COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

MÁSTER UNIVERSITARIO EN INGENIERÍA INDUSTRIAL

TRABAJO FIN DE MÁSTER

AGGREGATED ENERGY INTELLIGENCE: OPTIMIZING WIND POWER PLANTS – BATTERY ENERGY STORAGE SYSTEMS INTEGRATION IN MULTI-MARKET PARTICIPATION

Autor: Jorge Doménech Patón

Director: Ramadhani Kurniawan Subroto

Co-Director: Tomislav Dragicevic

Madrid

Agosto de 2024

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título

Aggregated energy intelligence: Optimizing wind power plants – battery energy storage systems integration in multi-market participation

en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el

curso académico 2023/24 es de mi autoría, original e inédito y

no ha sido presentado con anterioridad a otros efectos.

El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido tomada de otros documentos está debidamente referenciada.

Fdo.: Jorge Doménech Patón   Fecha: 28/08/2024

Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO

Fdo.: Ramadhani Kurniawan Subroto   Fecha: 28/08/2024

COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

MÁSTER UNIVERSITARIO EN INGENIERÍA INDUSTRIAL

TRABAJO FIN DE MÁSTER

AGGREGATED ENERGY INTELLIGENCE: OPTIMIZING WIND POWER PLANTS – BATTERY ENERGY STORAGE SYSTEMS INTEGRATION IN MULTI-MARKET PARTICIPATION

Autor: Jorge Doménech Patón

Director: Ramadhani Kurniawan Subroto

Co-Director: Tomislav Dragicevic

Madrid

Agosto de 2024

AGGREGATED ENERGY INTELLIGENCE: OPTIMIZING WIND POWER PLANTS – BATTERY ENERGY STORAGE SYSTEMS INTEGRATION IN MULTI-MARKET PARTICIPATION

Autor: Doménech Patón, Jorge.
Director: Kurniawan Subroto, Ramadhani.
Entidad Colaboradora: Universidad Técnica de Dinamarca (DTU)

RESUMEN DEL PROYECTO

Este proyecto se centra en el desarrollo de un algoritmo de estrategia de oferta para la participación y gestión de una planta de energía agregada (APP), compuesta por un parque eólico (WF) y un sistema de almacenamiento de energía con baterías (BESS), en varios mercados de electricidad y servicios complementarios en Dinamarca, concretamente en el área DK2.

Para ello, se integran modelos basados en datos históricos para la predicción de potencia eólica y precios de la electricidad con modelos de optimización para derivar estrategias de oferta óptimas en el mercado diario (DAM), y los dos servicios de regulación primaria de frecuencia (FCR-D y FCR-N). El proyecto tiene como objetivo proporcionar un marco práctico de cómo los modelos de predicción basados en datos y los modelos de optimización pueden integrarse de manera efectiva para definir estrategias de oferta y mejorar el rendimiento económico de las APP en múltiples mercados eléctricos.

El proceso implica el filtrado de datos y la selección de variables para las predicciones de generación eólica y precios, explorando modelos basados en regresión para la generación eólica y modelos de series temporales para la predicción de precios. Integrando estas predicciones con el modelo de optimización, se analizan las ofertas óptimas basadas en las predicciones, evaluando la fiabilidad del modelo mediante la comparación de los beneficios esperados y percibidos. Además, se prueba el modelo utilizando previsiones de información perfecta para analizar su impacto en la estrategia de oferta óptima y los ingresos esperados.

El análisis destaca la necesidad de previsiones precisas de precios y generación eólica y demuestra ciertos elementos recurrentes en las estrategias de oferta durante los días estudiados. El parque eólico ofrece consistentemente en el DAM y en reserva a bajar del servicio FCR-D, mientras que el BESS muestra una preferencia por los mercados de reserva sobre el DAM. Además, el estudio propone trabajos futuros para mejorar el modelo, como el desarrollo de técnicas de predicción más sofisticadas o la exploración de estrategias de control en tiempo real para mejorar la fiabilidad y el rendimiento económico de las APP.

Palabras clave: Plantas de energía agregada, energía eólica, BESS, servicios complementarios, regulación primaria, predicción.

1. Introducción

La energía eólica es fundamental para la descarbonización de los sistemas eléctricos y la transición hacia fuentes de energía renovable. La creciente adopción de la energía eólica en Europa está llevando a un aumento significativo de su participación en el mix energético de multitud de países. Esta creciente integración de la energía eólica plantea desafíos en la operación del sistema y en los mercados de electricidad debido a su

naturaleza variable. El reemplazo de sistemas de energía convencionales en favor de plantas de generación eólica supone que estas últimas deban asumir roles adicionales, como proporcionar servicios complementarios de regulación de frecuencia, lo que requiere una estimación precisa de la generación eólica. La integración de parques eólicos con sistemas BESS puede ofrecer una solución para abordar está situación.

Por otro lado, la evolución de los métodos basados en datos permite la digitalización de activos energéticos, lo que potencialmente desbloquea estrategias de control personalizadas y mejora su eficiencia. Bajo este este contexto general, este proyecto propone el desarrollo de un algoritmo de estrategia de oferta basado en un modelo de optimización para derivar estrategias óptimas de participación en varios mercados eléctricos y servicios complementarios. El modelo de optimización aprovecha técnicas avanzadas basadas en datos para integrar las previsiones de generación eólica y los pronósticos de precios, lo que permite una estrategia integral que tiene en cuenta tanto la variabilidad de la generación como la dinámica del mercado.

El modelo integrado se centra en una APP que consta de un parque eólico y un sistema BESS, con el objetivo de optimizar su participación combinada en varios mercados energéticos y servicios complementarios. Además, el estudio evalúa las expectativas de ingresos y la precisión con la que estos coinciden con los ingresos percibidos, y cómo las previsiones de información perfecta pueden mejorar las estrategias de oferta, aumentando la precisión de las proyecciones de ingresos.

2. Definición y objetivos del proyecto

El proyecto examina las estrategias de oferta basadas en las predicciones de energía eólica y precios, evalúa la fiabilidad del modelo comparando los beneficios esperados con los percibidos, y analiza cómo varían estos factores al utilizar previsiones con información perfecta sobre los precios y la generación eólica. El objetivo es proporcionar un marco práctico de cómo los modelos de pronóstico basados en datos y los modelos de optimización pueden integrarse de manera efectiva para desarrollar estrategias de oferta y mejorar el rendimiento económico de las plantas de energía agregadas en múltiples mercados eléctricos.

El trabajo aborda varios objetivos específicos, cada uno de los cuales contribuye al objetivo general del desarrollo del algoritmo. Estos objetivos específicos son los siguientes:

- Obtención y filtrado de los datos SCADA recopilados de múltiples turbinas eólicas.
- Implementación de técnicas de inteligencia artificial para la predicción de generación de energía eólica.
- Implementación de técnicas de inteligencia artificial para la predicción de precios de la electricidad.
- Formulación de un modelo de optimización para determinar las ofertas óptimas de la planta de energía agregada en diversos mercados de energía y servicios complementarios.
- Análisis de estrategias de participación en los mercados considerados, basadas en pronósticos de generación eólica y precios.

- Evaluación y comparación de las principales fuentes de ingresos alcanzables en los mercados considerados, utilizando los modelos de predicción desarrollados frente a pronósticos perfectos.

3. Descripción del modelo

El marco general del modelo desarrollado en esta tesis se ilustra a continuación. A partir de series temporales históricas de precios de electricidad y generación eólica, se desarrollarán modelos para pronosticar precios y generación eólica. Algunos de los métodos utilizados serán más complejos, incorporando variables exógenas obtenidas de pronósticos externos, mientras que otros emplearán modelos simplificados. Además, se utilizarán mediciones reales de la frecuencia de la red eléctrica para modelar la activación de reservas. Todos estos parámetros servirán como datos de entrada para un modelo de optimización MILP que tiene como objetivo maximizar los ingresos del DAM, los ingresos o costes por desviación en el mercado de balance (IMB) y los servicios auxiliares FCR para una APP ubicada en Dinamarca (DK2).
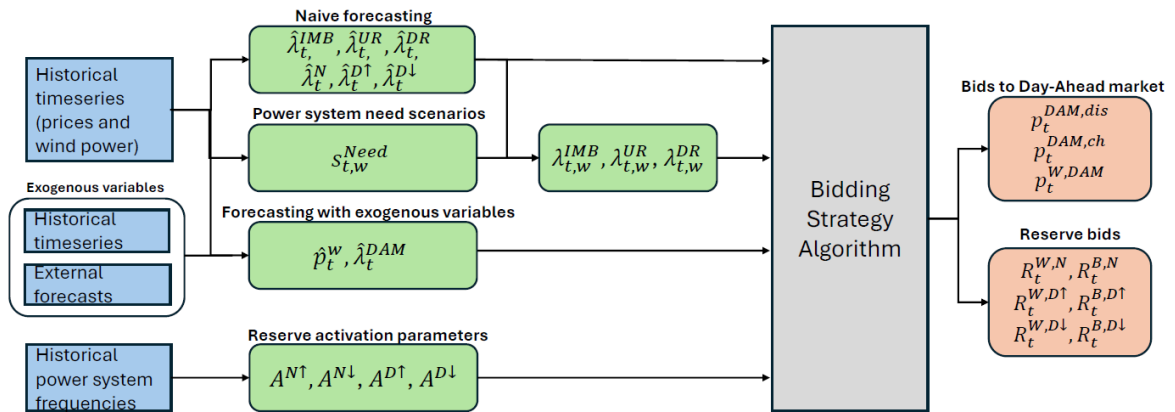

Ilustración 1. Descripción del modelo. Esquema y flujo de información.

4. Resultados

En cuanto a la previsión de la generación eólica, el análisis de los datos SCADA del parque eólico ha identificado la velocidad del viento como el factor principal. La fuerte correlación observada sugiere la exploración de modelos basados en regresión. En este sentido, se han desarrollado varios modelos, incluyendo regresión lineal, modelos autorregresivos y regresión ponderada localmente, siendo este último el que proporcionó los mejores resultados debido a su capacidad para generar una curva de ajuste más suave. A pesar de esto, la incorporación de variables externas introdujo errores de predicción notables, lo que resalta la importancia de utilizar previsiones adaptadas a las condiciones específicas del parque eólico.

La previsión de precios ha sido el factor más importante tanto para la planificación de las estrategias de participación en el mercado como para los ingresos finales obtenidos en los días estudiados, con desviaciones entre los ingresos esperados y los percibidos coincidiendo con desviaciones en las previsiones de precios. Se han evaluado varios métodos de previsión basados en series temporales para la predicción de precios del DAM, siendo SARIMAX el más preciso. Sin embargo, este modelo ha mostrado variabilidad significativa en el

rendimiento dependiendo del día estudiado, lo que indica la necesidad de modelos más complejos para una aplicación real. Por otro lado, la utilización de modelos de predicción simples para los precios de los mercados de reserva ha acabado resultando en errores de pronóstico significativos. A pesar de estos errores, la simplificación de estos modelos ha ayudado a mantener una comprensión general de las tendencias de precios y las estrategias de oferta.

En cuanto a las estrategias de oferta, la estrategia para el parque eólico se ha mantenido constante en todos los días estudiados. El parque eólico ofrece su predicción más precisa en el DAM y la misma cantidad como reserva a bajar en el mercado FCR-D. Con el nivel de precios actual, esto ocurre porque una turbina eólica tendría que operar bajo limitación de potencia si oferta reserva a subir en FCR-D o FCR-N para asegurar suficiente capacidad para aumentar la potencia en caso de una activación repentina de la reserva. Ante esta situación, es óptimo ofertar en el DAM y obtener beneficios adicionales de la reserva de regulación a bajar de FCR-D, ya que no requiere limitación de potencia en tiempo real a menos que se active. En cuanto al sistema BESS, se ha observado una mayor variabilidad en sus estrategias dependiendo de la configuración de precios del mercado. La batería generalmente prefiere participar en los mercados de reserva en lugar del DAM. Sin embargo, durante los períodos en los que es más óptimo ofrecer en el mercado FCR, aprovecha los máximos locales en los precios del DAM para descargar y aumentar la disponibilidad de reserva a subir y utiliza los mínimos locales para cargar y aumentar las reservas para la regulación a bajar. Las estrategias de oferta en función de los precios son estudiadas en más profundidad en la discusión de resultados del documento principal.

En cuanto a la rentabilidad de las ofertas, se han comparado los ingresos esperados con los percibidos. Tras analizar los ingresos por mercado, se observa que las mayores diferencias provienen de previsiones erróneas de los precios de reserva. Como se ha mencionado, estos precios se han obtenido de modelos de predicción simplistas, lo que explica los errores que llevaron a desviaciones en los ingresos. Por lo tanto, los ingresos esperados también se han analizado en comparación con los ingresos percibidos utilizando pronósticos de información perfecta, obteniendo una reducción en la volatilidad entre los ingresos esperados y los percibidos y demostrando la necesidad de pronósticos precisos para lograr escenarios de ingresos realistas.
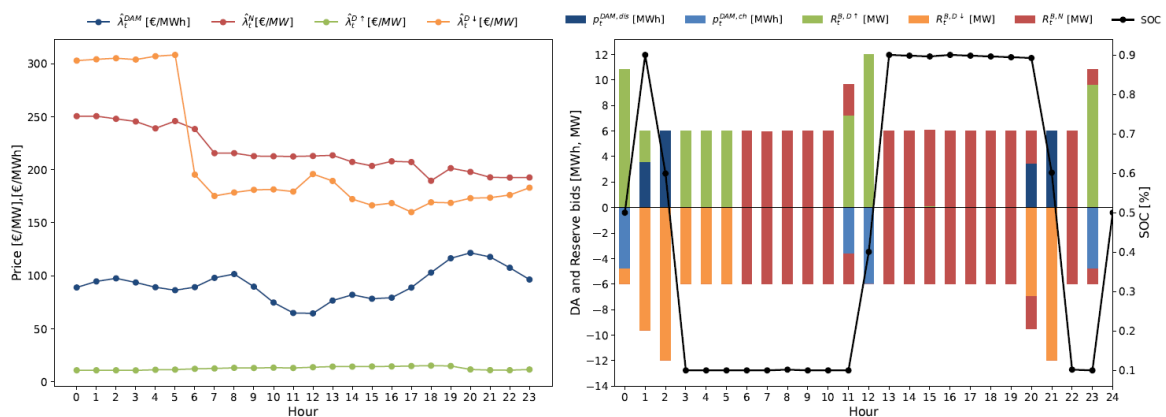

Ilustración 2. Configuración de precios y estrategia del sistema BESS para uno de los días estudiados.

	Día 1	Día 2	Día 3	Día 4
Resultados con los modelos de predicción desarrollados				
Beneficio esperado (€)	37,648	43,157	37,535	36,777
Beneficio percibido (€)	34,258	14,058	52,634	31,343
Resultados con predicciones de información perfecta				
Beneficio esperado (€)	32,484	14,736	52,615	31,846
Beneficio percibido (€)	33,408	14,882	52,615	32,265

Tabla 1. Comparación de beneficios esperados frente a percibidos utilizando las técnicas de predicción desarrolladas y previsiones con información perfecta.

5. Conclusiones

Este proyecto se ha centrado en el desarrollo de un algoritmo de estrategia de oferta para la participación y gestión de una planta de energía agregada (APP), compuesta por un parque eólico (WF) y un sistema de almacenamiento de energía con baterías (BESS), en mercados de electricidad y servicios auxiliares. A través de la consecución de sus objetivos específicos, el proyecto ha explorado la integración de modelos basados en datos para la previsión de precios de la electricidad y de la generación eólica con modelos de optimización para derivar estrategias óptimas de oferta en los mercados combinados del mercado diario (DAM) y los dos servicios de regulación primaria de frecuencia (FCR-D y FCR-N).

Dado que el objetivo de este trabajo no era realizar un estudio de viabilidad económica de los mercados, el modelo desarrollado ha sido probado para cuatro días de estudio con diferentes configuraciones de precios con el fin de evaluar diversas estrategias de participación en el mercado. Este enfoque ha permitido examinar estrategias óptimas de participación para la WF, revelando que la estrategia más efectiva consiste en ofertar su pronóstico más preciso en el DAM y participar en la reserva de regulación a bajar del servicio complementario FCR-D. Además, el análisis de las estrategias para el BESS ha indicado una preferencia por los mercados de reserva sobre el DAM. Por otro lado, se ha podido concluir que bajo las suposiciones realizadas, los ingresos adicionales por la activación de reservas en el mercado FCR-N generalmente no sirven como el criterio principal para seleccionar FCR-N sobre FCR-D.

Las estrategias desarrolladas se han analizado comparando los ingresos esperados con los ingresos percibidos. Este análisis incluye escenarios en los que se utilizaron pronósticos con información perfecta, en los cuales los precios reales de la electricidad y la generación eólica se conocían al momento de planificar la estrategia de oferta. Esta comparación ha destacado la necesidad de pronósticos más precisos para facilitar el desarrollo de planes de negocio realistas a largo plazo, mejorando así la viabilidad económica y la planificación estratégica de las APP en condiciones reales de mercado.

AGGREGATED ENERGY INTELLIGENCE: OPTIMIZING WIND POWER PLANTS – BATTERY ENERGY STORAGE SYSTEMS INTEGRATION IN MULTI-MARKET PARTICIPATION

Author: Doménech Patón, Jorge.
Supervisor: Kurniawan Subroto, Ramadhani.
Collaborating Entity: Technical University of Denmark (DTU)

ABSTRACT

This project focuses on developing a bidding strategy algorithm for the participation and management of an aggregated power plant (APP), consisting of a wind farm (WF) and battery energy storage system (BESS), in various electricity markets and ancillary services.

To address these challenges, data-driven models for wind power and electricity price forecasting are integrated with optimization models to derive optimal bidding strategies for the day-ahead market (DAM), frequency containment reserve for disturbance (FCR-D), and frequency containment reserve for normal operation (FCR-N) markets. The research aims to provide a comprehensive understanding of how data-driven forecasting and optimization models can be effectively integrated to define bidding strategies and improve the economic performance of APPs in multiple electricity markets.

The process involves data filtering and feature selection for wind power and price forecasts, exploring regression-based models for wind power and time series models for price forecasting. Once these forecasts are integrated with the optimization model, the study examines optimal bids based on forecasted prices, assessing the reliability of the model by comparing expected and realized profits. Additionally, the study tests the model using perfect information forecasts to analyze its impact on optimal bidding strategy and expected revenue.

The analysis highlights the requirement for precise price and wind power forecasts and demonstrates certain recurring elements in the bidding strategies over the days studied. The WF consistently bids in the DAM and FCR-D down regulation, while the BESS shows a preference for reserve markets over the DAM. Additionally, the study proposes future work to further improve the model, as the development of more sophisticated forecasting techniques or the exploration of real-time control strategies to improve the reliability and economic performance of APPs.

Keywords: Aggregated power plant, wind energy, BESS, forecasting, ancillary services, bidding strategy.

1. Introduction

Wind energy is crucial for decarbonizing electricity systems and transitioning to renewable energy sources. The growing adoption of wind power across Europe is leading to a significant increase in its share within the energy mix of several countries. However, the increased reliance on wind energy introduces challenges in system operation and electricity markets due to its variable nature. As WFs replace conventional power plants, they must take on additional roles, such as providing ancillary services, which requires

accurate wind power estimation. Integrating wind farms with BESS offers a potential solution to address uncertainties in wind power forecasts.

The evolution of data-driven methods enable the digitalization of assets, potentially unlocking customized control strategies and improving efficiency. Building on this general context, this thesis proposes the development of a bidding strategy algorithm based on an optimization model to derive optimal participation strategies in various electricity markets and ancillary services. The optimization model leverages advanced data-driven techniques to integrate wind power forecasts and price forecasts, enabling a comprehensive strategy that accounts for both generation variability and market dynamics.

The integrated model focuses on an APP that consists of a WF and a BESS, aiming to optimize their combined participation in various energy markets and ancillary services. Additionally, the study evaluates the revenue expectations and how closely these match the realized revenue and how perfect information forecasts can improve bidding strategies, improving the accuracy of revenue projections.

2. Project objectives

The project will examine bidding strategies based on forecasts, assess the reliability of the model by comparing expected and realized profits, and analyze how these factors vary when using forecasts with perfect information of prices and wind power. The research aims to provide a comprehensive understanding of how data-driven forecasting and optimization models can be effectively integrated to define bidding strategies and improve the economic performance of APPs in multiple electricity markets.

The work covers several specific objectives, each contributing to the overall goal of the algorithm development. These specific objectives are outlined below:

• Refinement and filtration of Supervisory Control and Data Acquisition (SCADA) data collected from multiple wind turbines (WT).

• Deployment of AI techniques to forecast wind power generation accurately.

• Deployment of AI techniques to forecast electricity prices accurately.

• Formulation of an optimization model to determine day-ahead schedules for WF participation in diverse energy markets and ancillary services, incorporating BESS aggregation.

• Analysis of participation strategies in the considered markets based on wind power and price forecasts.

• Evaluation and comparison of the main revenue streams achievable within the considered markets, using developed forecasts versus perfect forecasts.

3. Model description

The general framework of the model developed in this thesis is illustrated below. Based on historical time series of electricity prices and wind power generation, models for forecasting prices and wind power will be developed. Some of the methods used will be more complex, incorporating exogenous variables obtained from external forecasts, while others will use simplifications in the form of naive predictors. Additionally, real power grid frequency measurements will be used to model reserve activation. All these parameters will serve as inputs for a MILP optimization model that aims to maximize revenues from the DAM, imbalance settlement (IMB), and the FCR ancillary services for an aggregated power plant located in Denmark (DK2).
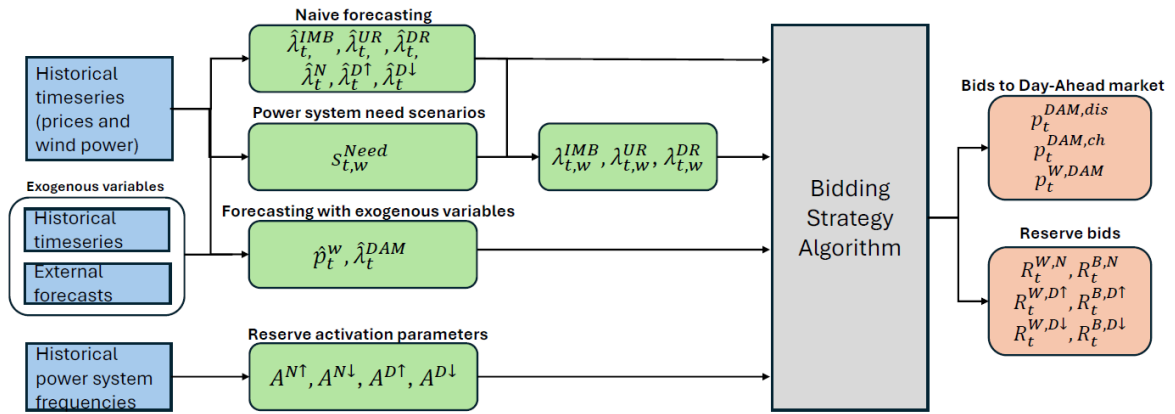
Figure 1. Model development. Framework and flow of information.

