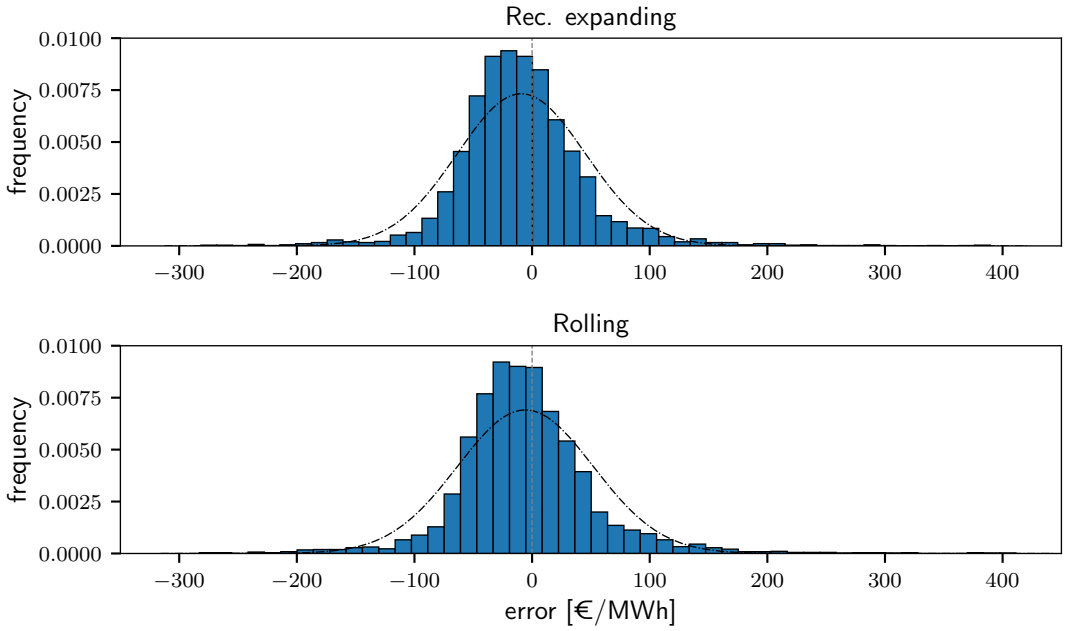


Figure 5.26: Distribution of BM price forecast errors for the "Rec. expanding" (top) and "Rolling" (bottom) approaches.

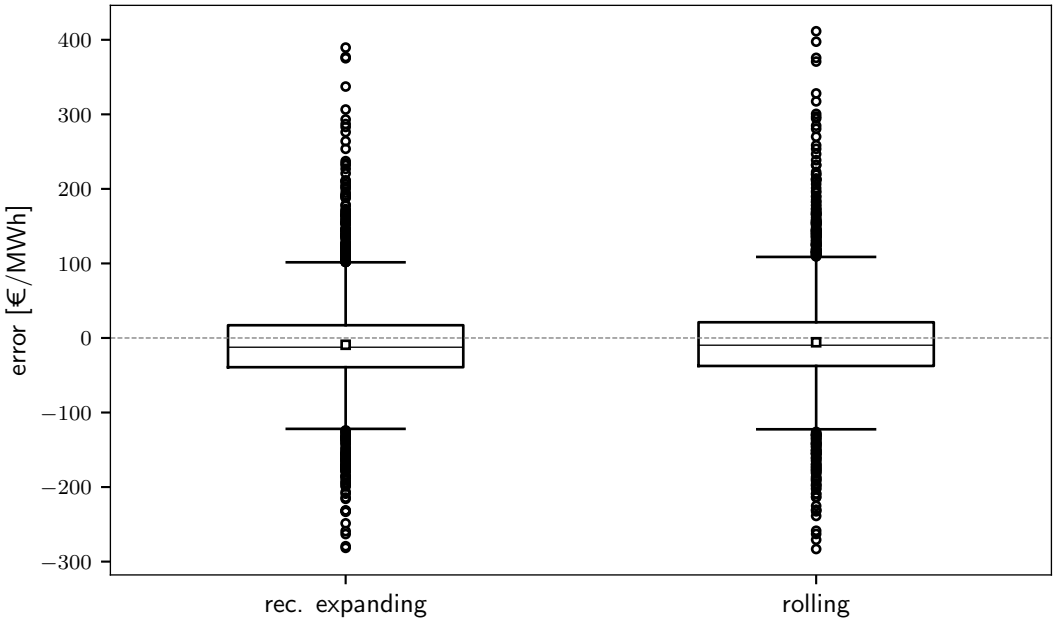


Figure 5.27: Boxplots of BM price forecast errors for each approach. The squares indicate the mean of the distribution (i.e. the forecast bias).

those obtained by spot price forecasts (see Table 5.4), it should be noted that while MAE and RMSE are noticeably higher for imbalance price forecasts, results are very similar in terms of scaled measures. The MASE and RMSSE are also calculated

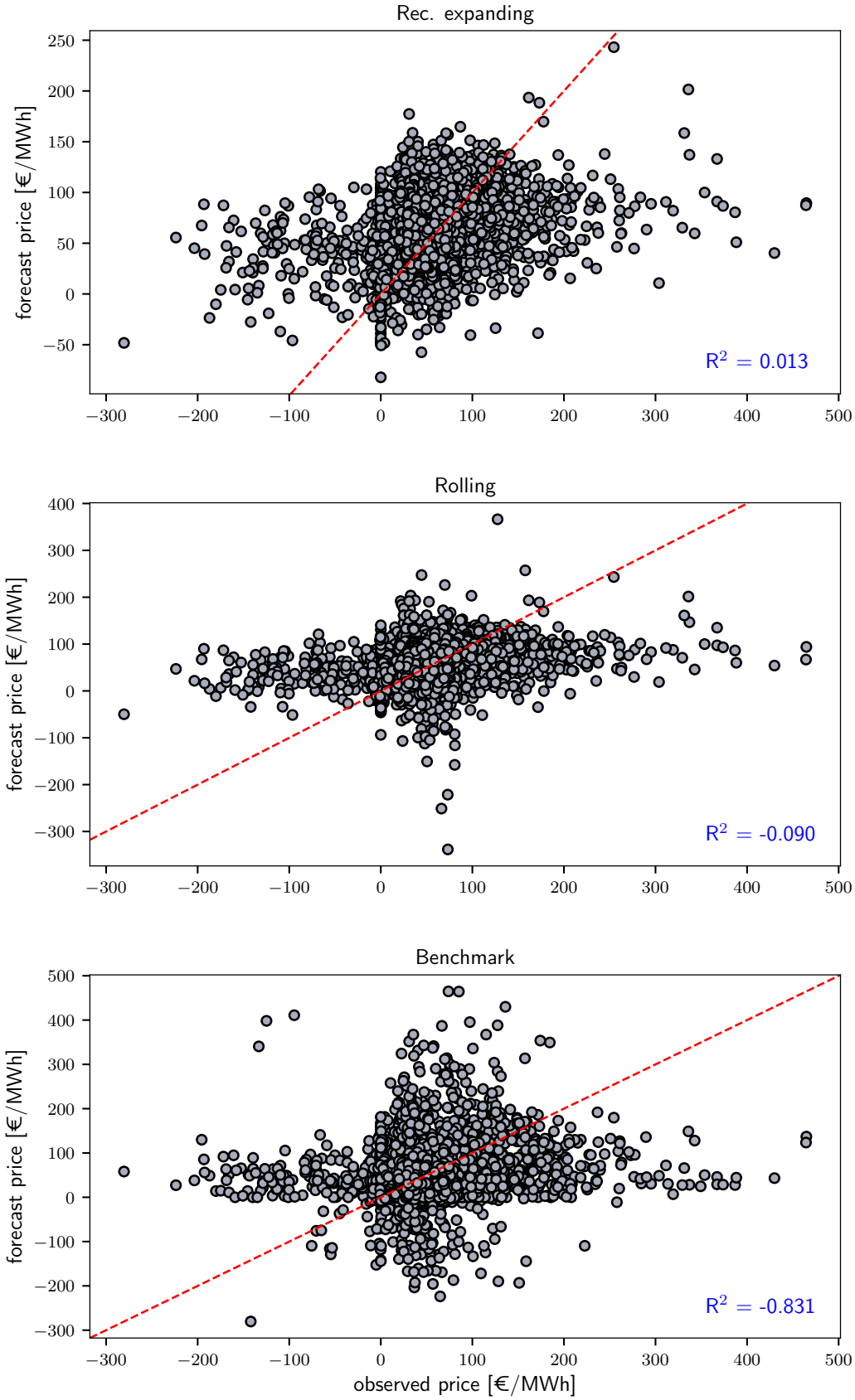


Figure 5.28: Scatter plot of observed vs forecast imbalance prices for each forecasting method. The coefficient of determination of each model is reported in blue. The red line corresponds to the perfect forecast.

Model	bias	MAE	RMSE	MASE	RMSSE
Rec. expanding	–9.12	39.57	55.23	0.66	0.90
Rolling	–5.75	41.04	58.02	0.68	0.93
Benchmark	–2.83	51.66	75.17	0.86	1.21

Table 5.8: Evaluation scores of the BM price point forecasts. All scores are expressed in €/MWh, except for MASE and RMSSE in per unit.

separately for each half-hour period, and the results are shown in Figures 5.29 and 5.30, respectively. The graphs confirm the two advanced methods perform similarly and both better than the benchmark across all periods. Moreover, errors from all three methods tend to slightly increase with the forecast horizon. Finally, the MAE and RMSE are calculated separately for each day of the week and results are shown in Figures 5.31 and 5.32. Again, the "Rec. expanding" and the "Rolling" methods perform consistently better than the benchmark. This time, forecasts perform more homogeneously across different days, although errors are still higher on the weekend.

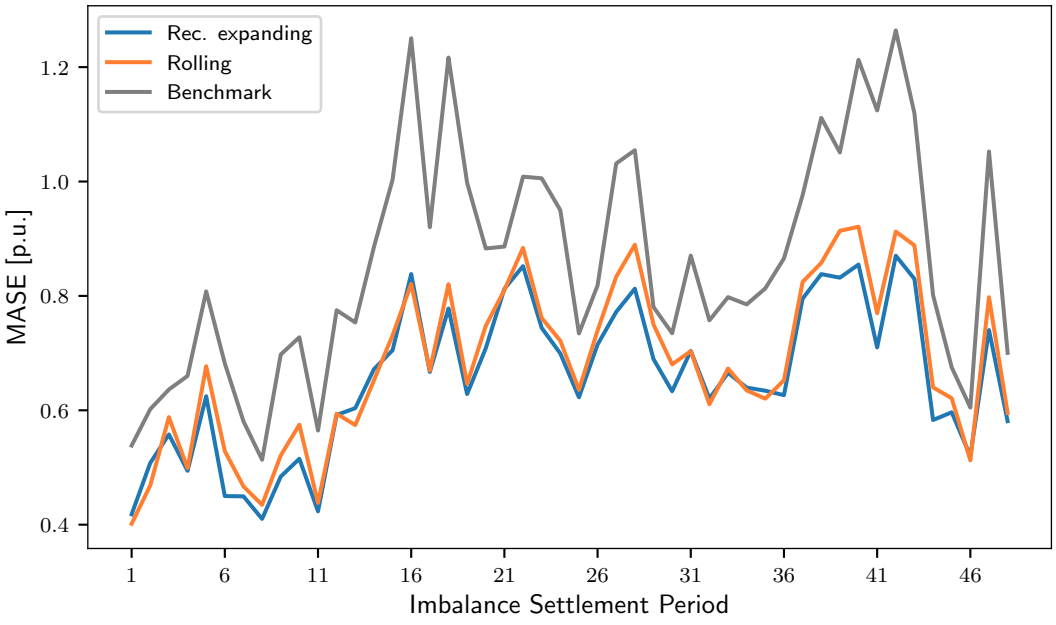


Figure 5.29: Mean Absolute Scaled Error of BM price forecasts for each half-hour.

The MAE of the two methods is calculated on a moving-window basis to investigate how the accuracy has changed over the test period. Figure 5.33 shows

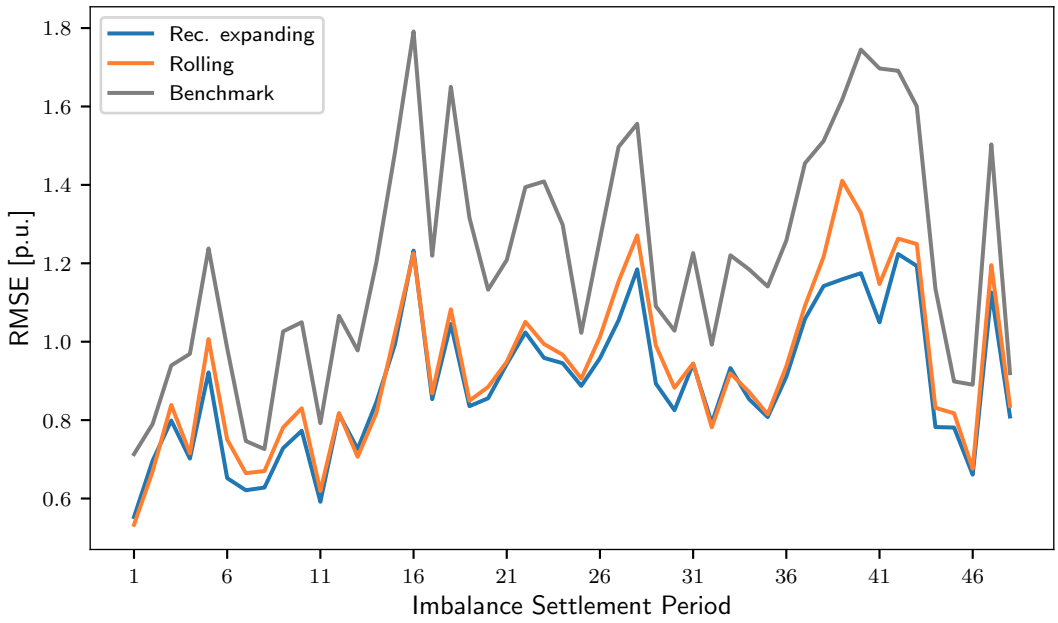


Figure 5.30: Root Mean Square Scaled Error of BM price forecasts for each half-hour.

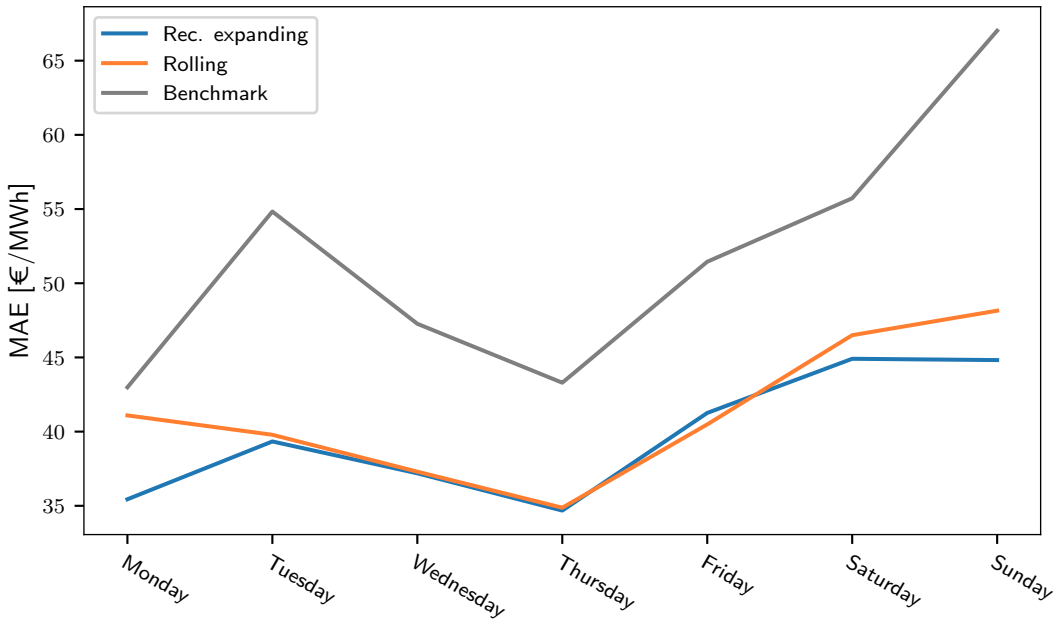


Figure 5.31: Mean Absolute Error of BM price forecasts for each day of the week.

the cumulative MAE of the two advanced forecasts, and Figure 5.34 the two-week rolling MAE. These graphs show that the "Rec. expanding" approach is consistently more accurate than the "Rolling" approach for the whole test period. It can also be noted that the accuracy of both methods improves over time, mainly as a

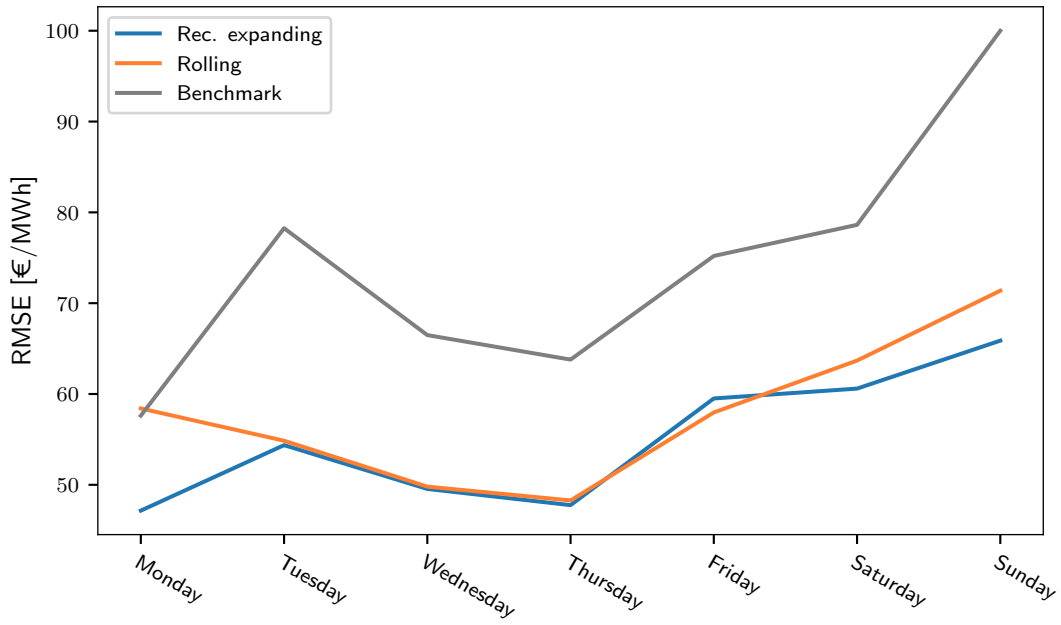


Figure 5.32: Root Mean Square Error of BM price forecasts for each day of the week.

consequence of the decreasing trend of average prices (cf. Figure (2.5)) and of their volatility (cf. Figure (2.6)) over these months.

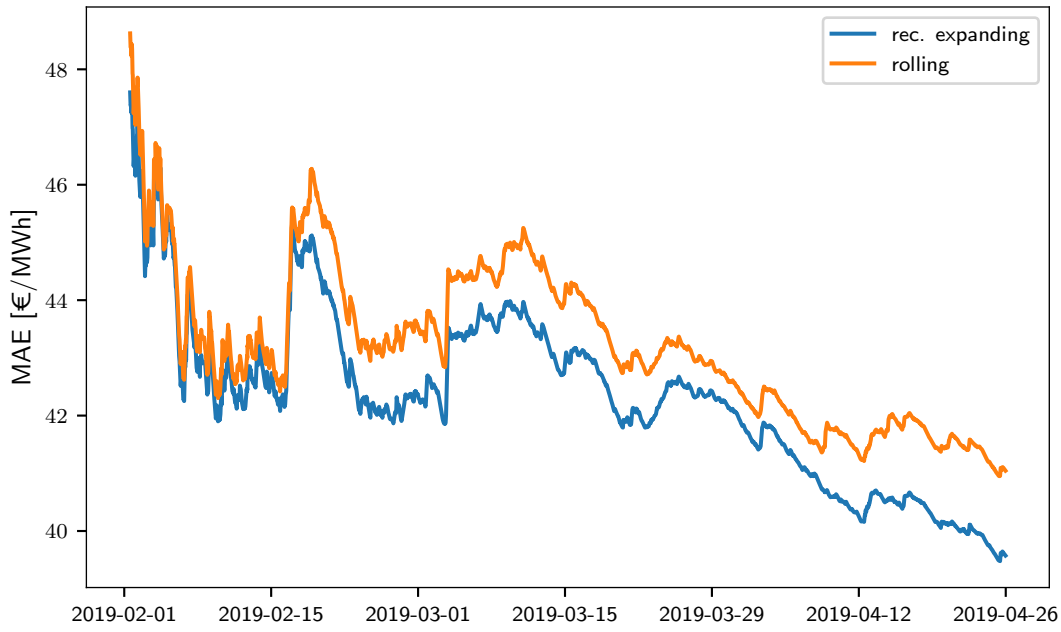


Figure 5.33: Cumulative Mean Absolute Error of BM price forecasts.

Following the same framework implemented for spot prices, the analysis on point forecasts of imbalance prices is concluded by applying DM tests to determine

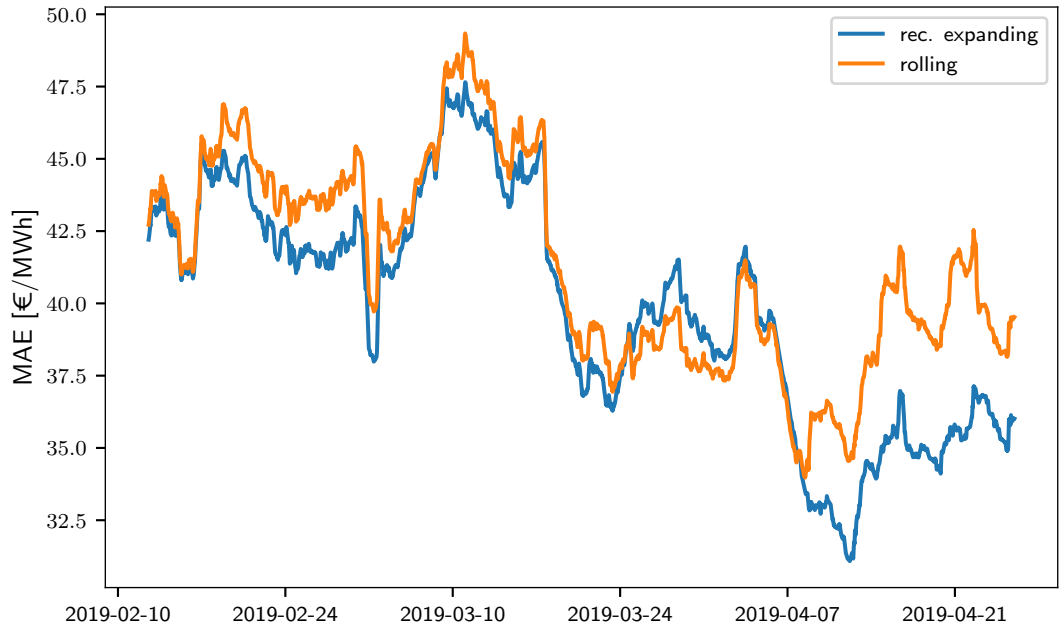


Figure 5.34: Two-week rolling Mean Absolute Error of BM price forecasts.

whether one of the two advanced methods is more accurate, having shown that both outperform the benchmark. The results presented in Table 5.9 show that the hypothesis of same accuracy is rejected in five periods out of 48, and in these periods p-values are always higher than the 5% significance level. This seems to indicate that the "Rec. expanding" method is slightly more accurate. However, given the small difference and the limited testing data set, it is not possible to state that this approach outperforms the "Rolling" approach. Therefore, it is concluded that the two advanced methods are indistinguishable for point predictions of BM prices..

Probabilistic forecasts

Next, the verification of the quantile forecasts is outlined. Figure 5.35 depicts the reliability diagram for the quantile forecasts from each method and empirical levels are tabulated in Table 5.10 , while the pinball loss scores are tabulated in Table 5.11. Both "Rec. expanding" and "Rolling" approaches are well calibrated for lower nominal levels, while they tend to overestimate the quantiles for higher nominal levels, with the "Rec. expanding" approach giving slightly larger biases.

hour	Rec. exp vs Roll	hour	Rec. exp vs Roll	hour	Rec. exp vs Roll
00:00	-	08:00	-	16:00	-
00:30	-	08:30	0.995469	16:30	-
01:00	-	09:00	-	17:00	-
01:30	-	09:30	0.978362	17:30	-
02:00	-	10:00	-	18:00	-
02:30	0.997108	10:30	-	18:30	-
03:00	-	11:00	-	19:00	-
03:30	-	11:30	-	19:30	-
04:00	-	12:00	-	20:00	-
04:30	0.998211	12:30	-	20:30	-
05:00	-	13:00	0.981038	21:00	-
05:30	-	13:30	-	21:30	-
06:00	-	14:00	-	22:00	-
06:30	-	14:30	-	22:30	-
07:00	-	15:00	-	23:00	-
07:30	-	15:30	-	23:30	-

Table 5.9: p-values returned by the DM tests applied to "Rec. expanding" models versus "Rolling" ones for BM price forecasting. The null hypothesis is that the first model is significantly more accurate than the second.

The benchmark, on the other hand, shows larger deviations in reliability and a tendency to over-forecast across all quantiles.

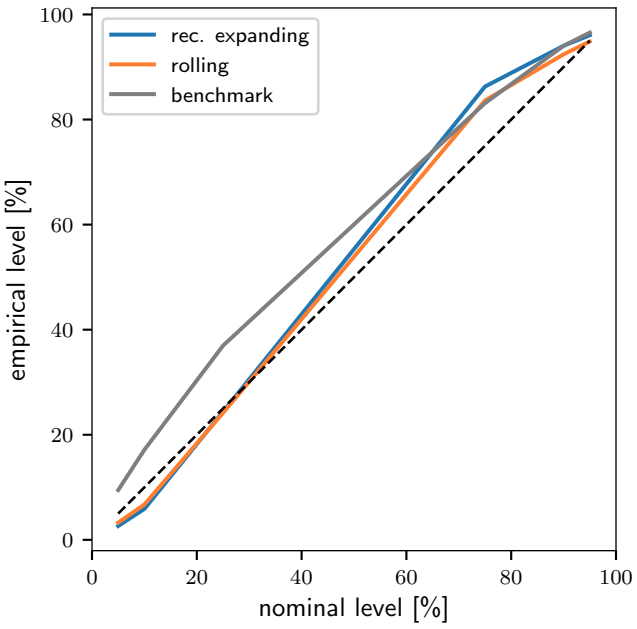


Figure 5.35: Reliability diagram of the BM price quantile forecasts. The dashed line is the diagonal indicating a perfectly calibrated forecast.

	nominal level					
Model	5	10	25	75	90	95
Rec. expanding	2.6	5.9	24.4	86.3	94.0	96.0
Rolling	3.3	6.7	24.2	83.6	92.4	94.9
Benchmark	9.5	17.1	37.0	83.1	94.0	96.5

Table 5.10: Empirical levels of the quantile forecasts of BM prices for each model.

Similarly to the case of DAM price forecasts discussed in the previous Section, the models' calibration could be improved in future work through statistical post-processing techniques. Adaptive bias-correction procedures can be implemented to reduce the probabilistic bias of the higher quantiles. As the data set length increases and past forecast performance information become available, MOS techniques can also be applied to produce improved forecasts. Smoothing procedures, on the other hand, may not be appropriate here as imbalance prices were shown to be extremely volatile.

In terms of sharpness, shown in Figure 5.36, the benchmark is the forecast providing the sharpest intervals, with the two advanced methods having similar but always wider intervals. This is also clearly visualised in Figure 5.37, where the sharpness values are calculated for each trading period (half-hour) separately. Similarly to what was observed for DAM prices, the intervals tend to be larger during the evening hours, reflecting the fact that uncertainties are larger with longer forecast horizons. In addition, comparing the sharpness of BM price quantile forecasts to that of DAM price (cf. Figure 5.22), it can be noted how the uncertainty

	nominal level					
Model	5	10	25	75	90	95
Rec. expanding	5.80	8.98	14.85	18.03	11.29	7.16
Rolling	6.09	9.47	15.71	18.39	11.53	7.40
Benchmark	6.44	9.38	15.02	17.28	11.43	7.58

Table 5.11: Pinball loss scores of the quantile forecasts of BM prices for each model.

associated with predictions of imbalance prices is considerably higher.