4. Results

Regarding wind power forecasting, the analysis of the WF's SCADA data available for this thesis has identified wind speed as the main factor. The strong correlation presented led to the exploration of regression-based models. Several models, including linear regression, autoregressive, and local weighted regression were tested, with the latter providing the best results due to its potential to generate a smoother fit curve. Despite this, incorporating external variables introduced notable prediction errors, highlighting the importance of using tailored forecasts specific to the conditions of the wind farm.

Price forecasting has been the most important factor both for planning market participation strategies and for the final revenue obtained on the studied days, with deviations between expected revenue and realized revenue coinciding with deviations in price forecasts. Various time series forecasting methods were evaluated for DAM price prediction, with SARIMAX delivering the most accurate results. However, there was still significant variability in performance depending on the day, indicating that further refinement with more sophisticated models could be beneficial for practical applications. While naive predictors were used for reserve market prices to simplify the methodology, they resulted in significant forecast errors. Despite these errors, they helped maintain a general understanding of price trends and potential bidding strategies.

As for bidding strategies, the strategy for the WF remained consistent across all study days. The wind farm offers its most accurate prediction in the DAM and the same amount as downward reserve in the FCR-D market. With the current price level, this occurs because a wind turbine would have to operate under curtailment if it were to offer in

FCR-D up-regulation or FCR-N to ensure sufficient capacity to increase power if a sudden reserve activation occurs. In this situation, the WF prefers to offer in the DAM and gain additional benefits from the FCR-D down-regulation reserve, as it does not require real-time curtailment unless activated. Regarding the BESS, more variability in its strategies is observed depending on the market price configuration. The battery generally prefers to participate in reserve markets rather than the DAM. However, during periods when it is more optimal to offer in the FCR market, it takes advantage of local maxima in DAM prices to discharge and increase reserve availability for up-regulation or utilizes local minima to charge and boost reserves for down-regulation. More price-specific strategies are discussed in the main text of the thesis.

Concerning the profitability of the results, expected revenues have been compared with realized revenues. After analyzing the revenues by market, it was observed that the largest differences come from erroneous forecasts of reserve prices. As mentioned, these prices were derived from naive forecasts, which explains the errors leading to revenue deviations. Consequently, the expected revenue was also analyzed against the realized revenue using perfect price forecasts, which demonstrated a reduction in the volatility between expected and realized revenues, underscoring the necessity for accurate forecasts to achieve realistic expected revenues.
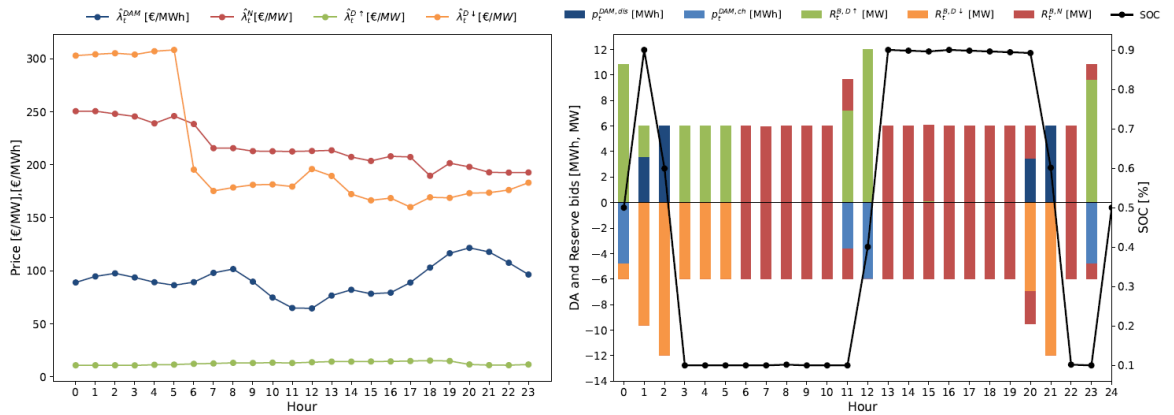

Figure 2. Price configuration and BESS bidding strategy for one of the selected days of study.

	Test Day 1	Test Day 2	Test Day 3	Test Day 4
Results with the developed forecast techniques				
Expected revenue (€)	37,648	43,157	37,535	36,777
Realized revenue (€)	34,258	14,058	52,634	31,343
Results with perfect information forecasts				
Expected revenue (€)	32,484	14,736	52,615	31,846
Realized revenue (€)	33,408	14,882	52,615	32,265

Table 1. Comparison of expected vs. realized profits using developed forecasting techniques and perfect information forecasts.

5. Conclusions

This thesis has focused on developing a bidding strategy algorithm for the participation and management of an aggregated power plant (APP), comprising a wind farm (WF) and a battery energy storage system (BESS), in electricity markets and ancillary services. Through achieving its specific objectives, this research has explored the integration of data-driven models for electricity price and wind power forecasting with optimization models to derive optimal bidding strategies in the combined markets of day-ahead market (DAM), FCR-D, and FCR-N.

Since the objective of this work was not to perform an economic viability study of the markets, the developed model was tested over four days with different price configurations to evaluate various market participation strategies. This approach enabled the examination of optimal participation strategies for the WF, revealing that the most effective strategy involves bidding its most accurate forecast into the DAM and participating in the FCR-D down-regulation reserve. Additionally, the analysis of strategies for the BESS indicated a preference for reserve markets over the DAM. It was also found that, under the assumptions made, FCR-N reserve activations generally do not serve as the primary criteria for selecting FCR-N over FCR-D.

The developed strategies were analyzed by comparing expected revenues with realized revenues. This analysis included scenarios where perfect information forecasts were used, in which actual electricity prices and wind power were known at the time of planning the bid offering strategy. This comparison highlighted the necessity for more precise forecasts to reduce the gap between expected and realized revenues. Achieving more accurate forecasts could facilitate the development of realistic long-term business scenarios, thereby enhancing the economic viability and strategic planning of APPs in actual market conditions.

Abstract

This thesis focuses on developing a bidding strategy algorithm for the participation and management of an aggregated power plant (APP), consisting of a wind farm (WF) and battery energy storage system (BESS), in various electricity markets and ancillary services.

To address these challenges, data-driven models for wind power and electricity price forecasting are integrated with optimization models to derive optimal bidding strategies for the day-ahead market (DAM), frequency containment reserve for disturbance (FCR-D), and frequency containment reserve for normal operation (FCR-N) markets. The research aims to provide a comprehensive understanding of how data-driven forecasting and optimization models can be effectively integrated to define bidding strategies and improve the economic performance of APPs in multiple electricity markets.

The process involves data filtering and feature selection for wind power and price forecasts, exploring regression-based models for wind power and time series models for price forecasting. Once these forecasts are integrated with the optimization model, the study examines optimal bids based on forecasted prices, assessing the reliability of the model by comparing expected and realized profits. Additionally, the study tests the model using perfect information forecasts to analyze its impact on optimal bidding strategy and expected revenue.

The analysis highlights the requierment for precise price and wind power forecasts and demonstrates certain recurring elements in the bidding strategies over the days studied. The WF consistently bids in the DAM and FCR-D down regulation, while the BESS shows a preference for reserve markets over the DAM. Additionally, the study proposes future work to further improve the model, as the development of more sophisticated forecasting techniques or the exploration of real-time control strategies to improve the reliability and economic performance of APPs.

Aggregated energy intelligence

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Ramadhani, for his excellent supervision, consistent follow-up, and valuable insights throughout the course of this work.

I also wish to thank my family and Julia for their support during these months.

Special thanks to Karolis and Martin for their help in obtaining the data necessary for the completion of this thesis.

Abbreviations

WF	Wind farm
BESS	Battery Energy Storage System
APP	Aggregated power plant
WT	Wind turbine
DAM	Day-Ahead market
FFR	Fast Frequency Reserve
FCR	Frequency Containment Reserve
aFRR	Automatic Frequency Restoration Reserve
mFRR	Manual Frequency Restoration Reserve
FCR-N	Frequency Containment Reserve - Normal operation
FCR-D	Frequency Containment Reserve - Disturbance
LER	Limited Energy Reservoir
NEM	Normal State Energy Management
SOC	State of charge
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
NRMSE	Normalized Root Mean Square Error
ACF	Autocorrelation Function
PACF	Partial Autocorrelation Function
WS	Wind Speed
WD-X	Wind Direction X-component
WD-Y	Wind Direction Y-component
T	Temperature
FFS	Forward Feature Selection
BFS	Backward Feature Selection
DMI	Danmarks Meteorologiske Institut
NWP	Numerical Weather Prediction
LR	Linear Regression
ARX	Autoregressive with Exogenous inputs
LWR	Locally Weighted Regression
RM	Reserve Market
EM	Energy Market
IM	Imbalance Market
BM	Balancing Market
AR	Autoregressive
MA	Moving Average
ARMA	Autoregressive Moving Average
ARIMA	Autoregressive Integrated Moving Average
SARIMA	Seasonal ARIMA
SARIMAX	Seasonal ARIMA with Exogenous variables
ENTSO-E	European Network of Transmission System Operators for Electricity
CCF	Cross-Correlation Function
SMA	Simple Moving Average Method
SOC	State of Charge
CET	Central European Time
CEST	Central European Summer Time

Nomenclature

General Theory

P_{max}	Maximum available wind power $[W]$
ρ	Air density $[kg/m^3]$
A_{wtr}	Rotor area $[m^2]$
V_0	Free wind speed $[m/s]$
C_p	Power coefficient [-]
Λ	Tip-speed ratio [-]
θ	Pitch angle $[°]$
w_{wtr}	Rotor rotational speed $[rad/s]$
R_{wtr}	WT blade radius $[m]$
N	Total number of observations in test set
y_t	Actual value at hour t of the variable to predict
$\hat{y}_t$	Forecasted value at hour t of the variable to predict
$y_{t,max}$	Maximum value in the test set of the variable to predict
$y_{t,max}$	Minimum value in the test set of the variable to predict
ρ_{ref}	Air density in standard conditions ($1.223 kg/m^3$)
T_{ref}	Reference temperature for air density calculation ($288.15K$)
T_{amb}	Ambient temperature i$[k]$
U_1	Wind speed at height Z_1 $[m/s]$
U_0	Wind speed measured at height Z_0 $[m/s]$
α	Wind shear coefficient [-]
f_t	Power system frequency at time t $[Hz]$
$a_t^{N,\uparrow}$	Activation function of FCR-N up-regulation at time t [-]
$a_t^{N,\downarrow}$	Activation function of FCR-N down-regulation at time t [-]
$a_t^{D,\uparrow}$	Activation function of FCR-D up-regulation at time t [-]
$a_t^{D,\downarrow}$	Activation function of FCR-D down-regulation at time t [-]

Optimization Model

Sets

$\mathcal{T}$	Set of time periods t
$\mathcal{W}$	Set of power system scenarios w
Ω	Set of variables
Ω^R	Subset of real variables
Ω^{R+}	Subset of positive real variables
Ω^b	Subset of binary variables

Decision Variables

p_t^{EM}	APP power produced at time t to the Energy Market $[MW]$
$p_t^{W,EM}$	WF power produced at time t to the Energy Market $[MW]$
$p_t^{B,EM}$	BESS power produced at time t to the Energy Market $[MW]$
$p_t^{W,DAM}$	WF power bid at time t to the DAM $[MW]$
$p_t^{DAM,dis}$	Discharge BESS power bid at time t to the DAM $[MW]$
$p_t^{DAM,ch}$	Discharge BESS power bid at time t to the DAM $[MW]$

ΔP_t	Power deviation at time t $[MW]$
ΔP_t^+	Positive power deviation at time t $[MW]$
ΔP_t^-	Negative power deviation at time t $[MW]$
b_t^{state}	Binary auxiliary variable to model BESS state at time t
$\delta_t^{W,D\uparrow}$	Binary auxiliary variable to bound WF reserve bid to FCR-D up at time t
$\delta_t^{W,D\downarrow}$	Binary auxiliary variable to bound WF reserve bid to FCR-D down at time t
$\delta_t^{W,N}$	Binary auxiliary variable to bound WF reserve bid to FCR-N at time t
$\delta_t^{B,D\uparrow}$	Binary auxiliary variable to bound BESS reserve bid to FCR-D up at time t
$\delta_t^{B,D\downarrow}$	Binary auxiliary variable to bound BESS reserve bid to FCR-D down at time t
$\delta_t^{B,N}$	Binary auxiliary variable to bound BESS reserve bid to FCR-N at time t
p_t^{dis}	Total BESS discharge power at time t $[MW]$
p_t^{ch}	Total BESS charge power at time t $[MW]$
SOC_t	State of charge at time t $[\%]$
$R_t^{W,D\uparrow}$	WF reserve bid to FCR-D up at time t $[MW]$
$R_t^{W,D\downarrow}$	WF reserve bid to FCR-D down at time t $[MW]$
$R_t^{W,N}$	WF reserve bid to FCR-N at time t $[MW]$
$R_t^{B,D\uparrow}$	BESS reserve bid to FCR-D up at time t $[MW]$
$R_t^{B,D\downarrow}$	BESS reserve bid to FCR-D down at time t $[MW]$
$R_t^{B,N}$	BESS reserve bid to FCR-N at time t $[MW]$
$p_t^{W,D\uparrow}$	WF FCR-D up reserve activation at time t $[MW]$
$p_t^{W,D\downarrow}$	WF FCR-D down reserve activation at time t $[MW]$
$p_t^{W,N\uparrow}$	WF FCR-N up reserve activation at time t $[MW]$
$p_t^{W,N\downarrow}$	WF FCR-N down reserve activation at time t $[MW]$
$p_t^{B,D\uparrow}$	BESS FCR-D up reserve activation at time t $[MW]$
$p_t^{B,D\downarrow}$	BESS FCR-D down reserve activation at time t $[MW]$
$p_t^{B,N\uparrow}$	BESS FCR-N up reserve activation at time t $[MW]$
$p_t^{B,N\downarrow}$	BESS FCR-N down reserve activation at time t $[MW]$
$S_{t,w}^{Need}$	Power system need at time t under scenario w
$\lambda_{t,w}^{UR}$	Up-regulation price at time t under scenario w $[\text{€}/MWh]$
$\lambda_{t,w}^{DR}$	Down-regulation price at time t under scenario w $[\text{€}/MWh]$
$\lambda_{t,w}^{IMB}$	Imbalance price at time t under scenario w $[\text{€}/MWh]$
$E_{t,w}^{EM}$	APP earnings from the EM at time t under scenario w $[\text{€}]$
$E_{t,w}^{FCR-N}$	APP earnings from FCR-N at time t under scenario w $[\text{€}]$
E_t^{FCR-D}	APP earnings from FCR-D at time t $[\text{€}]$

Parameters

$\hat{\lambda}_t^{DAM}$	Forecasted DAM price at time t $[\text{€}/MWh]$
$\hat{\lambda}_t^{UR}$	Forecasted up-regulation price at time t $[\text{€}/MWh]$
$\hat{\lambda}_t^{DR}$	Forecasted down-regulation price at time t $[\text{€}/MWh]$
$\hat{\lambda}_t^{D\uparrow}$	Forecasted FCR-D up price at time t $[\text{€}/MW]$
$\hat{\lambda}_t^{D\downarrow}$	Forecasted FCR-D down price at time t $[\text{€}/MW]$
$\hat{\lambda}_t^{N}$	Forecasted FCR-N price at time t $[\text{€}/MW]$
$\hat{p}_t^{W}$	Forecasted wind power at time t $[MW]$
$\overline{P}^B$	Maximum discharge and charge power of the BESS $[MW]$
$\overline{P}^W$	WF rated power $[MW]$
E_b	BESS nominal energy $[MW]$
Δt	One hour

η_{dis}	Discharging efficiency of the BESS [%]
η_{ch}	Charging efficiency of the BESS [%]
SOC_{init}	BESS SOC at the beginning of the day [%]
$A^{D\uparrow}$	Reserve activation parameter for FCR-D up [-]
$A^{D\downarrow}$	Reserve activation parameter for FCR-D down [-]
$A^{N\uparrow}$	Reserve activation parameter for FCR-N up [-]
$A^{N\downarrow}$	Reserve activation parameter for FCR-N down [-]

Contents

List of Figures

List of Tables

1 Introduction

1.1 Context and motivation

Wind energy plays a crucial role in the decarbonization of electricity systems and their shift towards renewable energy solutions. In 2023, 272 GW of wind power capacity was installed in Europe and 260 GW of new wind power capacity is expected for 2030. This increase in the installation of wind power generation represents a share of wind energy of more than 20% in the energy mix of ten European countries, with Denmark and Ireland leading the way with 56% and 36% respectively [1].

The increasing penetration of wind power in the energy mix brings with it multiple challenges in system operation and electricity markets due to the variable nature of the wind resource. Wind farms (WF) are replacing conventional power plants and, therefore, will have to assume the roles these plants had in ancillary services, making the accurate estimation of wind power a fundamental requirement. One possible solution to the detrimental effects induced by uncertainty in wind power forecasts is the aggregation of portfolios, particularly the integration of wind farms and battery energy storage systems (BESS).

Within the context of asset aggregation, the application of advanced modeling and optimal control techniques becomes essential for asset managers. In contrast to traditional physical models, which often rely on complex mathematical equations and physical laws to simulate asset behavior, the constant evolution of data-driven methods offers significant advantages. By using artificial intelligence models, these data-driven approaches enable the digitization of assets, potentially unlocking customized control strategies and improving efficiency. These models are developed solely from historical asset data, from which mathematical models are inferred to define the relationship between input variables and the asset's behavior. In addition to reducing complexity, these models offer the advantage of being computationally more efficient, making them a powerful tool for energy asset management.

Building on this foundation, this thesis proposes the development of a bidding strategy algorithm based on an optimization model to derive optimal participation strategies in various electricity markets. To achieve this, the creation of accurate wind power forecasts and electricity price forecasts is essential. These forecasts serve as critical inputs to the optimization model, guiding decision-making processes and enhancing the reliability and profitability of market participation.

The optimization model leverages advanced data-driven techniques to integrate wind power forecasts and price forecasts, enabling a comprehensive strategy that accounts for both generation variability and market dynamics. The integrated model focuses on an aggregated power plant (APP) that consists of a WF and a BESS, aiming to optimize their combined participation in various energy markets and ancillary services.

In summary, this thesis aims to demonstrate how data-driven models can be practically applied to electricity market participation. The model will allow for the analysis of optimal bids based on price and wind power forecasts, evaluating the revenue expectations generated and how closely these match the realized revenue. Additionally, the study will analyse how perfect information forecasts can improve bidding strategies, improving the accuracy of revenue projections.

1.2 Thesis objectives

The primary objective of this project is to develop a bidding strategy algorithm for the participation and management of an APP in electricity markets and ancillary services. However, this thesis covers several specific objectives, each contributing to the overall goal of the algorithm development. These specific objectives are outlined below:

- Refinement and filtration of Supervisory Control and Data Acquisition (SCADA) data collected from multiple wind turbines (WT).

- Deployment of AI techniques to forecast wind power generation accurately.

- Deployment of AI techniques to forecast electricity prices accurately.

- Formulation of an optimization model to determine day-ahead schedules for WF participation in diverse energy markets and ancillary services, incorporating BESS aggregation.

- Analysis of participation strategies in the considered markets based on wind power and price forecasts.

- Evaluation and comparison of the main revenue streams achievable within the considered markets, using developed forecasts versus perfect forecasts.

1.3 Thesis structure

In Chapter 2, a comprehensive literature review is conducted, leading to the definition of this thesis scope. Chapter 3 provides a theoretical framework, explaining the necessary background on wind turbines, the theoretical aspects of electricity markets, the uncertainties to consider, and the methods for evaluating the forecasts that will be developed later.

Chapter 4 details the acquisition, filtering, feature selection, and prediction models for wind power forecasting. Chapter 5 focuses on the development of the optimization model for deriving bidding strategies in the considered electricity markets.

In Chapter 6, the inputs to the optimization model are explained. These include electricity price forecasts, the creation of scenarios to model the state of the electrical system at each hour, and the creation of parameters to model reserve activations in ancillary services. Chapter 7 presents the results, covering the optimal bids obtained and the comparison between expected and realized revenue.

Chapter 8 discusses the results, providing an in-depth analysis of the findings. Finally, Chapter 9 concludes the project, mentioning potential future work to improve the model.

A section with the nomenclature and abbreviations used throughout this thesis is provided before the table of contents. Additionally, lists of figures and tables are included immediately after the table of contents.

2 State-of-the-Art

This chapter provides an overview of the current state-of-the-art in the integration of APPs into electricity markets. First, a literature review is conducted, exploring existing studies and methodologies related to data preparation and filtering, forecasting techniques for wind power and electricity prices, and optimization strategies for market participation. Following this, the specific contributions and scope of this thesis are explained.

2.1 Literature review

The initial step in modeling wind assets using artificial intelligence techniques is the collection and filtering of historical data from the WFs to eliminate outliers or samples that do not represent the regular behavior of the assets. Although data is sometimes filtered manually, the author in [2] provides a thorough review of different related studies in this field that employ machine learning models for outlier filtering. Among these studies, techniques such as K-means clustering, binary image analysis, binning of power curves, and k-nearest neighbors can be found. One model that appears to provide good results is the one proposed in [3], where the Least Mean of Squares (minimizing the sum of the squares over all the measurements) is employed to filter the points of the WT power curve. The results of this study reflect a 20% increase in computational time.

For the development of this thesis, both real SCADA data from different WTs and their operating ranges accepted by the manufacturer are available. These normal operating ranges provide a good basis for manual data filtering. If manual filtering proves to be insufficient, more complex techniques proposed in [2] and [3] will be explored.

After filtering the data, it is necessary to select the input variables for training the models, as well as the prediction techniques to be used. The literature shows a great variety in terms of input variables and models used, ranging from simple regression-based models to more complex deep learning techniques. Two specific studies reflecting this diversity are presented: In [4], a deep learning model based on a temporal Convolutional Neural Network (CNN) is used to predict power generation as a function of wind speed, direction, and theoretical power. Conversely, in [5], k-nearest neighbor regression is employed, with the set of variables reduced to only wind power and wind speed.

Regarding electricity price forecasting, there is an extensive literature, particularly for predicting day-ahead market prices. A comprehensive overview of existing methods is presented in [6]. The authors classify the typically used methods into three groups. First, there are time series models (such as ARIMA or dynamic regression), which relate current prices to past prices and current errors to past errors. Second, Artificial Neural Network (ANN) models are employed. Lastly, Wavelet Transform Analysis (WTA) based models are used. In the case of this paper, dynamic regression models provide the best results, followed by ARIMA and WTA, with ANN models performing the least effectively.

In [7], other models are compared across three different countries and two time periods. These models include Random Forest Regressor (RFR), CNN, Deep Neural Network (DNN), and Support Vector Regressor (SVR), incorporating external variables such as production and consumption forecasts. While the first two models do not yield good results, the latter two are able to extract valuable information from the features and reflect the impact of external variable patterns on the final price forecast.

In this thesis, different models will be explored for both wind power generation prediction and electricity price forecasting. The literature does not provide conclusive evidence that any particular model or set of features consistently outperforms others; rather, the best approach tends to depend on the specific case. Therefore, the techniques used in this thesis will depend on the prior data analysis. It is worth noting that instead of seeking the most precise forecasts possible, this thesis aims to exemplify different models for subsequent integration with the optimization model. These models will be explained in depth in later sections.

Regarding the integration of APPs into electricity markets, the literature presents various market considerations. For example, [8] develops an algorithm to maximize the net income of three WFs combined with BESS in the day-ahead and intraday markets, considering only the WF for day-ahead optimization.

Other works, such as [9], also consider reserve markets, specifically FCR-N. [10] examines the combined participation in the day-ahead market, FCR-N, and FCR-D, while [11] focuses solely on a BESS participating in FCR-N and FCR-D. The authors of these studies analyze the economic viability of participating in each scenario, both separately and in combination, and conclude that multi-market participation is the most profitable.

All the cited articles use MILP (Mixed-Integer Linear Programming) to model the optimization algorithms, which is the same technique that will be employed in this thesis.

It is worth mentioning that this thesis will not study the technical feasibility of WTs participating in reserve services, assuming no technical impediments. However, literature on how WTs can participate in reserve markets by using pitch-control to adjust its power output can be found in [12] and [13].

2.2   Contributions and thesis scope

Despite the existing literature on the integration of aggregated power plants (APP) into electricity markets, to the best of the author's knowledge, no studies have addressed the integration of data-driven models for both price and wind power forecasting, combined with optimization models to derive optimal bidding strategies in the joint markets of FCR-D, FCR-N, and day-ahead market. Additionally, this thesis will introduce the consideration of imbalance settlement, exploring whether wind farms can benefit by bidding different quantities than their forecasts.

This study will examine bidding strategies based on forecasts, assess the reliability of the model by comparing expected and realized profits, and analyze how these factors vary when using forecasts with perfect information of prices and wind power. The research aims to provide a comprehensive understanding of how data-driven forecasting and optimization models can be effectively integrated to define bidding strategies and improve the economic performance of APPs in multiple electricity markets.

Furthermore, this thesis does not focus on developing the most precise forecasts or conducting an exhaustive economic study on the integration of these plants into the market. Instead, it aims to demonstrate how the aforementioned techniques can be combined to plan bidding strategies in electricity markets. The main objective is to show how integrating these methods can be used to develop effective offer strategies, despite the constraints of time and complexity in creating highly detailed forecasts or evaluating all the possible electricity markets.

3 Theoretical Framework

3.1 Wind turbines

The scope of this thesis regarding the operation of WTs ranges from the acquisition of data from the SCADA system, their analysis and filtering, to their potential operation in different electricity markets. Therefore, this section describes the basic WT concepts necessary in the context of this thesis. These are power extraction from the wind resource, power curves and their different operating ranges.

Unless otherwise specified, the theory explained in this section is based on [14].

A WT captures the kinetic power of the wind resource and transform it into mechanical power using the aerodynamic characteristics of its rotor, which is then transformed into electrical power by the generator. The maximum wind power available for a WT with rotor area A_{wtr}, is calculated as

$$P_{max} = \frac{1}{2} \cdot \rho \cdot A_{wtr} \cdot V_0^3 \tag{3.1}$$

where ρ is the air density in kg/m^3 and V_0 is the free wind speed in m/s. Therefore, wind power increases with the cube of wind speed and the square of the rotor area. However, it is not physically possible to extract all the power available in the wind, so the power coefficient C_p is defined as the fraction that can be converted into mechanical power, being the limit at 0.593, defined as the Betz limit [14]. The final aerodynamic power equation is:

$$P = \frac{1}{2} \cdot \rho \cdot A \cdot V_0^3 \cdot C_p(\Lambda, \theta) \tag{3.2}$$

As can be seen in the previous equation, C_p actually depends on two variables, the tip-speed ratio (Λ) and the pitch angle (θ). In order to explain these parameters, the aerodynamic profile of a WT is introduced below.
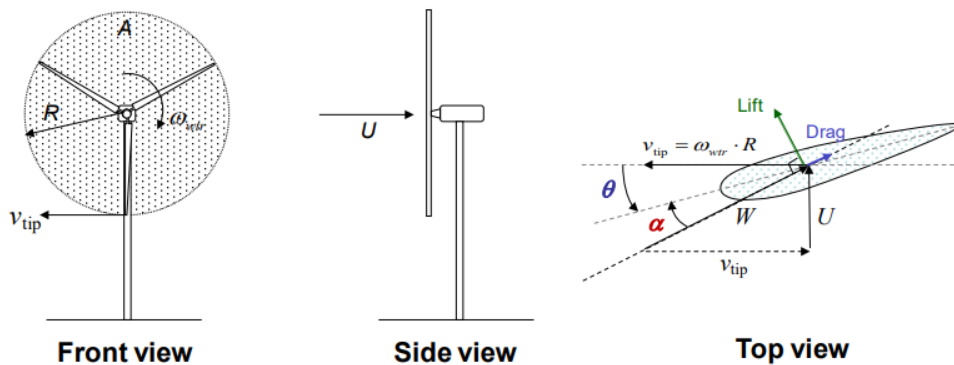

Figure 3.1: Top view, side view and airfoil view of a Wind Turbine. From: [15]

The pitch angle of the blade is the angle between the plane of rotation and the chord line, and it can be actively controlled by means of a pitch controller. Moreover, the tip-speed ratio is the ratio between the linear tip speed and the free wind speed, defined as

$$\Lambda = \frac{w_{wtr} \cdot R}{U_0} \tag{3.3}$$

where w_{wtr} is the rotor rotational speed in rad/s and R_{wtr} is the rotor radius in m. Thus, the maximum efficiency of a WT can be controlled by adjusting Λ and θ. In practice, there are WTs that operate at a fixed rotational speed and with a fixed pitch angle. In such cases, the optimal C_p is only achieved at a specific wind speed. However, the WTs studied in this thesis are VSPR (Variable Speed Pitch Regulation). With this type of WT, the optimal C_p corresponds to an optimal θ and Λ, so it can be achieved for a range of wind speeds by controlling the rotational speed.

The above theory has been included to introduce the power curve of a WT, and its different operating modes, since, as will be seen later, for a WT to be able to participate in reserve markets, it must be able to actively curtail its power, either to reserve available power and provide up-regulation reserve, or to reduce its power if it is necessary to provide down-regulation reserve.

The power curve of a WT relates the electrical power produced as a function of the wind speeds within its operating range, defining the cut-in wind speed as the minimum wind speed at which the blades start to rotate, and cut-out wind speed as the maximum operating speed achievable to avoid damage. Each turbine model has its own power curve. Figure 3.2 shows the electrical power, rotational speed, pitch and Cp as a function of wind speed for a VSPR WT, differentiating its operating modes.

A distinction can be made between two zones of operation: the optimisation zone (ABC) and the power limitation zone (CD). Between A and B the WT is operated at maximum efficiency by keeping the pitch angle constant and at its optimal value (normally close to zero), and the tip-speed ratio also constant and at its optimal value by controlling the rotational speed, increasing it as the wind speed increases. Between B and C, the pitch is still maintained at its optimal value. However, at B the nominal rotational speed is reached, which is kept constant, causing the tip-speed ratio to decrease as the wind speed increases. Therefore, the efficiency (C_p) begins to decrease. Finally, between C and D, once the nominal power of the WT is reached, it is necessary to start increasing the pitch of the blades to keep the power constant as the wind speed increases. In this zone there is a drop in efficiency [15].

By actively controlling both rotor speed and pitch angle, the turbine can at any time decrease its power to a given setpoint if it has to keep a certain reserve to participate in the markets considered in this thesis. On the other hand, if the wind speed is sufficient, and the turbine is operating under curtailment, it is also possible to increase the power setpoint.

3.2 Electricity markets

The bidding zone considered in this thesis is East Denmark or DK2. Electricity markets in DK2 can be broadly grouped as energy markets or reserve markets/ancillary services. Within energy markets, Day-Ahead market (DAM), Intraday Market and Balancing or Regulation market are included, while for ancillary services related to frequency, the services available are Fast Frequency Reserve (FFR), Frequency Containment Reserve (FCR), Automatic Frequency Restoration Reserve (aFRR) and Manual Frequency Restoration Reserve (mFRR). All of these markets differ in terms of closing gates, payment schemes and technical requirements.
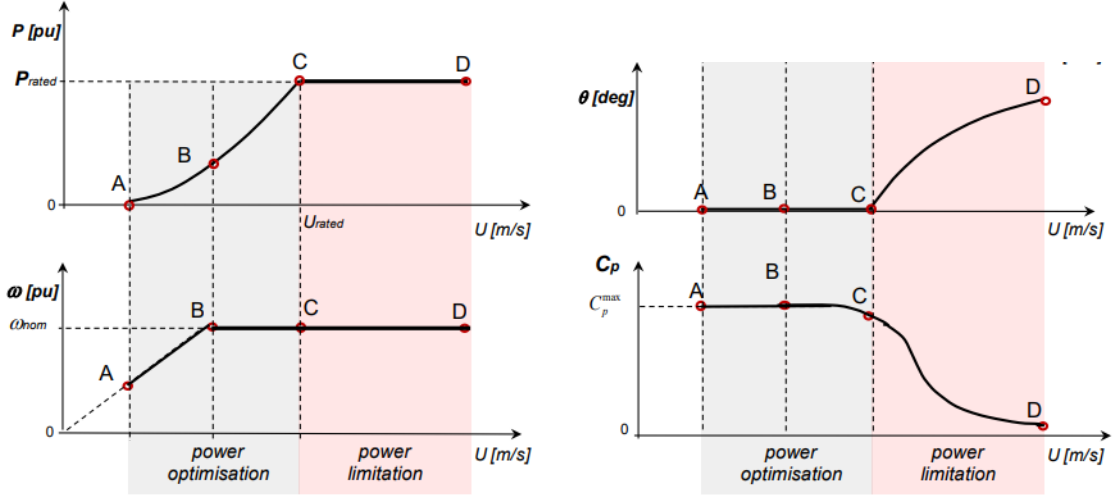
Figure 3.2: Operating zones of a VPR Wind Turbine. From: [15]

However, for the scope of this thesis not all the markets will be considered. The markets to be considered for the bidding strategy algorithm are the DAM and FCR, so their details will be summarized in this section. Although not actively participating in the balancing market, a term reflecting the imbalance cost for deviating from the original schedule will be taken into account, so how producers are penalized in this regard is also explained in this chapter.

3.2.1 Day-Ahead Market
The Day-Ahead market follows a market-clearing procedure and is managed by NordPool, where it is also known as Elspot. In this market, market participants have to submit their quantity-price bids before the market gate closure at 12:00 CET of day $D-1$ for every hour of the operating day, D. Once participants have submitted their bids all orders are aggregated into supply and demand curves for each delivery hour and the day-ahead price is calculated via algorithms with the objective function of maximizing social welfare, considering the Available Tranfer Capacity between bidding areas. The outcome is the market clearing price in €/MWh for each hour.

3.2.2 Imbalance settlement
The purpose of the balancing or regulation market, which is cleared by the TSO (Energinet) 45 minutes before actual delivery, is to balance generation and demand in real time.

For each hour, the clearing of the market results in two different prices corresponding with the price of mFRR activation of the dominant direction (up or down). These balancing energy prices are denoted as up-regulation price and down-regulation price.

The final imbalance price in the imbalance settlement depends on the pricing configuration. In DK2, with the previous dual-pricing model, two different imbalance prices were considered for production and consumption of imbalance. However, since 2021, a single-price scheme has been implemented, where only one imbalance price is calculated. In case of activation of balancing energy, the imbalance price is the mFRR price in the dominating direction, while it is the day-ahead price if there is no net activation [16].

Based on the above, the following relationships between regulation prices, imbalance price and day-ahead price can be established. If for an hour there is a power deficit (dominant direction of mFRR activation is up), the up-regulation price will be higher or equal than

the day-ahead price, while the down-regulation price will be equal to the day-ahead market price. In this case, the final imbalance price will be the up-regulation price. If, on the contrary, there is power excess (dominant direction of mFRR activation is down), the down-regulation price will be lower or equal to the day-ahead price, being the up-regulation price equal to the latter. In this case, the imbalance price will be the down-regulation price. Finally, if the net activation is 0, all three prices will be equal to the day-ahead price.

Denoting day-ahead market price, down-regulation price, up-regulation price and imbalance price as λ_t^{DAM}, λ_t^{DR}, λ_t^{UR} and λ_t^{IMB} respectively, the previous behaviour is depicted by the following equations.

$$\lambda_t^{UR} \begin{cases} = \lambda_t^{DAM} & \text{if power excess} \\ = \lambda_t^{DAM} & \text{if system is balanced} \\ \geq \lambda_t^{DAM} & \text{if power deficit} \end{cases} \tag{3.4}$$

$$\lambda_t^{DR} \begin{cases} \leq \lambda_t^{DAM} & \text{if power excess} \\ = \lambda_t^{DAM} & \text{if system is balanced} \\ = \lambda_t^{DAM} & \text{if power deficit} \end{cases} \tag{3.5}$$

$$\lambda_t^{IMB} \begin{cases} = \lambda_t^{DR} & \text{if power excess} \\ = \lambda_t^{DAM} & \text{if system is balanced} \\ = \lambda_t^{UR} & \text{if power deficit} \end{cases} \tag{3.6}$$

Finally, the price applied in the imbalance settlement to producers who have created an imbalance due to a power deviation with respect to the scheduled power in the day-ahead market is λ_t^{IMB}. The amount they will have to pay will depend on whether their imbalance, defined as produced power minus contracted power, is positive or negative and follows the following formula.

$$C_t^{IMB} = \begin{cases} (\lambda_t^{IMB} - \lambda_t^{DAM}) \cdot \Delta P_t & \text{if} \quad \Delta P_t \geq 0 \\ -(\lambda_t^{DA} - \lambda_t^{IMB}) \cdot \Delta P_t & \text{if} \quad \Delta P_t < 0 \end{cases} \tag{3.7}$$

From (3.7), it can be concluded that creating an imbalance does not always result in a payment, but can sometimes result in a revenue. In fact, if the imbalance of the plant occurs in the favourable direction of the system, it will be an income, while if it occurs in the unfavourable direction, it will be an expense. For example, in a power deficit situation where, as explained above, the imbalance price is higher than the DAM price, if the plant incurs in a positive imbalance (more energy produced than scheduled), the C_t^{IMB} term becomes positive.

Although the APP considered in this thesis does not actively bid in the balancing market, the above term is incorporated in the bidding strategy algorithm to address situations in which it may be favourable for the wind farm to incur an imbalance by bidding a different amount than the forecast amount of available wind power.

3.2.3 Frequency Containment Reserve
The frequency ancillary services considered and their market practicalities are summarized in this section.

Frequency Containment Reserve, also known as Primary Reserve, is used to stabilize the frequency in the power grid in the event of frequency jumps. In DK2, this service is

divided in Frequency Containment Reserve-Normal Operation (FCR-N) and Frequency Containment Reserve-Disturbance (FCR-D).

FCR-N is activated for deviations of +/-100mHz with respect to the nominal grid frequency of 50Hz. It consists of both up-regulation and down-regulation, but it is procured as a symmetrical product, meaning that the producer bid into the market is only one, being up and down regulation procured together.

On the other hand, FCR-D is asymmetrical, meaning that there is a market for up-regulation and one for down-regulation. Up regulation is activated if the frequency falls below 49.9 Hz, while up-regulation is activated if it surpases 50.1 Hz.

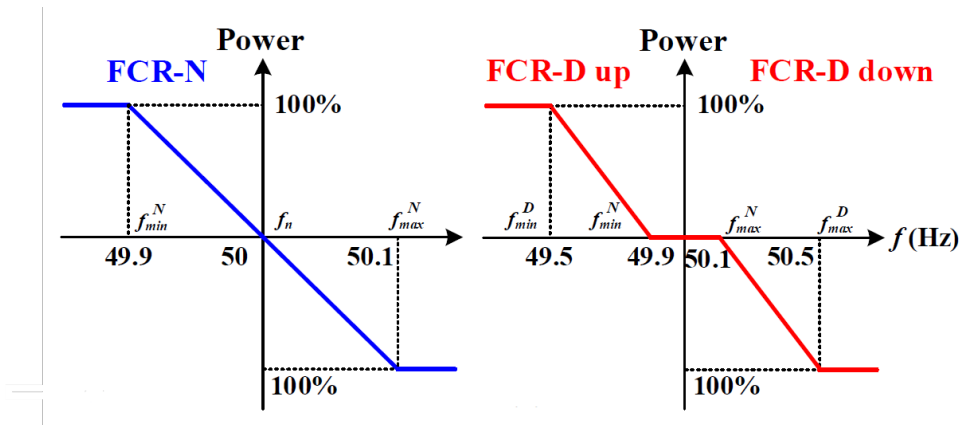The activation rate of these services follow the droop profiles presented in Figure 3.3.

Figure 3.3: Droop profiles on system level for FCR-N and FCR-D. From: [10]

Both FCR-N and FCR-D within the Nordic Synchronous area are a joint responsibility of all the TSOs in the area. For FCR-N the combined requirement is to provide 600MW, of which Energinet is obliged to provide a proportional share, being in 2024 +/- 18MW. On the other hand, from the total requirement of FCR-D in the Nordic area, the share corresponding to Energinet is +42MW in FCR-D up and -42MW in FCR-D down.

There is another difference between FCR-D up, FCR-D down and FCR-N in terms of payment schemes. The three services are paid for capacity reservation at the clearing price, in €/MW, of each service. However, while energy activations corresponding to FCR-N are paid at up-regulation price and down-regulation price depending of the direction of activation, FCR-D does not have any payment for energy supplied.

In terms of practicalities in the market, the three services behave in the same way. On the one hand, for each of them there are two auctions on which to bid. While both occur on D-1, the early auction closes at 00:30 CET while the late auction closes at 18:00 CET. In both auctions, bids must contain a price in €/MW with up to two decimal points and a minimum amount of 0.1MW.

The pricing scheme for reserve allocation is pay-as-clear, with the bids being sorted by price. However, the clearing of the market does not only consider bids price, but it also minimizes TSOs cost, so bids with prices lower than the marginal price might get out of the market. On the other hand, bids are always accepted in its totality or not at all. [17]

Finally, there are several exceptions for limited energy reservoir (LER) units, which are of special interest in this thesis as the integration of a BESS in this market will be considered.

Units that can not provide full energy supply for four consecutive hours are considered as LER. These units must have a Normal State Energy Management (NEM) system that allows the battery to change the baseline power to reduce the imbalance caused by state of charge (SOC) management [18]. The BESS in this thesis will be considered to have a NEM.

3.3 Uncertainties and forecasting methods

Both the available wind power capacity and the expected market prices are key elements to consider when establishing a strategy for participation in the markets described above. Since none of these elements are known in advance, there is a need to predict them as accurate as possible.

Although there are probabilistic forecasting methods that provide good results, based on the literature review, deterministic forecasts give good enough results to be used within the focus of this thesis, so it is decided that the methods studied will be of this type.

The difference is that while deterministic forecasts generate a single value for each horizon, probabilistic forecasts generate a range of possible values and allow a probability to be assigned to each of them. In this case, by using point forecasts, a price will be predicted for each market and a wind power available for each hour within the forecast horizon.

Since different forecasting methods will be studied depending on the variable to be predicted, they will be defined and described in their corresponding sections. In contrast, in this section the metrics used to evaluate the performance of each one of them are explained, as they will be common to all the variables predicted.

The selected metrics are Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Normalized Root Mean Squared Error (NRMSE). For all of them, the lower the value the better is the prediction.

First, RMSE and MAE have been used as metrics to assess performance in the same units as the original data. If the variable to be predicted is a price in €/MWh, the RMSE and MAE results will be in €/MWh, so these metrics give an idea of the performance of the models in physical quantities. The difference between the two is that RMSE calculates the squared error, so it is more sensitive to outliers, giving them more weight. MAE is less sensitive to outliers as it calculates only the absolute value of the error. It is calculated by following the equations below:

$$RMSE = \sqrt{\frac{1}{N} \cdot \sum_{t=1}^{N}(y_t - \hat{y}_t)^2} \tag{3.8}$$

$$MAE = \frac{1}{N} \cdot \sum_{t=1}^{N}|y_t - \hat{y}_t| \tag{3.9}$$

Where N is the total number of observations, y_t is the actual value and $\hat{y}_t$ is the predicted value at hour t.

However, if models are compared for different days and the data for these days have different scales, the use of MAE and RMSE does not provide as much information, as it is not the same to have a MAE or RMSE of 5 €/MWh on a day when prices ranged between 5 and 10 €/MWh, than for a day when they ranged between 100 €/MWh and 120 €/MWh. Therefore, it has been decided to evaluate the forecasts also using normalised metrics.

On the one hand NRMSE normalises the RMSE by dividing by the range of the original data $(y_{t,max} - y_{t,min})$, so NRMSE has no physical units. On the other hand, MAPE normalises the error by dividing by the actual error, so the result is in value per unit (p.u) or in percentage if multiplied by 100. They are calculated as follows:

$$MAPE = \frac{1}{N} \cdot \sum_{t=1}^{N} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \tag{3.10}$$

$$NRMSE = \frac{RMSE}{y_{t,max} - y_{t,min}} \tag{3.11}$$

4 Wind Power Forecasting

This section is dedicated to the prediction of the wind power available for the wind farm, information that will be used as input to the strategic bidding model that will be explained in the next section. Firstly, a description of the available data will be given, followed by its analysis, the selection of the variables to be used by the models and, finally, the prediction techniques used will be explained and the results will be analysed.