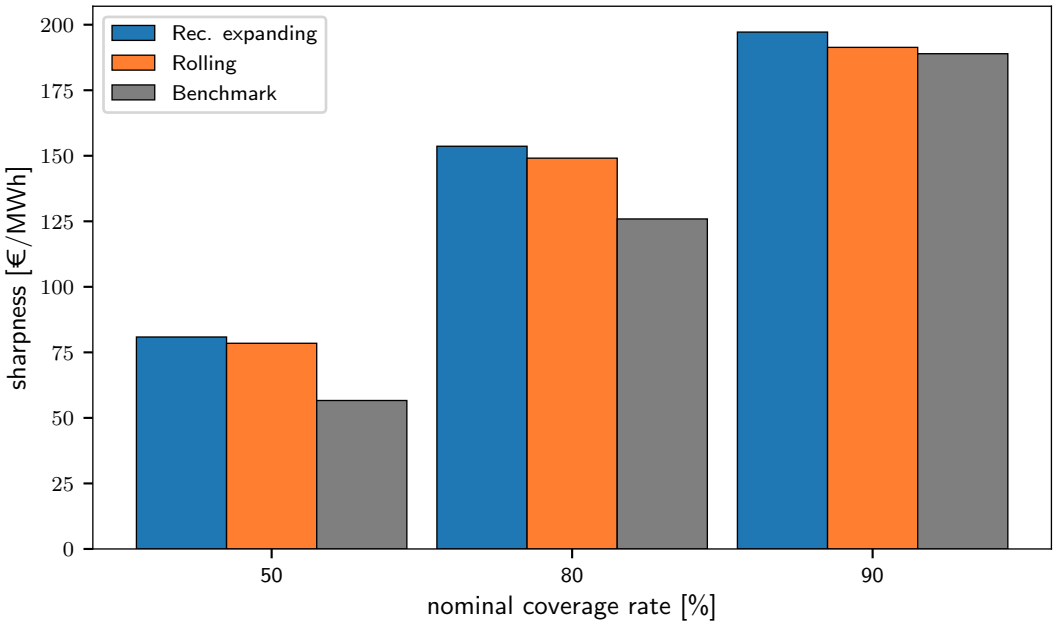


Figure 5.36: Sharpness of the BM price quantile forecasts.

Again, the benchmark can be discarded due to its poor calibration and the overall pinball loss score is calculated for the two advanced methods. This is equal to 11.020 for the "Rec. expanding" approach and to 11.431 for "Rolling". Similarly to the case of DAM prices, the difference in pinball loss score between the two methods is minimal and it is not possible to conclude which advanced approach provides better quantile forecasts.

Both advanced methods for predicting BM prices outperform the benchmark, but the difference in performance between them is very small, not allowing to conclude that either is superior. Nevertheless, the "Rec. expanding" approach is found to generate slightly more accurate forecasts over the testing period, with the advantage being more evident with regards to point predictions. This finding suggests that, given the particularly erratic behaviour of imbalance prices in the Irish market, it was more beneficial to use as much historical data as possible to train the models. Indeed, since the autoregressive and moving average components of the imbalance price process are very weak, most ARIMAX models are reduced

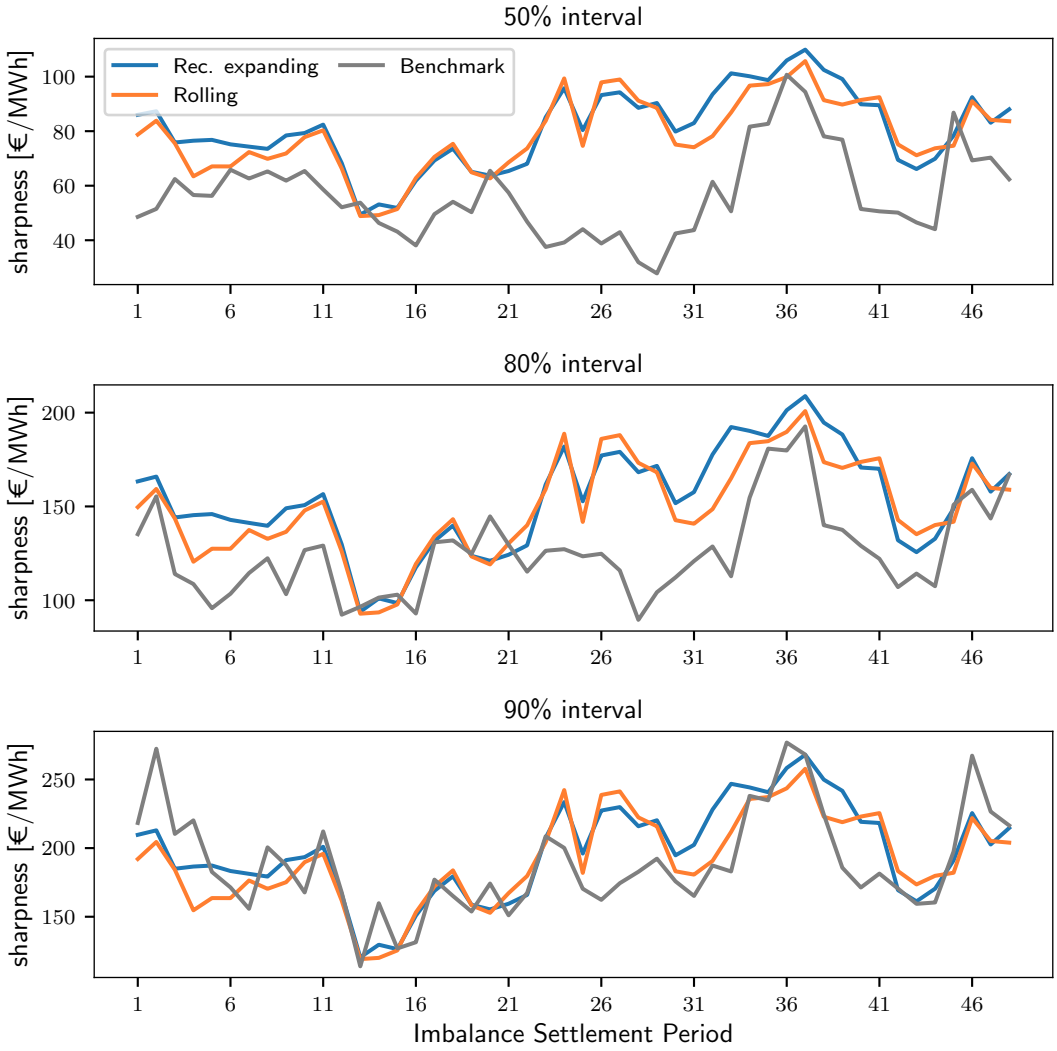


Figure 5.37: Half-hourly sharpness of the BM price quantile forecasts.

to multiple linear regression models and in such circumstances, the use of a larger training data set was advantageous. In conclusion, the "Rec. expanding" approach is selected as the forecasting method for BM prices and its predictions are used in the Monte Carlo simulations, whose results are presented in the next Section.

5.2.3 Monte Carlo Simulations

Having generated the forecasts of DAM and BM prices, Monte Carlo simulations are run to obtain the distribution of the price difference π^{diff} (cf. Equation (2.1)) for each trading period, as explained in Section 4.3.1.3. An example of the process is displayed in Figure 5.38, with the forecast quantiles of DAM and BM prices and

the quadratic interpolated functions shown in the top panel, and the distribution of the 10,000 randomly sampled values of price difference in the bottom panel. At this point, it is worth recalling that the variable of interest for the strategies developed in this work is the *sign* of π^{diff}, rather than its value. For each trading period, the forecast probability of positive price difference $\hat{\delta}$ (cf. Equation (4.20)) is obtained as the fraction of positive values in the distribution. For example, a value of $\hat{\delta}_{t+h}$ equal to 0.7 means that 70% of the values in the price difference distribution are larger than zero, hence there is a 70% expected probability that the price difference on trading period $t+h$ will be positive. It also follows that the value of 0.5 corresponds to equal probabilities (50%) of the price difference being positive or negative, indicating that the system is expected to be in balance.

The distribution of the probabilities $\hat{\delta}$ over the period under study is presented in Figure 5.39, which shows a tendency to have more probabilities above the 0.5 threshold, meaning positive price differences are expected more often than negative differences. The probability distributions are also analysed separately for each half-hour. The boxplots in Figure 5.40 indicate that values are quite evenly distributed around 0.5 for the early morning periods, while higher values of $\hat{\delta}$ become more frequent (in other words, positive price differences are expected more often) through the afternoon and in the evening hours. This is visualised in Figure 5.41, where the mean value of each half-hourly distribution is plotted.

The performance of the price difference forecast is compared to that of the "similar-day" benchmark method already discussed in the previous two Sections. The ROC curves for the two methods are illustrated in Figure 5.42, together with the diagonal line representing the performance of a random forecast. The graph illustrates the trade-off between hit (true-positive) rates and false alarm (false-positive) rates for each forecast. The Monte Carlo approach is considered as a probabilistic classifier since its output is a strict probability. The associated curve, which is generated from a finite but large set of instances, is in fact an interpolated step function. On the other hand, the benchmark method is a discrete classifier

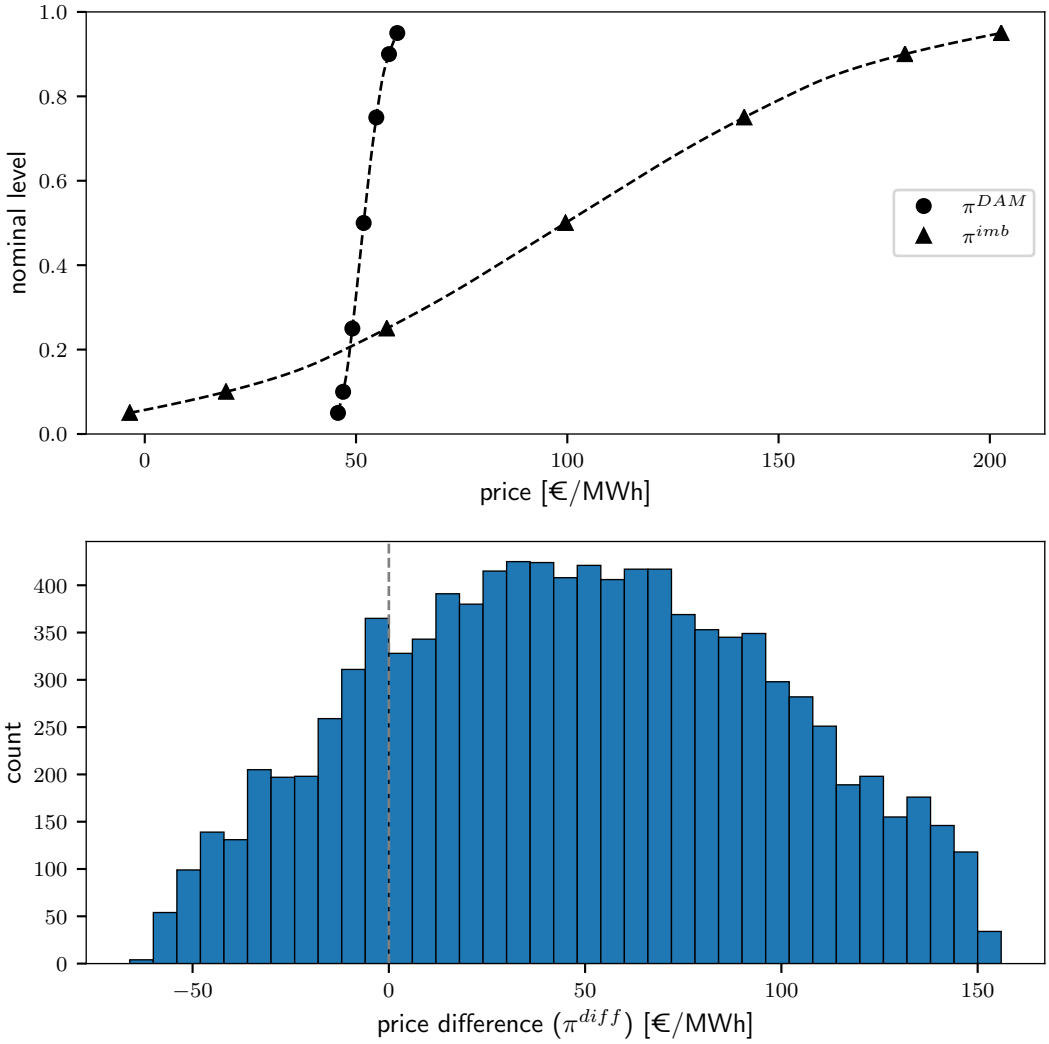


Figure 5.38: Example of the quadratic spline functions interpolated to the forecast quantiles of DAM and BM prices (top panel) and of the price difference distribution obtained from the Monte Carlo simulations (bottom panel) for a sample trading period (2019-03-12T20:00).

because the outputs of the point forecast are binary values, that is, if the price difference is positive or negative. Therefore, only one pair of true-positive and false-positive rates is computed and a single point is produced in the ROC space. The advanced method is more skillful, as its curve consistently lies above and to the left of the benchmark. The values of AUC are 0.557 for the Monte Carlo curve and 0.529 for the benchmark, confirming the superior performance of the first method. From the graph, it can also be observed that the Monte Carlo curve departs further from the diagonal in the lower left-hand side. This indicates that the method performs better in the more conservative region of the graph, i.e. when

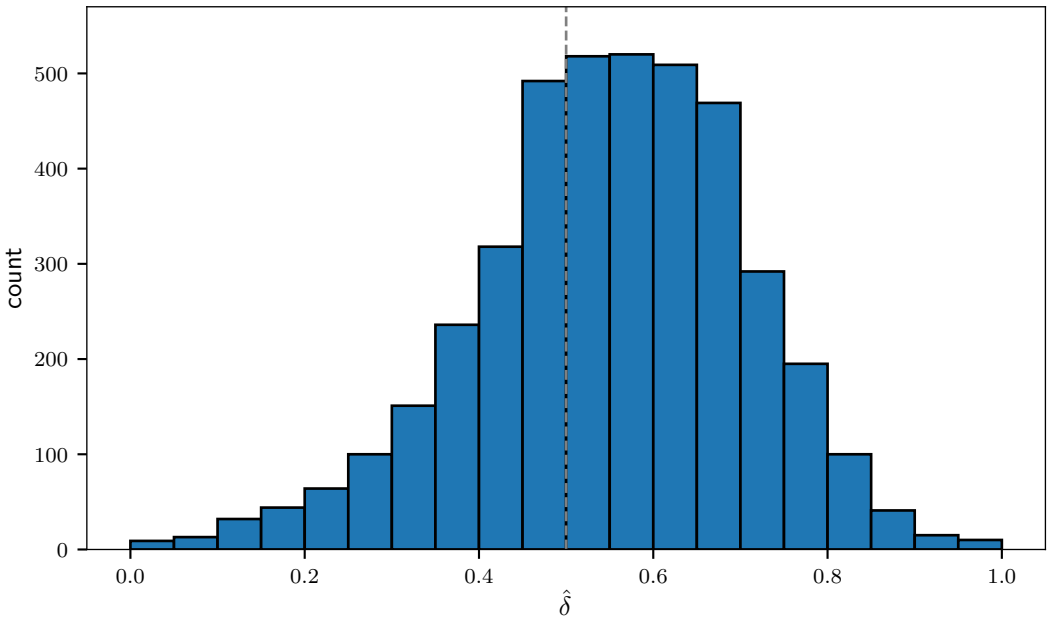


Figure 5.39: Distribution of forecast probabilities of positive price difference, $\hat{\delta}$.

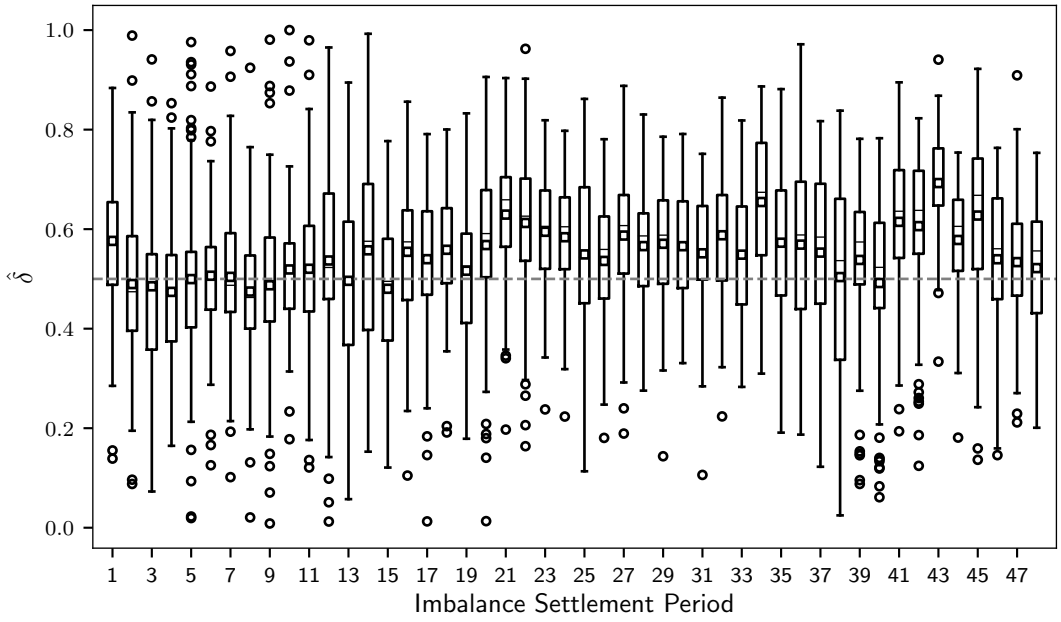


Figure 5.40: Boxplots of forecast probability of positive price difference, $\hat{\delta}$, for each half hour. The squares indicate the mean of the distribution.

strong evidence is needed to make a positive classification.

The accuracy of the forecast in terms of correct predictions (hits) of the price difference sign is then calculated. Note that in this context, the term "accuracy" refers to the metric calculated as $\frac{\text{true positives + true negatives}}{\text{total cases}}$ [65]. Overall, the advanced

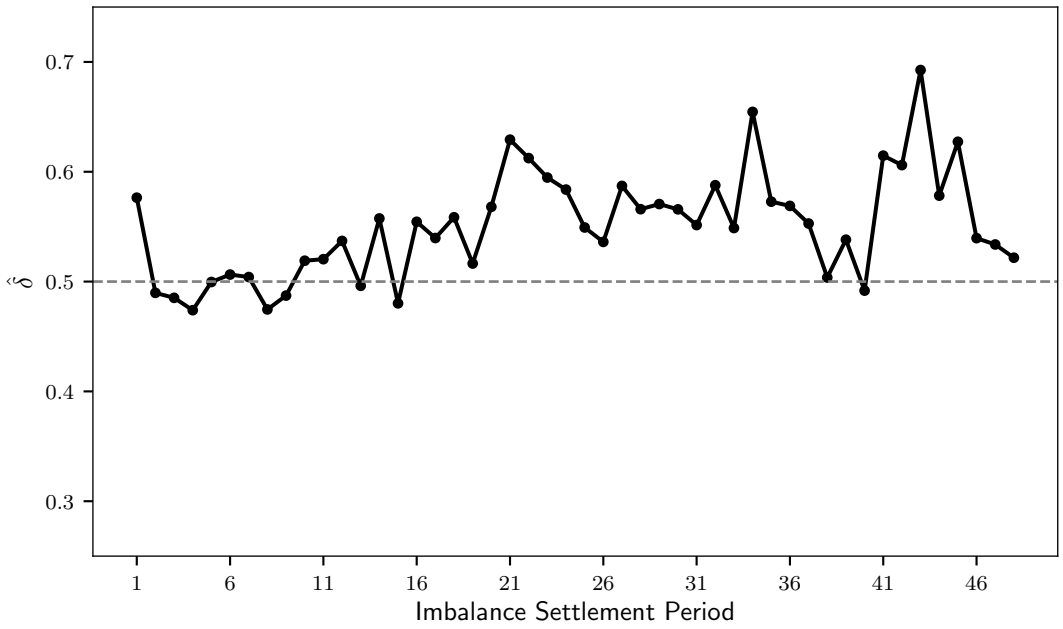


Figure 5.41: Average forecast probability of positive price difference for each half hour.

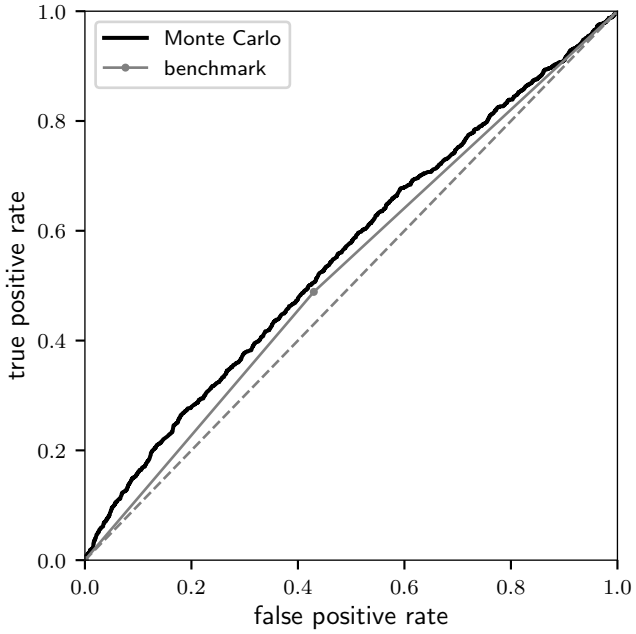


Figure 5.42: ROC curves for the price difference forecasts. The diagonal dashed line represents the performance of a random forecast.

method predicts the sign of price difference correctly 56.5% of the time, while the benchmark only 51.8%. The accuracy of the advanced method for each half-hour is illustrated in Figure 5.43. Predictions tend to be more accurate in the central part

of the day, although there is no clear hourly trend. Note that, for four periods, the accuracy is lower than that of a random guess, that is 50%.

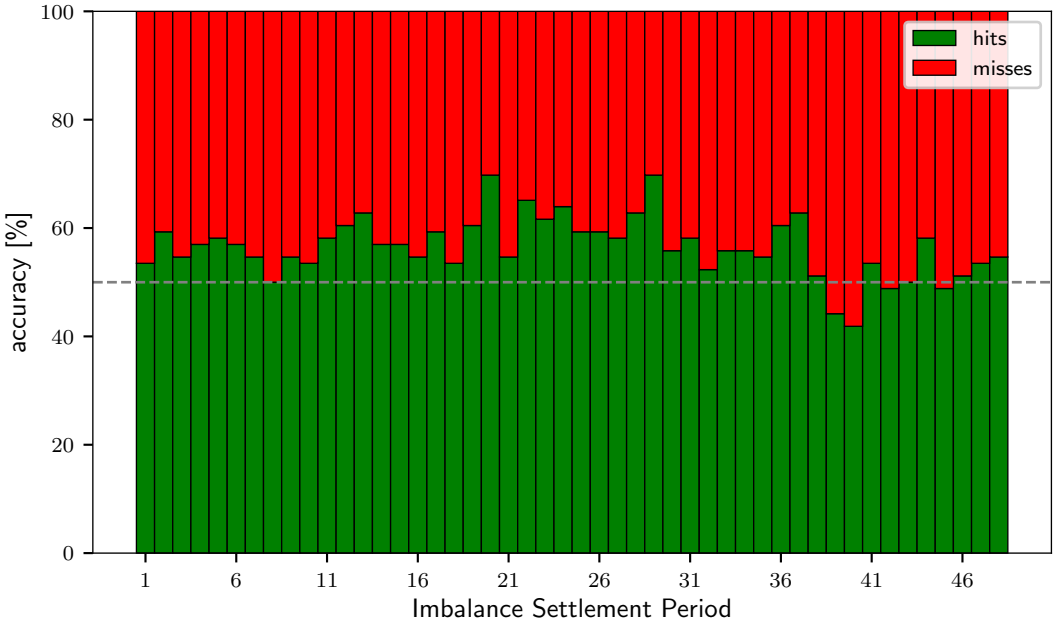


Figure 5.43: Half-hourly accuracy of the price difference forecast.

5.3 Imbalance Sign Forecast

The purpose of this Section is to evaluate the imbalance sign (i.e. the direction of system imbalance) forecasting methods developed in Section 4.3.2. The logistic regression models were trained over a four-month period from the 1st October 2018 to the 28th January 2019 and are evaluated on the 3-month test period from the 29th January 2019 to the 25th April 2019. The forecasting processes were implemented using `R` and run on an Intel Core i7 processor with 3.4 GHz CPU and 16 GB memory, with the total processing time being under 15 minutes.

Recall that four alternative approaches for the prediction of the sign of the system imbalance are developed:

- recursively expanding with model re-estimation ("Rec");
- recursively expanding with stepwise model re-selection and re-estimation ("Rec-Step");

- rolling with model re-estimation ("Roll"); and

- rolling with stepwise model re-selection and re-estimation ("Roll-Step").

The benchmark is given by the half-hourly historic proportion of the system being short calculated on the training set.

Moreover, recalling that the imbalance sign is described as a binary process, 0 corresponds to a negative sign and 1 to a positive sign. For each trading period $t + h$, the models predict the probability $\hat{\gamma}_{t+h} \in [0, 1]$ of the imbalance being positive (i.e. the system being short). Therefore, a value of 1 (0) indicates a perfectly confident forecast that the imbalance will be positive (negative), i.e. the system will be short (long). It also follows that a value of 0.5 corresponds to equal probabilities of the imbalance being positive or negative, that is, the system is expected to be in balance. An example of forecast probabilities generated for one day and corresponding observations is shown in Figure 5.44. Recall that a positive imbalance sign means that the system is short and there is need for up-regulation. Normally, this leads to the imbalance price being higher than the spot price, hence to a positive price difference. Therefore, it is postulated that forecasting a positive imbalance sign corresponds to forecasting a positive price difference (see Section 4.3.2).

The distribution of the probabilities $\hat{\gamma}$ generated by each logistic regression model is shown in Figure 5.45. It can be observed that the average predictions from the two "Rec"-based approaches are below 0.5, while the opposite is true for the two "Roll"-based approaches. The distributions from the two approaches with stepwise re-selection of the predictors, namely "Rec-Step" and "Roll-Step", are tighter, as indicated by the narrower inter-quantile ranges. With most values lying in the central region of the probability spectrum (i.e. around 0.5), this also shows that these predictions are more prudent. An in-depth investigation aimed at identifying the reason(s) for these differences in the distributions is beyond the scope of this research and is left for future work (see Section 6.2).

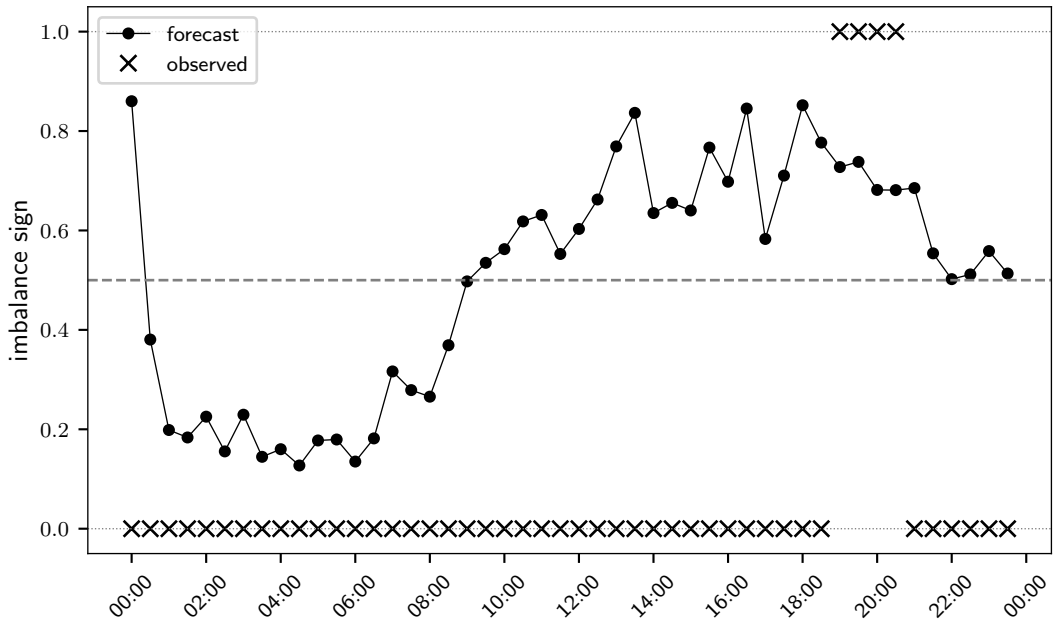


Figure 5.44: Example of forecast and observed imbalance signs for each trading period in one day (2019-02-23).

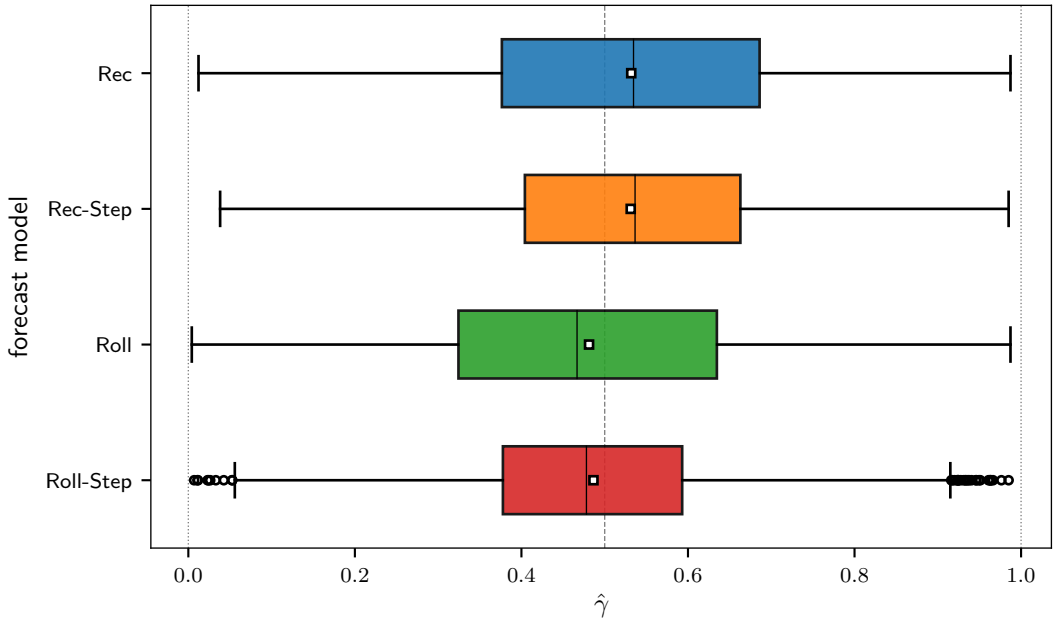


Figure 5.45: Boxplots of the forecast probabilities of positive imbalance, $\hat{\gamma}$, for each model. The squares indicate the mean of the distribution.