4.1 Data description

For the wind power forecasting, the actual data of a wind farm composed of 12 WTs rated at 2000kW each, located in the north-east of Germany, is used. Note that the aim of this thesis is to obtain optimal bids for an aggregated plant in Denmark, more precisely in DK2, so it will be assumed that this wind farm is actually located in Denmark. The necessary transformations will be explained in this section.

The data extracted from year 2020 to 2022 is obtained from the SCADA system with a resolution of 10 minutes. Thus, there are more than 157,000 data points of each parameter, with more than 50 different parameters. Among these parameters are the WT status (error or operational), variables related to wind speed measured by the anemometer and rotor rotation speed (average speed, maximum, minimum, standard deviation...), nacelle position, pitch angles, temperatures of different WT components, reactive power, voltages, currents and, the variables to forecast, average active power and generation.

4.2 Data analysis and feature selection

Since the objective is to forecast wind power, the time series for one WT in the park is illustrated in Figure 4.1. The time series should oscillate between 0kW and 2000kW, which is the rated power of the WT. However, some outliers are observed, which can be due to defective functioning, signal errors, or noise.

Figure 4.1: Time Series of wind power for a WT in the WF of study
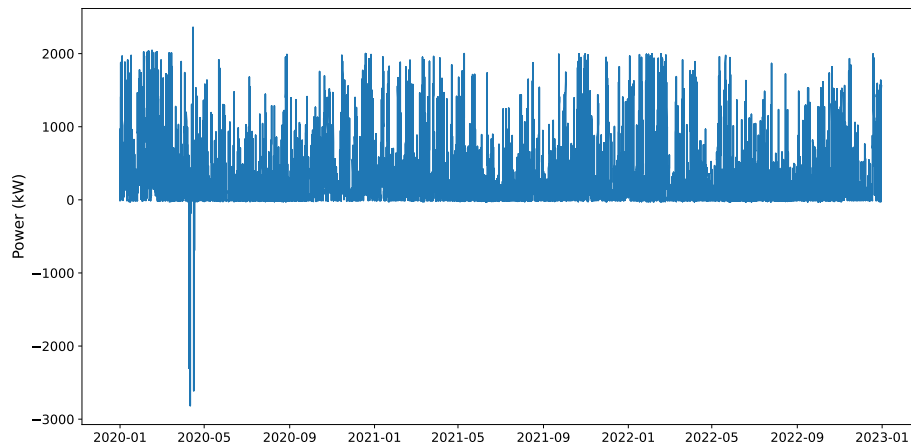
In order to perform a correct analysis of the data and not induce errors in the models developed later, the first step consists in the elimination of outliers from the dataset. There are numerous articles in the literature that use machine learning models for the elimination of outliers. However, another possible procedure is the use of the power curves of the WT

model studied, which is available in this thesis. As seen in the theoretical framework, the power curve shows the relationship between the wind speed measured at hub height and the active power generated by the WT. However, these power curves are only a guideline, since in reality the power of a WT will never follow the curve perfectly due to factors such as temperature, turbulence, air density, terrain, etc... For this reason, manufacturers also provide a range of validity for the power curve, within which the WT is considered to be under normal operation. In this thesis, both the power curve of the WT model and its validity range are available, so it has been decided to use them for the elimination of anomalous measurements, using only the WT normal operating data points

For this purpose, all the wind speed datapoints have been taken and interpolated with them on the maximum and minimum validity curve. If the actual power datapoint for a given wind speed is beyond the valid range, that datapoint is removed. Figures 4.2a and 4.2b show the theoretical and actual power curve before and after data filtering for one WT, although the same procedure is done for all WTs in the WF. It should be noted that before performing this procedure, the datapoints where the WT is under error status have been removed, so these points do not appear in the pre-filtering figure, where, however, values outside the correct operating range can still be observed. On the one hand, the observations that form a horizontal line correspond to the WT operating under curtailment. The other scattered points, which have not been filtered by eliminating the points with the WT in error state, correspond to power measurements in moments after the correction of an error, when the WT is trying to return to its optimal operating point.

(a) Before filtering.   (b) After filtering
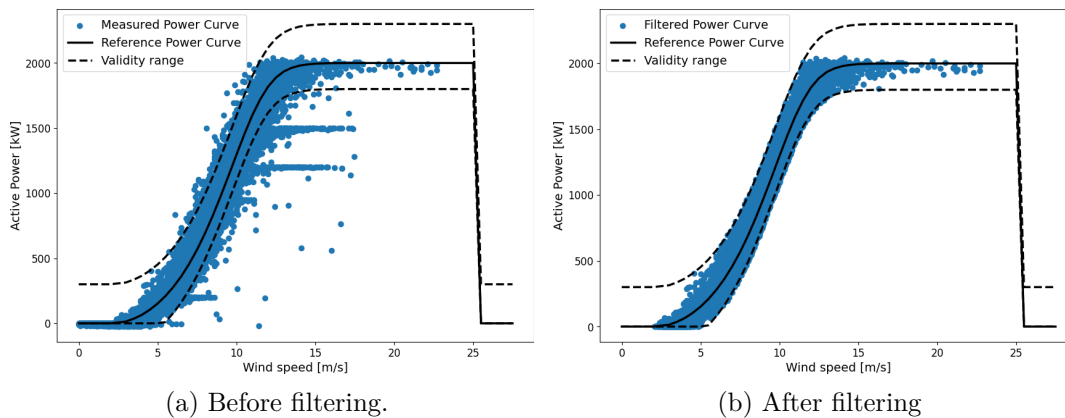
Figure 4.2: Power Curve for WT1, before and after filtering the outliers.

Once the data have been filtered, the next step is the selection of the explanatory variables for the prediction algorithms. For this purpose, from the total number of available variables, only those that are external to the operation of the WT have been selected for analysis. These are atmospheric variables such as wind speed, wind direction or temperatures, as well as the month of the year or the active power time series itself, which will be used to analyze whether past measurements have an influence on future predictions.

A first analysis of the explanatory variables can be to visualize the power curves of a WT under certain conditions. In this case, the power curves as a function of wind direction and month of the year are shown in Figures 4.3 and 4.4 respectively. It should be noted that this visual analysis is not based on finding out whether the WT reaches its maximum capacity under certain conditions, since that depends on the wind speed, but on how much the actual curve deviates from the manufacturer's curve. In addition, deviations can be

caused by other variables such as air density, which does vary by seasonality and therefore would be valid to explain the curves as a function of the month, but it distorts the analysis to some extent according to the wind direction.
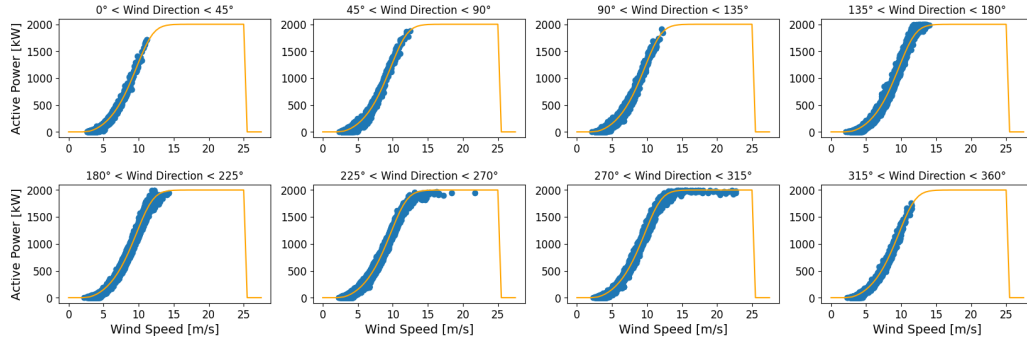
Figure 4.3: Power curve of WT1 per wind direction

Under ideal conditions, the WTs should perform the same regardless of wind direction as long as the nacelle can be correctly aligned with the prevailing direction. In reality, deviations may be observed due to factors such as orographic obstacles, or the existence of other WTs that cause wind turbulence for certain directions. In this case, no major differences are observed from Figure 4.3, although wind direction as a factor will continue to be considered in subsequent analyses.
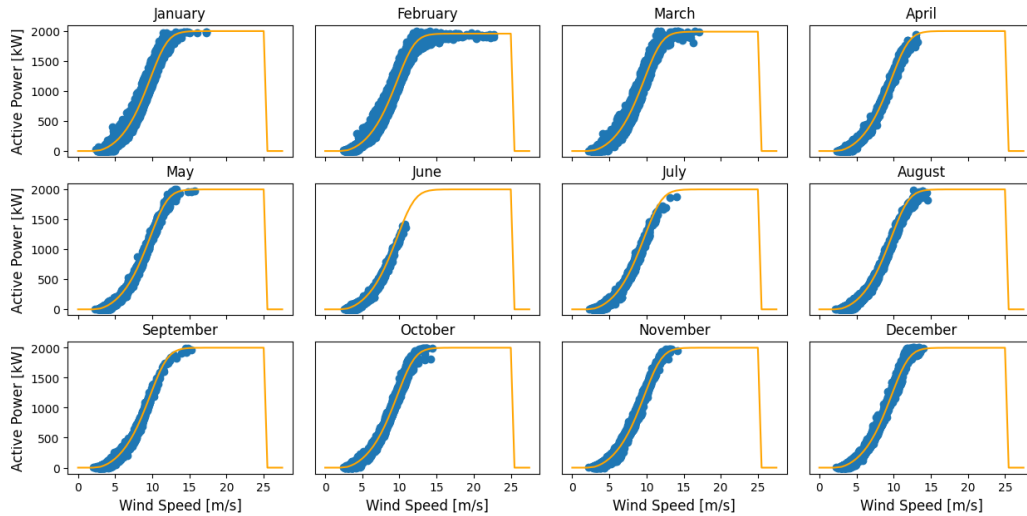
Figure 4.4: Power curve of WT1 per month

As for the month of the year, the differences in terms of deviations from the power curve are mainly due to the difference in temperatures. Higher temperatures mean lower air density, (4.1), which leads to lower power extraction according to (3.2) and therefore the power curves tend to be below the manufacturer's curve. In the case of WT1, it can be seen that in June, July and August the datapoints are not as centered around the nominal curve as they are for colder months, so the month of the year may be an important factor to consider.

$$\rho = \rho_{ref} \cdot \frac{T_{ref}}{T_{amb}} \tag{4.1}$$

The previous equation models the relation between air density and air temperature, where $\rho_{ref} = 1.223 kg/m^3$ is the air density in standard conditions, $T_{ref} = 288.15K$ is the reference temperature and T_{amb} is the actual temperature [19].

Continuing with the analysis of variables, the next step is the exploration of the correlation between the different variables. The variables chosen for exploration are wind power, wind speed and direction, temperature at hub height and month of the year. However, before performing the analysis, some changes are made.

First of all, the wind direction contains values between 0 and 360 degrees, so if the correlation with other variables is evaluated, a wind direction of 0 degrees and a wind direction of 360 degrees would give completely opposite information when, in practice, they are the same direction. Therefore, the wind direction will be evaluated as two different variables, using the cosine and sine of the actual measurement. Hereafter these variables will be referred to as "Wind Direction X" and "Wind Direction Y" respectively. On the other hand, the month of the year is a categorical variable, so in order not to work with label data directly and to eliminate the influence of its numerical value, one-hot encoding is performed.

A scatter pair plot is included in which the relationship between the variables is visually represented, Figure 4.5. In turn, Figure 4.6 shows the correlation coefficients between variables using both the Pearson coefficient and the Spearman coefficient.
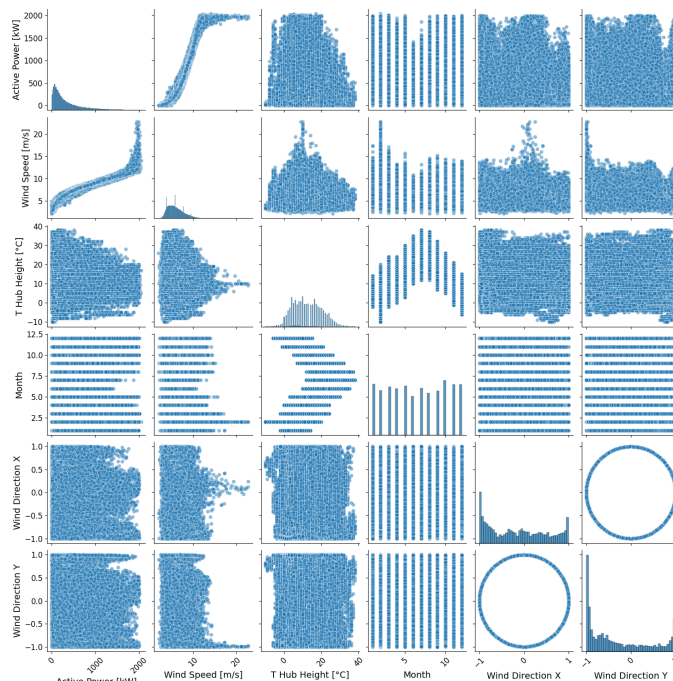

Figure 4.5: Scatter Pair Plots for explanatory wind power variables

Aggregated energy intelligence

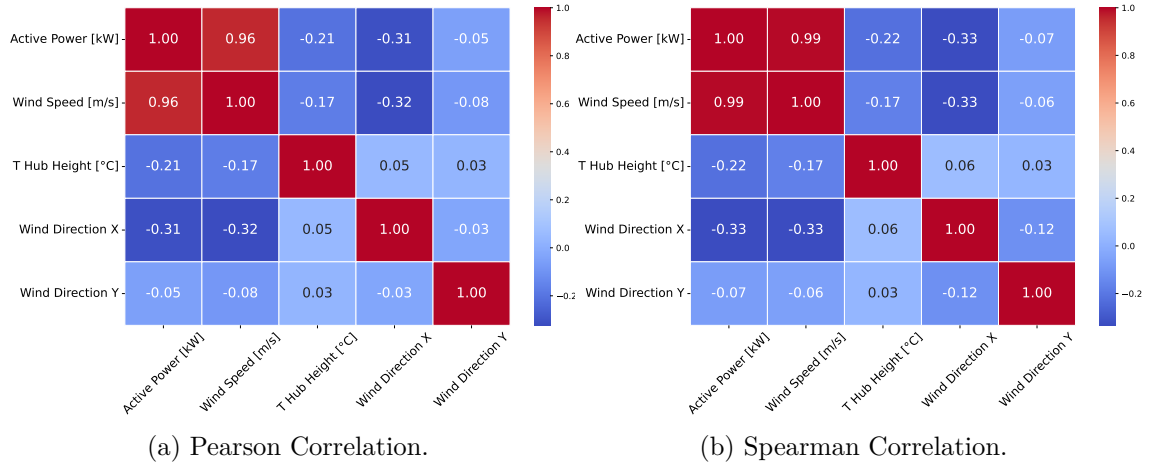(a) Pearson Correlation.   (b) Spearman Correlation.

Figure 4.6: Heatmaps for Pearson and Spearman correlation factors.

The choice of using two correlation factors is due to the fact that the Pearson correlation factor is very suitable for determining linear relationships between variables. However, if the relationship between variables is not linear, it has certain limitations and the Spearman coefficient is better [20]. While the Pearson coefficient is calculated as the covariance of two variables divided by the product of the individual standard deviations, the Spearman coefficient is based on the ranks of the variables rather than on the values themselves [21].

In addition, because months are categorical variables, the literature suggests that Pearson's or Spearman's coefficient are not the most optimal measures to calculate their correlation, but that other techniques such as point-biserial coefficient or ANOVA are more appropriate. Finally, it has been decided not to continue with months as an explanatory variable in order to simplify the prediction models that will be studied later. In this case, given the strong correlation between power and wind speed, such simplification may be correct since temperature is also considered and, as explained above, both factors can explain by themselves the variations relative to the time of the year.

The results indicate a strong correlation with wind speed, while temperature and wind direction X also show some correlation. Furthermore, it may also be interesting to study the influence of past power values on future values. For this purpose, autocorrelation (ACF) and partial autocorrelation (PACF) functions are calculated and shown in Figure 4.7. While ACF, measures the correlation between a time series and its lagged version over different time periods, PACF measures the same correlation but partialling out the effects of the values in-between [22].
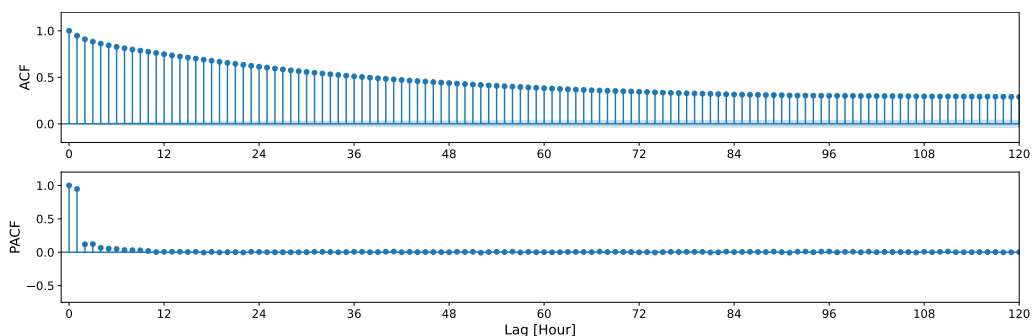

Figure 4.7: ACF and PACF of Wind Power.

From the PACF plot it can be concluded that only the first three lags show correlation, which indicates that the ACF values, can be explained by the first three. However, in this thesis the objective is to predict the wind power for the next day and the closing of the day-ahead market occurs at 12:00 of the day D-1, considering D as the operating day, so the horizon for the forecast would be 36 hours. In addition, due to market considerations and simplifications that will be explained later, the forecast will be made, at the latest, at 00:30 am on D-1, increasing the forecast horizon to 48 hours. Therefore, it must be considered that if it is decided to incorporate a lagged version up to 3 hours of the active power as an explanatory variable, this will be based on predicted and not real values.

To make the final selection of the variables, a dataset is created in which the explanatory variables to be studied are: wind speed in m/s (WS), temperature at hub height in °C (T), wind direction X ($WD-X$), wind direction Y ($WD-Y$) and the time series of wind power delayed by one ($P(t-1)$), two ($P(t-2)$) and three hours ($P(t-3)$). Therefore, there are a total of 7 possible explanatory variables.

Two sequential feature selection techniques, forward feature selection (FFS) and backward feature selection (BFS), are used to reduce the dimensionality of the problem. These models are based on starting from a subset of the set of independent variables and modifying it depending on whether better results are obtained. For both models, linear regression is considered as the prediction model and the coefficient of determination, R^2, as a metric of precision, which calculates the proportion of the variance of the variable to be predicted that is explained by the input variables. Furthermore, when evaluation each subset, cross-validation is performed using five k-folds.

Thus, FFS is started from a model M without features and evaluate all the subsets formed by a single feature. If any of these models gives a better result than model M, the feature in question is added to the set, creating the subset M_i. Then, all the subsets formed by the first chosen feature (i) and each of the remaining features (j) are evaluated. Again, if any model is better than M_i, the new feature is added, creating the subset $M_{i,j}$. The process continues until adding a new feature does not improve the previous subset. On the other hand, BFS is the reverse process, it starts from the total set of features and evaluates all the subsets created by eliminating each one of the features. If any new model is better, the feature corresponding to it is eliminated and the process continues until eliminating more features does not improve the model [23].

Forward Selection Results	Coefficients of determination					
Subsets	k-Fold 1	k-Fold 2	k-Fold 3	k-Fold 4	k-Fold 5	Avg Score
WS	0.877	0.917	0.908	0.922	0.899	0.905
WS + P(t-1)	0.922	0.956	0.949	0.953	0.948	0.946
WS + P(t-1) + T	0.922	0.956	0.949	0.953	0.948	0.946
WS + P(t-1) + T + P(t-2)	0.922	0.956	0.950	0.954	0.949	0.947

Table 4.1: Results for forward feature selection

Backward Selection Results	Coefficients of determination					
Subsets	k-Fold 1	k-Fold 2	k-Fold 3	k-Fold 4	k-Fold 5	Avg Score
WS + T + WD-X + WD-Y + P(t-1) + P(t-2) + P(t-3)	0.924	0.957	0.950	0.954	0.949	0.947
WS + T + WD-X + P(t-1) + P(t-2) + P(t-3)	0.925	0.957	0.950	0.954	0.949	0.947
WS + T + WD-X + P(t-1) + P(t-2)	0.925	0.957	0.950	0.954	0.949	0.947
WS + T + P(t-1) + P(t-2)	0.925	0.957	0.951	0.954	0.950	0.948

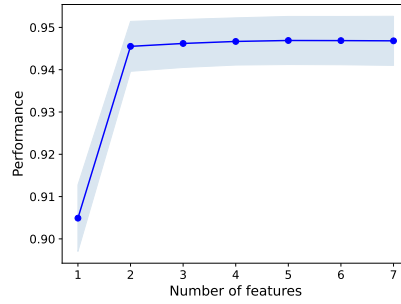Table 4.2: Results for backward feature selection

Figure 4.8: Sequential Feature Selection results, R^2

The results of these iterative processes are presented in Tables 4.1 and 4.2 . While according to FFS the optimal features are WS, P(t-1) and T, according to BFS, WD-X and P(t-2) should be included as well. However, the average scores are very similar between them after introducing more than two variables, showing that the feature that explains the most variance of the wind power is the wind speed. Figure 4.8 shows a graph with the best scores for each subset formed from 1 to 7 possible features. The light blue area shows the range of scores for the five k-folds considered for each subset, where the dark blue line is the average values. It can be seen how there is a jump from 1 to 2 features (although not very significant), and then stabilizes as more features are included. The algorithm calculates as optimal to use three features in FFS and four in BFS because it considers all the decimals of the R^2 values, but it can be concluded that only the first two, WS and P(t-1), are sufficient for forecasting, with the third variable, which would be T, contributing little value to the model.

Since the prediction models will work with time series, it is necessary to treat the missing values in the dataset. Missing data values from SCADA are mainly due to situations where WTs are under maintenance or if there are communication errors. Different methods of missing value treatment are evaluated below.