The performance of the models is first evaluated by examining their ROC curves, which are shown in Figure 5.46. All four logistic regression models consistently outperform the benchmark, whose performance resembles that of random

guessing. Brier scores are then calculated for each model — the observation o_i is equal to 1 if the imbalance is positive, and 0 otherwise (cf. Equation (3.30)) — and the results are presented in Table 5.12. These results confirm that all advanced models perform better than the benchmark. Moreover, the scores indicate that forecasts from the "Rec" and "Roll" approaches improve when coupled with stepwise re-selection of the predictors. Although the scores are very close, forecast errors from the advanced methods and the benchmark are differently distributed, as visible from Figure 5.47, where the boxplots illustrate the distribution of the squared forecast errors (cf. Equation 3.30) from each model. It can be observed that while the mean of the distribution (i.e. the Brier score) is highest for the benchmark, its distribution is tighter and there are no large errors. Finally, the lowest score is achieved by the "Roll-Step" approach. The predictions generated by this model are examined in more detail in the remainder of this Section, and will be used to implement the trading strategies discussed in Section 4.4.

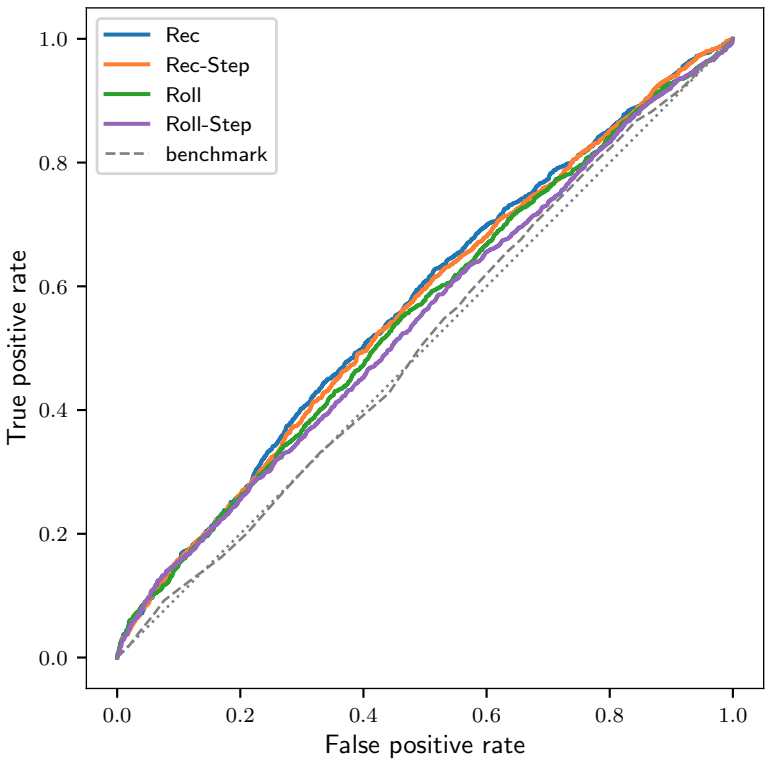


Figure 5.46: ROC curves for the imbalance sign forecasts. The diagonal dotted line represents the performance of a random forecast.

Model	Brier score
Rec	0.2720
Rec-Step	0.2680
Roll	0.2712
Roll-Step	0.2634
Benchmark	0.2758

Table 5.12: Brier scores of the imbalance sign forecasts.

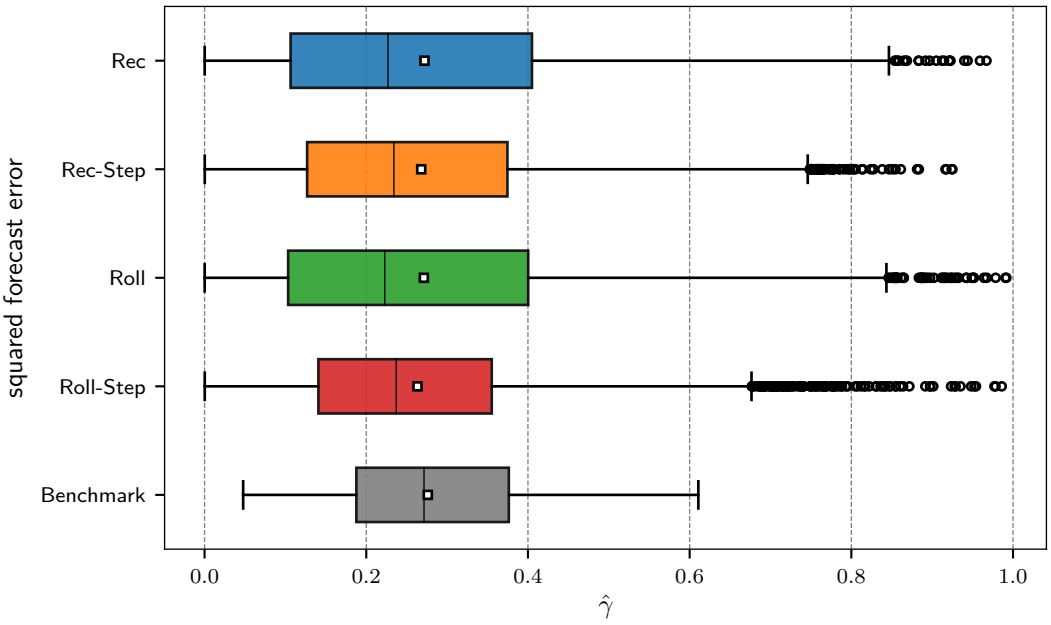


Figure 5.47: Boxplots of the squared errors for each imbalance sign forecasting model. The squares indicate the mean of the distribution, i.e. the Brier score.

As explained in Section 4.3.2, in the "Roll-Step" approach, the predictors are re-selected and models are re-estimated at each iteration. Figure 5.48 shows the occurrence of models with the same set of selected predictors, where:

- Demand = forecast demand;
- $\text{Wind}_{\text{point}}$ = point forecast of wind power generation;
- $\text{Wind}_{\text{interval}}$ = 80% interval forecast of wind power generation;
- Penetration = predicted wind power penetration; and
- Day = day of the week.

The results show that there is not one prevalent set of predictors. The two most frequent sets are {$\text{Wind}_{\text{point}}$, Demand} and {$\text{Wind}_{\text{point}}$, Demand, Penetration}, each accounting for 15.6% of the total. Interestingly, in 5.0% of the cases, all predictors are removed in the backward selection procedure. In this case (labelled with "1"), the prediction is simply given by $\hat{\gamma}_{t+h} = \frac{\exp(\hat{\beta}_0)}{1+\exp(\hat{\beta}_0)}$, where $\hat{\beta}_0$ is the estimated intercept in the model (cf. Equation (4.23)). Figure 5.49 shows how often a predictor is included in a model, by itself or in combination with other predictors. $\text{Wind}_{\text{point}}$ and Penetration are selected in more than half of the models (65.1% and 56.4%, respectively). On the other hand, Day is selected considerably less than the other predictors, appearing only in 6.5% of models.

The distribution of the probabilities $\hat{\gamma}$ generated by the "Roll-Step" model is displayed in Figure 5.50. The histogram highlights the presence of more values below the 0.5 threshold, indicating the tendency to expect the system to be long more often and therefore to have negative price differences more frequently. This trend is opposite to that found for the distribution of forecast price differences $\hat{\delta}$, as highlighted in Figure 5.51, where a comparison between the two distributions is offered. The distributions of forecast imbalance sign are also analysed for each half-hour separately in Figure 5.52. It can be observed that predictions are more cautious and distributions are tighter at night; forecasts are biased towards the negative imbalance sign (below 0.5) in the early morning, and towards the positive imbalance sign (above 0.5) in some evening periods.

The accuracy of the forecast in terms of correct predictions (hits) of the imbalance sign is then calculated. Overall, the model estimates the imbalance sign correctly in 53.5% of the cases, a result markedly better than the 46.0% obtained by the benchmark. The percentage of correct predictions for each half-hour is also illustrated in Figure 5.53. Predictions are more accurate in the morning and less in the evening and night, with a number of periods where accuracy is below 50%. However, some of the analyses performed on the first seven months of operation of the Irish market and presented in Section 2.2.2.3 have highlighted the considerable

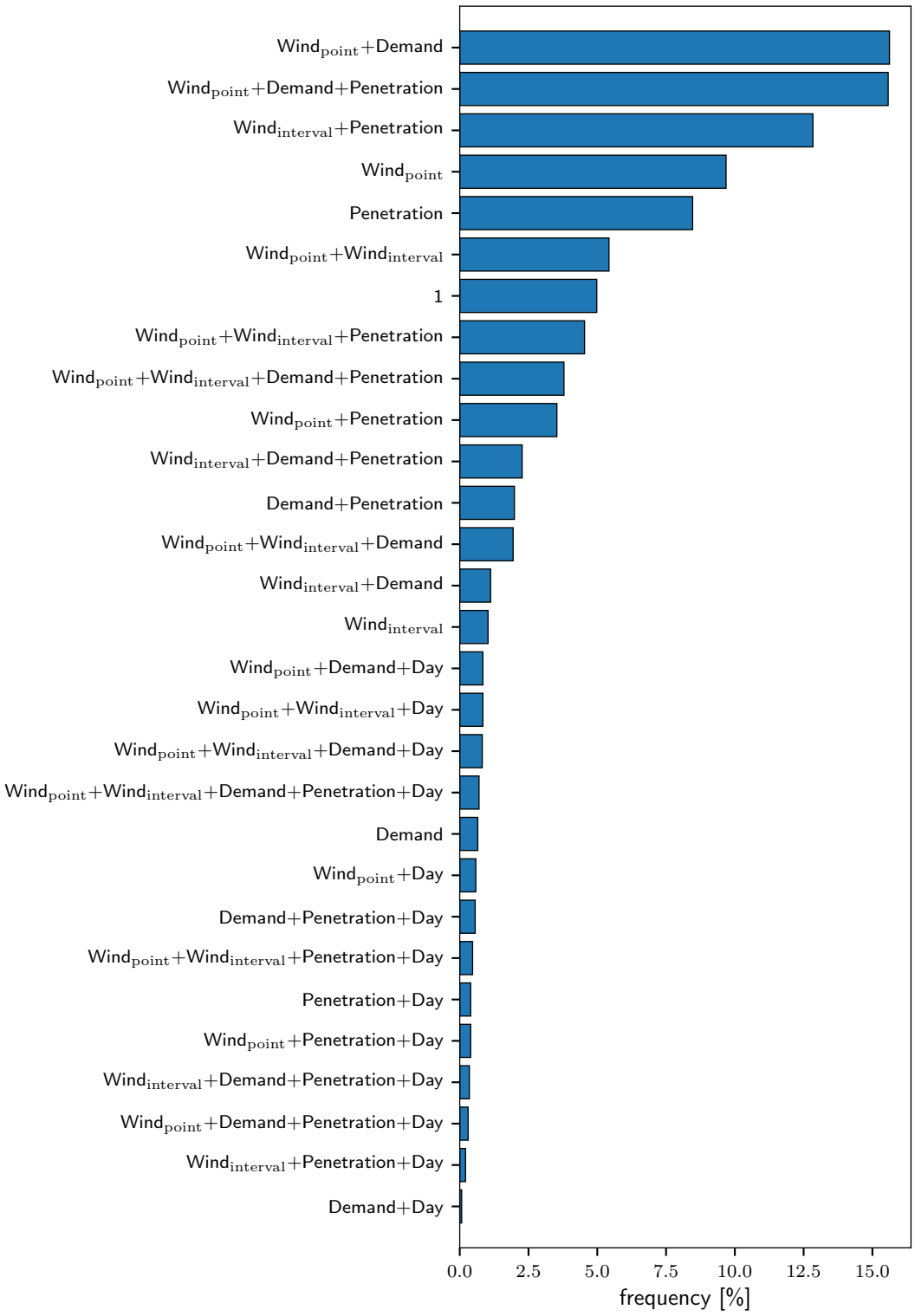


Figure 5.48: Frequency of imbalance sign forecasting models with the same set of predictors for the "Roll-Step" approach.

occurrence of counter-intuitive imbalance prices, whereby a positive imbalance sign does not always correspond to a positive price difference, and vice versa. Due to

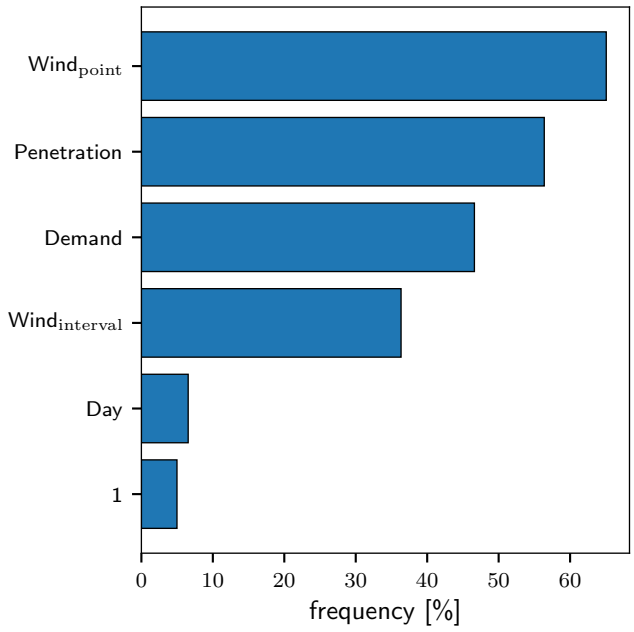


Figure 5.49: Frequency of predictors selected in the imbalance sign forecasting models for the "Roll-Step" approach.

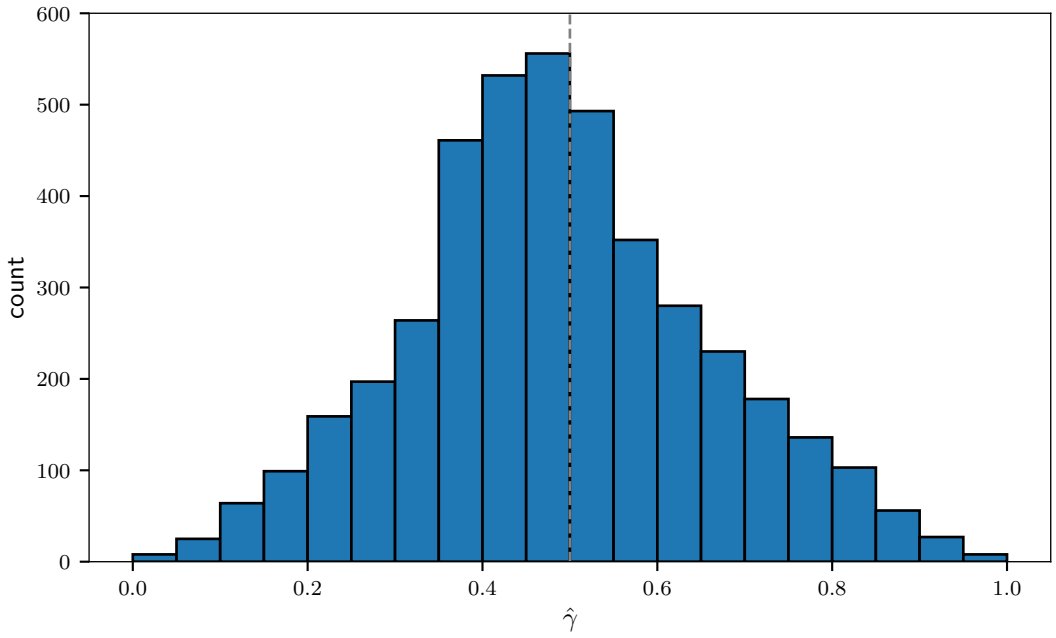


Figure 5.50: Distribution of forecast probabilities of positive imbalance sign, $\hat{\gamma}$, for the "Roll-Step" approach.

this anomaly, a correct prediction of the imbalance direction in this study does not necessarily correspond to a correct prediction of the price difference sign. For this reason, the accuracy of the imbalance sign forecast is also investigated in terms

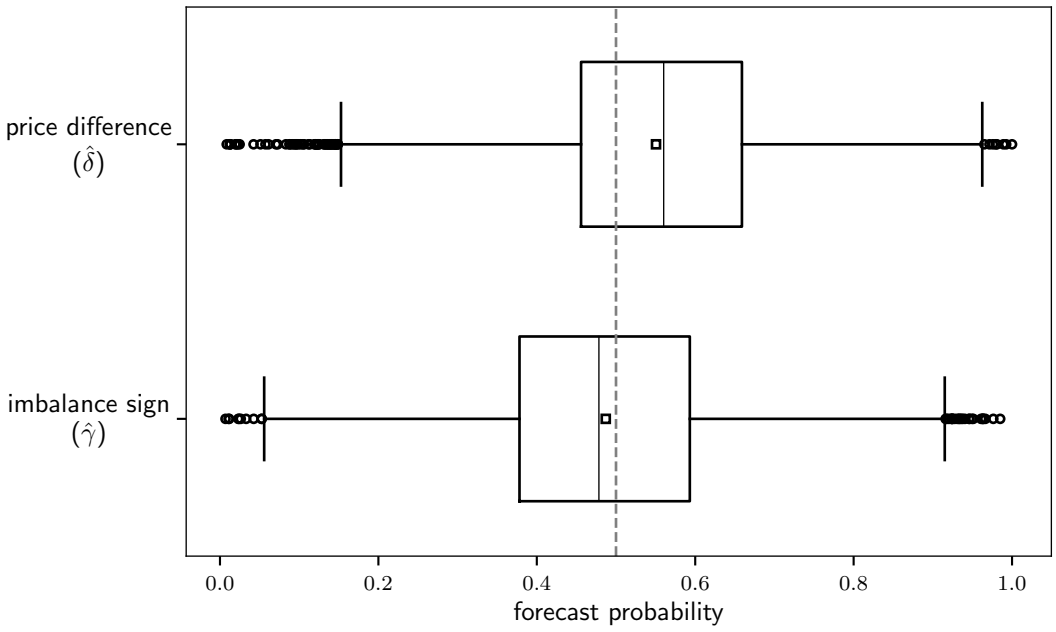


Figure 5.51: Distribution of forecast probabilities of positive price difference, $\hat{\delta}$, and of positive imbalance sign, $\hat{\gamma}$.

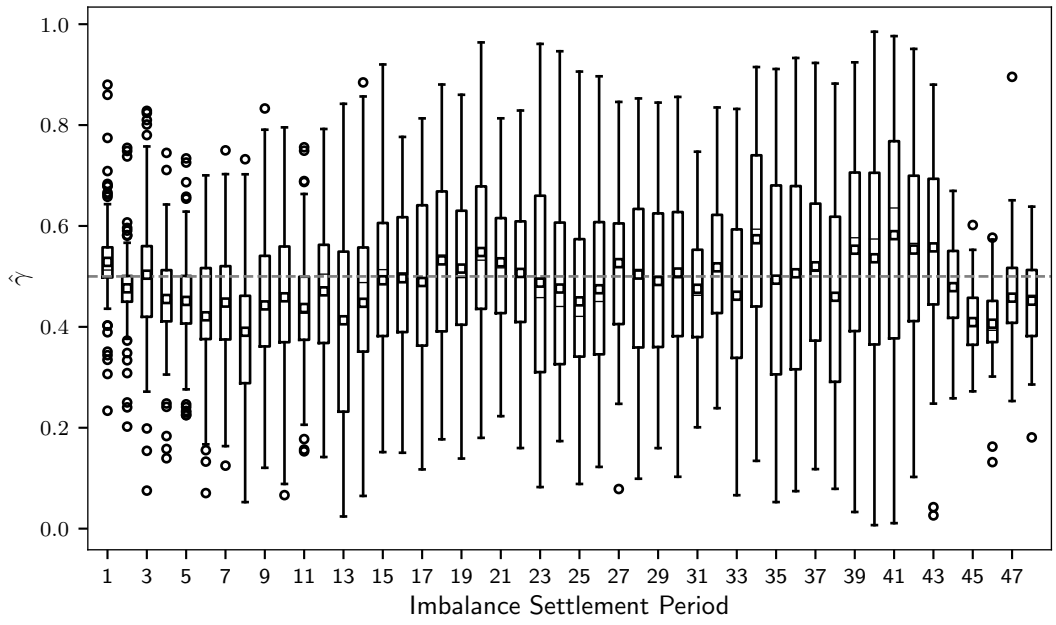


Figure 5.52: Boxplots of forecast probability of positive imbalance sign, $\hat{\gamma}$, for each half hour. The squares indicate the mean of the distribution.

of correct predictions of the price difference sign, and it is found that the accuracy degrades to 51.0%. The percentage of correct predictions of the sign of price difference for each half-hour is also illustrated in Figure 5.54. Compared to the

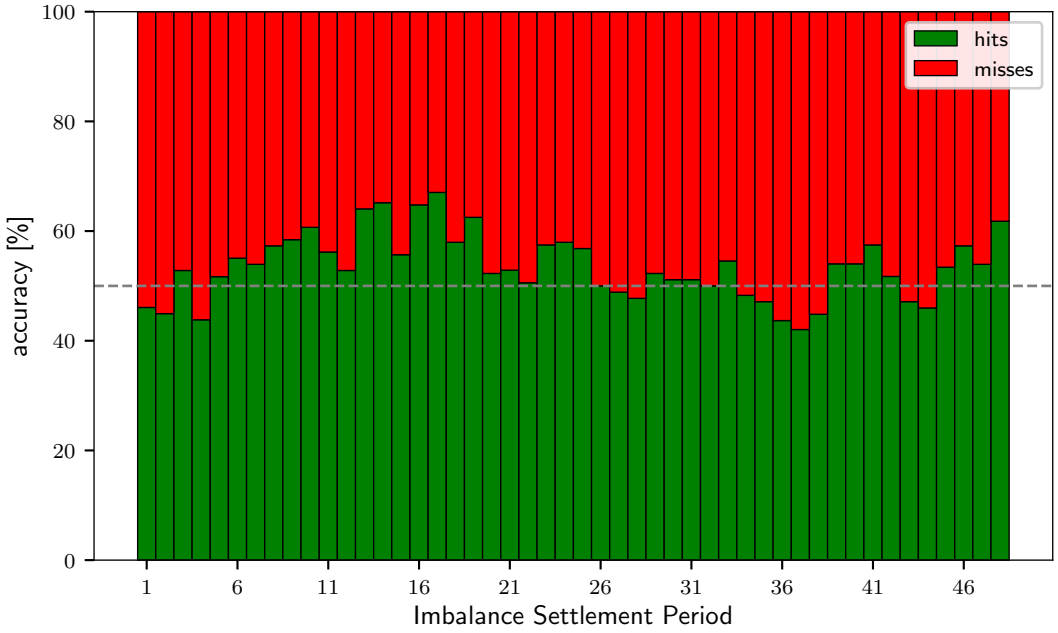


Figure 5.53: Half-hourly accuracy of the imbalance sign forecast in terms of correct predictions of the *imbalance sign*.

corresponding accuracy values in terms of imbalance sign (shown as dots), accuracies are lower across nearly all periods. Besides a few isolated periods during the day, accuracies are actually higher between ISP 37 and 43, where it was observed that price anomalies were slightly more frequent (cf. Figure 2.47).

5.4 Strategic Bidding

This Section presents the results from the application of the bidding strategies developed for trading wind energy in markets with a single imbalance pricing scheme and described in Section 4.4 . The strategies are aimed at participants who want to bid strategically in the day-ahead market to increase the value of the electricity generated at a wind farm, where the size of the wind farm is small in comparison to the average system imbalance volume (see the model assumptions in Section 4.1.1).

The models developed in this work employ forecast information on wind power generation and electricity market conditions in a probabilistic framework. Offers are then formulated to reflect the participant's appraisal of future uncertainties and

Figure 5.54: Half-hourly accuracy of the Imbalance Sign forecast in terms of correct predictions of the *price difference sign*. The dots correspond to the accuracy calculated in terms of imbalance sign shown in Figure 5.53.

risk appetite. The objective is to improve the value of wind power by increasing the revenue while controlling the exposure to risk. The proposed strategies are summarised in Table 5.13 together with the required types of forecasts. Both market quantity forecasts are tested through the implementation of the proposed strategies to quantify the value of forecast accuracy in monetary terms. The half-hourly accuracy of the two market quantity forecasts, already presented in Figures 5.43 and 5.54, is compared directly in Figure 5.55, from which it is evident that the price difference forecast $\hat{\delta}$ outperforms the imbalance sign forecast $\hat{\gamma}$ over most trading periods. In the following, when the price difference forecast $\hat{\delta}$ is used, the suffix "-price" is added to the strategy's name; when the imbalance sign forecast $\hat{\gamma}$ is used, then "-NIV" is added.

As previously discussed in Sections 4.4.3 and 4.4.4, the risk factor ρ and the adjustment function $f(x)$ in the Scal and Prop strategies need to be set by the participant beforehand, based on the target trade-off between risk and revenue. Here, results are provided for the range of risk factor values $\rho \in [0, 1]$ and examined

Strategy	Wind power forecast	Market quantity forecast
ImbMin	point	–
ImbMax	–	point
Scal	point	probabilistic
Prop	point	probabilistic
Quant	probabilistic	probabilistic

Table 5.13: Summary of the proposed bidding strategies and forecasts required for their implementation.

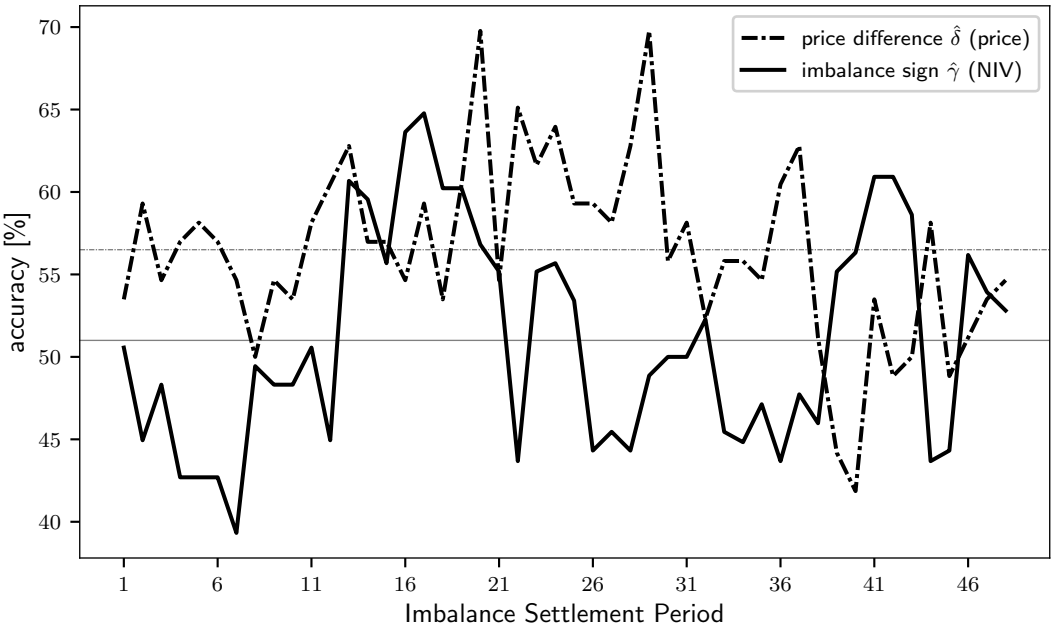


Figure 5.55: Accuracy in terms of correct predictions of the price difference sign of the two market quantity forecasts. The grey lines indicate the overall average accuracy of each method.

ex-post. On the other hand, since the adjustment function $f(x)$ can be designed heuristically in a number of ways to reflect the participant's sensitivity to market uncertainty, only the basic case of a step adjustment function (see Figure 4.11b) is selected for demonstration and presented here.

Based on an analogous rationale, the value of α, the function $g(x)$ and parameters a and b in the Quant strategy can be set in a variety of ways to reflect the desired trade-off between revenue and risk. Here, results are presented for the case of α being equal to 0.90 and the bid being a piecewise linear function of the market forecast quantity (see Figure 4.12b). While all possible combinations of parame-

ters a and b were tested, only the risk-minimising case (as measured by $\text{CVaR}_{1\%}$) is illustrated to reduce redundancy. The corresponding parameter values are:

- $a = 0.0$, $b = 0.8$ for the Quant-NIV strategy; and
- $a = 0.0$, $b = 0.9$ for the Quant-price strategy.

As discussed in Section 4.4, ImbMin and ImbMax will constitute the two benchmarks to examine the performance of the proposed novel strategies. To provide another useful reference, results are also presented for the hypothetical case where a perfect forecast of wind power generation is employed to bid in the market. This case, where no energy imbalances and therefore no balancing costs arise, will be referred to as "Perfect".

The proposed strategies are evaluated on the 3-month independent test period from the 29$^{\text{th}}$ January 2019 to the 25$^{\text{th}}$ April 2019 by simulating the participation of the wind farm under study in the I-SEM and using actual data from the market. During this period, the capacity factor of the wind farm was 31.2%. The average price difference was 31.13 €/MWh when the system was short and –26.05 €/MWh when it was long, indicating that on average price differences were not symmetric. The observed mean absolute price difference for each half-hour is illustrated in Figure 5.56, which shows the size of π^{diff} is on average lower in the morning, increases in the afternoon, and is largest in the evening hours. The system was left long more frequently than short, as shown in Table 5.14, with NIV being negative 49.6% of the time and positive 42.9%. As discussed in Chapter 2, a long system is expected to lead to a negative price difference, and vice versa for a short system. However, due to the significant occurrence of price anomalies already highlighted in Section 2.2.2.3, proportions are altered and price differences were negative 54.6% of the time and positive 43.9%.

The performance of the various bidding strategies is evaluated in terms of bid quantities, revenue, energy imbalances, $\text{CVaR}_{1\%}$ and $\text{CVaR}_{5\%}$. Recall that in all the analyses, wind power values and related quantities are normalised by the wind

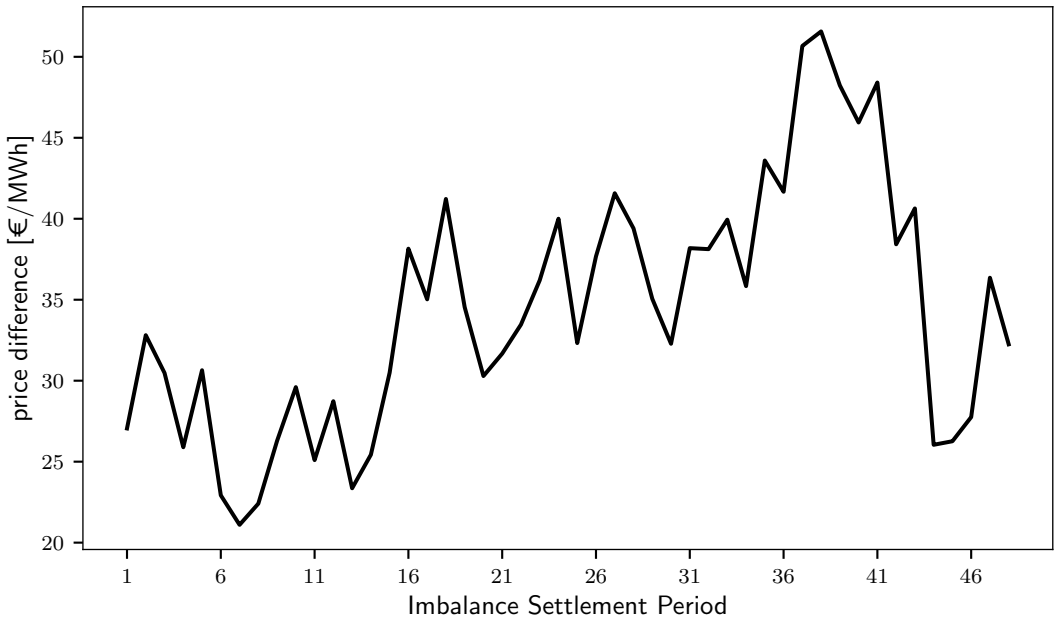


Figure 5.56: Half-hourly observed mean absolute price difference during the test period.

	Frequency		
Model	positive	negative	zero
NIV	42.9	49.6	7.3
π^{diff}	43.9	54.6	1.5

Table 5.14: Frequency of positive, negative and zero net imbalance volume (NIV) and price difference (π^{diff}) during the 3-month test period between the 29$^{\text{th}}$ January 2019 and the 25$^{\text{th}}$ April 2019. Values are expressed as percentages.

farm's installed capacity P_{nom}. Consequently, monetary results (i.e. revenue and CVaR_α) are expressed in € per installed MW. In the following graphs, results from ImbMin are indicated with a cross, while those from ImbMax-price and ImbMax-NIV are depicted as black dots. Note that where the variable of interest is plotted against the risk factor ρ, results for the two Quant strategies are shown outside the range of possible values of ρ since risk factor is not involved in their formulation.

First, the variation of bid quantities for each strategy is examined by looking at the average bids in Figure 5.57. For both Quant cases, less energy is offered on average compared to ImbMin. For the Scal and Prop strategies, the average bid increases monotonically for increasing values of ρ when the "-NIV" forecast is em-

ployed; indeed, the system is forecast long more often than short (see Figure 5.50) and therefore the tendency is to increase the energy forecast. Note that values are considerably higher for Scal-NIV due to the strategy's formulation.

On the other hand, when the "-price" forecast is employed, trends differ between Scal and Prop. Recalling that price difference is forecast positive more often than negative (see Figure 5.39), the tendency is to decrease the energy forecast. For Prop-price, this translates to the bid decreasing monotonically as ρ increases. For Scal-price, however, the average bid decreases for $\rho \leq 0.20$ and then increases for larger values. This happens as the average forecast energy is relatively low (30.5% of P_{nom}), so once the size of the adjustment becomes significant, negative corrections are limited at zero, whereas positive corrections have margin to be applied. This is visualised in Figure 5.58, where the distributions of bid quantities from the two approaches are compared for increasing values of ρ. In the Scal-price case, it can be noted that although the median consistently decreases as risk factor increases, the mean becomes larger.

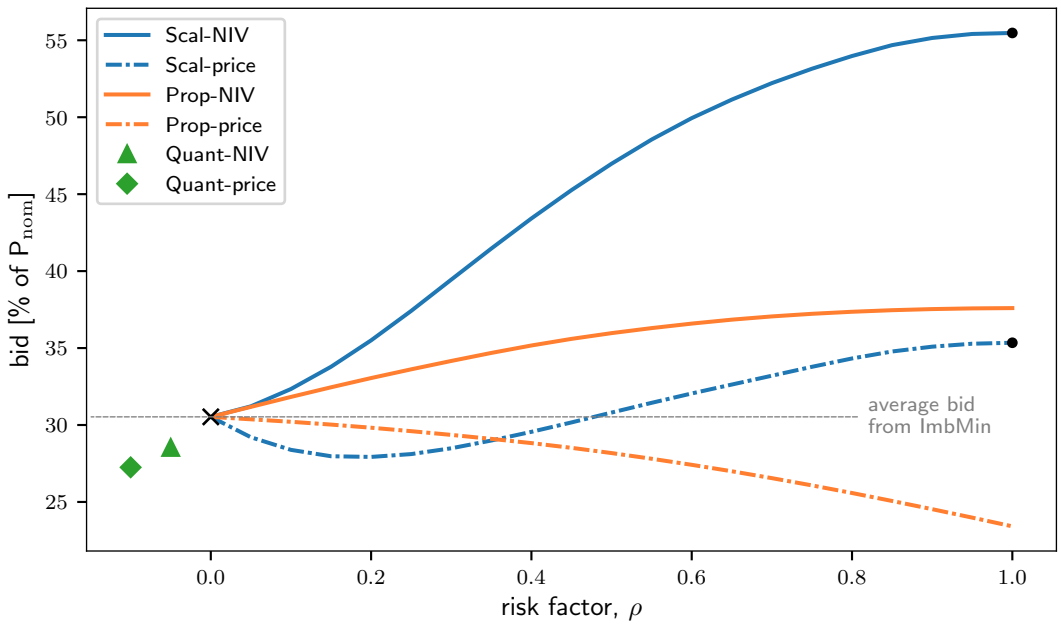


Figure 5.57: Average bid quantity from each bidding strategy. The cross and the dots represent the ImbMin and ImbMax strategies, respectively.

The total revenue yielded from each strategy is shown in Figure 5.59 and tabu-

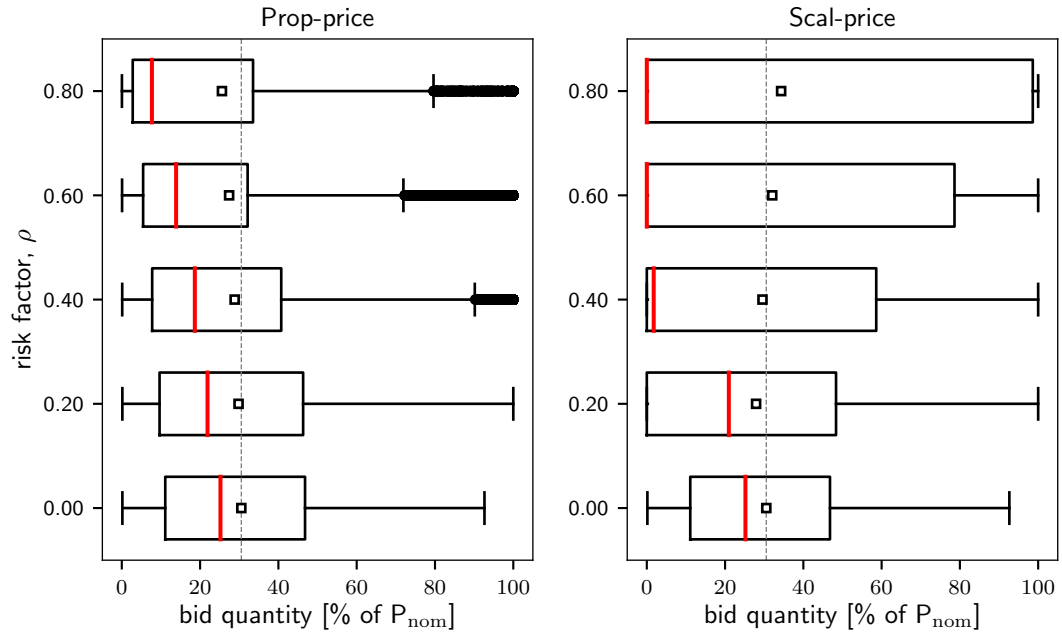


Figure 5.58: Distribution of bid quantities from Prop-price (left) and Scal-price (right) strategies for increasing risk factor. The red lines indicate the median of the distribution and the squares the mean, while the dashed grey line indicates the average forecast energy.

lated in Table 5.15. The revenue is lowest for ImbMin and increases for all the advanced strategies, with Scal-price consistently outperforming the other approaches for increasing values of ρ. It can be noted that for each strategy type, the "-price" case is always better than the "-NIV" homologous. This is particularly noticeable in the two Scal cases. For increasing values of the risk factor, both revenues initially increase, but values for Scal-NIV are consistently lower due to the lower accuracy of the market quantity forecast, which causes adjustments to be made in the wrong direction more often. For $\rho > 0.60$, as the size of the adjustments becomes substantial, these inaccuracies cause the total revenue to invert trend and decrease.

It can be noted how some strategies are able to yield a revenue higher than Perfect: this occurs for Scal-price when $\rho \geq 0.25$, Prop-price when $\rho \geq 0.65$, Prop-NIV when $\rho \geq 0.80$, and ImbMax-price. Note also that the highest revenue is not obtained by ImbMax-price (i.e. Scal-price$_{[\rho=1.00]}$), but by Scal-price$_{[\rho=0.90]}$ and is equal to € 31,713, which is 15.9% higher than ImbMin and 6.9% higher than

ImbMin	ImbMax-price	ImbMax-NIV	Quant-price	Quant-NIV	Perfect
27,352	31,693	29,128	28,019	27,692	29,673

ρ	Scal-price	Scal-NIV	Prop-price	Prop-NIV
0.20	29,376	28,630	28,150	27,972
0.40	30,488	29,322	28,954	28,644
0.60	31,299	29,620	29,618	29,225
0.80	31,660	29,375	30,159	29,692
1.00	31,693	29,128	30,674	30,087

Table 5.15: Revenue from each bidding strategy. Values are expressed in € per installed MW.

Perfect.

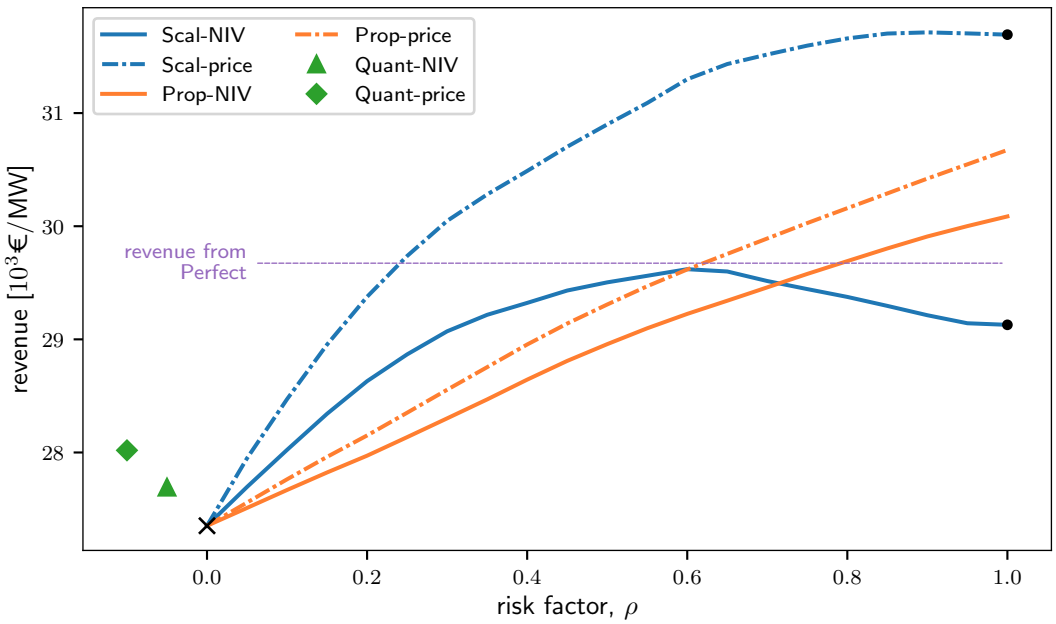


Figure 5.59: Revenue from each bidding strategy. The cross and the dots represent the ImbMin and ImbMax strategies, respectively.

The strategies are then evaluated in terms of the size of the energy imbalances they cause. The mean absolute energy imbalance (MAI) for each strategy calculated as

$$MAI = mean\left(|E_{t+h}^{DAM} - E_{t+h}|\right)$$

is shown in Figure 5.60 and tabulated in Table 5.16. The results show that, as

ImbMin	ImbMax-price	ImbMax-NIV	Quant-price	Quant-NIV
8.4	38.3	42.3	9.8	9.4

ρ	Scal-price	Scal-NIV	Prop-price	Prop-NIV
0.20	16.9	17.3	9.8	9.7
0.40	27.2	28.6	13.4	13.2
0.60	33.8	36.4	17.8	17.2
0.80	37.2	40.7	22.2	21.0
1.00	38.3	42.3	26.5	24.6

Table 5.16: Mean absolute energy imbalance from each bidding strategy. Values are expressed as % of P_{nom}.

expected, the ImbMin strategy does minimise the size of imbalances[3] and the ImbMax maximise them, while the MAI is always increasing for increasing values of the risk factor. Scal strategies cause significantly larger imbalances compared to Prop and Quant, regardless of the market quantity forecast employed; this makes intuitive sense given how the strategies are formulated.

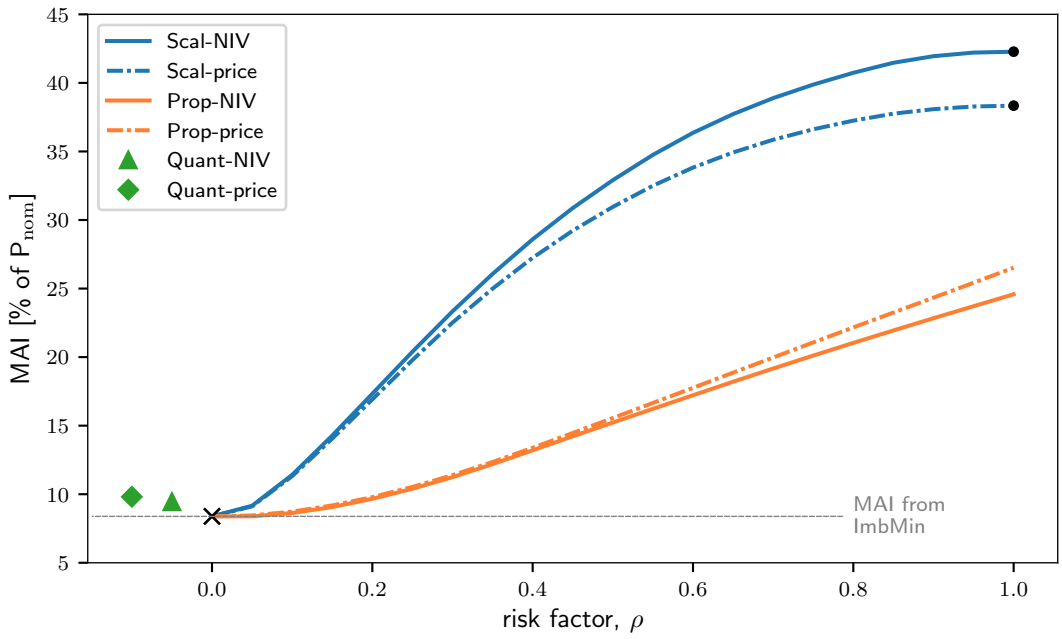


Figure 5.60: Mean absolute energy imbalance from each bidding strategy. The cross and the dots represent the ImbMin and ImbMax strategies, respectively.

[3]The MAI for ImbMin is in fact the MAE of the point forecast employed in the strategy.

Next, the exposure to risk as measured by $\text{CVaR}_{1\%}$ and $\text{CVaR}_{5\%}$ resulting from each strategy is evaluated. Recall that $\text{CVaR}_{1-\alpha}$ is defined as the expected revenue in the worst $(1-\alpha)\times 100\%$ of cases (cf. Equation (4.11)). Therefore, decreasing CVaR values correspond to increasing risk, since the participant is exposed to lower profits or, when values are negative, larger losses. The values obtained in the case study are shown in Figure 5.61 and tabulated in Table 5.17 for $\text{CVaR}_{1\%}$, and Figure 5.62 for $\text{CVaR}_{5\%}$. It can be noted how the $\text{CVaR}_{1\%}$ from Perfect is considerably better than any of the other strategies and that, even though no balancing costs are incurred in this strategy, the value is negative as a result of the instances of negative DAM prices. Overall, the Scal approaches cause the worst $\text{CVaR}_{1-\alpha}$ for increasing values of ρ due the higher energy imbalances they cause (cf. Figure 5.60). Furthermore, for all approaches the $\text{CVaR}_{1-\alpha}$ for the "-NIV" case is consistently worse than the "-price" homologous as a result of the different accuracy of the two market quantity forecasts, with the lowest (hence worst) $\text{CVaR}_{1-\alpha}$ obtained by ImbMax-NIV (i.e. Scal-$\text{NIV}_{[\rho=1.00]}$).

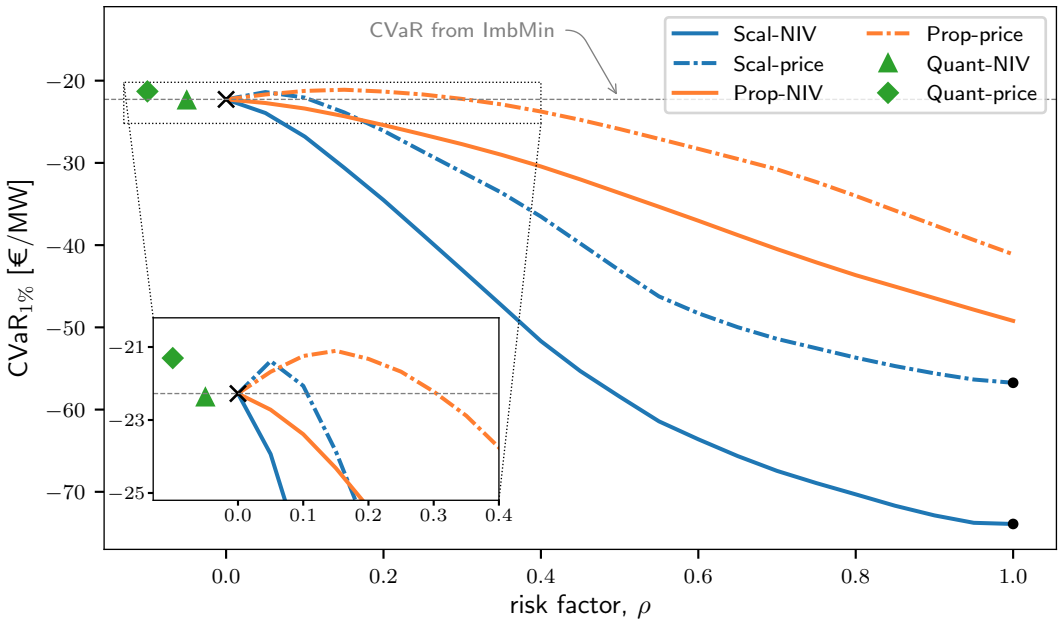


Figure 5.61: $\text{CVaR}_{1\%}$ from each bidding strategy. The cross and the dots represent the ImbMin and ImbMax strategies, respectively.

The most remarkable result, however, is that within certain operating regions of

ImbMin	ImbMax-price	ImbMax-NIV	Quant-price	Quant-NIV	Perfect
–22.28	–56.74	–73.91	–21.30	–22.37	–1.85

ρ	Scal-price	Scal-NIV	Prop-price	Prop-NIV
0.20	–26.09	–34.52	–21.33	–25.40
0.40	–36.53	–51.67	–23.75	–30.43
0.60	–48.29	–63.61	–28.29	–37.03
0.80	–53.69	–70.31	–34.01	–43.65
1.00	–56.74	–73.91	–41.12	–49.21

Table 5.17: $\text{CVaR}_{1\%}$ from each bidding strategy. Values are expressed in € per installed MW.

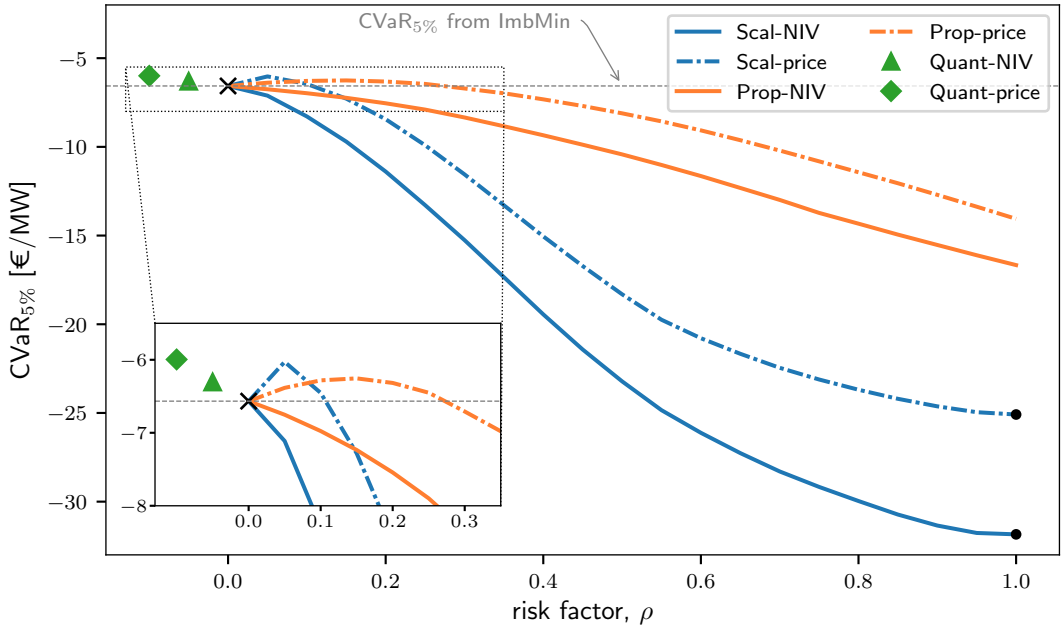


Figure 5.62: $\text{CVaR}_{5\%}$ from each bidding strategy. The cross and the dots represent the ImbMin and ImbMax strategies, respectively.

the "-price" strategies, the $\text{CVaR}_{1-\alpha}$ increases in comparison to ImbMin, meaning that exposure to risk is reduced. Improvements are observed for both $\text{CVaR}_{1\%}$ and $\text{CVaR}_{5\%}$ but as the former corresponds to a higher aversion to risk, the remainder of the analysis will focus on this measure. Risk exposure as measured by $\text{CVaR}_{1\%}$ reduces for Scal-price when $\rho \leq 0.10$, Prop-price when $\rho \leq 0.30$, and Quant-price. The best $\text{CVaR}_{1\%}$ is obtained by $\text{Prop-price}_{[\rho=0.15]}$ and is equal to € –21.10, a 5.3% improvement with respect to the ImbMin case. Interestingly, the reduced exposure

to risk does not correspond to a reduction of energy imbalances. This is visualised in Figure 5.63, which shows that while $\text{CVaR}_{1\%}$ improves (see the zoom inset in the graph), the MAI is still increasing up to almost 12% of P_{nom}.

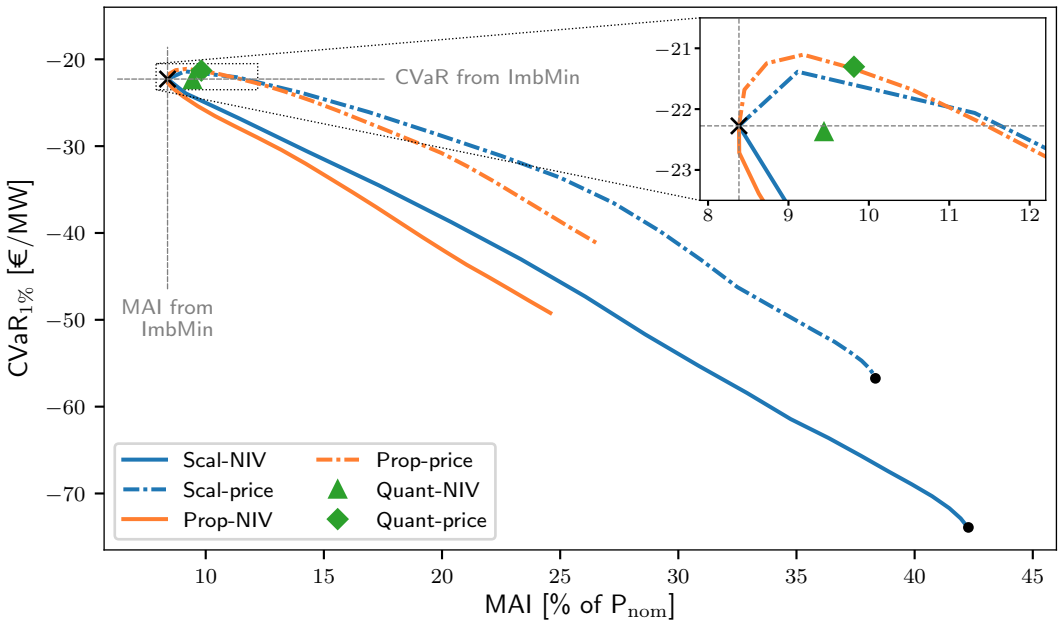


Figure 5.63: Mean absolute energy imbalance vs $\text{CVaR}_{1\%}$ from each bidding strategy. The cross and the dots represent the ImbMin and ImbMax strategies, respectively.

To visually analyse the trade-off between expected profit and risk resulting from each approach, the revenue is shown against the $\text{CVaR}_{1\%}$ in Figure 5.64. The zoom inset in the graph highlights the region where, in comparison to ImbMin, certain advanced strategies allow to increase the revenue while simultaneously reducing or leaving unchanged the risk for the participant: revenues are increased up by 4.4% for Prop-price, 4.1% for Scal-price, and 2.4% for Quant-price. For Quant-NIV, the $\text{CVaR}_{1\%}$ is only slightly worse than ImbMin, and revenue is increased by 1.2%.

As previously noted for the cases of revenue and $\text{CVaR}_{1-\alpha}$, for each strategy type, employing the price difference instead of the imbalance sign forecast results in a consistently better performance. Indeed, for a given level of risk (i.e. equal $\text{CVaR}_{1\%}$), revenues are always higher for the "-price" case. Likewise, the risk associated with a given level of profit (i.e. equal revenue) is always lower for the

"-price" case. This emphasizes how the accuracy of the market quantity forecast employed has a crucial impact on the strategy's performance.