The first and simplest option considered for dealing with missing values is to use linear interpolation. This method assumes a linear relationship between known values, connects the two adjacent points with a straight line and calculates the points in between. This method may work well when the interval of missing values is not too large, but leads to loss of detail in the time series otherwise. To exemplify how it works, a missing value section has been extracted from the original time series and linear interpolation has been applied, illustrated in Figure 4.9.
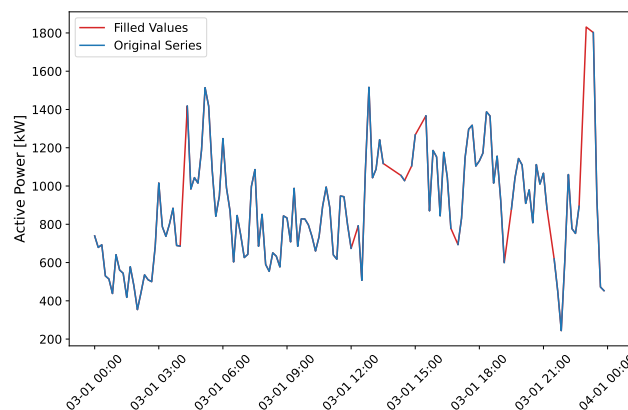
Figure 4.9: Results of filling missing datapoints with linear interpolation

The next method used is known as moving average or rolling mean, it is used to try to capture the variations of the time series in a better way than linear interpolation. This method consists of assigning to each missing datapoint the average of the observations falling within the specified window. Although it generally works better than linear interpolation, if the specified window is very small, it may not be able to assign values to the missing datapoints. On the other hand, if the window is too large, the ability to reflect fluctuations is lost. The same time period as in the previous case is illustrated in Figure 4.10
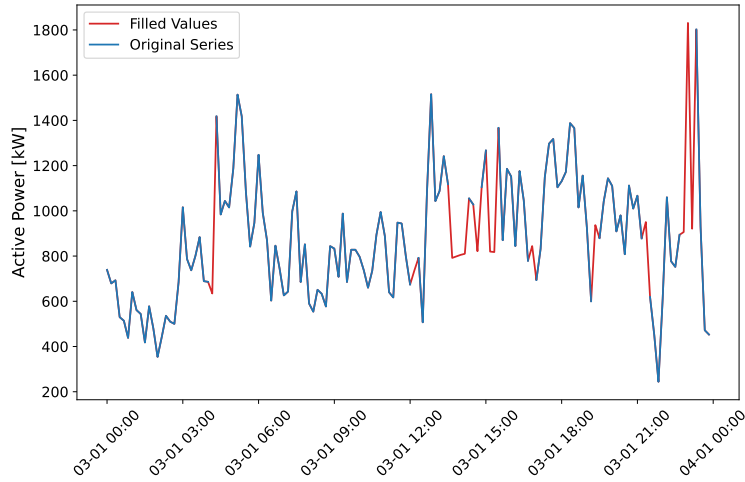
Figure 4.10: Results of filling missing datapoints with rolling mean

Finally, a missing value imputation method based on kNN (k-Nearest Neighbours) models is tested, which calculates and assigns the mean of the k datapoints closest to the missing datapoint [24]. Different measures can be used to calculate the distance, in this case Euclidean distance has been used. The result of the model can be seen in Figure 4.11. Although the result for this date range is similar to the moving average, in general it works better in capturing fluctuations without leaving unassigned values.
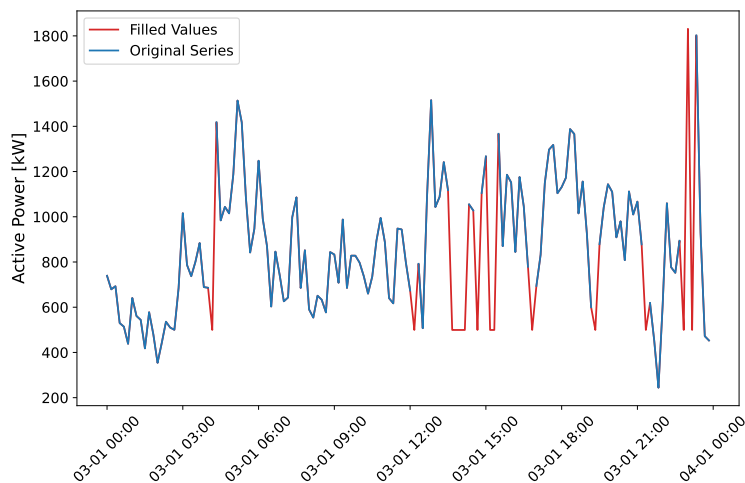
Figure 4.11: Results of filling missing datapoints with kNN

Before analyzing the forecasting methods, it should be noted that the bids in the electricity market have hourly resolution, while the data available come from SCADA, which has 10

minutes resolution. Therefore, resampling has been performed by grouping the data by hours and averaging.

4.3 Forecasting models

This section explains the forecast models considered for wind prediction. It should be noted that the aim is not to develop the best possible or the most complex forecasting method, since this is not the central aspect of this thesis, but to evaluate and exemplify different available options. For a real application, it may be necessary to explore more sophisticated methods.

The strong correlation seen between wind power and wind speed in the previous section suggests that simpler models such as regression can provide good prediction results. However, as mentioned above, the prediction will take place with a horizon of up to 48 hours, so the actual value of any of the input variables is not known. Therefore, the input variables used, wind speed and temperature, will be obtained in the form of external forecasts.

The external forecasts are obtained from the database of the Danish Meteorological Institute (DMI), DMI Open Data [25]. In particular, the platform offers access to the numerical weather prediction (NWP) model HARMONIE (Hirlam Aladin Research towards Mesoscale Operational NWP in Europe), which produces deterministic forecasts of up to 90 different variables for Denmark, Iceland, the Netherlands and Ireland [26]. The forecast results can be obtained from the API provided by DMI by entering the required coordinates and dates. However, the database requires Lambert Conformal Conic projection as the coordinate system, so transformation is required to convert the system coordinates. In addition, the wind speed forecasts are for a height of ten meters above ground level, so it is necessary to transform them to the corresponding wind speeds at hub level. For this purpose, the wind profile power-law as defined in (4.2) is employed.

$$U_1 = U_0 \cdot \left(\frac{Z_1}{Z_0}\right)^\alpha \tag{4.2}$$

The wind profile power-law estimates the wind speed U_1 at height Z_1 above the ground from the wind speed U_0 measured at height Z_0 [27]. In the equation, α is the wind shear coefficient, an empirically derived constant that varies depending on the stability of the atmosphere. In this case it has been decided to use the corresponding to neutral conditions, thus $\alpha = 0.143$ [28]. Lastly, the hub height of the WTs is 100 m.

The selected models to be used are Linear Regression (LR), an Auto-Regressive with eXogenous input (ARX) model, and a Locally Weighted Regression (LWR). These are not the most complex models in the literature, but they are chosen given the data available and the high correlation shown between the dependent and independent variables.

It should be noted that although the features used are external forecasts as explained above, it is decided to test these models using real data first, which would represent perfect forecasts of the input variables. The best performing model will then be evaluated using the actual external forecasts. This is done in order to be able to visualise more clearly how much of the error in the final forecast is induced by the model itself, and how much by the errors of the external forecasts.

Linear Regression
For its simplicity, the first of the models studied is a multiple linear regression in which wind speed and temperature are used as features.

A multiple linear regression follows (4.3), where y is the predicted value of the dependant variable (wind power), β_0 is the intercept, the set $\beta_1, \beta_2, \beta_n$ are the regressors of each feature and ϵ is the random error component assumed to have mean zero and unknown variance. To find the regression coefficients, the least squares method is used, which minimizes the sum of the square of the differences between observations and predictions[29].

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n + \epsilon \tag{4.3}$$

The model is trained and tested for a period of seven days. The final mathematical formulation of the model, with the calculated coefficients, is given in (4.4). On the other hand, the visual representation of the forecast and the results based on the metrics described in the previous chapter are included in Figure 4.12 and table Table 4.3 respectively.

$$\hat{p}_t^W = -0.176 + 1.307 \cdot WS_t + 0.017 \cdot T_t \tag{4.4}$$

In the above equation $\hat{p}_t^W$ is the predicted wind power for hour t, as a function of wind speed and temperature at hour t, WS_t and T_t respectively. It can be seen that the temperature input variable has a significantly lower weight than the wind speed, as expected from the results of the feature selection.
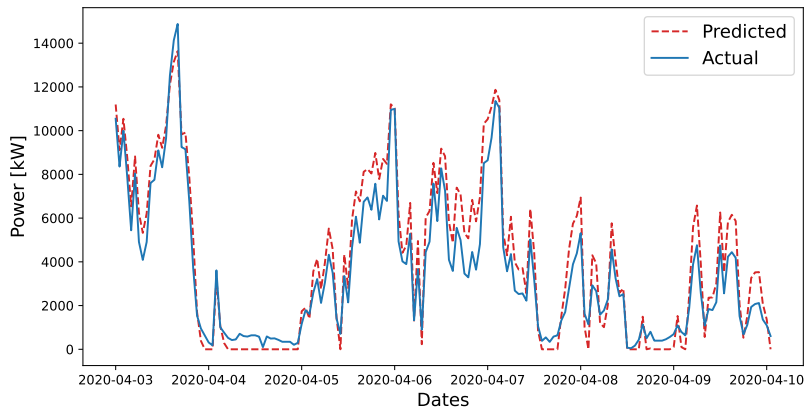
Figure 4.12: Wind power forecasting results with Linear Regression

	RMSE [kW]	MAE [kW]	MAPE [%]	NRMSE[p.u]
Linear Regression	1040	882	46.9	0.07

Table 4.3: Wind power forecasting results with Linear Regression

Analysing the results, the model performs poorly for low wind speeds that cause the power to be close to zero. In fact, for those time periods with power close to zero, the model predicts negative values, so they have been actively modified to be null. Furthermore, if the predictions for these dates are removed from the calculation of the metrics, the MAPE falls drastically from the original 46.9% to 22.79%.

The prediction has been made for the total wind farm, consisting of 12 WTs of 2000kW, so that an RMSE of 1040kW and a MAE of 882kW means an average RMSE of 87kW and an average MAE of 74kW per WT. While these errors are a larger percentage of the actual instantaneous value, as can be seen from the MAPE, the errors with respect to the

rated power of the WT are around 4%, so while it is not an extremely accurate model, the errors are not that significant for its simplicity.

ARX model
In addition to wind speed and temperature, the feature selection algorithms in the previous section indicated the possibility of using delays of the power itself for prediction, so the next model studied is an auto-regressive model with exogenous inputs (ARX).

An ARX model is similar to a LR model but it includes past values of the variable to predict as a feature. Therefore, its mathematical representation is as follows [30]:

$$y(t) + a_1 \cdot y(t-1) + ... + a_n \cdot y(t-n) = c + b_0 \cdot x(t) + b_1 \cdot x(t-1) + ... + b_m \cdot x(t-m) + \epsilon(t) \quad (4.5)$$

Being $y(t)$ the variable to predict and $x(t)$ the input variable, in the previous equation the coefficients a_n and b_m are those referring to the lags of the variable to predict and the input features respectively, being n and m the total number of lags considered.

From the analysis of features selection of the previous section, only the first two lags of the wind power will be considered. For the exogenous variables, wind speed and temperature, the external forecasts are used and no lags considered. However, the wind power values of D-1 are not known at the moment of forecasting the wind power for day D. Therefore, the power lags to be considered are in turn power forecasts for day D-1, for which linear regression similar to the previous model is used.

After training the model, the final mathematical formulation of the ARX model with the obtained coefficients is:

$$\hat{p}_t^W = -0.147 + 1.06 \cdot WS_t + 0.013 \cdot T_t + 0.199 \cdot p_{t-1}^W - 0.019 \cdot p_{t-2}^W \quad (4.6)$$

Again, the wind speed coefficient is the highest as expected. In addition, the coefficients of the power delays become more important than the temperature. The graphical representation of the prediction for the test set is included in Figure 4.13, and the metrics in Table 4.4.
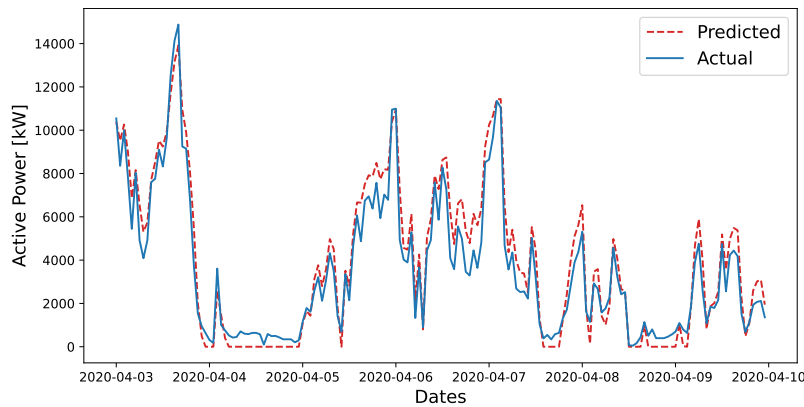

Figure 4.13: Wind power forecasting results with ARX model

	RMSE [kW]	MAE [kW]	MAPE [%]	NRMSE [p.u]
ARX	877	731	42.5	0.06

Table 4.4: Wind power forecasting results with ARX model

Based on the metrics shown in Table 4.4, it is found that this method is better than the linear regression studied above, both in relative and absolute terms. However, the model presents the same problem, predicting as negative the power in periods of low wind speeds. Recalculating the MAPE without accounting for these periods, a result of 18.8% is obtained, which again leads to the conclusion that most of the error is due to the low power periods. As for Figure 4.13, the result is very similar to that of the linear regression, although an improvement can be seen in the peaks of the curve.

Locally Weighted Regression
The two models above are based on linear regression. However, as seen in the theoretical section, the relation between wind speed and power is not linear, but power grows with the cube of wind speed. Therefore, the next model used is a Locally Weighted Regression (LWR) model.

In a LWR model, a regression function is estimated for each query point (x_0) of the test set using only observations close to that query point, which results in a smoother regression function that fits the data better than if only a straight line were used, as in the case of linear regression. To this end, a model is estimated for each query point that optimises the loss function of (4.7), where w_i is a weighting function or kernel that assigns weights to the rest of the datapoints according to their distance to the query point. By converting to matrix notation and assigning weights to the diagonal of the matrix W, the coefficients for the LWR have the solution of (4.8). Finally, the prediction for each query point is obtained with (4.9) [31].

$$J(\theta(x_0)) = \frac{1}{2} \sum_{i=1}^{N} w_i(x_0, x_i)[y_i - x_i^T \cdot \theta(x_0)]^2 \tag{4.7}$$

$$\Theta = (X^T W X)^{-1} X^T W y \tag{4.8}$$

$$\hat{y}_t = (x_0)_t^T \theta \tag{4.9}$$

The function used as a kernel for the predictions is the Gaussian Kernel (4.10), where τ is called bandwidth and determines how fast the assigned weights decay with the distance to the query point, so the higher the value, the more points are considered [32].

At this point, to obtain the best possible model it would be necessary to optimise the value of τ. However, as finding the most accurate forecast possible is not the aim of this thesis, it has been decided to use a fixed value of 0.1.

$$w_i = \exp\left(-\frac{\|x_i - x_{\text{query}}\|^2}{2\tau^2}\right) \tag{4.10}$$

Once the prediction has been made using the LWR model, it can be seen from the results in Figure 4.14 and Table 4.5 that the model predicts with considerably less error than those studied previously. In addition to predicting peak power more accurately, this method does not have the same problem as the previous ones with low values.
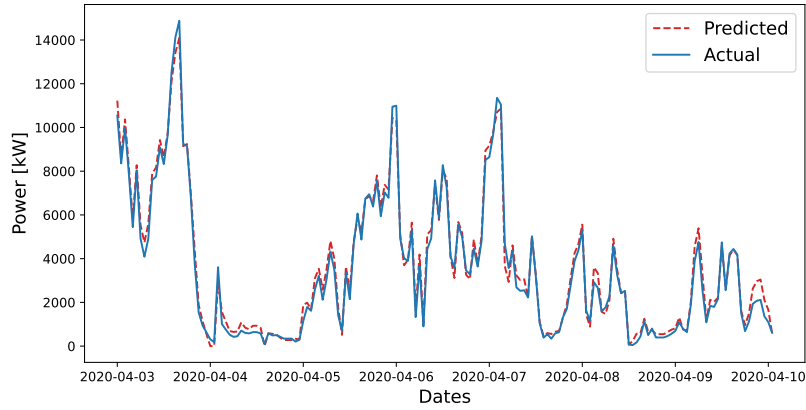
Figure 4.14: Wind power forecasting results with LWR.

	RMSE [kW]	MAE [kW]	MAPE [%]	NRMSE [p.u]
LWR	363	288	20.01	0.02

Table 4.5: Wind power forecasting results with LWR.

It should be noted again that so far the models have been evaluated with the actual values of the features, simulating a perfect forecast scenario of them. However, given that LWR is the best method, it is decided to test the model using the external forecasts obtained from DMI as described above.

In this case the results are quite poor as can be seen in Figure 4.15 and Table 4.6. When analysing the forecasts obtained, it is calculated that the MAPE for the wind speed forecast is around 14%, while the temperature forecast is around 7%. These errors are due to several factors. On the one hand, it is not possible to obtain the forecasts at the exact location of the WTs, but they are taken for a bounding box of coordinates and the average value is calculated. On the other hand, the wind speed predictions are at 10 metres above the ground and, although the conversion has been made according to the height with the profile wind power law, there are other more complex factors that come into play, such as wind stability or orographic factors. Therefore, given this situation, in a real application it would be convenient to be able to develop own forecasts of the exogenous variables to be considered when predicting the wind power.
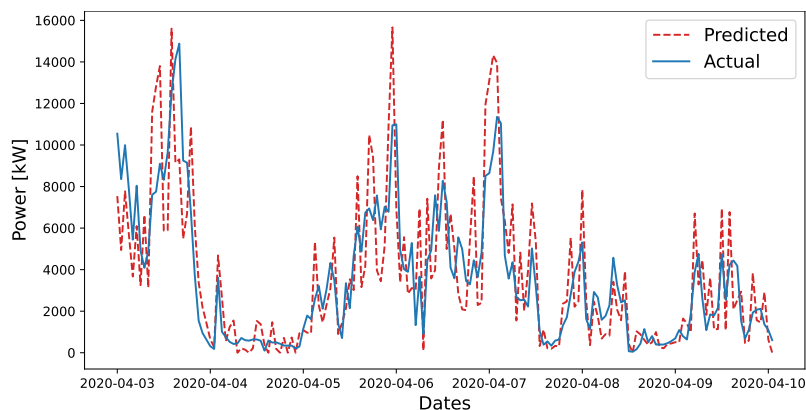
Figure 4.15: Wind power forecasting results with LWR using external forecasts.

	RMSE [kW]	MAE [kW]	MAPE [%]	NRMSE [p.u]
LWR	2053	1631	67.9	0.14

Table 4.6: Wind power forecasting results with LWR using external forecasts.

5 Bidding Strategy Formulation

This chapter introduces the deterministic optimization model developed to obtain the optimal bidding of the APP (WF + BESS) in both energy and reserve markets.

5.1 Considerations

The model is developed considering the APP as a price-taker actor, meaning that the APP has a small enough market share not to influence the market equilibrium and therefore does not have the ability to change the market clearing outcomes in its favor. In practice, this means that it will not be necessary to formulate market prices as a function of the bidding quantities of the plant. On the other hand, the bidding strategy is formulated as a Mixed-Integer Linear Programming model.

The APP is considered to participate in the Day-Ahead (DAM) and reserve markets FCR-N and FCR-D. Therefore, regarding energy markets (EM), the APP participation in the Intraday Market (IM) and Balancing Market (BM) is disregarded in this thesis. However, although the plant will not be considered to bid in the BM to reduce its power imbalance with respect to the offered quantity in the DAM, the cost associated with such deviations will be taken into account. Regarding FCR-N and FCR-D, the plant only participates in the early auction.

5.2 Objective Function

$$OBJ = \max_{\Omega} \sum_{t \in \mathcal{T}} (E_t^{EM} + E_t^{FCR-D} + E_t^{FCR-N}) \tag{5.1}$$

The objective function (5.1) maximizes the daily revenue of the APP, being E_t^{EM}, E_t^{FCR-N} and E_t^{FCR-D} the earnings from participating in the EM, FCR-N and FCR-D markets respectively.

5.2.1 Energy Markets

The earnings from the EM come from the participation of the APP in the DAM and the income/cost from the power deviation with respect to it.

$$E_t^{EM} = (p_t^{EM} \cdot \hat{\lambda}_t^{DAM} + C_t^{IMB}) \cdot \Delta t \quad \forall t \in \mathcal{T} \tag{5.2}$$

In (5.2), p_t^{EM} is the total APP power scheduled to the EM at hour t in MW, $\hat{\lambda}_t^{DAM}$ is the forecasted DAM price at hour t in EUR/MWh and C_t^{IMB} is the imbalance income/cost, which following (3.7) can be expressed as in (5.3).

$$C_t^{IMB} = (\hat{\lambda}_t^{IMB} - \hat{\lambda}_t^{DAM}) \cdot \Delta P_t^+ + (\hat{\lambda}_t^{DAM} - \hat{\lambda}_t^{IMB}) \cdot \Delta P_t^- =$$
$$(\hat{\lambda}_t^{IMB} - \hat{\lambda}_t^{DAM}) \cdot (\Delta P_t^+ - \Delta P_t^-) \quad \forall t \in \mathcal{T} \tag{5.3}$$

where $\hat{\lambda}_t^{IMB}$ is the forecasted imbalance price in EUR/MWh, and ΔP^+ and ΔP^- are the power deviation, in MW, with respect to the scheduled DAM output when such deviation is positive or negative respectively.

5.2.2 Reserve Markets

The FCR-D service is divided into FCR-D down-regulation and FCR-D up-regulation. The revenue is only based on the reserve availability payment of the APP and there is no payment for energy supplied.

$$E_t^{FCR-D} = \hat{\lambda}_t^{D\uparrow} \cdot R_t^{D\uparrow} + \hat{\lambda}_t^{D\downarrow} \cdot R_t^{D\downarrow} \quad \forall t \in \mathcal{T} \tag{5.4}$$

The earnings from the FCR-D service are calculated in (5.4), where $\hat{\lambda}_t^{D\uparrow}$ and $\hat{\lambda}_t^{D\downarrow}$ are the forecasted reserve prices for up and down regulation in EUR/MW, and $R_t^{D\uparrow}$, $R_t^{D\downarrow}$ are the reserve power bids in MW.