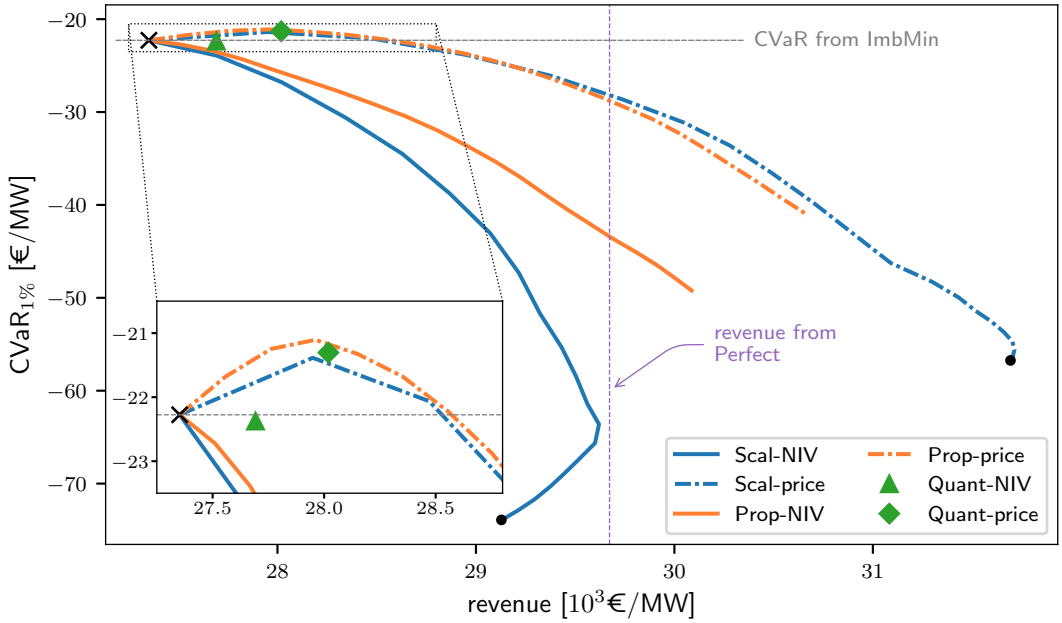


Figure 5.64: Revenue vs $\text{CVaR}_{1\%}$ from each bidding strategy. The cross and the dots represent the ImbMin and ImbMax strategies, respectively.

Figure 5.64 also shows how the strategies that are able to yield revenues equal or even higher than those obtained from using a perfect wind power forecast, do so at the expense of an increased exposure to risk. From a practical perspective, the relevance of this finding must be noted, as it shows that a risk-seeking participant may consider it more valuable to invest in a forecast of market quantities rather than getting a new, more accurate wind power forecast. An example is illustrated in Figure 5.65, where the net cumulative revenues of a number of strategies are shown. The *net cumulative revenue* is defined as the difference between the cumulative revenue obtained by the strategy and the cumulative revenue obtained by ImbMin (see Zugno et al. in [193]). The Figure shows that Prop-NIV$_{[\rho=0.80]}$ and Prop-price$_{[\rho=0.60]}$ generate around the same total revenue as Perfect by the end of the test period, but do so through visibly different paths. The size of the adjustments is larger in the case of Prop-NIV$_{[\rho=0.80]}$ (the energy forecast is increased or decreased by 80%), but the accuracy of the market quantity forecast is lower, resulting in

the steep jumps and drops observed in the solid curve. On the other hand, the adjustment size is smaller for Prop-price$_{[\rho=0.60]}$ (60% of the forecast energy) but the market quantity forecast is more accurate, resulting in the smoother, generally rising dash-dotted curve. The higher volatility of Prop-NIV$_{[\rho=0.80]}$ is in fact reflected by its CVaR$_{1\%}$ (€ –43.65) being substantially worse than that of Prop-price$_{[\rho=0.60]}$ (€ –28.29).

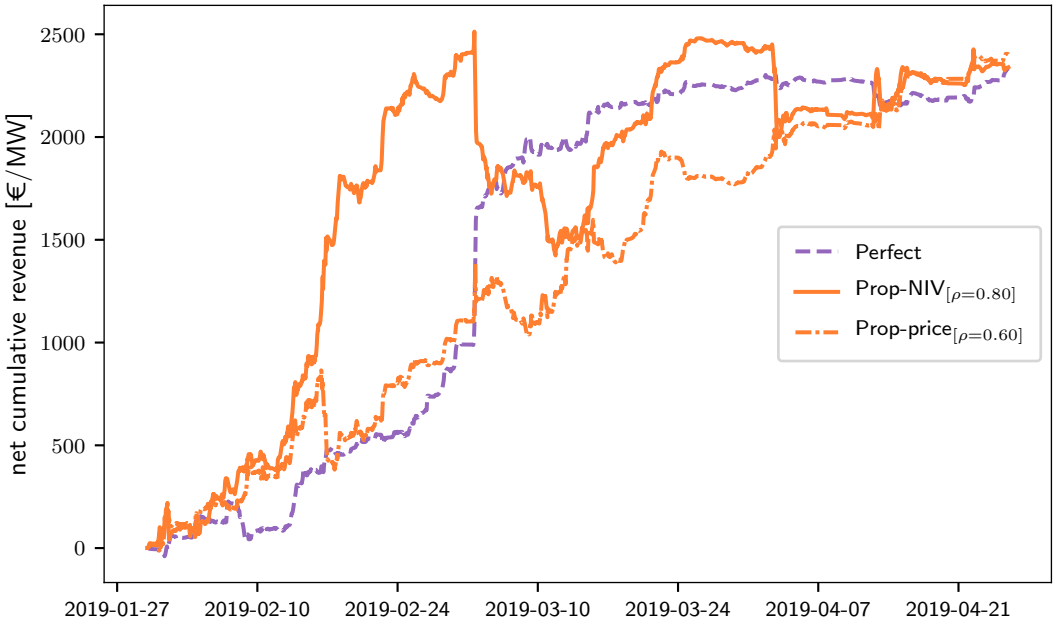


Figure 5.65: Net cumulative revenues of different bidding strategies yielding the same total revenue as Perfect at the end of the test period.

The ability of certain strategies to generate more revenue than a perfect wind power forecast is further analysed in Figure 5.66. Here, the distributions of daily revenue are shown from Perfect and from the Prop-price strategies for increasing risk factors. While the total revenue yielded by Prop-price$_{[\rho=0.75]}$ and Prop-price$_{[\rho=1.00]}$ is higher than that of Perfect (cf. Table 5.15), the volatility of daily revenue increases substantially: the range of values expands visibly, with revenues increasing while possible losses become larger and more frequent.

In Figure 5.67, the revenues obtained from each strategy are shown against the MAI they cause. The graph illustrates how the average size of imbalances becomes larger for increasing revenues, indicating that the objective of the participant, who

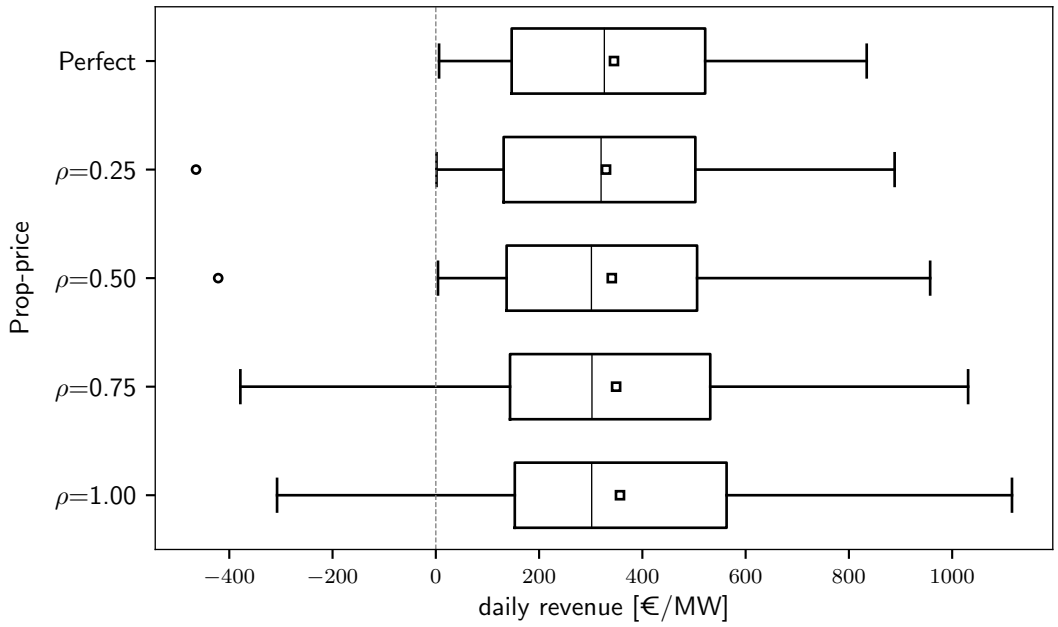


Figure 5.66: Distribution of daily revenues using a perfect wind power forecast (Perfect) and with the Prop-price strategy for increasing risk factors.

wants to increase their revenue, is in fact in conflict with that of the system operator, who wants imbalances to be minimised. Prop-price causes the lowest MAI for a given level of revenue, and yields the highest revenue for a given level of MAI. Note that to obtain the same revenue as Perfect (where no imbalances are caused), the size of imbalances increases more than twofold in comparison to ImbMin and is halved in comparison to ImbMax-NIV and ImbMax-price: MAI is equal to 18.2% of P_{nom} with the Prop-price strategy, 19.3% with Scal-price, and 20.9% with Prop-NIV.

A summary of the key results for some of the advanced bidding strategies developed in this work is illustrated in Figure 5.68. The scatter plot shows the *average* revenue yielded by the Scal-price and Prop-price strategies against the corresponding $\text{CVaR}_{1\%}$. The size of each marker is proportional to the MAI caused by the strategy, while the intensity of the colour increases with the risk factor. ImbMin and ImbMax are indicated with black crosses and constitute industry and academia reference strategies, respectively. With regards to ImbMax, for example, it can be noted that in 1 trading period out of 100, a participant adopting this strategy incurs

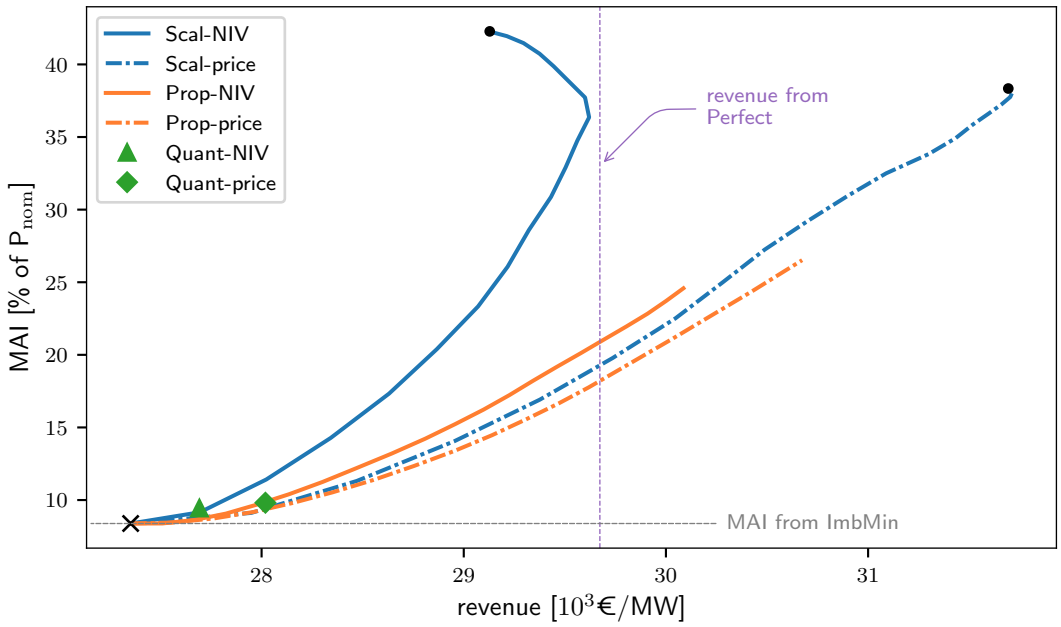


Figure 5.67: Revenue vs Mean absolute energy imbalance from each bidding strategy. The cross and the dots represent the ImbMin and ImbMax strategies, respectively.

an average loss of 57 € per installed MW, when the average revenue is less than 8 € per installed MW. The revenue obtained with a perfect wind power forecast is also marked in the graph. From the Figure, it can be observed how exposure to risk can be reduced by adopting the proposed advanced trading approaches with lower risk tolerance levels. As discussed earlier, these strategies can outperform Perfect in terms of revenue for higher values of the risk factor, at the expenses of bigger losses on a single period (i.e. lower $\text{CVaR}_{1\%}$) and larger energy imbalances.

It must be reiterated that the strategies developed in this work are only valid for wind farms whose size is small relative to the power system. Considering that the strategies presented here can introduce average imbalances up to 40% of the wind farm's nominal capacity, limitations exist on their applicability. To provide some context, the average size of imbalances in Ireland was 21.56 MWh with a standard deviation of 28.24 MWh during the first seven months of operation of the I-SEM (cf. Table 2.6). Indeed, a participant who intentionally causes imbalances of significant size would impact the market's behaviour and in turn affect the price

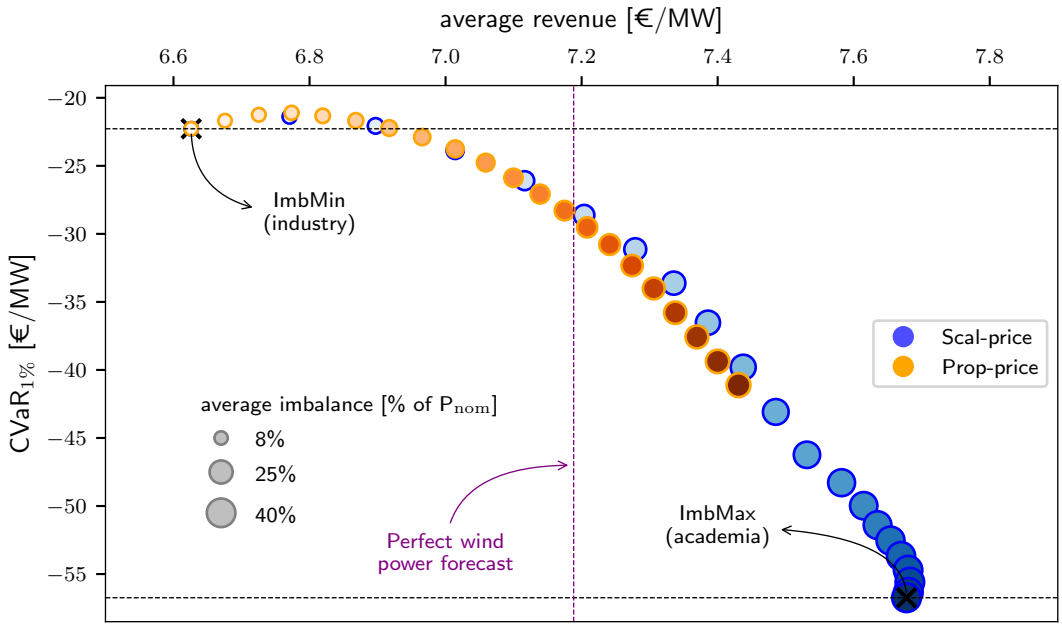


Figure 5.68: Average revenue vs $CVaR_{1\%}$ from Scal-price and Prop-price. The colour intensity increases with the risk factor, while the marker size is proportional to the MAI caused.

formation processes, hence violating the price-taker assumption (see Section 4.1.1) and making the value forecasting model inconsistent. Moreover, such systematic, large deviations could be deemed unacceptable by the regulators, with consequent penalties for the trader.

Chapter 6

Conclusions

6.1 Summary of Conclusions

The aim of this research was to develop new probabilistic models forecasting the short-term value of wind power and to investigate their use in bidding in electricity markets adopting a single-imbalance pricing scheme. The perspective adopted here is that of a participant who wants to bid strategically in the day-ahead market to increase the value (measured in terms of revenue and risk exposure) of the energy generated by a wind farm. The proposed models combined short-term predictions of wind power and relevant electricity market quantities in a probabilistic framework. Offers were then formulated to reflect the participant's risk profile, conditioned upon the uncertainty in future wind power generation and electricity market conditions. The aim of the research was achieved, as the proposed strategies successfully improved the value of wind power for the participant by increasing their revenue while simultaneously controlling exposure to risk.

In order to define the general context of the Thesis, wholesale electricity markets and their operation were described in Chapter 2. First, the general features and mechanisms of electricity markets were discussed, focussing on the main energy trading platforms of interest for a wind farm operator. The specific case of the I-SEM, the new Irish electricity market, was then studied. After outlining

its structure, a comprehensive picture was provided of the market data available at the time this work was conducted. A number of market processes had to be modelled and forecast to implement the value forecasting models. For this reason, exploratory and statistical analyses were performed on the available I-SEM market data to gain insight into these processes and improve the modelling quality and forecasting performance. The outcomes of these analysis provided the motivation for the modelling approaches adopted within the methodology of this work.

In the first part of Chapter 3, the theoretical background necessary to develop and evaluate time series forecasting models was given. This was followed by a review of the academic literature of relevance to this work. The state-of-the-art in short-term forecasting of wind power was described. It was highlighted how in recent years the focus of academic research has moved towards probabilistic forecasts, although business practice is still rooted in the use of point predictions.

Next, an analogous review was given for short-term forecasting of electricity prices. The literature on day-ahead market price forecasting was examined. Computational intelligence and statistical models emerged as the two most popular approaches. However, the focus of this work was restricted to the latter due to the limited amount of historical data available for the study. It was concluded that of the statistical models, ARMAX tends to achieve the best performance, and modelling each load period separately improves the forecast accuracy in comparison to having a single specification for all time periods. On the other hand, the literature on short-term forecasts of balancing market quantities was found to be very limited and mostly concerned with Scandinavian markets, and the evidence for one class of models outperforming the others was scarce. This research advances the discussion on the subject by proposing a new method to generate probabilistic forecasts of the difference between day-ahead and balancing market prices combining ARMAX models and Monte Carlo simulations.

In the final part of Chapter 3, the academic literature on short-term value forecasting of wind power for energy trading was reviewed. It was observed that

strategic bidding in the day-ahead market has been investigated with growing interest in recent years. Nevertheless, a gap was highlighted, as the literature on power systems with single-price imbalance settlement was found to be extremely limited. It was evidenced that the most basic approach to trading, that is, directly bidding the point forecasts of wind generation, also constitutes the current industry practice among most utilities in Europe and the USA. On the other hand, the Zero/Max strategy represents the state-of-the-art in academic research, despite its practical implementability being arguable. This both justified and motivated the objective of developing novel value forecasting models, as this research advances the discussion on strategic bidding in these types of markets: the new proposed strategies use forecasts of both wind power and market quantities in a probabilistic framework while integrating explicitly the participant's attitude towards risk.

Chapter 4 concentrated on the methodology used in developing the value forecasting models. First, the problem of trading wind energy in day-ahead and single-price balancing markets was presented together with the assumptions made. It was underlined how the two variables of primary interest for a market participant are wind power generation and the sign of the difference between day-ahead and balancing market prices. The characteristics of the wind power forecasts available for the wind farm under study were then described.

Next, the models developed to forecast the sign of price difference were explained. Given the occurrence of imbalance price anomalies highlighted in the analyses of I-SEM data, two alternative approaches were used. First, a novel method was proposed, whereby day-ahead and balancing market prices were forecast individually using ARMAX processes and the resulting quantile predictions combined with Monte Carlo simulations to estimate the probability distribution of price differences. In the second method, the imbalance sign was modelled as a binomial distribution and forecast with logistic regression models; the probability of a positive price difference was then deduced directly from this forecast. The multi-model approach (whereby a separate model is fitted to each load period)

was implemented in both methods. Moreover, within each method, two alternative techniques — namely, "Recursively expanding" and "Rolling" — were adopted to train the models in order to accommodate for the time-varying nature of the processes.

Three novel bidding strategies were then formulated and tested against two benchmarks to examine possible improvements. These two benchmarks are the business practice and the research state-of-the-art, and correspond to the extreme cases where the objective is to minimise or maximise the energy imbalances, respectively. In these two approaches, offers are based on either the sole point predictions of wind power or the estimated sign of price difference. The three novel advanced strategies were then described. Both wind power and market quantity probabilistic forecasts were used to define the bids and the objective was to increase the revenue while controlling the exposure to risk. It was highlighted how these bids are formulated to reflect the participant's risk profile, conditional on the uncertainty in future wind power generation and electricity market states.

The methodology was applied to a case study where the participation of a wind farm in the Irish electricity market was simulated and the results were presented and discussed in Chapter 5. The wind power forecasts for the wind farm under study were examined. The model with significantly better accuracy was identified and selected for use in the trading strategies.

In the second part of the Chapter, the electricity price forecasting methods were evaluated. For day-ahead prices, both proposed models clearly outperformed the benchmark. Although the difference with the "Rec. expanding" approach was very small, the "Rolling" approach was found to generate slightly more accurate point predictions as well as better quantile forecasts. As spot prices were characterised by marked seasonalities and a strong dependence on the level of demand (which in turn varies with the time of the year), training the models using only the more recent observations helped to better capture these seasonal variations. Hence, the "Rolling" method was selected as the forecasting method for DAM prices and its

predictions used as inputs to the Monte Carlo simulations. For imbalance prices, the vast majority of the fitted models were multiple linear regression models for both the "Rec. expanding" and "Rolling" approaches. This finding confirmed how the autoregressive and seasonal components in the imbalance price process were very weak, indicating that the prediction of future values relied primarily on the exogenous variables. Again, both proposed models clearly outperformed the benchmark. This time, the "Rec. expanding" approach was found to be slightly superior, with the advantage being more evident in terms of point predictions. Given the particularly erratic behaviour of imbalance prices in the Irish market, the use of as much historical data as possible to train the models proved to be more beneficial. Hence, the "Rec. expanding" approach was selected as the forecasting method for BM prices and its predictions used in the Monte Carlo simulations. The forecast probabilities of positive price differences were obtained from these simulations, where a tendency to predict positive price differences more frequently was observed. Predictions were found to be more accurate than the benchmark, with the sign of price difference being correctly estimated 56.5% of the time.

In the third part of the Chapter, results from the evaluation of imbalance sign forecasting methods were presented. All four logistic regression models consistently outperformed the benchmark. The best performance was achieved by the approach using a rolling training window and stepwise re-selection of the predictors, and the corresponding predictions were employed in the trading strategies. For this approach, point predictions of wind power and predicted wind power penetration were included in more than half of the models, indicating a correlation between net imbalance volumes and system-level wind generation and demand. Negative price differences were forecast more frequently, and the imbalance sign was predicted correctly in 53.5% of the cases. However, due to the occurrence of anomalous prices in the I-SEM balancing market, a correct prediction of the imbalance direction did not always correspond to a correct prediction of the price difference sign, and vice versa. In fact, the accuracy of the imbalance sign forecast in terms of correct

predictions of the sign of price differences was found to be lower at 51.0%.

In the last part of Chapter 5, the performances of the various novel bidding strategies were analysed. The most important result was that, through the use of the advanced strategies described above, revenues were increased without affecting, and even reducing the risk for the participant. The benefits to the participant were significant, with revenues improving by up to 4.4% and exposure to risk (as measured by $\text{CVaR}_{1\%}$) decreasing by 5.3%. As expected, higher revenues were obtained as a result of larger energy imbalances; interestingly, reduced exposure to risk did not correspond to a reduction of energy imbalances. This highlighted how the objective of the participant is in conflict with that of the system operator, who would prefer imbalances to be minimised. Therefore, the importance of limiting the proposed strategies to wind farms of small size was emphasized. Another relevant finding was that some strategies were able to yield a revenue higher (up by 6.9%) than that obtained with a perfect forecast of wind power generation. This result was achieved at the expense of a significantly higher exposure to risk. Therefore, it was noted that for a risk-seeking participant, investing in a market quantity forecast would be more valuable than getting a new, more accurate wind power forecast.

A central conclusion was that the proposed strategies are highly sensitive to the accuracy of the market quantity forecasts employed. Indeed, strategies using the electricity price forecast reported a substantially better performance than their homologous using the imbalance sign forecast. The following should be noted. One novel and one informed state-of-the-art statistical models were implemented to forecast market quantities in a new and volatile electricity market with limited available historical data. As highlighted in the literature review, forecasting balancing market quantities before the day-ahead gate closure is particularly challenging. Nevertheless, it should be emphasized that the results from the case study demonstrated that both methods implemented here were successfully employed in the strategic bidding framework.

6.2 Future Work

During this research programme, several analyses were performed and methods developed to implement the proposed value forecasting models. Yet, their performance and usefulness to the academic community and end-users can obviously be improved with a number of possible extensions and advancements. Some perspectives for future work are given in this Section.

Forecasting market quantities before the day-ahead market closure was challenging but the forecasts developed in this thesis were successfully employed in the proposed bidding strategies. However, the models' performances were mediocre at times, leaving significant room for improvement.

Systematic errors emerged from the evaluation of price forecast models. The negative bias in point predictions should be removed and the calibration of quantile predictions needs improving. Statistical post-processing methods could be applied to address these issues, such as short-term bias-correction (e.g. Sweeney et al. in [107]), smoothing procedures (e.g. Maciejowska and Nowotarski in [217]) and, as the record of data increases, Model Output Statistics (MOS) (e.g. Glahn and Lowry in [103]).

Most models used for the prediction of imbalance prices were reduced to multiple linear regression models, as the autoregressive and moving average components of the process were found to be very weak. If more information on the power system (for instance, planned outages, grid constraints, forecast curtailment) was made available, it could be included as explanatory variables in the models and potentially improve the forecast performance.

An initial condition problem was also highlighted in the price forecasting models, as errors were found to be higher on Mondays and at weekends (cf. Figures 5.17 and 5.31). This could be addressed by including more recent information in the training data set. For example, the models estimating Monday prices could include Sunday historical data; similarly, models estimating weekend prices could include Friday historical data.

An in-depth investigation should be carried out on the imbalance sign forecast models to identify the reasons why the approaches with stepwise re-selection of predictors generated tighter and more prudent estimates (cf. Figure 5.45). For example, distributions could be separately analysed for each half-hour to verify if variations are more or less pronounced depending on the time of the day; in turn, the relevant predictors could be analysed in detail during those trading periods to find their impact on imbalance sign estimates.

A significant drawback arose from price anomalies in the I-SEM in terms of accuracy of the imbalance sign forecast. Further investigations on the reasons behind their occurrence might be able to address these drawbacks. For example, tie-line congestions and wind penetration were found to contribute to the creation of counter-intuitive prices; including system-level forecasts for the Republic of Ireland and Northern Ireland separately in the models could help to account for this. Given the price anomalies, it could also be beneficial to directly model the sign of price difference as a binomial distribution and use logistic regression models to forecast future values.

Overall, given the sensitivity of the bidding strategies to forecast accuracy, advancements in market quantity forecasts would be very valuable to end-users and potential commercial applications would certainly motivate research efforts in this area.

In parallel, it would be useful to extend the bidding strategies to provide further insight into the forecast uncertainty. This could be achieved by including more statistics on future wind power generation and market quantities. For example, the standard deviation of the distribution of forecast price differences could contribute to conditioning the optimal energy bid, with adjustments becoming more limited as variance increases. As the predictive power of the models was found to change noticeably across trading periods (see Figure 5.55), another possible extension would be to introduce time-varying adjustments in the strategies. For instance, adjustment functions could be re-designed in order to include the time of the day in their

formulation.

Given the limited length of data record, sensitivity analysis could be conducted to verify some of the predictive model responses and features in order to build confidence in the developed models. Future work could also continue to implement the methodology presented in this paper by testing the proposed strategies on a longer test period and achieve even more conclusive results. Data could be used from other compatible power systems or, as more I-SEM data become available with time, from the Irish market again.

To conclude, it would be of particular interest to extend the framework for the assessment of forecast value by generalising what was formulated in this work. To achieve optimal benefits, prediction models need to be tailored to the end-user requirements. Therefore, integrating the forecast user's appraisal of value (for example, the decision maker's risk profile) is essential to give a consistent evaluation. The importance of defining such a framework ought to be realised, as it would provide the basis to gain clear and objective insight into the characteristics of rival approaches, quantify their relative benefits, and justify possible new developments. For this purpose, enhanced dialogue and collaboration between forecasters and end-users will be necessary in the future.

Bibliography

[1] United Nations, "Paris Agreement," 2015. [Online]. Available: https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement [Accessed: 2020-04-25]

[2] Global Wind Energy Council (GWEC), "Global Wind Report 2019," 2019. [Online]. Available: https://gwec.net/global-wind-report-2019/ [Accessed: 2020-04-25]

[3] REN21, "Renewables 2019 Global Status Report," Paris: REN21 Secretariat, 2019. [Online]. Available: http://www.ren21.net/gsr-2019/ [Accessed: 2020-04-25]

[4] BloombergNEF, "New Energy Outlook 2019," 2019. [Online]. Available: https://about.bnef.com/new-energy-outlook/ [Accessed: 2020-04-25]

[5] J. Manwell, J. McGowan, and A. Rogers, *Wind Energy Explained: Theory, Design and Application*, 2010.

[6] Z. Chen, J. M. Guerrero, and F. Blaabjerg, "A review of the state of the art of power electronics for wind turbines," 2009.

[7] H. Holttinen and R. Hirvonen, "Power System Requirements for Wind Power," in *Wind Power in Power Systems*, 2005.