In addition, FCR-N is a symmetrical service where up and down regulation reserves are provided as a single product. For this service, the revenue stream includes payment for reserve availability at forecasted FCR-N price ($\hat{\lambda}_t^N$) and payment for energy supplied to up-regulation and down-regulation service ($p_t^{N\uparrow}$, $p_t^{N\downarrow}$) at up-regulation and down-regulation forecasted prices respectively ($\hat{\lambda}_t^{UR}$, $\hat{\lambda}_t^{DR}$).

$$E_t^{FCR-N} = \hat{\lambda}_t^N \cdot R_t^N + (\hat{\lambda}_t^{UR} \cdot p_t^{N\uparrow} - \hat{\lambda}_t^{DR} \cdot p_t^{N\downarrow}) \cdot \Delta t \quad \forall t \in \mathcal{T} \tag{5.5}$$

In (5.4) and (5.5), the reserves and energy activations account for the combined contribution of both the WF and the BESS.

5.2.3 Scenario Indexation

The previous subsections explained the objective function of the optimization problem considering that the regulation and imbalance prices ($\hat{\lambda}_t^{UR}$, $\hat{\lambda}_t^{DR}$, $\hat{\lambda}_t^{IMB}$) are the final prices of the regulation market. However, these prices depend on the system imbalance at each hour of the operating day.

In reality it is of great complexity to forecast accurately the system state in each hour, so in order to model the behavior described above it has been decided to use an indexation of scenarios for the prices relative to regulation. Thus, under scenario w, at time t the power system is consider to be in state $S_{t,w}^{Need}$. The definition of states, creation of scenarios, and a more precise explanation of this methodology is included in the later chapter 6, subsection 6.5.2. In this section, the final objective function considering these scenarios is presented.

$$OBJ = \max_{\Omega} \sum_{t \in \mathcal{T}} (E_{t,w}^{EM} + E_{t,w}^{FCR-N} + E_t^{FCR-D}) \tag{5.6}$$

$$E_{t,w}^{EM} = (p_t^{EM} \cdot \hat{\lambda}_t^{DAM} + \sum_{w \in \mathcal{W}} \pi_w[(\lambda_{t,w}^{IMB} - \hat{\lambda}_t^{DAM}) \cdot (\Delta P^+ - \Delta P^-)]) \cdot \Delta t \quad \forall t \in \mathcal{T} \tag{5.7}$$

$$E_t^{FCR-D} = \hat{\lambda}_t^{D\uparrow} \cdot R_t^{D\uparrow} + \hat{\lambda}_t^{D\downarrow} \cdot R_t^{D\downarrow} \quad \forall t \in \mathcal{T} \tag{5.8}$$

$$E_t^{FCR-N} = \hat{\lambda}_t^N \cdot R_t^N + \sum_{w \in \mathcal{W}} \pi_w[(\lambda_{t,w}^{UR} \cdot p_t^{N\uparrow} - \lambda_{t,w}^{DR} \cdot p_t^{N\downarrow})] \cdot \Delta t \quad \forall t \in \mathcal{T} \tag{5.9}$$

It should be noted that the prices are no longer considered as forecasts, but derived from them, so the symbol "^" is eliminated in the notation. On the other hand, π_w represents the probability of each scenario. This will be explained in the next chapter.

5.3 Constraints

5.3.1 Energy Market constraints

The total contribution of the APP to the EM is the combined contribution of the WF and the BESS.

$$p_t^{EM} = p_t^{W,EM} + p_t^{B,EM} \quad \forall t \in \mathcal{T} \tag{5.10}$$

$$p_t^{W,EM} = \hat{p}_t^W - (R_t^{W,N} + R_t^{W,D\uparrow}) \quad \forall t \in \mathcal{T} \tag{5.11}$$

$$p_t^{B,EM} = p_t^{DAM,dis} - p_t^{DAM,ch} \quad \forall t \in \mathcal{T} \tag{5.12}$$

In real time, the WF will be operated under curtailment if a bid for up-regulation reserve is placed in the FCR-D auction, leaving enough capacity available to provide the service if activated. On the other hand, if for a certain hour a bid for down-regulation is placed, the WF will produce the realized level of wind power and will be curtailed if needed to provide the service. Regarding FCR-N, since this is a symmetrical product, the WF will also be operated under curtailment if a bid is placed. This behaviour is represented by constraint (5.11). It should be noted that when the bidding strategy algorithm is run, it is not possible to know what the WF realization will be. Therefore, the forecasted wind power ($\hat{p}_t^W$) is used instead at this stage. When the outcome of the algorithm is evaluated after the operating hour and the actual revenue is calculated, the forecasted wind power should be replaced by the actual realization of the WF at that hour.

The BESS can place a bid in the DAM for either charging ($p_t^{DAM,ch}$) or discharging ($p_t^{DAM,dis}$). In the first case, the charging power will be charged at DAM price, while discharging supposes an income for the APP. $p_t^{B,EM}$ can be understood as the baseline power for the BESS.

5.3.2 Imbalance constraints

The imbalance income/cost has been introduced in the section dedicated to the objective function. The imbalance accounts for the power deviation of the WF with respect to the scheduled output levels contracted in the DAM.

$$\Delta P_t = p_t^{W,EM} - p_t^{W,DAM} \quad \forall t \in \mathcal{T} \tag{5.13}$$

$$\Delta P_t = \Delta P_t^+ - \Delta P_t^- \quad \forall t \in \mathcal{T} \tag{5.14}$$

$$\Delta P_t^+, \Delta P_t^- \geq 0 \quad \forall t \in \mathcal{T} \tag{5.15}$$

where ΔP_t is the total imbalance of the WF at hour t. Contraint (5.13) accounts for the difference between the realized power output and the contracted power in the DAM auction, denoted as $p_t^{W,DAM}$.

The total imbalance can be positive or negative, so two non-negative variables are included, ΔP_t^+ and ΔP_t^-. These variables are the ones used to account for the final imbalance income/cost of the APP. When the total imbalance is negative, ΔP_t^- will take the absolute value of the imbalance and ΔP_t^+ will be 0. The same, but reversed, applies if the imbalance is positive.

5.3.3 Bids limitation constraints

The DAM bid of the BESS has to be either for charging and discharging and it can not exceed the maximum charging or discharging rate. This is ensured by (5.16) and (5.17), where $\overline{P}^B$ is the maximum discharge and charge power of the BESS and b_t^{state} is introduced as an auxiliary binary variable. When $b_t^{state} = 1$, the BESS will be scheduled to charge in the day-ahead auction. Likewise, $b_t^{state} = 0$ will indicate discharging.

$$0 \leq p_t^{DAM,dis} \leq \overline{P}^B \cdot (1 - b_t^{state}) \quad \forall t \in \mathcal{T} \tag{5.16}$$

$$0 \leq p_t^{DAM,ch} \leq \overline{P}^B \cdot b_t^{state} \quad \forall t \in \mathcal{T} \tag{5.17}$$

The WF bid to the DAM is bounded between zero and the day-ahead power forecast, ensured by (5.18).

$$0 \leq p_t^{W,DAM} \leq \hat{p}_t^W \quad \forall t \in \mathcal{T} \tag{5.18}$$

On the other hand, FCR participation requires all the reserve bids to be higher than 0.1 MW, so (5.19) to (5.24) are introduced in this regard.

$$0.1 \cdot \delta_t^{W,D\uparrow} \leq R_t^{W,D\uparrow} \leq \delta_t^{W,D\uparrow} \cdot \overline{P}^W \quad \forall t \in \mathcal{T} \tag{5.19}$$

$$0.1 \cdot \delta_t^{W,D\downarrow} \leq R_t^{W,D\downarrow} \leq \delta_t^{W,D\downarrow} \cdot \overline{P}^W \quad \forall t \in \mathcal{T} \tag{5.20}$$

$$0.1 \cdot \delta_t^{W,N} \leq R_t^{W,N} \leq \delta_t^{W,N} \cdot \overline{P}^W \quad \forall t \in \mathcal{T} \tag{5.21}$$

$$0.1 \cdot \delta_t^{B,D\uparrow} \leq R_t^{B,D\uparrow} \leq \delta_t^{B,D\uparrow} \cdot 2 \cdot \overline{P}^B \quad \forall t \in \mathcal{T} \tag{5.22}$$

$$0.1 \cdot \delta_t^{B,D\downarrow} \leq R_t^{B,D\downarrow} \leq \delta_t^{B,D\downarrow} \cdot 2 \cdot \overline{P}^B \quad \forall t \in \mathcal{T} \tag{5.23}$$

$$0.1 \cdot \delta_t^{B,N} \leq R_t^{B,N} \leq \delta_t^{B,N} \cdot \overline{P}^B \quad \forall t \in \mathcal{T} \tag{5.24}$$

The binary variables denoted as δ are used as auxiliary variables to ensure that at each hour, the reserve bids are either null or larger than 0.1 MW for both the WF and the BESS and for every FCR service. The upper limit for the reserve bids of the WF is considered to be the rated power of the WF, although the constraints introduced below limit this behaviour.

It should also be noted that the upper limit for the BESS FCR-D reserve bids is twice the maximum power of the BESS. This is because reserves are activated on top of the baseline power of the BESS ($p_t^{B,EM}$), so there are situations where the battery can go from a baseline of full discharge to a reserve activation of full charging power, or vice versa. However, the limits of the APP are to be fulfilled, so the following equations are introduced.

$$p_t^{B,EM} + R_t^{B,D\uparrow} + R_t^{B,N} \leq \overline{P}^B \quad \forall t \in \mathcal{T} \tag{5.25}$$

$$-p_t^{B,EM} + R_t^{B,D\downarrow} + R_t^{B,N} \leq \overline{P}^B \quad \forall t \in \mathcal{T} \tag{5.26}$$

$$p_t^{W,DAM} + R_t^{W,D\uparrow} + R_t^{W,N} \leq \hat{p}_t^W \quad \forall t \in \mathcal{T} \tag{5.27}$$

$$R_t^{W,D\downarrow} + R_t^{W,N} \leq p_t^{W,DAM} \quad \forall t \in \mathcal{T} \tag{5.28}$$

Constraints (5.25) and (5.26) ensure the total power bids of the BESS to be within limits. In the case of the WF, the upper limit for power bids is the forecasted power, ensured by

constraint (5.27), while (5.28) prevents the down regulation reserve bids to be lower than the day-ahead bid. Theoretically, the WF could offer to reduce all its capacity, but since the total power it will be able to decrease is the power at which it is operating in real time when the reserve is activated, it is limited to the amount offered in the DAM.

Since FCR-N reserve is a symmetrical product, its equivalent variable is included in both up and down limitation constraints.

5.3.4 Battery constraints

For the BESS is not enough for the bids to be within its maximum discharge and charge capacity, but it is also needed to ensure the SOC to be within limits at every hour of the operating day. Therefore, it is necessary to account for the total charging and discharging power of the BESS at every time step.

$$p_t^{dis} = p_t^{DAM,dis} + p_t^{B,D\uparrow} + p_t^{B,N\uparrow} \quad \forall t \in \mathcal{T} \tag{5.29}$$

$$p_t^{ch} = p_t^{DAM,ch} + p_t^{B,D\downarrow} + p_t^{B,N\downarrow} \quad \forall t \in \mathcal{T} \tag{5.30}$$

$$p_t^{dis} \leq \overline{P}^B \quad \forall t \in \mathcal{T} \tag{5.31}$$

$$p_t^{ch} \leq \overline{P}^B \quad \forall t \in \mathcal{T} \tag{5.32}$$

The total discharge (p_t^{dis}) and charge (p_t^{ch}) power are defined in 5.29 and 5.30 as the sum of the scheduled DAM power and the BESS activated power for up regulation and down regulation in FCR-N ($p_t^{B,N\downarrow}$, $p_t^{B,N\uparrow}$) and FCR-D ($p_t^{B,D\downarrow}$, $p_t^{B,D\uparrow}$). As it has been done with the power bids, the total charging and discharging power is limited to the maximum charging and discharging power capacity of the BESS with constraints (5.31) and (5.32).

The SOC of the BESS at hour t is defined as the previous SOC and percentage energy change due to total charging and discharging. This is modeled by constraint (5.33) and (5.34), where charging and discharging efficiencies are also introduced and E_b represents the nominal energy of the BESS in MWh. In addition, as it has been already mentioned above, the BESS has to be at 50% (SOC_{init}) of charge at the end of each day to have the same capacity of providing up and down regulation in the first hour of the next day. Constraint (5.35) is introduced in this regard.

$$SOC_t = SOC_{init} + \frac{\Delta t}{E_b} \cdot (p_t^{ch} \cdot \eta_{ch} - p_t^{dis}/\eta_{dis}) \quad t = 0 \tag{5.33}$$

$$SOC_t = SOC_{t-1} + \frac{\Delta t}{E_b} \cdot (p_t^{ch} \cdot \eta_{ch} - p_t^{dis}/\eta_{dis}) \quad \forall t \in \{1, 2, ..., |\mathcal{T}|\} \tag{5.34}$$

$$SOC_t = SOC_{init} \quad t = |\mathcal{T}| \tag{5.35}$$

Finally, the characteristics of the BESS are such that no full FCR-response can be sustained for four concurrent hours, so it is considered a unit with Limited Energy Reservoir (LER). As far as this algorithm is concerned, this means that the SOC limit at each instant of time can be limited to the minimum and maximum, without considering that a full activation of a reserve would lead to exceeding the limits. In case of not being LER, it would be necessary to take into account the reserve offers and consider that they could be fully activated when limiting the SOC. Therefore, the SOC limits follow the constraint (5.36).

$$SOC_{min} \leq SOC_t \leq SOC_{max} \quad \forall t \in \mathcal{T} \tag{5.36}$$

5.3.5 Reserve constraints

In the section dedicated to the objective function, the reserve and energy activation variables account for the overall APP values. Those variables are the sum of the WF and BESS contributions.

$$R_t^{D\uparrow} = R_t^{W,D\uparrow} + R_t^{B,D\uparrow} \quad \forall t \in \mathcal{T} \tag{5.37}$$

$$R_t^{D\downarrow} = R_t^{W,D\downarrow} + R_t^{B,D\downarrow} \quad \forall t \in \mathcal{T} \tag{5.38}$$

$$R_t^{N} = R_t^{W,N} + R_t^{B,N} \quad \forall t \in \mathcal{T} \tag{5.39}$$

$$p_t^{N\downarrow} = p_t^{W,N\downarrow} + p_t^{B,N\downarrow} \quad \forall t \in \mathcal{T} \tag{5.40}$$

$$p_t^{N\uparrow} = p_t^{W,N\uparrow} + p_t^{B,N\uparrow} \quad \forall t \in \mathcal{T} \tag{5.41}$$

Finally, the relation between the energy activated for FCR-D or FCR-N and the scheduled reserves is determined by the activation variables. This activation variables are a function of the grid frequency, as explained in the literature review, and determine how much of the contracted reserve needs to be delivered. Since in practice is not possible to forecast the frequency of the system, the activation variables are replaced by activation parameters, which will be given as an input to the model. This will be explained further in subsequent sections of this thesis.

The previous behaviour is modeled by introducing the following constraints, where $A^{D\uparrow}$, $A^{D\downarrow}$, $A^{N\uparrow}$ and $A^{N\downarrow}$ are the activation parameters for FCR-D up-regulation, FCR-D down-regulation, FCR-N up-regulation and FCR-N down-regulation respectively.

$$p_t^{W,D\uparrow} = A^{D\uparrow} \cdot R_t^{W,D\uparrow} \quad \forall t \in \mathcal{T} \quad (5.42) \qquad p_t^{B,D\uparrow} = A^{D\uparrow} \cdot R_t^{B,D\uparrow} \quad \forall t \in \mathcal{T} \quad (5.46)$$

$$p_t^{W,D\downarrow} = A^{D\downarrow} \cdot R_t^{W,D\downarrow} \quad \forall t \in \mathcal{T} \quad (5.43) \qquad p_t^{B,D\downarrow} = A^{D\downarrow} \cdot R_t^{B,D\downarrow} \quad \forall t \in \mathcal{T} \quad (5.47)$$

$$p_t^{W,N\uparrow} = A^{N\uparrow} \cdot R_t^{W,N} \quad \forall t \in \mathcal{T} \quad (5.44) \qquad p_t^{B,N\uparrow} = A^{N\uparrow} \cdot R_t^{B,N} \quad \forall t \in \mathcal{T} \quad (5.48)$$

$$p_t^{W,N\downarrow} = A^{N\downarrow} \cdot R_t^{W,N} \quad \forall t \in \mathcal{T} \quad (5.45) \qquad p_t^{B,N\downarrow} = A^{N\downarrow} \cdot R_t^{B,N} \quad \forall t \in \mathcal{T} \quad (5.49)$$

5.4 Decision Variables

This section summarizes all the variables of the model presented in the previous sections. The complete set of variables is denoted as Ω, where Ω^{R+} is the subset of positive real variables, Ω^R is the subset of real variables and Ω^b is the subset of binary variables:

$$\begin{aligned} \Omega = \{ &p_t^{EM}, p_t^{W,EM}, p_t^{B,EM}, p_t^{W,DAM}, p_t^{DAM,dis}, p_t^{DAM,ch}, \Delta P_t, \Delta P_t^+, \Delta P_t^-, b_t^{state} \\ &\delta_t^{W,D\uparrow}, \delta_t^{W,D\downarrow}, \delta_t^{W,N}, \delta_t^{B,D\uparrow} \delta_t^{B,D\downarrow}, \delta_t^{B,N}, p_t^{dis}, p_t^{ch}, SOC_t \\ &R_t^{W,D\uparrow}, R_t^{W,D\downarrow}, R_t^{W,N}, R_t^{B,D\uparrow}, R_t^{B,D\downarrow}, R_t^{B,N} \\ &p_t^{W,D\uparrow}, p_t^{W,D\downarrow}, p_t^{W,N\uparrow}, p_t^{W,N\downarrow}, p_t^{B,D\uparrow}, p_t^{B,D\downarrow}, p_t^{B,N\uparrow}, p_t^{B,N\downarrow}, S_{t,w}^{Need} \} \end{aligned} \tag{5.50}$$

$$\begin{aligned} \Omega^{R+} = \{ &p_t^{W,EM}, p_t^{W,DAM}, p_t^{DAM,dis}, p_t^{DAM,ch}, \Delta P_t^+, \Delta P_t^-, p_t^{dis}, p_t^{ch}, SOC_t, \\ &R_t^{W,D\uparrow}, R_t^{W,D\downarrow}, R_t^{W,N}, R_t^{B,D\uparrow}, R_t^{B,D\downarrow}, R_t^{B,N}, \\ &p_t^{W,D\uparrow}, p_t^{W,D\downarrow}, p_t^{W,N\uparrow}, p_t^{W,N\downarrow}, p_t^{B,D\uparrow}, p_t^{B,D\downarrow}, p_t^{B,N\uparrow}, p_t^{B,N\downarrow} \} \end{aligned} \tag{5.51}$$

$$\Omega^R = \{ p_t^{EM}, p_t^{B,EM}, \Delta P_t \} \tag{5.52}$$

$$\Omega^b = \{ b_t^{state}, \delta_t^{W,D\uparrow}, \delta_t^{W,D\downarrow}, \delta_t^{W,N}, \delta_t^{B,D\uparrow}, \delta_t^{B,D\downarrow}, \delta_t^{B,N} \} \tag{5.53}$$

6 Model Inputs

In this chapter, the inputs to the bidding strategy algorithm developed in Chapter 5 are described. These inputs include the wind power forecasts, forecasts for day-ahead, regulation and reserve market prices, the set of power system need scenarios (power deficit, power excess or balanced) and the frequency parameters for energy activation in FCR-D and FCR-N markets. A graphical representation of the proposed system is shown in Figure 6.1.
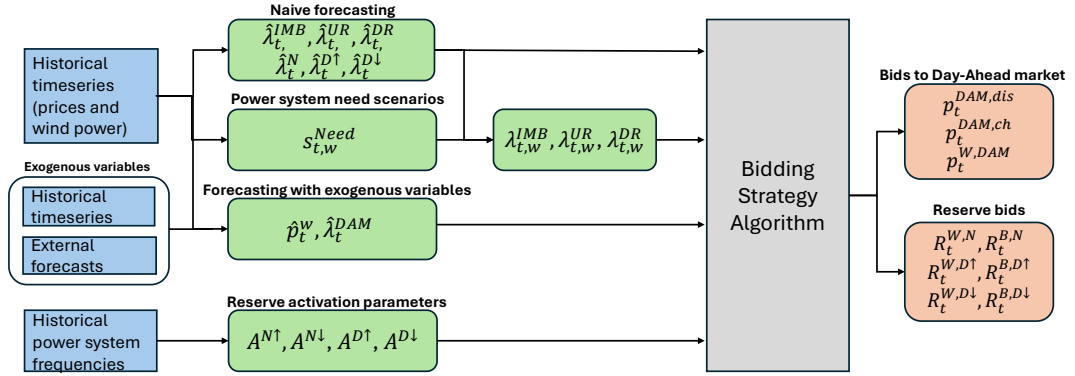

Figure 6.1: Flow of inputs to the bidding strategy algorithm.

First, although already introduced in previous chapters, an overview of the forecasting horizons with respect to the APP operating day is provided. Then, data collection, processing and implementation of the forecasting techniques are explained. Particular emphasis is given to the forecasting of energy and reserve prices, since wind power forecasting has been explained in a dedicated chapter, Chapter 4.

6.1 Time Framework

The bidding strategy model developed in this thesis considers the APP to participate in DAM, FCR-D and FCR-N, so this section elaborates on the time frameworks of these markets and the available data for the operator to perform the forecasting of prices.

Referring to the operating day as D and the decision day, when the forecasts are performed and the bids are placed, as $D-1$, an operator has until 12:00 CET of $D-1$ to submit the final DAM bids to NordPool. Once the market is cleared, NordPool typically announces the hourly clearing prices at 12:45 CET. Therefore, when submitting the offer bids to the DAM, the operator has knowledge of the actual clearing prices of decision day, $D-1$, due to the outcome of $D-2$ auction.

On the other hand, FCR-D and FCR-N markets are divided into an early auction and a late auction. In this thesis, the APP is considered to bid only in the early auctions. To participate in theearly auctions, operators are required to submit their bids to Energinet no later than 00:30 CET of $D-1$. As it is the case with DAM prices, the operator has knowledge of the reserve prices of the current day $(D-1)$ from the clearing outputs of $D-2$.

The case is however different for up-regulation, down-regulation and imbalance market prices since the imbalance market is cleared 45 minutes prior to the hour of operation.

Therefore, the data closest in time available to the operator at the beginning of $D-1$ are those corresponding to day $D-2$.

For simplicity, since FCR-D and FCR-N reserve markets are the first of those considered to close, all the bidding decisions are considered to take place prior to 00:30 a.m of $D-1$. Therefore, even though the DAM closes later in time, the optimal bid offers obtained from the optimisation algorithm for DAM are considered definitive and not changed closer to DAM closing gate. On the other hand, no participation in late auctions are considered for FCR-N and FCR-D. As a result of this decision, the data used for the forecasting of prices explained below is that available at 00:30 a.m of $D-1$.

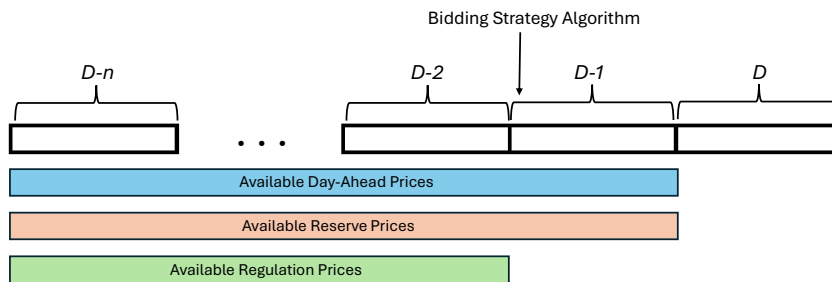A graphical overview of the explained timeline is included in Figure 6.2

Figure 6.2: Prices known by the operator at the time of bidding.

6.2 Wind Power Forecasts

The hourly wind power forecast algorithm developed in Chapter 4 is executed prior to the bidding strategy model and the wind power point forecasts are recorded for every hour of operation day D. The forecasted wind power for hour t of day D is denoted as $\hat{p}_t^W$ in the bidding strategy algorithm.

The forecasting method employed is LWR, as it was the one showing the best results in Chapter 4. As it was explained, the exogenous data for the forecasting is obtained from external forecasts developed by DMI.

6.3 Day-Ahead Market Price

6.3.1 Overview of historical prices

Historical DK2 day-ahead prices are obtained from Energinet via an API to the Energi Data Service platform ([33]) and the time-series for the current year (2024) is presented in Figure 6.3. Since data is available in Danish time, the last Sunday in March has 23 hours due to the switch from CET (Central European Time) to CEST (Central European Summer Time). This means that in the linear timeline used for the data, the time step 31st March 2024, 2:00 am is missing. For the presented time-series it has been decided to assume the DAM price for the missing hour to be the one of the previous hour. A similar issue will happen for the last Sunday of October, when time is switched from CEST to CET at 3:00 am, so in the linear timeline used for the data this day will have 25 hours.

Figures 6.3 and 6.4 show how the prices of DAM in 2024 have oscillated mainly in the range of 0 to 200€/MWh, with an average of just under 100€/MWh. Some exceptions can be observed, such as some price peaks in January or negative prices in April and June, with a maximum of 526€/MWh and a minimum of -54€/MWh. The price peaks in January were due to high electricity demand due to the cold weather, while the negative prices are generally caused by low demand and high renewable energy market penetration.
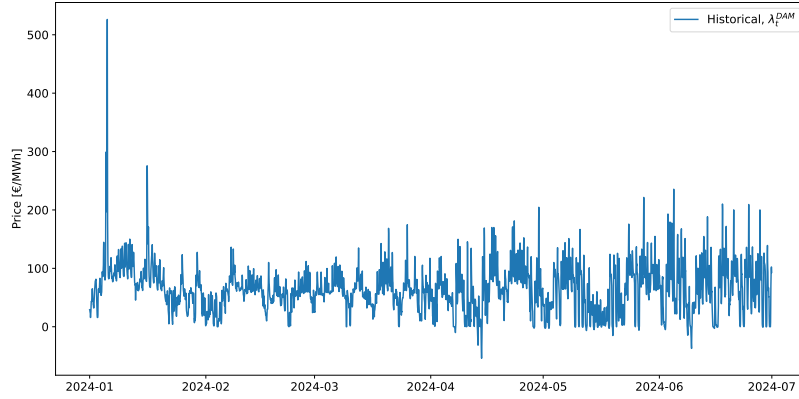
Figure 6.3: Historical DK2 DAM prices from Jan 2024 to July 2024. Data from: [33].
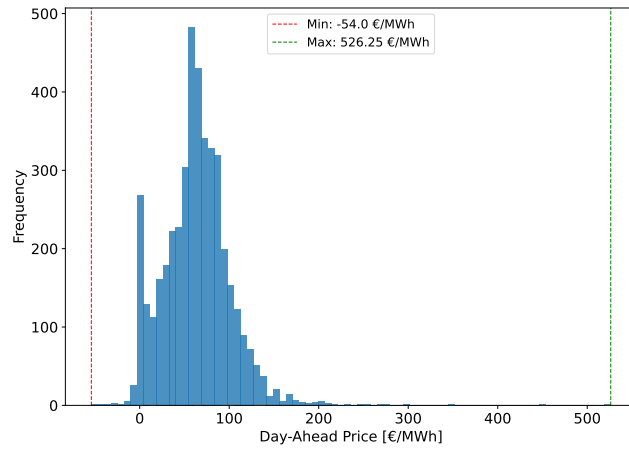

Figure 6.4: Histogram of DK2 DAM prices from Jan 2024 to July 2024. Data from: [33].

6.3.2 Forecasting

In order to make an optimal offer, it is necessary to make an accurate prediction of DAM prices. A good price forecast can lead operators to take better financial decisions and maximize payoffs. While this topic could be the subject of a thesis in itself, this section proposes and compares some methods to obtain a price prediction that will serve as an input to the optimisation algorithm. As will be explained below, price prediction with more sophisticated models will be exemplified with DAM forecasting. However, the prediction of regulation prices and reserve prices will be simplified by the use of naive predictors.

DAM prices are characterised by high seasonality, which can be explained by many factors such as demand, temperature or renewable generation. Although these explanatory variables will be studied in more detail in the last forecasting model used in this thesis (SARIMAX), the high seasonality may indicate that the development of models based solely on the seasonality of the data may be effective in predicting prices. The two models chosen for this purpose in this thesis are Prophet and SARIMA, although as already reflected in the literature review section, there are numerous studies using other methods. The choice of Prophet and SARIMA in this thesis is mainly due to the ease of implementation of the former and the large existing literature on the latter.

Before explaining the prediction models used, figures 6.5 and 6.6 are included to illustrate the seasonality of the data. Figure 6.5 shows how the price profile over the hours of the day is similar for all the months studied. The seasonality of prices is calculated using the

statsmodel library over the whole dataset. However, for presentation purposes, only 5 day are represented in Figure 6.6, where black dotted lines are placed at the beginning of each day to appreciate how the data has daily seasonality.
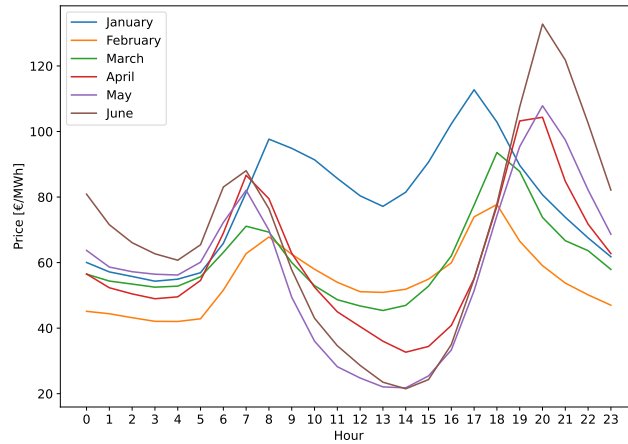

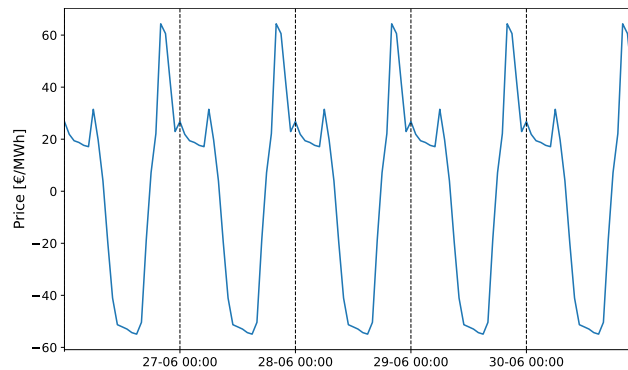Figure 6.5: Hourly averaged DAM prices per month. Data from: [33].

Figure 6.6: Day-Ahead market prices seasonality

Prophet Model

Prophet is a time series forecasting model developed by Facebook Data Science Team [34]. The equation on which the model is based is Equation 6.1, created by decomposing the time series into three main components: trend, seasonality and holidays. The trend component $(g(t))$ and seasonality component $(s(t))$ model changes in the time series that are non-periodic and periodic respectively, while the holidays term $(h(t))$ represents the effect of holidays that occur on a irregular basis. Finally, ϵ_t represents changes in the time series not captured by the previous terms [35].

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \tag{6.1}$$

In this thesis, the forecasting is performed following the Prophet python procedure guide published by Meta [36]. The necessary input to the model is a dataframe with two columns, 'ds' being a column dedicated to the timestamps and 'y' the variable to forecast. In this case, data in column 'y' are the historical DK2 day-ahead prices from [33].

To evaluate the performance of this forecasting method, Prophet models have been tested on six randomly selected days. Keeping these days fixed, price predictions have been made

for each of them by varying the amount of data entered for fitting the models and the RMSE, MAE, MAPE and NRMSE metrics have been calculated for each time horizon, as listed in Table 6.1. These metrics are shown as the resulting average over the six test days, and it should be noted that since the data used spans for the first 6 months of 2024, none of the days tested belong to January or February in order to fit the model with up to two months of data.

Fitting days	1	3	5	10	20	40	60
Average RMSE [€/MWh]	42.22	26.00	18.73	32.67	26.70	20.89	18.18
Average MAE [€/MWh]	32.29	20.35	14.97	26.38	22.98	17.81	14.49
Average MAPE [%]	69.07	29.91	21.13	35.78	36.42	29.22	25.55
Average NRMSE [p.u]	0.31	0.19	0.14	0.24	0.19	0.15	0.13

Table 6.1: Prophet forecasting results over 6 testing days

From the results it is difficult to draw a conclusion on which horizon would be better to use for the Prophet model, as the total mean of the metrics is very dependent on not having large forecasting errors for one of the days studied. This is reflected in the results obtained when using five days as training horizon, obtaining better metrics than with the rest of the horizons studied. In fact, when plotting the prediction results for each of the days studied, three of them show similar results regardless of the training horizon, while the other two days show large fluctuations. In general, the model shows more accuracy the longer the horizon, but in the case of five fitting days there is an exception, and that is that the prediction for the two days mentioned is better than in the rest of the cases, improving the global metric.

Figure 6.7 shows the forecast results for the six days tested and using a five-day training horizon for demonstration purposes, since it is the horizon with better results. In general, the model captures the trend but peak prices are not predicted correctly, specially for the first three days studied.
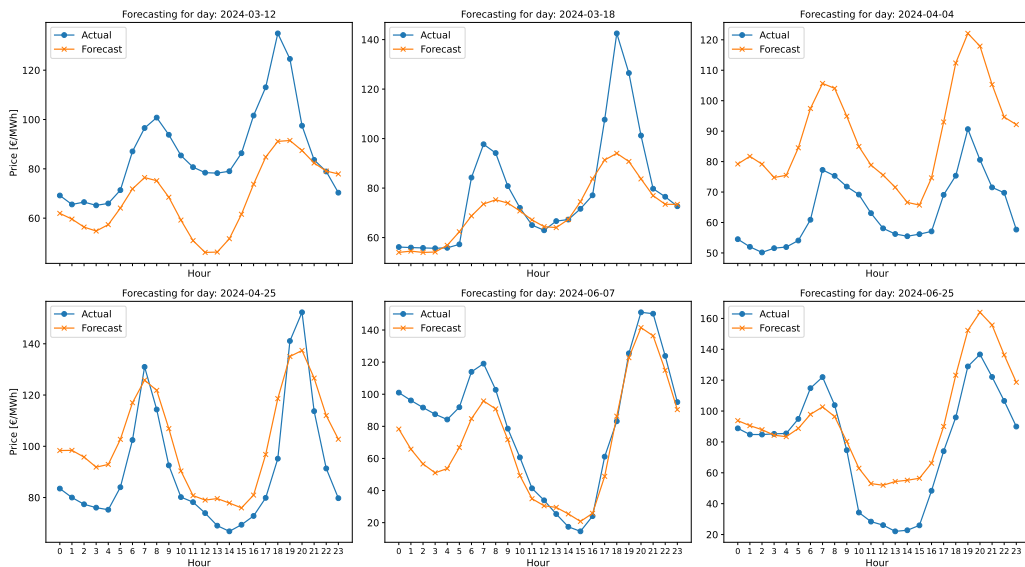

Figure 6.7: Prophet forecasting results over 6 testing days

SARIMA Model

Due to the extensive literature available and the feasibility of these models for forecasting electricity prices over short horizons, the next times-series forecasting method employed is SARIMA.

A SARIMA model is an extension of an ARIMA model that includes seasonal effects, which in turn is a combination of autoregressive (AR) and moving average (MA) models [37]. An AR model is one in which the value of the series is predicted by modelling its relationship with its past p values, while a MA model of order q models the relationship of the series with the q past prediction errors [22]. Combining an AR model with a MA model would yield an ARMA (Autoregressive Moving Average) model. However, an integrating process is added to convert non-stationary time series into stationary series. This "integrating" process actually consists of subtracting previous values from the current value of the series. The number of differences necessary to convert the series into a stationary series determines the order d of the "integrating" process [38]. Combining all the above together, an ARIMA (Autoregressive Integrated Moving Average) model is built. In addition, if the time series presents seasonality as in the case of day-ahead market prices, a SARIMA (Seasonal Autoregressive Integrated Moving Average) model is built, including additional terms related to this seasonality [39]. The general notation for ARIMA and SARIMA models is:

$$ARIMA(p,d,q) \quad , \quad SARIMA(p,d,q)(P,D,Q,s) \tag{6.2}$$

Where:

- p is the non-seasonal AR order.

- d is the non seasonal integrating/differencing order.

- q is the non-seasonal MA order.

- P is the seasonal AR order.

- D is the seasonal integrating/differencing order.

- Q is the seasonal MA order.

- s is the seasonal period.

The seasonal parameters follow the same reasoning as the non-seasonal parameters explained above, but the lags follow seasonal intervals. The mathematical representation of a SARIMA model follows (6.3) [39], [40].

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^S)^D y_t = \mu + \theta_q(B)\Theta_Q(B^S)\varepsilon_t \tag{6.3}$$

In the previous equation, y_t is the observed time-series at time t, B is the backshift operator $(B^l \cdot y_t = y_{t-l})$, ε_t represents noise with zero mean and μ is a constant term. The polinomial operators are defined as (6.4), where ϕ, Φ, θ and Θ are the AR, seasonal AR, MA and seasonal MA coefficients respectively.

$$\phi_p(B) = 1 - \phi_1(B) - \phi_2(B^2)... - \phi_p(B^p) \tag{6.4a}$$

$$\Phi_P(B^s) = 1 - \Phi_1(B^s) - \Phi_2(B^{2s})... - \Phi_P(B^{Ps}) \tag{6.4b}$$

$$\theta_q(B) = 1 - \theta_1(B) - \theta_2(B^2)... - \theta_q(B^q) \tag{6.4c}$$

$$\Theta_Q(B^s) = 1 - \Theta_1(Q^s) - \Theta_2(B^{2s})... - \Theta_q(B^{Qs}) \tag{6.4d}$$

Using the same data as in the previous models, the first SARIMA model for day-ahead market price prediction is performed by using the *auto_arima* functionality available in the python *pmdarima* library [41]. By default, the function optimizes by minimizing AIC (Akaike´s Information Criterion) via the Kwiatkowski-Phillips-Schmidt-Shin test [42]. The result of this function are the SARIMA parameters p, d, q, P, D and Q. However, the seasonality of the data s, defined as the number of periods in each season, has to be entered manually. From Figure 6.6 it is worth to set $s = 24$.

After executing the *auto_arima* function, the optimal parameters obtained for the model are $p = 2$, $d = 1$, $q = 1$, $P = 1$, $D = 0$, $Q = 0$. In order to evaluate the performance of the model, it is tested for the same six days as in the case of the Prophet model, and the average performance metrics are shown in Table 6.2. Results show that SARIMA model performs worst than Prophet following this methodology, so it is decided to tune the parameters manually following the information provided by the ACF and PACF in Figure 6.8.
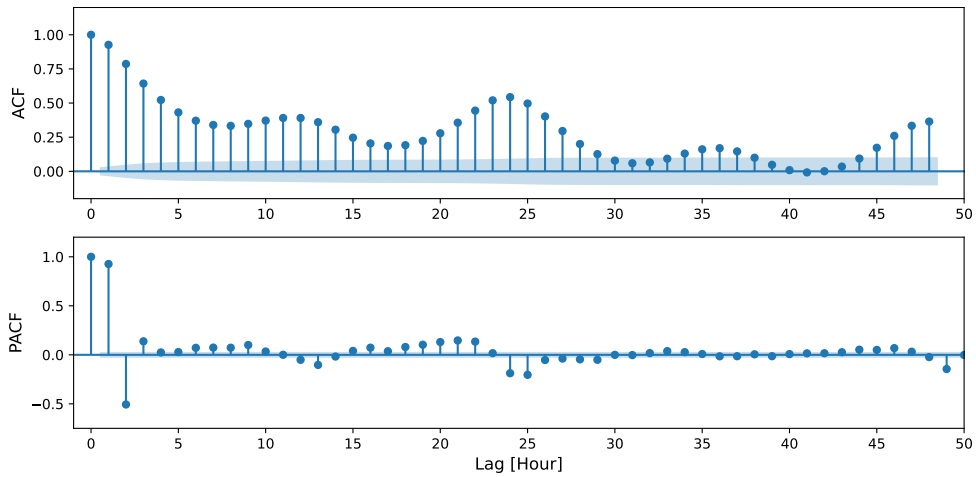

Figure 6.8: ACF and PACF of DAM prices. Data from: [33].

Previously introduced, both ACF and PACF are functions generally used to identify orders for Moving Average (MA) and Autoregressive (AR) models [43]. The blue horizontal cone in the figures represents the 95% confidence level, so lags with an ACF inside the cone are considered to not have significant autocorrelation.

The previous figure shows ACF and PACF up to 48 lags (hours). The PACF values give an idea of the p order of the autoregressive models, while the ACF plot can help determine the q order of the moving average model [22]. Therefore, reasonable parameters to test would be $p = 3$ and $q = 4$, since the largest PACF values within the 24-lag season are given for the first three lags, while, if only ACF values greater than 0.5 are considered for simplicity of the model, the largest are given for the first four lags. On the other hand, as for the seasonality parameters, it would be reasonable to consider $P = 1$ or $P = 2$, since the PACF values for first and second multiples of the seasonality (24 and 48 hours) are higher than the rest. Finally, considering again values of ACF greater than 0.5, it is reasonable to consider $Q = 1$. Starting from these values, different iteration tests are performed on the parameters, obtaining the best result for $p = 5$, $d = 1$, $q = 1$, $P = 2$, $D = 0$, $Q = 1$ in terms of the average performance metrics for the six days studied.

The results for the best performing models using automatic and manual tuning of the parameters are shown in Table 6.2, while the fitting of the manual tuning, which yields

better results, are illustrated in Figure 6.9

SARIMA tuning	Auto	Manual
(p, d, q)(P, D, Q, s)	(2, 1, 1)(1, 0, 0, 24)	(5, 1, 1)(2, 0, 1, 24)
Average RMSE [€/MWh]	21.68	14.72
Average MAE [€/MWh]	17.99	11.04
Average MAPE [%]	29.42	17.95
Average NRMSE [p.u]	0.16	0.11

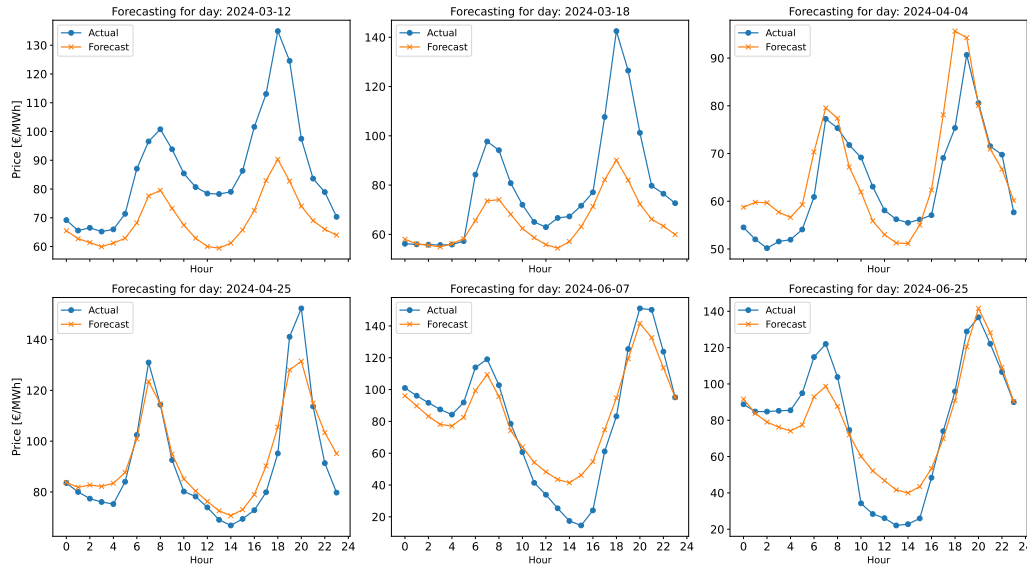Table 6.2: SARIMA forecasting. Parameters and performance metrics

Figure 6.9: SARIMA forecasting results over 6 testing days

SARIMAX Model

The two prediction models studied before have focused on the seasonality of the data itself for the predictions. This is why, in this last model, the possibility of incorporating external explanatory variables in order to make a correct prediction of prices is studied. In this regard, the model studied is SARIMAX, which can be defined as a SARIMA model but including exogenous explanatory variables. Keeping the same notation as for SARIMA, the mathematical formulation of SARIMAX follows (6.5) [40].

$$\phi_p(B)\Phi_P(B^s)(1 - B)^d(1 - B^S)^D y_t = \mu + \theta_q(B)\Theta_Q(B^S)\varepsilon_t + \mathcal{X}_t\beta \tag{6.5}$$

with

$$\mathcal{X}_t\beta = \beta_1 x_{1,t} + \beta_2 x_{2,t}... + \beta_k x_{k,t} \tag{6.6}$$

In (6.6), $\mathcal{X}_t = [x_{1,t}, x_{2,t}, ..., x_{k,t}]$ is the set of exogenous variables and $\beta = [\beta_1, \beta_2, ..., \beta_k]$ the set of exogenous coefficients. The rest of the terms from (6.5) are already defined in (6.4).