[8] G. Giebel and G. Kariniotakis, *Wind power forecasting-a review of the state of the art*. Elsevier Ltd, 2017, no. 2007.

[9] R. J. Bessa, C. Möhrlen, V. Fundel, M. Siefert, J. Browell, S. Haglund El Gaidi, B. M. Hodge, U. Cali, and G. Kariniotakis, "Towards improved understanding of the applicability of uncertainty forecasts in the electric power industry," *Energies*, vol. 10, no. 9, pp. 1–48, 2017.

[10] B. M. Hodge, C. Brancucci Martinez-Anido, Q. Wang, E. Chartan, A. Florita, and J. Kiviluoma, "The combined value of wind and solar power forecasting improvements and electricity storage," *Applied Energy*, vol. 214, no. June 2017, pp. 1–15, 2018.

[11] Q. Wang, H. Wu, A. R. Florita, C. Brancucci Martinez-Anido, and B. M. Hodge, "The value of improved wind power forecasting: Grid flexibility quantification, ramp capability analysis, and impacts of electricity market operation timescales," *Applied Energy*, vol. 184, pp. 696–713, 2016.

[12] Q. Wang, C. Brancucci Martinez-Anido, H. Wu, A. R. Florita, and B. M. Hodge, "Quantifying the Economic and Grid Reliability Impacts of Improved Wind Power Forecasting," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 4, pp. 1525–1537, 2016.

[13] "EirGrid Library." [Online]. Available: http://www.eirgridgroup.com/library/ [Accessed: 2020-04-25]

[14] "Celtic Interconnector - Connecting the electricity grids of Ireland and France." [Online]. Available: https://www.celticinterconnector.eu/ [Accessed: 2020-04-25]

[15] EirGrid, "Celtic Interconnector - The Project." [Online]. Available: http://www.eirgridgroup.com/the-grid/projects/celtic-interconnector/the-project/ [Accessed: 2020-04-25]

[16] SEAI, "Energy in Ireland 2019 Report," 2019. [Online]. Available: https://www.seai.ie/publications/ [Accessed: 2020-04-25]

[17] ——, "Energy Statistics in Ireland." [Online]. Available: https://www.seai.ie/data-and-insights/seai-statistics/ [Accessed: 2020-04-25]

[18] EirGrid, "System & Renewable Summary Report," 2019. [Online]. Available: http://www.eirgridgroup.com/how-the-grid-works/renewables/ [Accessed: 2020-04-25]

[19] "National Renewable Energy Action Plans." [Online]. Available: https://ec.europa.eu/energy/topics/renewable-energy/national-renewable-energy-action-plans-2020{_}en [Accessed: 2020-04-25]

[20] EirGrid, "DS3 Programme." [Online]. Available: http://www.eirgridgroup.com/how-the-grid-works/ds3-programme/ [Accessed: 2020-04-25]

[21] ——, "Record renewable energy on power grid." [Online]. Available: http://www.eirgridgroup.com/newsroom/record-renewable-energy-o/ [Accessed: 2020-04-25]

[22] SEM-O, "Market Overview," 2017. [Online]. Available: https://www.sem-o.com/training/modules/market-overview/ [Accessed: 2020-04-25]

[23] SEM Committee, "ISEM Project." [Online]. Available: https://www.semcommittee.com/isem-project [Accessed: 2020-04-25]

[24] EirGrid, "Industry Guide to the I-SEM," 2017. [Online]. Available: https://www.sem-o.com/documents/general-publications/I-SEM-Industry-Guide.pdf [Accessed: 2020-04-25]

[25] C. Möhrlen, R. J. Bessa, M. Barthod, G. Goretti, and M. Siefert, "Use of Forecast Uncertainties in the Power Sector : State-of-the – Art of Business Practices," *15th Wind Integration Workshop*, 2016.

[26] R. J. Bessa, V. Miranda, A. Botterud, and J. Wang, "Good or bad wind power forecasts: A relative concept," *Wind Energy*, 2011.

[27] Murphy, "What Is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting," pp. 281–293, 1993.

[28] P. Pinson, C. Chevallier, and G. N. Kariniotakis, "Trading wind generation from short-term probabilistic forecasts of wind power," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1148–1156, 2007.

[29] D. Kirschen and G. Strbac, *Fundamentals of Power System Economycs*. Chichester: John Wiley & Sons, Ltd, 2004.

[30] F. Ocker, S. Braun, and C. Will, "Design of European balancing power markets," *International Conference on the European Energy Market, EEM*, vol. 2016-July, 2016.

[31] S. Stoft, *Power System Economics: Designing Markets for Electricity*. New York: Wiley-IEEE Press, 2002.

[32] S. Hunt and G. Shuttleworth, *Competition and choice in electricity.*, 1996, vol. 34, no. 4.

[33] E. Ela, M. Milligan, A. Bloom, J. Cochran, A. Botterud, A. Townsend, and T. Levin, "Overview of Wholesale Electricity Markets," in *Electricity Markets with Increasing Levels of Renewable Generation: Structure, Operation, Agent-based Simulation, and Emerging Designs. Studies in Systems, Decision and Control.*, F. Lopes and H. Coelho, Eds. Springer International Publishing, 2018, vol. 144, pp. 3–21.

[34] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014.

[35] C. Weber, "Adequate intraday market design to enable the integration of wind energy into the European power systems," *Energy Policy*, vol. 38, no. 7, pp. 3155–3163, 2010.

[36] SEM Committee, "Trading and Settlement Code," 2017. [Online]. Available: https://www.semcommittee.com/trading-and-settlement-code [Accessed: 2020-04-25]

[37] SEMO, "Mod_06_19 Determination of the Marginal Energy Action Price where no energy is available in the Net Imbalance Volume," 2019. [Online]. Available: https://www.sem-o.com/rules-and-modifications/balancing-market-modifications/ [Accessed: 2020-04-25]

[38] ——, "Mod_09_19 Removal of Locational Constraints from Imbalance Pricing Calculation," 2019. [Online]. Available: https://www.sem-o.com/rules-and-modifications/balancing-market-modifications/ [Accessed: 2020-04-25]

[39] F. Sensfuß, M. Ragwitz, and M. Genoese, "The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany," *Energy Policy*, vol. 36, no. 8, pp. 3086–3094, 2008.

[40] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewables in Electricity Markets*. Springer, 2014, vol. 205.

[41] R. E. Walpole, R. H. Myers, S. L. Myers, and K. Ye, *Probability & statistics for engineers & scientists*, 2016, vol. 6.

[42] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 2nd ed. Melbourne, Australia: OTexts, 2018. [Online]. Available: https://otexts.com/fpp2/

[43] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Statistics Surveys*, vol. 4, pp. 40–79, 2010.

[44] M. A. Hernández and S. J. Stolfo, "Real-world data is dirty: Data cleansing and the merge/purge problem," *Data Mining and Knowledge Discovery*, vol. 2, no. 1, pp. 9–37, 1998.

[45] R. Y. Pipino, Leo L.; Lee, Yang W. ; and Wang, "Data quality assessment," *Communications of the ACM*, vol. 45, no. 4, pp. 211–218, 2002.

[46] E. Rahm and H. H. Do, "Data Cleaning: Problems and Current Approaches," *IEEE Data Eng. Bull.*, vol. 23, no. 4, pp. 3–13, 2000.

[47] G. Goretti and A. Duffy, "Evaluation of Wind Energy Forecasts : the Undervalued Importance of Data Preparation," *2018 15th International Conference on the European Energy Market (EEM)*, pp. 1–5, 2018.

[48] H. Madsen, P. Pinson, and G. Kariniotakis, "Standardizing the performance evaluation of shortterm wind power prediction models," *Wind Engineering*, vol. 29, no. 6, pp. 475–489, 2005.

[49] I. T. Jolliffe and D. B. Stephenson, *Forecast Verification. A Practicioner's Guide in Atmospheric Science*, 2006, vol. 22, no. 2.

[50] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, 2006.

[51] P. Goodwin and R. Lawton, "On the asymmetry of the symmetric MAPE," *International Journal of Forecasting*, vol. 15, no. 4, pp. 405–408, 1999.

[52] D. A. Swanson, J. Tayman, and C. F. Barr, "A note on the measurement of accuracy for subnational demographic estimates," *Demography*, vol. 37, no. 2, pp. 193–201, 2000.

[53] F. X. Diebold and R. S. Mariano, "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, vol. 13, pp. 253–265, 1995.

[54] F. X. Diebold, "Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold–Mariano Tests," *Journal of Business and Economic Statistics*, vol. 33, no. 1, 2015.

[55] T. Gneiting, F. Balabdaoui, and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness," *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, vol. 69, no. 2, pp. 243–268, 2007.

[56] S. Bentzien and P. Friederichs, "Decomposition and graphical portrayal of the quantile score," *Quarterly Journal of the Royal Meteorological Society*, vol. 140, no. 683, pp. 1924–1934, 2014.

[57] H. A. Nielsen, T. S. Nielsen, H. Madsen, G. Giebel, J. Badger, L. Landberg, K. Sattler, L. Voulund, and J. Tøfting, "From wind ensembles to probabilistic information about future wind power production - Results from an actual application," in *2006 9th International Conference on Probabilistic Methods Applied to Power Systems, PMAPS*, 2006.

[58] T. M. Hamill, "Reliability diagrams for multicategory probabilistic forecasts," *Weather and Forecasting*, vol. 12, no. 4, pp. 736–741, 1997.

[59] J. B. Bremnes, "Probabilistic Forecasts of Precipitation in Terms of Quantiles Using NWP Model Output," *Monthly Weather Review*, 2004.

[60] P. Pinson, H. A. Nielsen, J. K. Møller, H. Madsen, and G. N. Kariniotakis, "Non-parametric probabilistic forecasts of wind power: Required properties and evaluation," *Wind Energy*, vol. 10, no. 6, pp. 497–516, 2007.

[61] J. Bröcker and L. A. Smith, "Scoring probabilistic forecasts: The importance of being proper," *Weather and Forecasting*, vol. 22, no. 2, pp. 382–388, 2007.

[62] R. L. Winkler, "Rewarding Expertise in Probability Assessment," in *Decision Making and Change in Human Affairs*, 1977, pp. 127–140.

[63] T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, and R. J. Hyndman, "Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond," *International Journal of Forecasting*, vol. 32, no. 3, pp. 896–913, 2016.

[64] J. A. Swets, "Measuring the Accuracy of Diagnostic Systems," *Science*, vol. 240, no. 4857, pp. 1285–1293, 1988.

[65] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

[66] G. W. Brier, "Verification of Forecasts Expressed in Terms of Probability," *Monthly Weather Review*, vol. 78, no. 1, pp. 1–3, 1950.

[67] A. H. Murphy, "A New Vector Partition of the Probability Score," *Journal of Applied Meteorology*, vol. 12, no. 4, pp. 595–600, 1973.

[68] D. B. Stephenson, C. A. Coelho, and I. T. Jolliffe, "Two Extra Components in the Brier Score Decomposition," *Weather and Forecasting*, vol. 23, no. 4, pp. 752–757, 2008.

[69] L. Landberg, "Short-term prediction of local wind conditions," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 89, no. 3-4, pp. 235–245, 2001.

[70] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, "The State-Of-The-Art in Short-Term Prediction of Wind Power A Literature Overview," *Technical Report, ANEMOS.plus*, pp. 1–109, 2011.

[71] U. Jensen, E. Pelgrum, and H. Madsen, "The development of a forecasting model for the prediction of wind power production to be used in central dispatch centres," in *Proceedings of the EWEC '94 in Tessaloniki, 10-14 October*, Tessaloniki, 1994, pp. 353–356.

[72] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, and G. Conzelmann, "Wind Power Forecasting: State-of-the-Art 2009," Decision and Information Sciences, Argonne National Laboratory, and INESC Porto, Argonne, IL (United States), Tech. Rep., 2009.

[73] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, "Current methods and advances in forecasting of wind power generation," *Renewable Energy*, vol. 37, no. 1, pp. 1–8, 2012.

[74] G. Kariniotakis, J. Halliday, R. Brownsword, I. Marti, A. Palomares, I. Cruz, H. Madsen, T. S. Nielsen, H. A. Nielsen, U. Focken, M. Lange, G. Kallos, P. Louka, N. Hatziargyriou, P. Frayssinet, H. P. Waldl, F. Dierich, G. Giebel, R. Barthelmie, J. Badger, J. Usaola, I. Sanchez, D. Heinemann, J. Tambke, J. Moussafir, G. Descombes, M. Calleja, T. Jouhanique, J. Tøfting, P. O'Donnell, M. Ryan, J. Indorf, C. Barquero, P. Seguardo, A. Gigantidou, M. Thalassinakis, G. Gonzales, L. Bryans, and D. Garrett, "Next generation short-term forecasting of wind power - Overview of the ANEMOS project," in *European Wind Energy Conference and Exhibition 2006, EWEC 2006*, vol. 3, 2006, pp. 2417–2426.

[75] G. Giebel, J. Badger, I. Martí Perez, P. Louka, G. Kallos, A. M. Palomares, C. Lac, and G. Descombes, "Short-term forecasting using advanced physical modelling - The results of the anemos project: Results from mesoscale, microscale and CFD modelling," in *European Wind Energy Conference and Exhibition 2006, EWEC 2006*, vol. 3, 2006, pp. 2431–2459.

[76] M. Lange and U. Focken, *Physical approach to short-term wind power prediction*, 2006.

[77] L. Paiva, C. Veiga Rodrigues, and J. Palma, "Determining wind turbine power curves based on operating conditions," *Wind Energy*, vol. 17, no. 10, 2014.

[78] International Electrotechnical Commission, "IEC 61400-12-1 - Part 12-1: Power performance measurements of electricity producing wind turbines," 2005. [Online]. Available: https://www.iec.ch/

[79] F. Trivellato, L. Battisti, and G. Miori, "The ideal power curve of small wind turbines from field data," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 107-108, pp. 263–273, 2012.

[80] G. Goretti, A. Duffy, and T. Lie, "The impact of power curve estimation on commercial wind power forecasts - An empirical analysis," in *International Conference on the European Energy Market, EEM*, 2017.

[81] M. Jafarian, A. Soroudi, and M. Ehsan, "The effects of enviromental parameters on wind turbine power PDF curve," in *Canadian Conference on Electrical and Computer Engineering*, 2008, pp. 1193–1197.

[82] J. Collins, J. Parkes, and A. Tindal, "Forecasting for utility-scale wind farms – the power model challenge," *2009 CIGRE/IEEE PES Joint Symposium Integration of Wide-Scale Renewable Resources Into the Power Delivery System*, pp. 247–257, 2009.

[83] M. Xu, P. Pinson, Z. Lu, , Y. Qiao, and Y. Min, "Adaptive robust polynomial regression for power curve modeling with application to wind power forecasting," *Wind Energy*, vol. 19, pp. 2321–2336, 2016.

[84] M. Raj, M. Alexander, and M. Lydia, "Modeling of wind turbine power curve," in *2011 IEEE PES International Conference on Innovative Smart Grid Technologies-India, ISGT India 2011*, 2011.

[85] J. Gottschall and J. Peinke, "How to improve the estimation of power curves for wind turbines," *Environmental Research Letters*, vol. 3, no. 1, p. 015005, 2008.

[86] E. Anahua, S. Barth, and J. Peinke, "Markovian power curves for wind turbines," *Wind Energy*, vol. 11, no. 3, pp. 219–232, 2008.

[87] T. Jin and Z. Tian, "Uncertainty analysis for wind energy production with dynamic power curves," *2010 IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems*, pp. 745–750, 2010.

[88] M. Lydia, A. I. Selvakumar, S. S. Kumar, and G. E. P. Kumar, "Advanced algorithms for wind turbine power curve modeling," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 3, pp. 827–835, 2013.

[89] S. Gill, B. Stephen, and S. Galloway, "Wind Turbine Condition Assessment Through Power Curve Copula Modeling," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 1, pp. 94–101, 2012.

[90] V. Thapar, G. Agnihotri, and V. K. Sethi, "Critical analysis of methods for mathematical modelling of wind turbines," *Renewable Energy*, vol. 36, no. 11, pp. 3166–3177, 2011.

[91] T. Üstüntaş and A. D. Şahin, "Wind turbine power curve estimation based on cluster center fuzzy logic modeling," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 96, no. 5, pp. 611–620, 2008.

[92] M. Lydia, S. S. Kumar, A. I. Selvakumar, and G. E. Prem Kumar, "A comprehensive review on wind turbine power curve modeling techniques," *Renewable and Sustainable Energy Reviews*, vol. 30, pp. 452–460, 2014.

[93] C. Carrillo, A. F. Obando Montaño, J. Cidrás, and E. Díaz-Dorado, "Review of power curve modelling for windturbines," *Renewable and Sustainable Energy Reviews*, vol. 21, pp. 572–581, 2013.

[94] U. Focken, M. Lange, K. Mönnich, and H. Waldl, "Reduction of wind power prediction error by spatial smoothing effects," in *Proceedings of the European Wind Energy Conference, Copenhagen, Denmark*, 2001, pp. 822–825.

[95] U. Focken, M. Lange, K. Mönnich, H. P. Waldl, H. G. Beyer, and A. Luig, "Short-term prediction of the aggregated power output of wind farms — A statistical analysis of the reduction of the prediction error by spatial smoothing effects," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 90, no. 3, pp. 231–246, 2002.

[96] B. Hasche, "General statistics of geographically dispersed wind power," *Wind Energy*, 2010.

[97] J. J. Miettinen and H. Holttinen, "Characteristics of day-ahead wind power forecast errors in Nordic countries and benefits of aggregation," *Wind Energy*, 2017.

[98] Y. Han and L. Chang, "A study of the reduction of the regional aggregated wind power forecast error by spatial smoothing effects in the Maritime Canada," in *2nd International Symposium on Power Electronics for Distributed Generation Systems, PEDG 2010*, 2010, pp. 942–947.

[99] J. Miettinen, H. Holttinen, and B. M. Hodge, "Simulating wind power forecast error distributions for spatially aggregated wind power plants," *Wind Energy*, no. July, pp. 1–18, 2019.

[100] X. Wang, P. Guo, and X. Huang, "A Review of Wind Power Forecasting Models," *Energy Procedia*, vol. 12, pp. 770–778, 2011.

[101] L. Landberg, "A mathematical look at a physical power prediction model," *Wind Energy*, vol. 1, no. 1, pp. 23–28, 1998.

[102] F. Castellani, M. Burlando, S. Taghizadeh, D. Astolfi, and E. Piccioni, "Wind energy forecast in complex sites with a hybrid neural network and CFD based method," in *Energy Procedia*, vol. 45, 2014, pp. 188–197.

[103] H. R. Glahn and D. A. Lowry, "The Use of Model Output Statistics (MOS) in Objective Weather Forecasting," *Journal of Applied Meteorology*, vol. 11, no. 8, pp. 1203–1211, 1972.

[104] P. Louka, G. Galanis, N. Siebert, G. Kariniotakis, P. Katsafados, I. Pytharoulis, and G. Kallos, "Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 96, no. 12, pp. 2348–2362, 2008.

[105] S. Salcedo-Sanz, Á. M. Pérez-Bellido, E. G. Ortiz-García, A. Portilla-Figueras, L. Prieto, and F. Correoso, "Accurate short-term wind speed prediction by exploiting diversity in input data using banks of artificial neural networks," *Neurocomputing*, vol. 72, no. 4-6, pp. 1336–1341, 2009.

[106] H. A. Nielsen, T. S. Nielsen, H. Madsen, M. J. San Isidro Pindado, and I. Marti, "Optimal combination of wind power forecasts," *Wind Energy*, vol. 10, no. 5, pp. 471–482, 2007.

[107] C. P. Sweeney, P. Lynch, and P. Nolan, "Reducing errors of wind speed forecasts by an optimal combination of post-processing methods," *Meteorological Applications*, vol. 20, no. 1, pp. 32–40, 2013.

[108] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 4, pp. 915–920, 2009.

[109] J. Jung and R. P. Broadwater, "Current status and future advances for wind speed and power forecasting," *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 762–777, 2014.

[110] M. Hibon and T. Evgeniou, "To combine or not to combine: Selecting among forecasts and their combinations," *International Journal of Forecasting*, vol. 21, no. 1, pp. 15–24, 2005.

[111] A. Botterud, Z. Zhou, J. Wang, J. Sumaili, H. Keko, J. Mendes, R. J. Bessa, and V. Miranda, "Demand dispatch and probabilistic wind power forecasting in unit commitment and economic dispatch: A case study of Illinois," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 1, pp. 250–261, 2013.

[112] J. Wang, A. Botterud, R. Bessa, H. Keko, L. Carvalho, D. Issicaba, J. Sumaili, and V. Miranda, "Wind power forecasting uncertainty and unit commitment," *Applied Energy*, vol. 88, no. 11, pp. 4014–4023, 2011.

[113] M. A. Matos and R. J. Bessa, "Setting the operating reserve using probabilistic wind power forecasts," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 594–603, 2011.

[114] R. J. Bessa, M. A. Matos, I. C. Costa, L. Bremermann, I. G. Franchin, R. Pestana, N. Machado, H. Waldl, and C. Wichmann, "Reserve setting and steady-state security assessment using wind power uncertainty forecast: A case study," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 4, pp. 827–836, 2012.

[115] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, 2014.

[116] J. Dobschinski, R. Bessa, P. Du, K. Geisler, S. E. Haupt, M. Lange, C. Mohrlen, D. Nakafuji, and M. De La Torre Rodriguez, "Uncertainty Forecasting in a Nutshell: Prediction Models Designed to Prevent Significant Errors," *IEEE Power and Energy Magazine*, vol. 15, no. 6, pp. 40–49, 2017.

[117] M. Lange, "On the uncertainty of wind power predictions - Analysis of the forecast accuracy and statistical distribution of errors," *Journal of Solar Energy Engineering, Transactions of the ASME*, vol. 127, no. 2, pp. 177–184, 2005.

[118] H. Bludszuweit, J. A. Domínguez-Navarro, and A. Llombart, "Statistical analysis of wind power forecast error," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 983–991, 2008.

[119] N. Zhang, C. Kang, Q. Xia, and J. Liang, "Modeling conditional forecast error for wind power in generation scheduling," *IEEE Transactions on Power Systems*, 2014.

[120] K. Bruninx and E. Delarue, "A statistical description of the error on wind power forecasts for probabilistic reserve sizing," *IEEE Transactions on Sustainable Energy*, 2014.

[121] B. M. Hodge, A. Florita, K. Orwig, D. Lew, and M. Milligan, "A comparison of wind power and load forecasting error distributions," *World Renewable Energy Forum, WREF 2012, Including World Renewable Energy Congress XII and Colorado Renewable Energy Society (CRES) Annual Conferen*, vol. 3, no. July, pp. 1850–1857, 2012.

[122] P. Pinson and G. Kariniotakis, "Conditional prediction intervals of wind power generation," *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1845–1856, 2010.

[123] J. B. Bremnes, "Probabilistic wind power forecasts using local quantile regression," *Wind Energy*, vol. 7, no. 1, pp. 47–54, 2004.

[124] R. J. Bessa, V. Miranda, A. Botterud, J. Wang, and E. M. Constantinescu, "Time adaptive conditional kernel density estimation for wind power forecasting," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 4, pp. 660–669, 2012.

[125] G. Sideratos and N. D. Hatziargyriou, "Probabilistic wind power forecasting using radial basis function neural networks," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1788–1796, 2012.

[126] G. Giebel, L. Landberg, J. Badger, K. Sattler, H. Feddersen, T. Nielsen, and H. Madsen, "Using ensemble forecasting for wind power," in *Proceedings of the 2003 European Wind Energy Conference and Exhibition*, Madrid, 2003.

[127] J. W. Taylor, P. E. McSharry, and R. Buizza, "Wind power density forecasting using ensemble predictions and time series models," *IEEE Transactions on Energy Conversion*, vol. 24, no. 3, pp. 775–782, 2009.

[128] E. B. Iversen, J. M. Morales, J. K. Møller, and H. Madsen, "Short-term probabilistic forecasting of wind speed using stochastic differential equations," *International Journal of Forecasting*, vol. 32, no. 3, pp. 981–990, 2016.

[129] P. Pinson and G. Kariniotakis, "On-line assessment of prediction risk for wind power production forecasts," *Wind Energy*, vol. 7, no. 2, pp. 119–132, 2004.

[130] P. Pinson, H. A. Nielsen, H. Madsen, and G. Kariniotakis, "Skill forecasting from ensemble predictions of wind power," *Applied Energy*, vol. 86, no. 7-8, pp. 1326–1334, 2009.

[131] P. Pinson, H. Madsen, H. A. Nielsen, G. Papaefthymiou, and B. Klöckl, "From probabilistic forecasts to statistical scenarios of short-term wind power production," *Wind Energy*, vol. 12, no. 1, pp. 51–62, 2009.

[132] J. M. Morales, R. Mínguez, and A. J. Conejo, "A methodology to generate statistically dependent wind speed scenarios," *Applied Energy*, vol. 87, no. 3, pp. 843–855, 2010.

[133] T. S. Nielsen, A. Joensen, H. Madsen, L. Landberg, and G. Giebel, "A new reference for wind power forecasting," *Wind Energy*, vol. 1, no. 1, pp. 29–34, 1998.

[134] L. Landberg, "Short-term prediction of the power production from wind farms," *Journal of Wind Engineering and Industrial Aerodynamics*, 1999.

[135] I. Martí, G. Kariniotakis, P. Pinson, I. Sanchez, T. S. Nielsen, H. Madsen, G. Giebel, J. Usaola, A. M. Palomares, R. Brownsword, J. Tambke, U. Focken, M. Lange, G. Sideratos, and G. Descombes, "Evaluation of advanced wind power forecasting models - Results of the anemos project," in *European Wind Energy Conference and Exhibition 2006, EWEC 2006*, vol. 3, 2006, pp. 2473–2481.

[136] G. Kariniotakis, I. Martí, D. Casas, P. Pinson, T. S. Nielsen, H. Madsen, G. Giebel, J. Usaola, and I. Sanchez, "What performance can be expected by short-term wind power prediction models depending on site characteristics ?" *Forecast*, pp. 22–25, 2004.

[137] M. Lange and U. Focken, "State-of-the-art in Wind Power Prediction in Germany and International Developments," 2005.

[138] M. Lange and D. Heinemann, "Relating the uncertainty of short-term wind speed predictions to meteorological situations with methods from synoptic climatology," in *Proceedings of the European Wind Energy Conference EWEC, Madrid (2003)*, 2003, p. 7.

[139] S. Sperati, S. Alessandrini, P. Pinson, and G. Kariniotakis, "The "Wire" benchmarking exercise on short-term forecasting of renewable power forecasting," *Energies*, vol. 8, pp. 1–19, 2015.

[140] S. K. Aggarwal, L. M. Saini, and A. Kumar, "Short term price forecasting in deregulated electricity markets: A review of statistical models and key issues," *International Journal of Energy Sector Management*, vol. 3, no. 4, pp. 333–358, 2009.

[141] C. M. Ruibal and M. Mazumdar, "Forecasting the mean and the variance of electricity prices in deregulated markets," *IEEE Transactions on Power Systems*, vol. 23, no. 1, pp. 25–32, 2008.

[142] M. Burger, B. Graeber, and G. Schindlmayr, *Managing Energy Risk: An Integrated View on Power and Other Energy Markets*, 2007.

[143] R. Carmona and M. Coulon, "A survey of commodity markets and structural models for electricity prices," in *Quantitative Energy Finance: Modeling, Pricing, and Hedging in Energy and Commodity Markets*, 2014, pp. 41–83.

[144] F. E. Benth, J. S. Benth, and S. Koekebakker, *Stochastic Modeling of Electricity and Related Markets*, vol. 11 ed. Singapore: World Scientific, 2008.

[145] K. F. Chan, P. Gray, and B. van Campen, "A new approach to characterizing and forecasting electricity price volatility," *International Journal of Forecasting*, 2008.

[146] R. Becker, S. Hurn, and V. Pavlov, "Modelling spikes in electricity prices," *Economic Record*, 2007.

[147] R. Weron, *Modeling and forecasting electricity loads and prices: A statistical approach*. Chichester: Wiley, 2006.

[148] S. K. Aggarwal, L. M. Saini, and A. Kumar, "Electricity price forecasting in deregulated markets: A review and evaluation," *International Journal of Electrical Power and Energy Systems*, vol. 31, no. 1, pp. 13–22, 2009.

[149] J. Janczura, S. Trück, R. Weron, and R. C. Wolff, "Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling," *Energy Economics*, vol. 38, pp. 96–110, 2013.

[150] A. J. Conejo, J. Contreras, R. Espínola, and M. A. Plazas, "Forecasting electricity prices for a day-ahead pool-based electric energy market," *International Journal of Forecasting*, vol. 21, no. 3, pp. 435–462, 2005.

[151] A. Cruz, A. Muñoz, J. L. Zamora, and R. Espínola, "The effect of wind generation and weekday on Spanish electricity spot price forecasting," *Electric Power Systems Research*, vol. 81, no. 10, pp. 1924–1935, 2011.

[152] G. E. Box, G. M. Jenkins, and G. C. Reinsel, *Time series analysis: Forecasting and control*, 4th ed., 2013.

[153] H. Tong and K. S. Lim, "Threshold autoregression, limit cycles and cyclical data," in *Exploration of a Nonlinear World: An Appreciation of Howell Tong's Contributions to Statistics*, 2009, pp. 9–56.

[154] R. A. Stine and H. Tong, "Non-linear Time Series: A Dynamical System Approach." *Journal of the American Statistical Association*, vol. 87, no. 419, p. 903, 1992.

[155] C. Lucheroni, "Tarx Models for Spikes and Antispikes in Electricity Markets," *SSRN Electronic Journal*, 2012.

[156] ——, "SETARX models for spikes and antispikes in electricity prices," in *Proceedings - International Symposium: Modern Electric Power Systems, MEPS'10*, 2010.

[157] R. F. Engle, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, vol. 50, no. 4, p. 987, 1982.

[158] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 01, pp. 307–327, 1986.

[159] J. C. Cuaresma, J. Hlouskova, S. Kossmeier, and M. Obersteiner, "Forecasting electricity spot-prices using linear univariate time-series models," *Applied Energy*, vol. 77, no. 1, pp. 87–106, 2004.

[160] A. J. Conejo, M. A. Plazas, R. Espínola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1035–1042, 2005.

[161] C. García-Martos, J. Rodríguez, and M. J. Sánchez, "Mixed models for short-run forecasting of electricity prices: Application for the Spanish market," *IEEE Transactions on Power Systems*, vol. 22, no. 2, pp. 544–552, 2007.

[162] J. Contreras, R. Espínola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1014–1020, 2003.

[163] F. J. Nogales, J. Contreras, A. J. Conejo, and R. Espínola, "Forecasting next-day electricity prices by time series models," *IEEE Transactions on Power Systems*, vol. 17, no. 2, pp. 342–348, 2002.

[164] F. J. Nogales and A. J. Conejo, "Electricity price forecasting through transfer function models," *Journal of the Operational Research Society*, vol. 57, no. 4, pp. 350–356, 2006.

[165] C. R. Knittel and M. R. Roberts, "An empirical examination of restructured electricity prices," *Energy Economics*, vol. 27, no. 5, pp. 791–817, 2005.

[166] R. C. Garcia, J. Contreras, M. van Akkeren, and J. B. C. Garcia, "A GARCH forecasting model to predict day-ahead electricity prices," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 867–874, 2005.

[167] R. Weron and A. Misiorek, "Short-term electricity price forecasting with time series models: A review and evaluation," in *Complex Electricity Markets*, 2006, pp. 231–254.

[168] ——, "Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models," *International Journal of Forecasting*, vol. 24, no. 4, pp. 744–763, 2008.

[169] T. Kristiansen, "Forecasting Nord Pool day-ahead prices with an autoregressive model," *Energy Policy*, vol. 49, pp. 328–332, 2012.

[170] R. Weron and A. Misiorek, "Forecasting Spot Electricity Prices with Time Series Models," *International Conference "The European Electricity Market EEM-05"*, p. 8, 2005.

[171] A. Misiorek, S. Trueck, and R. Weron, "Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models," *Studies in Nonlinear Dynamics and Econometrics*, vol. 10, no. 3, 2006.

[172] T. Jónsson, P. Pinson, H. A. Nielsen, H. Madsen, and T. S. Nielsen, "Forecasting electricity spot prices accounting for wind power predictions," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 1, pp. 210–218, 2013.

[173] N. V. Karakatsani and D. W. Bunn, "Forecasting electricity prices: The impact of fundamentals and time-varying coefficients," *International Journal of Forecasting*, vol. 24, no. 4, pp. 764–785, 2008.

[174] E. Raviv, K. E. Bouwman, D. Van Dijk, E. Raviv, K. E. Bouwman, and D. Van Dijk, "Forecasting day-ahead electricity prices: utilizing hourly prices European workshop on electricity price forecasting Forecasting day-ahead electricity prices," *Tinbergen Institute*, vol. 68, no. III, 2013.

[175] D. W. Bunn, "Forecasting loads and prices in competitive power markets," *Proceedings of the IEEE*, vol. 88, no. 2, pp. 163–169, 2000.

[176] M. Shahidehpour, H. Yamin, and Z. Li, *Market operations in electric power systems: forecasting, scheduling, and risk management*. John Wiley & Sons, 2002.

[177] S. Fan and R. J. Hyndman, "Short-term load forecasting based on a semi-parametric additive model," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 134–141, 2012.

[178] G. Klæboe, A. L. Eriksrud, and S. E. Fleten, "Benchmarking time series based forecasting models for electricity balancing market prices," *Energy Systems*, vol. 6, no. 1, pp. 43–61, 2013.

[179] Y. Ma, P. Luh, K. Kasiviswanathan, and E. Ni, "A neural network-based method for forecasting zonal locational marginal prices," in *2004 IEEE Power Engineering Society General Meeting*, vol. 1, 2004, pp. 296–302.

[180] M. Olsson and L. Söder, "Modeling real-time balancing power market prices using combined SARIMA and Markov processes," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 443–450, 2008.

[181] M. O. Brolin and L. Söder, "Modeling Swedish real-time balancing power prices using nonlinear time series models," in *2010 IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems, PMAPS 2010*, 2010, pp. 358–363.

[182] S. Jaehnert, H. Farahmand, and G. L. Doorman, "Modelling of prices using the volume in the Norwegian regulating power market," in *2009 IEEE Bucharest PowerTech: Innovative Ideas Toward the Electrical Grid of the Future*, 2009.

[183] K. Skytte, "The regulating power market on the Nordic power exchange Nord Pool: An econometric analysis," *Energy Economics*, vol. 21, no. 4, pp. 295–308, 1999.

[184] Y. Ji, J. Kim, R. J. Thomas, and L. Tong, "Forecasting real-time locational marginal price: A state space approach," in *Conference Record - Asilomar Conference on Signals, Systems and Computers*, 2013, pp. 379–383.

[185] T. K. Boomsma, N. Juul, and S. E. Fleten, "Bidding in sequential electricity markets: The Nordic case," *European Journal of Operational Research*, vol. 238, no. 3, pp. 797–809, 2014.

[186] T. Jónsson, P. Pinson, H. A. Nielsen, and H. Madsen, "Exponential smoothing approaches for prediction in real-time electricity markets," *Energies*, vol. 7, no. 6, pp. 3710–3732, 2014.

[187] I. Dimoulkas, M. Amelin, and M. R. Hesamzadeh, "Forecasting balancing market prices using Hidden Markov Models," in *International Conference on the European Energy Market, EEM*, vol. 2016-July, 2016.

[188] G. Kouemou and D. Dymarski, "History and Theoretical Basics of Hidden Markov Models," in *Hidden Markov Models, Theory and Applications*. IntechOpen, 2011.

[189] P. Pinson, "Estimation of the Uncertainty in Wind Power Forecasting," Ph.D. dissertation, 2006.

[190] P. Artzner, F. Delbaen, J. M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

[191] R. Clemen and T. Reilly, "Making hard decisions with DecisionTools," *South-Western*, p. 848, 2014.

[192] A. J. Conejo, M. Carrión, and J. M. Morales, *Decision making under uncertainty in electricity markets*. New York: Springer, 2010.

[193] M. Zugno, T. Jónsson, and P. Pinson, "Trading wind energy based on probabilistic forecasts of wind generation and market quantities," *Wind Energy*, vol. 16, no. 6, pp. 909–926, 2012.

[194] J. Matevosyan and L. Söder, "Minimization of imbalance cost trading wind power on the short-term power market," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1396–1404, 2006.

[195] C. J. Dent, J. W. Bialek, and B. F. Hobbs, "Opportunity cost bidding by wind generators in forward markets: Analytical results," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1600–1608, 2011.

[196] M. Rahimiyan, J. M. Morales, and A. J. Conejo, "Evaluating alternative offering strategies for wind producers in a pool," *Applied Energy*, vol. 88, no. 12, pp. 4918–4926, 2011.

[197] N. Mazzi and P. Pinson, "Wind power in electricity markets and the value of forecasting," *Renewable Energy Forecasting: From Models to Applications*, no. 2007, pp. 259–278, 2017.

[198] A. Botterud, Z. Zhou, J. Wang, R. J. Bessa, H. Keko, J. Sumaili, and V. Miranda, "Wind power trading under uncertainty in LMP markets," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 894–903, 2012.

[199] J. Browell, "Risk constrained trading strategies for stochastic generation with a single-price balancing market," *Energies*, vol. 11, no. 6, 2018.

[200] Department of Communications; Climate Action & Environment, "REFIT Schemes and Supports." [Online]. Available: https://www.dccae.gov.ie/en-ie/energy/ [Accessed: 2020-04-25]

[201] C. Möhrlen, J. Zack, J. Messner, and J. Browell, "Expert Group Report on Recommended Practices for Selecting Renewable Power Forecasting Solutions - Part 3: Evaluation of Forecasts and Forecast Solutions," 2019. [Online]. Available: http://www.ieawindforecasting.dk/publications/recommendedpractice [Accessed: 2020-04-25]

[202] G. E. P. Box and D. R. Cox, "An Analysis of Transformations," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 26, no. 2, pp. 211–243, 1964.

[203] R. J. Hyndman and Y. Khandakar, "Automatic time series forecasting: The forecast package for R," *Journal of Statistical Software*, vol. 27, no. 3, pp. 1–22, 2008.

[204] D. Kwiatkowski, P. C. Phillips, P. Schmidt, and Y. Shin, "Testing the null hypothesis of stationarity against the alternative of a unit root," *Journal of Econometrics*, vol. 54, no. 1-3, pp. 159–178, 1992.

[205] M. Stone, "An Asymptotic Equivalence of Choice of Model by Cross-Validation and Akaike's Criterion," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 44–47, 1977.

[206] P. J. Brockwell and R. A. Davis, "Introduction to Time Series and Forecasting," *Biometrics*, vol. 54, no. 3, p. 1204, 1998.

[207] S. N. Lahiri, "Resampling Methods for Dependent Data." New York, USA: Springer, 2003, pp. 1–16.

[208] R. L. Winkler, "Evaluating Probabilities: Asymmetric Scoring Rules," *Management Science*, vol. 40, no. 11, pp. 1395–1405, 1994.

[209] N. Metropolis and S. Ulam, "The Monte Carlo Method," *Journal of the American Statistical Association*, vol. 44, no. 247, p. 335, 1949.

[210] R. Liesenfeld and J. F. Richard, "Monte Carlo Methods and Bayesian Computation: Importance Sampling," in *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*, 2015, pp. 758–762.

[211] T. E. Oliphant, "Python for Scientific Computing Python Overview," *Computing in Science and Engineering*, pp. 10–20, 2007.

[212] W. McKinney, "Data Structures for Statistical Computing in Python," *Proceedings of the 9th Python in Science Conference*, vol. 1697900, no. Scipy, pp. 51–56, 2010.

[213] A. J. Dobson, *An Introduction to Generalized Linear Models*. London: Chapman and Hall, 1990.

[214] W. N. Venables and B. D. Ripley, *Modern Applied Statistics with S*, New York, 2002, vol. 53, no. March.

[215] R Core Team, "R: A Language and Environment for Statistical Computing," Vienna, Austria, 2019. [Online]. Available: https://www.r-project.org/

[216] T. Gneiting, "Quantiles as optimal point forecasts," *International Journal of Forecasting*, vol. 27, no. 2, pp. 197–207, 2011.

[217] K. Maciejowska and J. Nowotarski, "A hybrid model for GEFCom2014 probabilistic electricity price forecasting," *International Journal of Forecasting*, vol. 32, no. 3, pp. 1051–1056, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.ijforecast.2015.11.008

[218] F. X. Diebold, J. Hahn, and A. S. Tay, "Multivariate density forecast evaluation and calibration in financial risk management: High-frequency returns on foreign exchange," *Review of Economics and Statistics*, 1999.

[219] SEM-O, "Imbalance Pricing," 2017. [Online]. Available: https://www.sem-o.com/training/modules/imbalance-pricing/ [Accessed: 2020-04-25]

[220] SEMO, EirGrid, and SONI, "Report on the Imbalance Prices calculated on 24/01/2019," 2019. [Online]. Available: https://www.sem-o.com/publications/general-publications [Accessed: 2020-04-25]

Appendix A

Imbalance Pricing

The "flagging-and-tagging" approach (similar to that employed in the GB market) is implemented for imbalance pricing in the I-SEM [219]. The process has the following objectives:

- a single imbalance settlement price for imbalances in all directions and all energy balancing actions should be determined;
- the price should be marginal and based on the actions taken by the TSO to balance the system; and
- non-energy actions should be settled pay-as-bid.

A different price is calculated for each 5-minute Imbalance Pricing Period (IPP), and the average of the six 5-minute prices within a half hour is set as the price for the Imbalance Settlement Period. The flagging-and-tagging process consists of the following steps:

1. Ranking;
2. Flagging;
3. Marginal Energy Action Price;
4. Tagging;

5. Imbalance Price; and

6. Market Back-Up Price.

Each of these steps will be explained next, together with a worked example of how the Imbalance Price is determined for a set of accepted bids and offers. The example is based on the one in Appendix A of the "Industry Guide to the I-SEM" document [24].

1. Ranking

All accepted bids and offers (BOA) are ranked in order of economic merit:

- First, all accepted bids (also called *decs*, i.e. generation/supply offers to decrease output) are ranked in order of increasing price from lowest to highest.
- Next, all accepted offers (also called *incs*, i.e. generation/supply offers to increase output) are ranked in order of increasing price from lowest to highest.
- All actions are assigned a rank number k starting at 1 from the lowest priced dec action.

Example

All accepted bids and offers are ranked by price.

Rank	**Type**	**Price** [€/MWh]	**QBOA** [MWh]
15	Offer	100	10
14	Offer	90	20
13	Offer	80	1
12	Offer	70	5
11	Offer	60	7.5
10	Offer	50	5
9	Offer	40	2.5
8	Offer	30	5
7	Offer	20	10
6	Bid	50	–5
5	Bid	40	–10

4	Bid	30	–1.5
3	Bid	20	–10
2	Bid	10	–10
1	Bid	5	–5

2. Flagging

The flagging process identifies the BOA that are caused by system or unit constraints using two methods. The *System Operator* (SO) flagging identifies and flags non-energy actions taken for system reasons, such as binding network or operational constraints. The aim is to minimise the extent of non-energy actions influencing or setting the final price. The *Non-Marginal* flagging determines if the unit's ability to respond to changing demand is limited by its physical output limits and flags the corresponding BOA if that is the case. For example, this can occur when the unit is at its minimum operating limit, at its maximum capacity, or its output is restricted by its ramp limits.

The initial value for each flag is 1, indicating that the BOA is not excluded from setting the imbalance price. The value is switched to 0, i.e. the BOA is "flagged", when the BOA satisfies the conditions of the flag. The imbalance price flag (IPF) then is calculated as the product of the SO and Non-Marginal flags, and BOA with an IPF value of 1 is said to be "unflagged". The aim is to minimise the extent of non-energy actions and non-marginal actions influencing or setting the final price.

Example (continued)

Actions are SO and Non-Marginal flagged, and the IPF is calculated as the product of these two flags.

Rank	**Type**	**SO Flag**	**Non-Marginal Flag**	**Imbalance Price Flag**	**Reason for flagging**
15	Offer	0	1	0	Reserve
14	Offer	0	1	0	Congestion
13	Offer	1	1	1	
12	Offer	1	1	1	
11	Offer	0	1	0	Congestion

10	Offer	1	1	1	
9	Offer	1	1	1	
8	Offer	1	0	0	
7	Offer	1	0	0	At max output
6	Bid	1	1	1	
5	Bid	0	1	0	Congestion
4	Bid	1	1	1	
3	Bid	1	0	0	Ramp limited
2	Bid	0	1	0	Reserve
1	Bid	1	1	1	

3. Marginal Energy Action Price

The Net Imbalance Volume Quantity (QNIV) is calculated as the sum of all accepted bid and offer quantities, with bid quantities being negative and offer quantities being positive. If QNIV is positive, the system was *short*: there was too little generation so the TSO had to take more positive inc actions than negative dec actions. If QNIV is negative, the system was *long*: there was too much generation so the TSO had to take more negative dec actions than positive inc actions.

The Marginal Energy Action Price (PMEA) is the most *expensive* unflagged action in the ranked set. This means that:

- when QNIV is positive, PMEA is the highest priced unflagged offer;
- when QNIV is negative, PMEA is the lowest priced unflagged bid.

This value sets the reference price for the period. All actions whose prices are more expensive than PMEA have their reference price set equal to PMEA. For all other actions, the reference price is equal to the original price. This ensures that the price of the actions which were not in merit for setting the marginal price are not considered further in the process. Consequently, the final price cannot be set less economic than PMEA:

- when QNIV is positive, the price can be set equal or lower than PMEA;
- when QNIV is negative, the price can be set equal or higher than PMEA.

Note that through the PMEA and replacement price process, it is possible for an action which was in the opposite direction to the QNIV to set the imbalance price, if it was the marginal action. For example, if the QNIV is positive but all accepted offers were flagged, the process will select the highest priced accepted bid to be the Marginal Energy Action.

Example (continued)

The QNIV is positive and equal to 24.5 MWh. Offer 13 is the highest priced unflagged offer and therefore is the marginal energy action. The corresponding price of 80 €/MWh sets the PMEA and all higher prices are capped at this value (see offers 14 and 15).

Rank	**Type**	**Price** [€/MWh]	**QBOA** [MWh]	**Imbalance Price Flag**	**Reference Price** [€/MWh]
15	Offer	100	10	0	80
14	Offer	90	20	0	80
13	Offer	80	1	1	80
12	Offer	70	5	1	70
11	Offer	60	7.5	0	60
10	Offer	50	5	1	50
9	Offer	40	2.5	1	40
8	Offer	30	5	0	30
7	Offer	20	10	0	20
6	Bid	50	–5	1	50
5	Bid	40	–10	0	40
4	Bid	30	–1.5	1	30
3	Bid	20	–10	0	20
2	Bid	10	–10	0	10
1	Bid	5	–5	1	5

4. Tagging

The tagging process identifies which BOA are excluded from the imbalance price calculation. Two tagging methods are used: *net imbalance volume* (NIV) tagging and *price average reference* (PAR) tagging.

The *NIV tagging* has the purpose of ensuring that there are sufficient untagged

actions to meet the QNIV. Initially, all NIV tag values are set equal to 1. First, all actions in the opposite direction to the QNIV are tagged, i.e. their value is switched from 1 to 0. Next, all actions in the same direction to the QNIV which were previously flagged are tagged. At this point, the *Residual Tagged Quantity* (QRTAG) is calculated as the sum of the volume of currently "tagged" actions, changed of sign. Depending on the signs of QNIV and QRTAG, a volume of actions equal to QRTAG is tagged or untagged. If QRTAG is negative, more actions need to be untagged, with the least expensive getting untagged first until the volume of the actions untagged equals QRTAG:

- if QNIV > 0, the lowest priced incs get untagged first.
- if QNIV < 0, the highest priced decs get untagged first.

If QRTAG is positive, more actions need to be tagged, with the most expensive getting tagged first until the volume of the actions tagged equals QRTAG:

- if QNIV > 0, the highest priced incs get tagged first.
- if QNIV < 0, the lowest priced decs get tagged first.

If a fraction of the offer is tagged or untagged to reach QRTAG, the corresponding value between 0 and 1 is assigned to the NIV tag.

As a result of the NIV tagging, the volume of untagged actions now equals QNIV.

Example (continued)

Because QNIV is positive, the Initial NIV tag is set equal to 0 for all bids and all price flagged offers. The QRTAG is then calculated and found equal to –11 MWh, so more actions need to be untagged. As QNIV is positive, the lowest priced offers (incs) get untagged first until the untagged actions sum to QNIV. This results in offer 7 and, partially, offer 8 being untagged.

Rank	Type	Reference Price [€/MWh]	Imbalance Price Flag	QBOA [MWh]	Initial NIV Tag	NIV Tag
15	Offer	80	0	10	0	0
14	Offer	80	0	20	0	0
13	Offer	80	1	1	1	1
12	Offer	70	1	5	1	1
11	Offer	60	0	7.5	0	0
10	Offer	50	1	5	1	1
9	Offer	40	1	2.5	1	1
8	Offer	30	0	5	0	0.2
7	Offer	20	0	10	0	1
6	Bid	50	1	–5	0	0
5	Bid	40	0	–10	0	0
4	Bid	30	1	–1.5	0	0
3	Bid	20	0	–10	0	0
2	Bid	10	0	–10	0	0
1	Bid	5	1	–5	0	0

The *PAR Tagging* has the purpose of determining what actions which are not NIV tagged, and what volume of those actions, are considered in the final calculation of the imbalance price. The *Price Average Reference Quantity* (QPAR) for an Imbalance Pricing Period has been defined equal to 10 MWh by market design [219]. This is equivalent to 60 MWh for an Imbalance Settlement Period.

Initially, all PAR tag values are set equal to the NIV tag. If QNIV > 0, the highest priced untagged BOA offers whose sum is equal to QPAR are identified. Their PAR tag is changed to a value between 0 and 1 to indicate the fraction of volume required to have a total volume of QPAR. The PAR tag for all lower priced BOA offers is set to 0. If QNIV < 0, the lowest priced untagged BOA bids whose sum is equal to QPAR are identified. Their PAR tag is changed to a value between 0 and 1 to indicate the fraction of volume required to have a total volume of QPAR. The PAR tag for all higher priced BOA offers is set to 0.

Once the NIV and PAR tags are defined, the *Imbalance Price tag* (TIP) is calculated as the product between the two tags. Therefore, if a BOA is either NIV

tagged or PAR tagged it will have no impact on the imbalance price calculation.

Example (continued)

Because QNIV is positive, the highest priced 10 MWh offers with NIV tag of 1 are PAR tagged. In this case, offers 13, 12, and 10 partially. The Imbalance Price tag is then calculated as the product of the NIV and PAR tags.

Rank	**Type**	**Ref Price** [€/MWh]	**NIV Tag**	**QBOA** [MWh]	**Untagged quantities**	**PAR Tag**	**Imbalance Price Tag**
15	Offer	80	0	10	0	0	0
14	Offer	80	0	20	0	0	0
13	Offer	80	1	1	1	1	1
12	Offer	70	1	5	5	1	1
11	Offer	60	0	7.5	0	0	0
10	Offer	50	1	5	5	0.8	0.8
9	Offer	40	1	2.5	2.5	0	0
8	Offer	30	0.2	5	1	0	0
7	Offer	20	1	10	10	0	0
6	Bid	50	0	–5	0	0	0
5	Bid	40	0	–10	0	0	0
4	Bid	30	0	–1.5	0	0	0
3	Bid	20	0	–10	0	0	0
2	Bid	10	0	–10	0	0	0
1	Bid	5	0	–5	0	0	0

5. Imbalance Price

The Imbalance Price is calculated as the volume-weighted average price of the untagged actions. One of the main reasons why PAR tagging is weighted towards an average price approach rather than a marginal price approach is to mitigate volatility in the prices. The volume weight for each BOA is calculated as the product of the QBOA by the TIP. These volumes are then used to weight the corresponding BOA prices and calculate the final Imbalance Price (PIMB). In mathematical notation:

$$PIMB = \frac{\sum price_i \cdot QBOA_i \cdot TIP_i}{\sum QBOA_i \cdot TIP_i}$$

This price is then used for settlement unless the price cap or price floor are exceeded. If the imbalance price is greater than the price cap, which was set equal to 10,000 €/MWh, then PIMB is set equal to the price cap. On the contrary, if the imbalance price is lower than the price floor, which was set equal to –1,000 €/MWh, PIMB is set equal to the price floor. Finally, the Imbalance Settlement Price is calculated as the average of the six PIMB in each (5-minute) Imbalance Pricing Period.

Example (continued)

The Imbalance Price is calculated as the volume-weighted average of the untagged prices and found equal to 63 €/MWh (= 630/10).

Rank	**Type**	**Ref Price** [€/MWh]	**QBOA** [MWh]	**Imbalance Price Tag**	**RP*QBOA*TIP**	**QBOA*TIP**
15	Offer	80	10	0	0	0
14	Offer	80	20	0	0	0
13	Offer	80	1	1	80	1
12	Offer	70	5	1	350	5
11	Offer	60	7.5	0	0	0
10	Offer	50	5	0.8	200	4
9	Offer	40	2.5	0	0	0
8	Offer	30	5	0	0	0
7	Offer	20	10	0	0	0
6	Bid	50	–5	0	0	0
5	Bid	40	–10	0	0	0
4	Bid	30	–1.5	0	0	0
3	Bid	20	–10	0	0	0
2	Bid	10	–10	0	0	0
1	Bid	5	–5	0	0	0

6. Market Back-Up Price

For each ISP, the Market Operator calculates the Market Back-Up Price (PMBU) as the quantity-weighted average price associated with each day-ahead traded quantity and intraday traded quantity for all units. The PMBU is used in case of:

- QNIV=0;
- failure in the pricing system;

- administered settlement;
- settlement where a curtailment instruction has been issued.

Appendix B

Imbalance Prices on 24 January 2019

On the 24th January 2019, the 5-minute Imbalance Prices exceeded the strike price of €500 in 11 Imbalance Pricing Periods. This resulted in the half-hour Imbalance Settlement Price being at levels above the strike price in four Imbalance Settlement Periods. EirGrid, SONI, and SEMO have undertaken a detailed examination of the events leading up to and during the 24th January which are laid out in detail in a report published on the 21st February [220]. A summary of this report is presented below, providing an outline of the drivers behind the operation of the Balancing Market on that date, the actions taken by the TSOs, how these actions were treated in the flagging-and-tagging process, and how they eventually led to the Imbalance Price and Imbalance Settlement Price calculations.

The ex-ante auctions operated by SEMOpx functioned normally for Trading Day 24/01/2019. Given the wind forecast versus load forecast, the ex-ante markets closed with more production cleared than consumption in a number of trading periods. Such generation surplus drives exports from the I-SEM to adjacent coupled markets across the Moyle and EWIC interconnectors. Due to the application of losses in the EUPHEMIA algorithm, this generally results in the Moyle interconnector being scheduled first until the price spread between the I-SEM and any

coupled markets exceeds the loss value on both interconnectors.

Because the North-South tie line is modelled as an unlimited inter-area in the ex-ante markets, no location-based restrictions on exports from the I-SEM are applied on them. This resulted in high levels of export being scheduled on the Moyle interconnector (see Figure B.1), effectively representing an export of wind production. However, the wind on the power system was mainly based in the ROI region with very little generation available in NI across the lunchtime period to meet the scheduled exports (see Figure B.2).

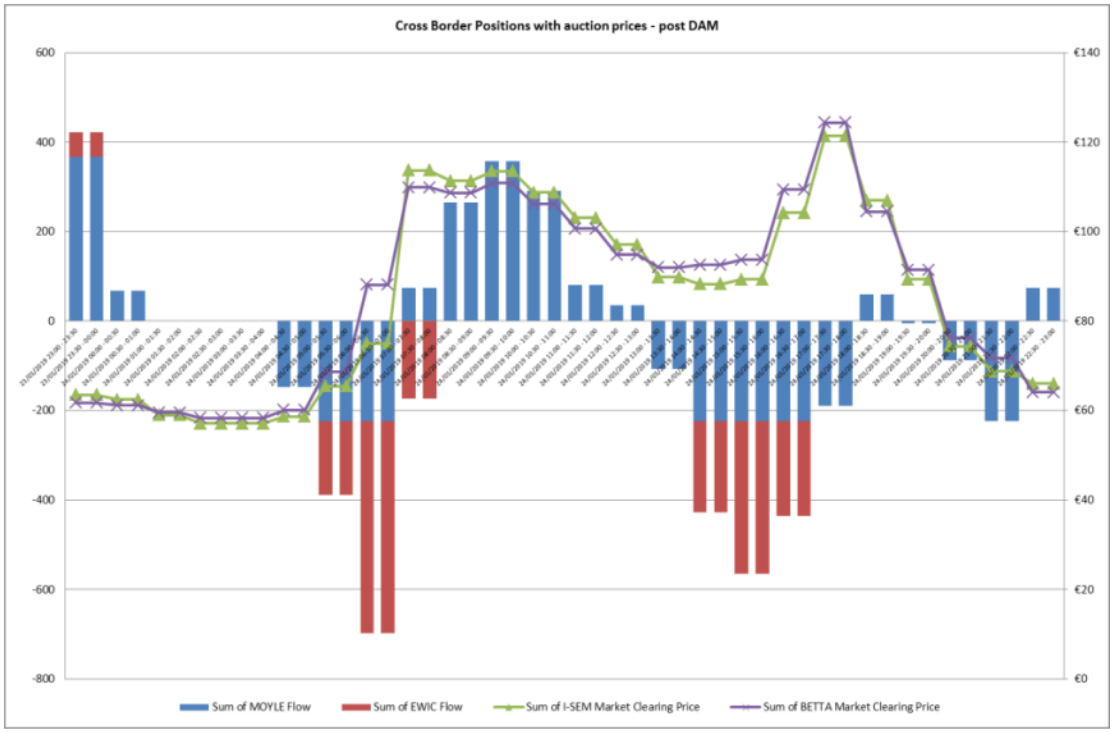


Figure B.1: Scheduled cross border flows on the Moyle (blue) and EWIC (red) interconnectors after DAM clearing. Positive values indicate import, negative indicate export. DAM clearing prices are also indicated for the I-SEM (green) and BETTA (purple) markets. Source: "Report on the Imbalance Prices calculated on 24/01/2019" [220].

In the early morning of the 23rd January, the unit GU_500040 was declared unavailable due to a technical issue, and it would subsequently remain unavailable for three days. The total output of the wind farms in NI on the 23rd was low. The Moyle interconnector was importing from GB for most of the day. With these operating conditions, all available conventional NI generation was scheduled to

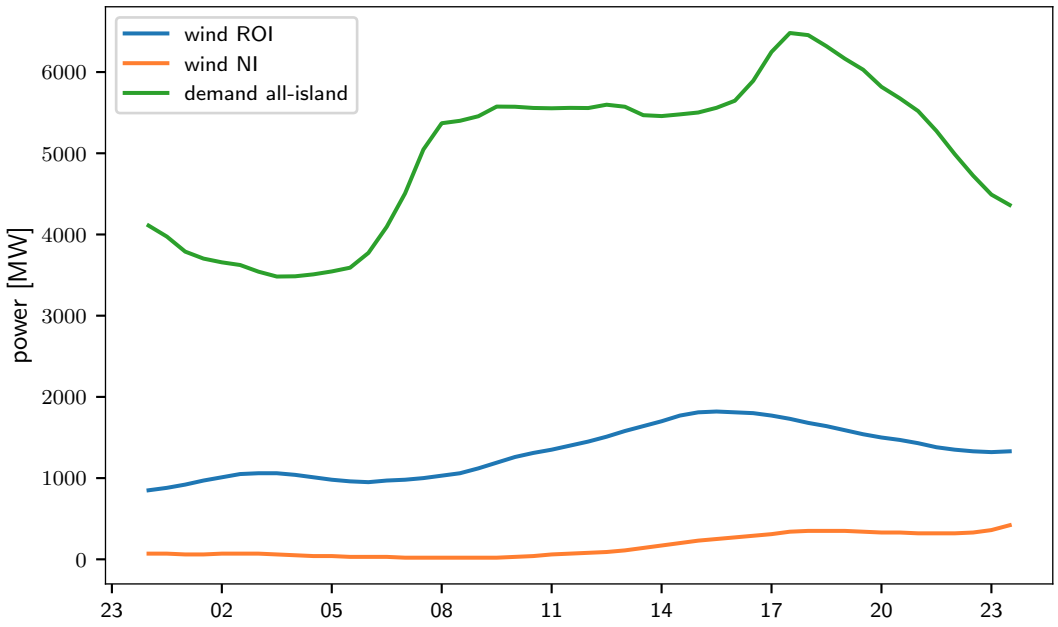


Figure B.2: Forecast wind generation in Republic of Ireland (blue) and Northern Ireland (orange), and forecast demand for the all-island system (green) on 24/01/2019.

meet the NI demand. On the 24$^{\text{th}}$ January, the NI wind remained low in the early part of the day as per forecast and was even below forecast for most of the morning. The wind in ROI was above forecast for most of the day but could not be used to help support the Moyle export or the demand in NI due to the tie line operational limits. The ROI to NI flow on the tie line schedules indicated that the maximum achievable south-to-north stable operational flows were occurring for large periods of the day. In order not to breach the stable operational limits on the tie line, the indicative operational schedules indicated the need to start the majority of NI fast-start units when Moyle was at its maximum export position during the day.

These conditions on the system, coupled with the system being long, resulted in the extreme fluctuation in the imbalance price throughout the day. This fluctuation was caused by two key components of the Trading and Settlement Code [36]: the price of the marginal energy action and the replaced bid offer price. Depending on whether the market is short or long, the most expensive or the cheapest unit that is neither SO nor Non-marginal flagged is set as the marginal energy action. The price of this unit then defines the Marginal Energy Action Price (PMEA) and

units that are less economic than this have their price replaced by the PMEA (cf. Appendix A).

The imbalance prices and net imbalance volumes for the period of interest on the 24th Januray are shown in Figure B.3. As a result of the System conditions in NI, the System Operators brought on two fast acting open cycle units (GU_500283 and GU_500284). These units had simple COD in at 6,341 €/MWh and 5,636 €/MWh, respectively. Both these units were issued instructions to come on at their lower operating limit at 10:46. The first impact of the two units on pricing can be seen immediately after they come off their lower operating limit, becoming marginal at 11:35. During this period, the all-island market was short and GU_500283 set the PMEA.

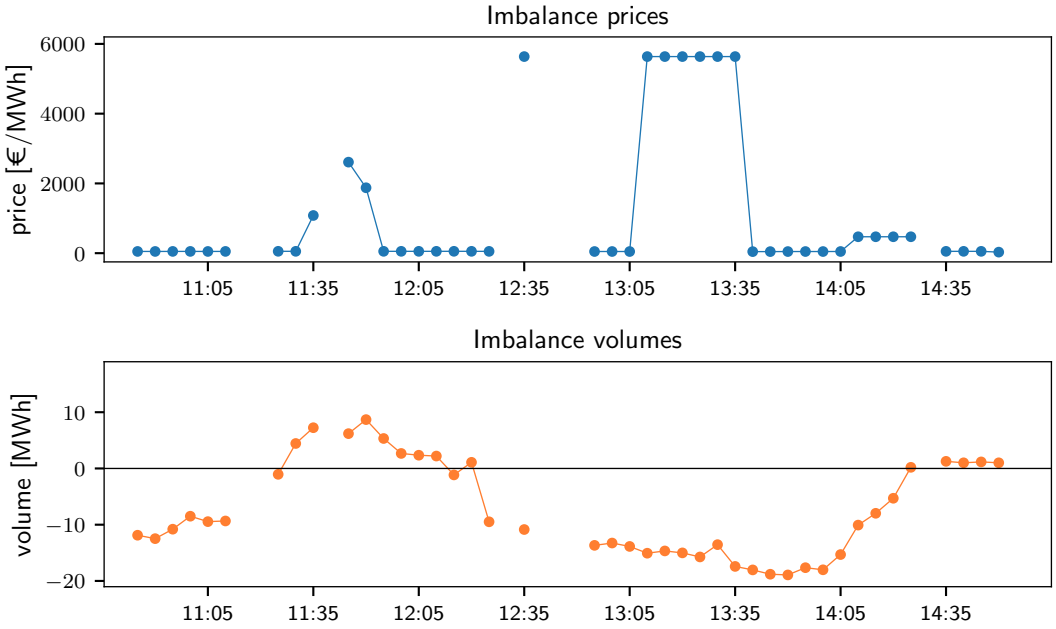


Figure B.3: Five-minute imbalance prices (top) and imbalance volumes (bottom) on 24/01/2019.

Shortly after the initial high price event, the all-island market switched from being short to long. During the following 40 minutes, GU_500283 and GU_500284 were non-marginal flagged in the indicative schedule as a result of a ramp constraint against these units. Although the units are fast acting, the schedule was trying to move them to a level above their dwell time breakpoint. The units must remain

at this level for a 5-minute period as part of ramping to a higher output level and were therefore unable to get to the higher level intended by the schedule, resulting in them being non-marginally flagged for ramping. Over this period, all ROI units were flagged out because of the tie-line constraint, while all NI units were being run at their operating limit and being flagged as non-marginal.

At 12:35, the indicative schedule moved GU_500824 away from its dwell time breakpoint for a 5 minute period. The bid stack had remained almost identical except for this unit becoming marginal. As a result, GU_500284 set the PMEA and the entire imbalance price at € 5,636.

For the next 30 minutes, GU_500284 was either at its higher operating limit or ramp constrained, again becoming non marginal-flagged in the indicative schedule. The physical conditions on the power system remained very consistent throughout this event. All ROI units remained SO flagged because of the limitations of the North-South tie line, while all NI units were non-marginal due to being run at their operating limit. At 13:10, GU_500824 was moved off its operating limit for the following 30 minutes in the indicative schedule, before being desynced at 13:35. This resulted in the GU_500824 becoming marginal and being the only unflagged unit in the bid stack for this window, thus setting the PMEA. This caused GU_500824 to set the entire 5-minute imbalance price at 5,636 €/MWh for the next six pricing periods. These six periods affected two imbalance settlement prices resulting in prices of € 3,773.69 and € 1,909.45 respectively.

These two units were desynced just after 13:30. Pricing remained volatile, with prices being much higher than a typical day and just below the strike price.

The detailed review of the flagging-and-tagging of units in the system outputs, which can be found in Section 8 of Reference [220], has concluded that flags and tags were applied correctly and none of the known defects in the Imbalance Pricing algorithm had an impact on the relevant periods.

Appendix C

Source Codes

All the `python` and `R` codes presented in the following Sections are available at `https://github.com/ggoretti/phd-thesis`; the website can also be accessed scanning the QR code below.

Scan this QR code

C.1 Electricity Price Forecasting

C.1.1 Day-ahead and imbalance price forecasting with ARMAX models

The input data used to forecast electricity prices are stored in distinct folders based on price and day-type. Historical data are separated between day-ahead and imbalance prices, and then again between weekday and weekend; the folder structure is illustrated in Figure C.1. Note that because the multi-model approach

is implemented — whereby each trading period is modelled separately — each .csv file corresponds to a different half-hour (that is, one file for all prices at H00:00, one for H00:30, etc.).

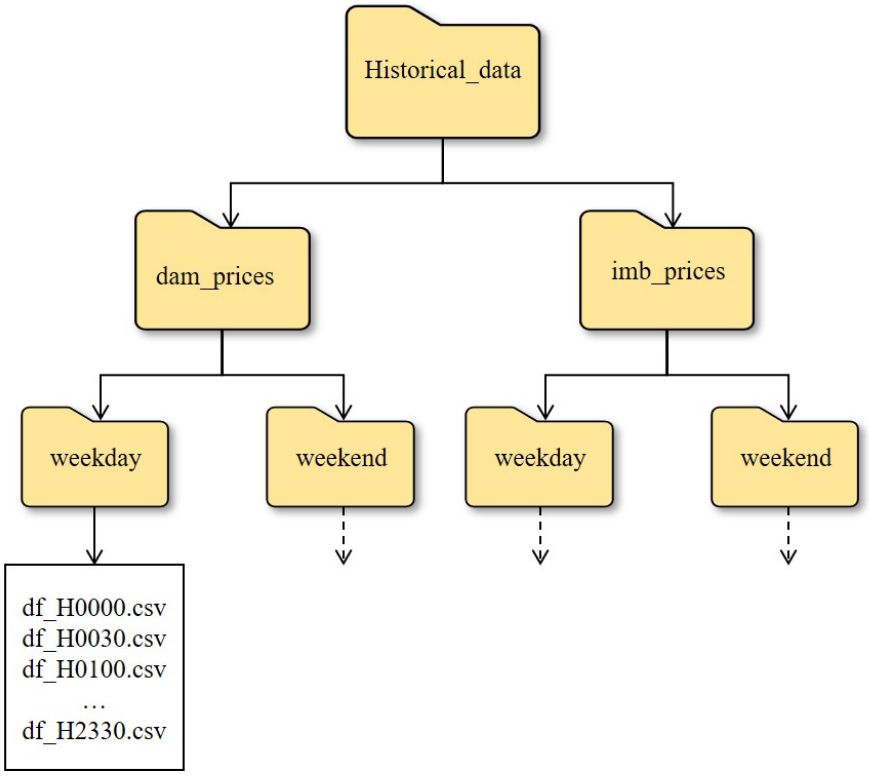


Figure C.1: Folder structure for price forecasting input data.

Recursively expanding approach

The following `R` script generates forecasts of *weekday* day-ahead prices adopting the "Recursively expanding" approach. To generate forecasts of *weekend* prices, the same code is used with `pattern` in line 8 set to `"weekend/*.csv"` and `freq` in line 10 set equal to 2. To generate forecasts of imbalance prices, the same code is used, changing the working directory in line 2 to `".../Historical_data/imb_prices"`

Note that the keys `"Demand_Fcst"`, `"p50_All"`, `"interval_All"`, `"penetration"` correspond to the predictors (i.e. the exogenous variables) used in the regression part of the ARIMAX models.

```r
# Set working directory
setwd(".../Historical_data/dam_prices")

# Import packages
library(fpp2, tsibble, readr, purrr)

# Import prices from half-hourly .csv files
temp = list.files(pattern="weekday/*.csv")
myfiles = lapply(temp, read.csv)
freq = 5    # period of seasonal component (weekday=5, weekend=2)

# Create empty lists to store outputs
dflist = list()
fit_list = list()
coef_list = list()

# Loop through the half-hourly files
for(j in 1:length(myfiles)){
  df_file <- data.frame(myfiles[j])
  df <- ts(df_file, frequency=freq)    # convert to Time series object

  # split data into initial Training and Test sets
  df_train <- subset(df, end=nrow(df)*0.5716)
  df_test <- subset(df, start=nrow(df)*0.5716 + 1)


  # - - - One-step forecast with re-selection and re-estimation of ARIMA model - - -
  h <- 2    # forecast horizon
  n <- nrow(df_test) - 1    # number of forecasts to issue
  fcmat <- matrix(0, nrow=n, ncol=14)    # empty matrix to store forecast values
  colnames(fcmat) <- c("point_1", "lower50_1", "lower80_1", "lower90_1",
                       "upper50_1", "upper80_1", "upper90_1",
                       "point_2", "lower50_2", "lower80_2", "lower90_2",
                       "upper50_2", "upper80_2", "upper90_2")

  fit_df = data.frame(aic=numeric(),
                      p=integer(), d=integer(), q=integer(),
                      P=integer(), D=integer(), Q=integer())    # empty df for AIC value and model order
  coef_df = list()    # empty list for coefficients of fitted model

  # *RECURSIVELY EXPANDING* Training set
  lambda <- BoxCox.lambda(df_train[,"dam_price"])    # BoxCox transformation lambda for Training data
  for(i in 1:n){
    x_i <- subset(df, end=nrow(df)*0.5716+i-1)    # increase x by (i-1)
    xreg_i <- subset(df, start=nrow(df)*0.5716+i)    # shift regressors by i
```

```
46
47      # fit ARIMA model to updated x and xreg
48      fit_i <- auto.arima(x_i[,"dam_price"],
49                          xreg=x_i[,c("Demand_Fcst", "p50_All",
                            ↪  "interval_All", "penetration")],
50                          lambda=lambda)
51
52      fit_df[i,1] <- fit_i$aic
53      fit_df[i,2:7] <- arimaorder(fit_i)[1:6]
54      coef_df <- rbind(coef_df, fit_i$coef)
55
56      # forecast with updated model
57      fcst <- forecast(fit_i, h=h,
58                       xreg=xreg_i[,c("Demand_Fcst", "p50_All",
                         ↪  "interval_All", "penetration")],
59                       level=c(50, 80, 90))
60
61      # forecast horizon: h = 1
62      fcmat[i,1] <- fcst$mean[1]
63      fcmat[i,2:4] <- fcst$lower[1,]
64      fcmat[i,5:7] <- fcst$upper[1,]
65
66      # forecast horizon: h = 2
67      fcmat[i,8] <- fcst$mean[2]
68      fcmat[i,9:11] <- fcst$lower[2,]
69      fcmat[i,12:14] <- fcst$upper[2,]
70    }
71
72    df_fcst <- df_file[(nrow(df_file)*0.5716+1):(nrow(df_file)-1),]
73    dflist[[j]] <- cbind(df_fcst, as.data.frame(fcmat))    # list of
      ↪  dataframes; each df includes forecast and observed values for a
      ↪  single load period
74
75    fit_list[[j]] <- fit_df    # list of dataframes with AIC and model
      ↪  order values
76
77    coef_list[[j]] <- as.data.frame(coef_df)    # list of dataframes with
      ↪  model coefficients
78  }
79
80  # Store results in separate data.frames
81  df_tot <- do.call(rbind, dflist)
82  fit_tot <- map_df(fit_list, ~as.data.frame(.x), .id="id")    # merge
    ↪  into data.frame keeping id of each list
83  coef_tot <- map_df(coef_list, ~as.data.frame(.x), .id="id")    # merge
    ↪  into data.frame keeping id of each list
84  is.na(coef_tot) <- coef_tot == "NULL"     # replace NULL with NA
85  coef_tot <- as.data.frame(sapply(coef_tot, unlist))    # unlist each
    ↪  column and convert to data.frame
```

Rolling approach

To generate price forecasts adopting the "Rolling" approach, the previous `R` script is used with the inner `for` loop starting at line 41 changed as follows.

```
41  # *ROLLING* Training set
42  for(i in 1:n){
43    x_i <- subset(df, start=i, end=nrow(df)*0.5716+i-1)    # shift x by
      ↪  i
44    xreg_i <- subset(df, start=nrow(df)*0.5716+i)    # shift regressors
      ↪  by i
45    lambda <- BoxCox.lambda(x_i[,"dam_price"])    # BoxCox
      ↪  transformation lambda for Training data
46
47    # fit ARIMA model to updated x and xreg
48    fit_i <- auto.arima(x_i[,"dam_price"],
49                        xreg=x_i[,c("Demand_Fcst", "p50_All",
                          ↪  "interval_All", "penetration")],
50                        lambda=lambda)
51
52    fit_df[i,1] <- fit_i$aic
53    fit_df[i,2:7] <- arimaorder(fit_i)[1:6]
54    coef_df <- rbind(coef_df, fit_i$coef)
55
56    # forecast with updated model
57    fcst <- forecast(fit_i, h=h,
58                     xreg=xreg_i[,c("Demand_Fcst", "p50_All",
                       ↪  "interval_All", "penetration")],
59                     level=c(50, 80, 90))
60
61    # forecast horizon: h = 1
62    fcmat[i,1] <- fcst$mean[1]
63    fcmat[i,2:4] <- fcst$lower[1,]
64    fcmat[i,5:7] <- fcst$upper[1,]
65
66    # forecast horizon: h = 2
67    fcmat[i,8] <- fcst$mean[2]
68    fcmat[i,9:11] <- fcst$lower[2,]
69    fcmat[i,12:14] <- fcst$upper[2,]
70  }
```

C.1.2 Monte Carlo simulations

The following `python` script performs the Monte Carlo simulations to generate the probability distribution of price difference ($\pi^{diff} = \pi^{imb} - \pi^{DAM}$) from the quantile forecasts of day-ahead and imbalance prices. For each time interval:

- two quadratic functions are interpolated to the quantile forecasts (0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95) of day-ahead and imbalance prices, respectively.
- the day-ahead and imbalance price interpolating functions are calculated on two randomly generated numbers (between 0.05 and 0.95), and the difference between the two resulting prices is calculated.
- the previous step is repeated 10,000 times (with the same interpolated functions but different random numbers) to generate a sample population of price differences.

The process is run on each time interval in the test period.

```
# Import libraries
import pandas as pd
from scipy.interpolate import interp1d
import random


# - - - Import data - - -
df = pd.read_csv("data.csv", index_col=0, parse_dates=True)
df.sort_index(inplace = True)
df = df.resample('30min').mean()


#=========================================================================
#    MONTE CARLO SIMULATION
#=========================================================================

# Define function that performs Monte Carlo simulation
def rand_diff():
    '''Monte Carlo simulation of price difference on a single time
    period. It calculates the difference between randomly generated
    imbalance and day-ahead prices 10,000 times.

    Variables
```

```
    ----------
    f_imb, f_dam : interp1d
        Functions interpolated to the quantile forecasts of imbalance
        and day-ahead prices, respectively.

    Returns
    ----------
    df_diff : pandas.DataFrame
       One-column DataFrame with 10,000 rows.
       The column name corresponds to the time period.
    '''
    diff_rand = []
    for i in range(10000):
        y_diff = f_imb(random.uniform(0.05, 0.95)) -
        ↪  f_dam(random.uniform(0.05, 0.95))
        diff_rand.append(y_diff)
    df_diff = pd.DataFrame(diff_rand, columns=[y_dam.name])
    return (df_diff)


# Day-ahead price forecast quantiles
df_dam = df[['lower90_dam', 'lower80_dam', 'lower50_dam', 'point_dam',
             'upper50_dam', 'upper80_dam', 'upper90_dam']]

# Imbalance price forecast quantiles
df_imb = df[['lower90_imb', 'lower80_imb', 'lower50_imb', 'point_imb',
             'upper50_imb', 'upper80_imb', 'upper90_imb']]

# List of corresponding probabilities
x = [0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95]


# Run Monte Carlo simulation on each time period in the test set
df_mc = pd.DataFrame()
for i in range(len(df_dam)):
    y_dam = df_dam.iloc[i]
    y_imb = df_imb.iloc[i]

    # Interpolate quadratic function to quantiles
    f_dam = interp1d(x, y_dam, kind='quadratic')
    f_imb = interp1d(x, y_imb, kind='quadratic')

    # Run Monte Carlo simulation
    df_mc = pd.concat([df_mc, rand_diff()], axis=1)


#============================================================================
#    DATA ANALYSIS
#============================================================================

# - - - Descriptive statistics for each distribution - - -
```

```
mc_prob = [(df_mc[col] > 0).sum()/float(len(df_mc[col]))
           for col in df_mc.columns]    # probability of positive price
           ↪ difference
mc_mean = [df_mc[col].mean() for col in df_mc.columns]    # mean
mc_median = [df_mc[col].median() for col in df_mc.columns]    # median
mc_std = [df_mc[col].std() for col in df_mc.columns]    # std. dev.

# 5, 10, 25, 75, 90, 95% quantiles
mc_q05 = [df_mc[col].quantile(0.05) for col in df_mc.columns]
mc_q10 = [df_mc[col].quantile(0.10) for col in df_mc.columns]
mc_q25 = [df_mc[col].quantile(0.25) for col in df_mc.columns]
mc_q75 = [df_mc[col].quantile(0.75) for col in df_mc.columns]
mc_q90 = [df_mc[col].quantile(0.90) for col in df_mc.columns]
mc_q95 = [df_mc[col].quantile(0.95) for col in df_mc.columns]


# Store statistics in DataFrame
mc_res = pd.DataFrame({'prob_positive':mc_prob, 'mc_mean':mc_mean,
                       'mc_median':mc_median, 'mc_std':mc_std,
                       'q05':mc_q05, 'q10':mc_q10, 'q25':mc_q25,
                       'q75':mc_q75, 'q90':mc_q90, 'q95':mc_q95
                       },
                       index = df.index)
```

C.1.3 Benchmark price forecasting model

The following `python` script implements the benchmark forecasting model for day-ahead and imbalance prices. The model is a variation of the "similar-day" method proposed by Conejo et al. in [150], and a description of its rationale is presented in Section 4.3.1.2.

```
# Import libraries
import pandas as pd


# - - - Import data - - -
df = pd.read_csv("data.csv", index_col=0, parse_dates=True)
df.sort_index(inplace = True)
df = df.resample('30min').mean()


#=====================================================================
```

```
#    BENCHMARK MODEL
#============================================================================
# - - - Monday, Saturday, Sunday - - -
# Shift forward of one week (48*7 periods) the 'price' column
df['bench'] = df['price'].shift(7*48)


# - - - Wednesday, Thursday, Friday - - -
# Shift forward of one day hours before 10:00
df['bench'].mask((df.index.hour<10) &
 ↪  (df.index.dayofweek.isin([2,3,4])),
                 df['price'].shift(1*48), inplace=True)

# Shift forward of two days hours after 10:00
df['bench'].mask((df.index.hour>=10) &
 ↪  (df.index.dayofweek.isin([2,3,4])),
                 df['price'].shift(2*48), inplace=True)


# - - - Tuesday - - -
# Shift forward of one day hours before 10:00
df['bench'].mask((df.index.hour<10) & (df.index.dayofweek.isin([1])),
                 df['price'].shift(1*48), inplace=True)

# Shift forward of four days hours after 10:00 (previous Friday)
df['bench'].mask((df.index.hour>=10) & (df.index.dayofweek.isin([1])),
                 df['price'].shift(4*48), inplace=True)
```

C.2 Imbalance Sign Forecasting

The imbalance sign is forecast using logistic regression models and adopting the multi-model approach. Therefore, input data are separated by load period, with a separate .csv file corresponding to each half-hour. The following `R` script generates the forecast of the imbalance sign using the recursively expanding approach ("Rec").

```
# Set working directory
setwd(".../Historical_data/NIV")

# Import packages
```

```
library(fpp2, tsibble, readr)

# Import data from half-hourly .csv files
temp = list.files(pattern="*.csv")
myfiles = lapply(temp, read.csv)

dflist = list()    # empty list to store outputs

# Loop through the half-hourly files
for(j in 1:length(myfiles)){
  df = data.frame(myfiles[j])

  # split data into Training and Test sets
  df_train <- df[1 : (nrow(df)*0.5716-1),]
  df_test <- df[(nrow(df)*0.5716) : nrow(df),]

  # - - - One- and two-step forecasts with re-estimation of GLM model - - -
  h <- 2    # forecast horizon
  n <- nrow(df_test) - 1    # number of forecasts to issue

  # empty lists to store forecast outputs
  NIV_list_1 = list()
  NIV_list_2 = list()
  formula_list = list()

  # *RECURSIVELY EXPANDING* training set
  for(i in 1:n){
    train_i <- df[1 : (nrow(df)*0.5716-1+i-1),]    # increase x by (i-1)
    test_i <- df[(nrow(df)*0.5716+i-1) : nrow(df),]    # shift regressors by i

    # fit GLM model to updated Training set, train_i
    fit_i <- glm(NIV ~ p50_All + interval_All + Demand_Fcst + penetration + factor(day),
                 data=train_i, family=binomial())

    #fit_i <- step(fit_i, direction="backward", trace=0)    # *STEPWISE SELECTION* of model with lowest AIC
    formula_list[i] <- deparse(fit_i$formula)    # formula of selected model

    pred_i = predict(fit_i, newdata=test_i, se.fit=T)  # predict on updated Test set, test_i

    NIV_fcst_i = exp(pred_i$fit)/(1+exp(pred_i$fit))  # convert to probability

    NIV_list_1[i] <- NIV_fcst_i[1]    # first (h=1) forecast value
    NIV_list_2[i] <- NIV_fcst_i[2]    # second (h=2) forecast value
```

```
48    }
49
50    df_fcst <- df_test[1:(nrow(df_test)-1),]
51    df_fcst$NIV_fcst_1 <- NIV_list_1   # attach forecast data to Test set
52    df_fcst$NIV_fcst_2 <- NIV_list_2
53    df_fcst$formulas <- unlist(formula_list)
54    dflist[[j]] <- df_fcst
55
56  }
57
58
59  # Store results in df_tot data.frame
60  df_tot <- do.call(rbind, dflist)
61  df_tot$NIV_fcst_1 <- unlist(df_tot$NIV_fcst_1)
62  df_tot$NIV_fcst_2 <- unlist(df_tot$NIV_fcst_2)
```

The rolling approach ("Roll") is implemented using the same script except for line 32 being changed as follows.

```
32  train_i <- df[i : (nrow(df)*0.5716-1+i-1),]    # shift x by i
```

To extend the "Rec" and "Roll" approaches with the *backward selection* of the model predictors (resulting in the "Rec-Step" and "Roll-Step" approaches, respectively), the `R` script is used with line 39 uncommented.

```
39    fit_i <- step(fit_i, direction="backward", trace=0)    # *STEPWISE
      ↪  SELECTION* of model with lowest AIC
```

List of Publications

C. Möhrlen, R. J. Bessa, M. Barthod, G. Goretti, and M. Siefert, "Use of Forecast Uncertainties in the Power Sector : State-of-the-Art of Business Practices," *15th International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Farms*, Vienna, Austria, 2016, doi: 10.21427/D7FJ7D.

G. Goretti, A. Duffy, and T. T. Lie, "The impact of power curve estimation on commercial wind power forecasts — An empirical analysis," *2017 14th International Conference on the European Energy Market (EEM17)*, Dresden, Germany, 2017, pp. 1-4, Publisher: IEEE, doi: 10.1109/EEM.2017.7981885.

G. Goretti and A. Duffy, "Evaluation of Wind Energy Forecasts : the Undervalued Importance of Data Preparation," *2018 15th International Conference on the European Energy Market (EEM)*, Łodz, Poland, 2018, pp. 1-5, Publisher: IEEE, doi: 10.1109/EEM.2018.8469845.

In Preparation:

G. Goretti and A. Duffy, "Risk-Aware Strategies for Short-term Wind Energy Trading in Markets with Single Imbalance Pricing," *Working Paper*, 2020.

G. Goretti and A. Duffy, "State-of-the-art of Balancing Market Price Forecasting before Day-ahead Closure," *Working Paper*, 2020.

G. Goretti and A. Duffy, "The New Wholesale Electricity Market in Ireland: I-SEM, an Empirical Study," *Working Paper*, 2020.

G. Goretti and A. Duffy, "Evaluation of Wind Power Forecasts: Guidelines for Data Preparation," *Working Paper*, 2021.

Training

List of modules undertaken during the research work as part of the Ph.D. curriculum.

Employability Skills

- Statistical Analysis for Engineers (STAT1950) – 5 ECTS
- Research Methods (GRSO1001) – 5 ECTS
- Writing in Science and Engineering (PH6022) – 5 ECTS
- Entrepreneurship for Engineers (ENTR1950) – 5 ECTS

Discipline Specific Skills

- Wind Energy for Electricity Supply (WEES1701) – 5 ECTS
- Advanced Topics in Research - Energy (ENEH1008) – 5 ECTS
- Modern Applied Statistical Modelling (MATH9952) – 10 ECTS

Others

- Time Series Analysis, Technical University of Denmark (DTU) – 2.5 ECTS