With the mathematical formulation presented, the objective is to find a set of explanatory or exogenous variables for improving the forecasting of prices.

The factors to be considered in this thesis as explanatory variables are the consumer demand and renewable generation, both wind and solar. In addition to the existing studies

in the literature about the influence that these factors have on DAM prices, they have been chosen because of the ease of obtaining historical data. Historical data is available online in the transparency platform of ENTSO-E (European Network of Transmission System Operators for Electricity) [44] , which collects and publishes historical data of electricity generation, transportation and consumption for the pan-Europen market. With these data, two approaches will be explored. On the one hand, SARIMAX will be tested with historical data of these variables as exogenous variables. On the other hand, instead of historical data, the model will be tested using external forecasts of the factors, which are also provided by the same platform.

The influence that the selected exogenous features have on prices is analyzed in order to determine whether the DAM price timeseries is affected by all or some specific lags of the demand, solar generation or wind generation timeseries. For this purpose it is decided to use the Cross-Correlation Function (CCF) to identify which lags of the exogenous variables might be useful to predict the prices by estimating the correlation between the objective variable y_t and lagged versions of the explanatory/exogenous variables x_{t+k} [45]. Figure 6.10 illustrates the CCFs between the DAM prices and the three studied exogenous variables.
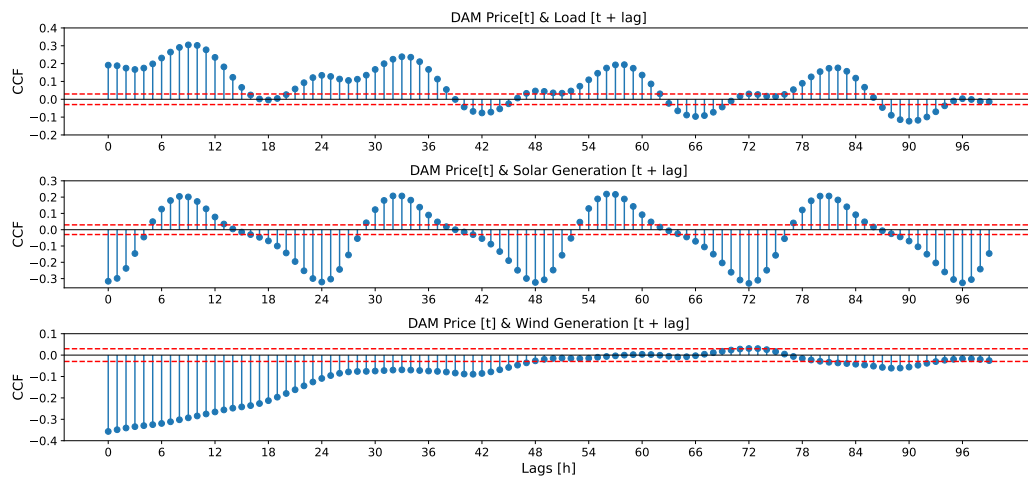

Figure 6.10: Cross-Correlation Functions between historical DAM prices and exogenous variables. From top to bottom: Historical Load [46], Historical Solar Generation [47], Historical Wind Generation [47]

In the graph above, each blue bar represents the CCF factor, while the red lines represent a 95% confidence interval representing whether the CCF is significant (bars outside red range), or not significant (bars inside red range). Several conclusions can be drawn from this picture. On the one hand, it can be seen that there is a certain seasonality in the cross-correlation between prices and load and solar generation, being more pronounced for solar generation, since in the case of demand, the peak of each pattern is declining as the lags increase, while that of solar generation remains stable. However, while demand has a positive CCF for almost all lags, solar generation oscillates between positive and negative. On the other hand, the CCF with wind generation does not show these patterns. While electricity demand, due to working hour patterns, and solar generation, due to weather conditions or the position of the sun, present seasonality themselves in the short term, wind generation has a more arbitrary character. Therefore, prices are impacted by the wind resource in the short term (which do not vary as much), but the influence decreases as the lags increase. In terms of absolute values, the highest correlation with prices is the wind generation during the first 12 lags.

As mentioned above, the first approach is to use historical data on exogenous factors. This idea arises because introducing external forecasts into the predictions can increase the uncertainty and error of the models. However, following the timeline explained above, bidding takes place on D-1 at 00:30am, so unlike prices, which are known for all the hours of D-1 due to the auction on D-2, the latest demand and renewable generation data available for forecasting are those corresponding to D-2. Therefore, despite the above conclusions, the closest lags to real-time operation that can be used for the model are those corresponding to $D - 2$. In this way, an input dataset is constructed for the model in which, for each hour to be predicted, the D-2 lagged values of demand and renewable generation are considered as regressors. However, after training and testing the model in the same way as with the previous models (using six test days and calculating the average of the evaluation metrics used), the results are worse than using the SARIMA model without exogenous variables. A reason could be found in that introducing exogenous variables with not enough correlation to the objective variables can lead to overfitting the model.

In a real application, the bidding for the DAM could be done until 12:00 at the latest, the closing time of this market, so that the real exogenous variables corresponding to the first 12 hours of day D-1 could be used, potentially leading to better forecasts. However, with the simplifications of the bidding strategy algorithm developed in this thesis, that is not possible, so the option of using the external forecasts available in ENTSO-E platform is also considered.
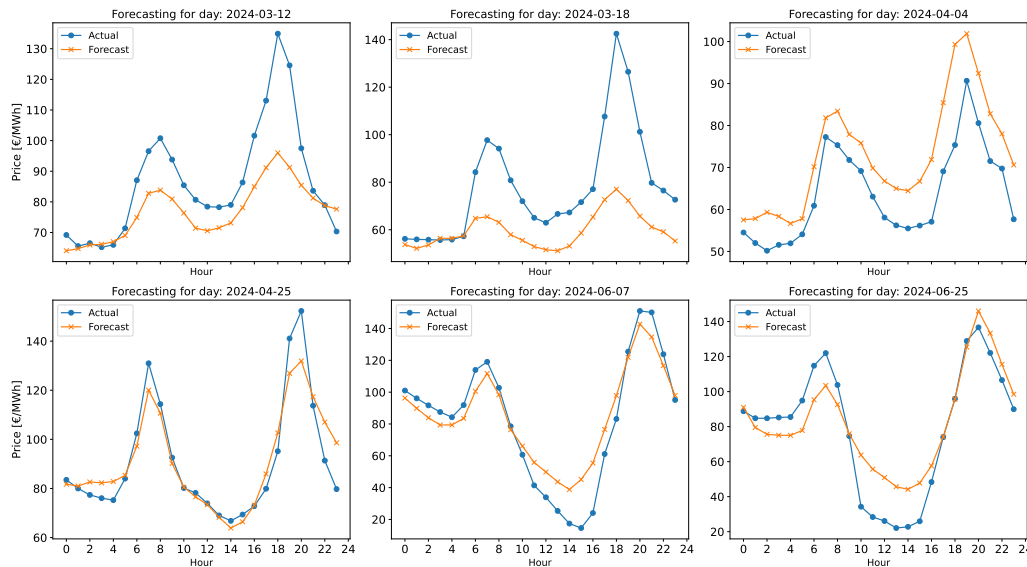

Figure 6.11: SARIMAX forecasting results over 6 testing days using $D - 1$ forecasts as exogenous variables

ENTSO-E provides hourly day-ahead forecasts of solar, wind onshore and wind offshore generation [48] and demand [46], which are published at 18:00pm of D-1. Therefore, the forecasts for day of operation are not known when the bidding takes place, so the forecasts for D-1 will be used for predictions instead. After testing the model for the same six days that have been used so far, results show that the model performs better than Prophet but worst than SARIMA. Forecasts are illustrated in Figure 6.11. Again, the fact that the model performs worse than with respect to SARIMA may be due to overfitting of the model by introducing complexity with exogenous variables that do not show as much correlation with the target variable. In this case, the correlation is lost due to the fact

that external forecast for day D are not available and, as seen in Figure 6.10, lags above 24 have poor CCF between the prices and the exogenous variables. To study the impact of having external forecasts for day D, it is decided to use the model to test the same days but assuming that forecasts are available at the moment of developing the offering strategy, and not published after the placing of the bids as it is the case in reality. The result of this model is depicted in Figure 6.12, where it can be seen that having the day D forecasts significantly improves the model.

The metrics for SARIMAX with historical data, SARIMAX with predictions and SARIMAX with day D forecasts are shown in Table 6.3. Finally, Table 6.4 compares the metrics of the three prediction algorithms explored, concluding that the best model is SARIMAX, followed by SARIMA and, finally, Prophet.
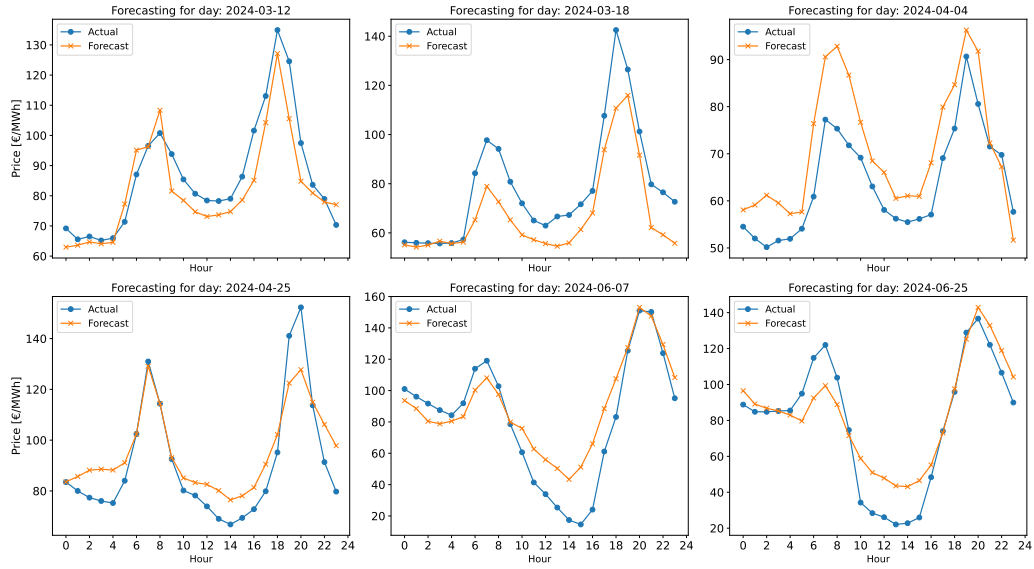

Figure 6.12: SARIMAX forecasting results over 6 testing days using D forecasts as exogenous variables

Exogenous data source	Historical Data	D-1 Forecasts	D Forecasts
Average RMSE [€/MWh]	20.05	15.38	12.79
Average MAE [€/MWh]	14.37	11.39	10.05
Average MAPE [%]	23.85	18.97	18.54
Average NRMSE [p.u]	0.15	0.11	0.09

Table 6.3: SARIMAX (5, 1, 1)(2, 0, 1, 24) forecasting. Performance metrics

Model	Prophet	SARIMA	SARIMAX
Average RMSE [€/MWh]	18.73	14.72	12.79
Average MAE [€/MWh]	14.97	11.04	10.05
Average MAPE [%]	21.13	17.95	18.54
Average NRMSE [p.u]	0.14	0.11	0.09

Table 6.4: Forecasting results. Comparison of models

6.4 Reserve Prices

6.4.1 Overview of historical prices

Historical data from DK2 FCR-N and FCR-D services is also available in the Energi Data Service platform ([49]). The time-series for the current year (2024) is presented in Figure 6.13 in hourly resolution. It should be noted that the prices presented in this sections are the ones related to the early auctions, since the APP is not considered to participate in the late auctions.

Several observations can be drawn from Figure 6.13. First, it can be observed that FCR-D up-regulation prices have been generally lower than down-regulation prices in the first half of 2024, although down-regulation prices show higher volatility with more extreme price peaks. The difference in price volatility can be seen most clearly in Figure 6.14, with a difference between the maximum and minimum value in the order of 60€/MW for FCR-D up-regulation versus the higher volatility of FCR-D down-regulation, with a maximum price of around 800€/MW. On the other hand it can be observed that the reserve prices for FCR-N are generally higher than those of FCR-D with the exception of the peaks of FCR-D down-regulation occurred in the month of April and some other peaks distributed throughout the year.
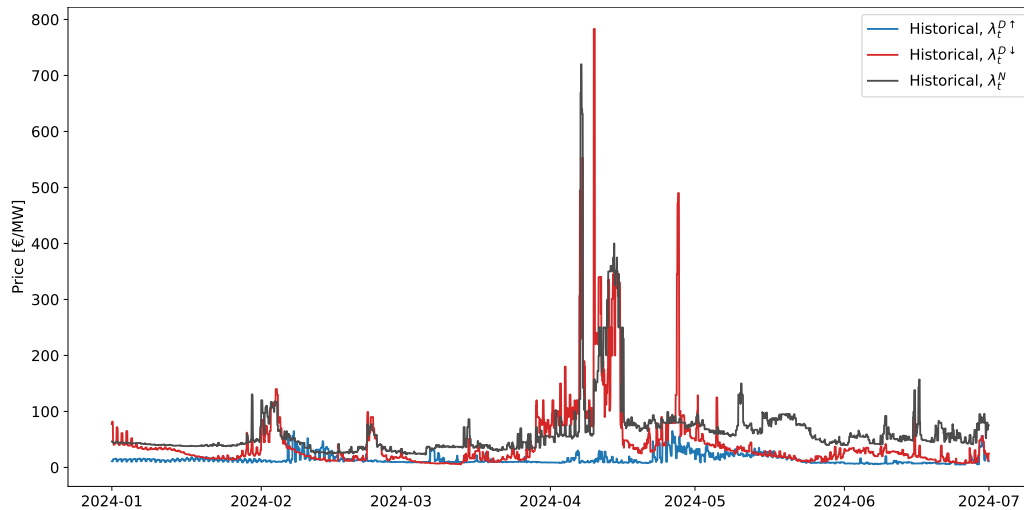

Figure 6.13: Historical DK2 FCR-D and FCR-N early-auction prices from Jan 2024 to July 2024. Data from: [49].
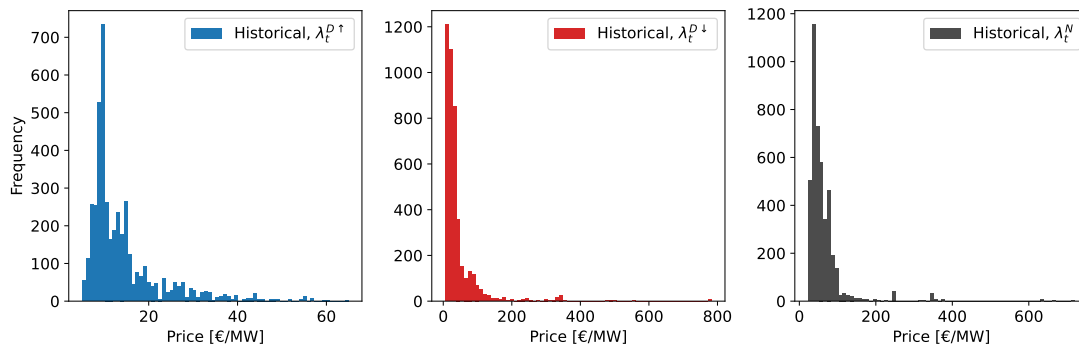

Figure 6.14: Histogram of DK2 FCR-N and FCR-D early-auction prices from Jan 2024 to July 2024. Data from: [49].

6.4.2 Forecasting

Regarding the bidding strategy algorithm, FCR-D and FCR-N reserve prices ($\hat{\lambda}_t^{UR}$, $\hat{\lambda}_t^{DR}$) should also be predicted for each of the operating hours prior to executing the bidding strategy algorithm. However, since a more complex forecasting model has already been exemplified in the case of DAM prices, it has been chosen to simplify the procedure and develop a naive method for the prediction of reserve prices.

The chosen approach is to use a Simple Moving Average method (SMA) defined as (6.7), where the price at each hour is predicted to be the hourly average of the previous N_{pred} days to day of operation D. The same method is applied to FCR-D up, FCR-D down and FCR-N.

$$\hat{\lambda}_t^{RM} = \frac{1}{N_{pred}} \sum_{j=1}^{N_{pred}} \lambda_{t,D-j}^{RM} \quad RM \in \{D\uparrow, D\downarrow, N\} \quad \forall t \in \mathcal{T} \tag{6.7}$$

Using $N_{pred} = 5$, Figures 6.15, 6.16a and 6.16b illustrate the results of forecasting reserve prices for test day 07-06-2024, showcasing the need for more complex prediction models, as using a simplistic predictor can yield large errors, as it is the case for FCR-N prices for the test day represented
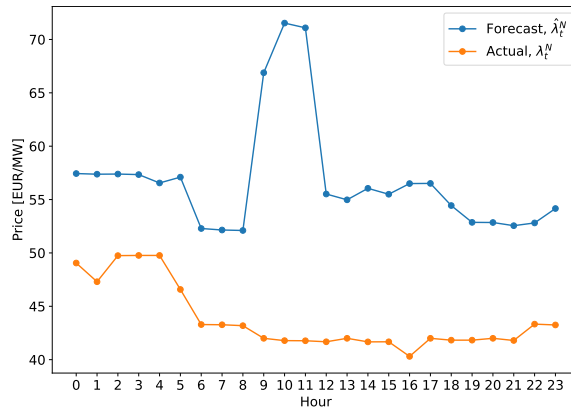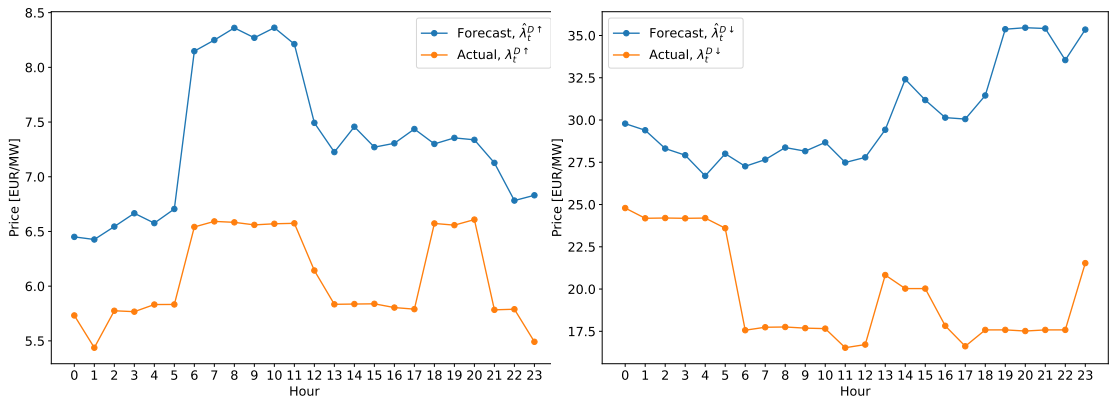

Figure 6.15: $FCR - N$ price forecast. 07-06-2024

(a) $FCR - D \uparrow$ price forecast. 07-06-2024

(b) $FCR - D \downarrow$ price forecast. 07-06-2024

Figure 6.16: Forecasting of FCR-D prices

6.5 Regulation Prices

6.5.1 Overview of historical prices

Finally, historical DK2 regulation prices are obtained from the Energi Data Service Platform ([50]) and the time-series for the current year (2024) is presented in Figure 6.17 in hourly resolution, where up-regulation prices show greater volatility with price peaks throughout the year, while in the case of down-regulation there is only one notable price peak, which occurs in the month of April. It is also possible to see how for each hour, the up-regulation price is never lower than the down-regulation price. This is to be expected since, as discussed in the theoretical framework, the up-regulation price will always be higher or equal to the DAM price while the down-regulation price will be equal or lower. As for the price distribution, this is illustrated in Figure 6.18a, where the maximum and minimum prices recorded during the year are also indicated, and it can be seen that the up-regulation prices have a greater dispersion.

Finally, the difference between the final imbalance price and DAM price is calculated. The histogram of the resulting values is illustrated in Figure 6.18b, where missing values have been removed. Therefore, a positive value indicates an hour when the system was in power deficit and up-regulation was activated, while a negative value indicates power excess and down-regulation activation. In terms of occurrences, for the 4367 hours recorded, up-regulation was activated in 1121 hours and down-regulation in 1842 hours, while no regulation was activated in 1404 hours.
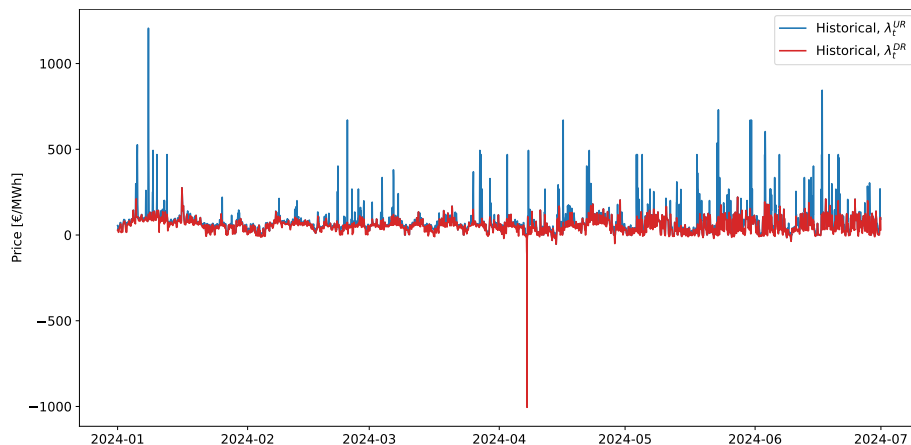

Figure 6.17: Historical DK2 regulation prices from Jan 2024 to July 2024. Data from: [50].

6.5.2 Forecasting

As in the case of reserve prices, it is also decided to use a naive prediction method for up-regulation ($\hat{\lambda}_t^{UR}$), down-regulation ($\hat{\lambda}_t^{DR}$) and imbalance prices ($\hat{\lambda}_t^{IMB}$). For a real application, it would be necessary to explore more sophisticated forecasting models.

In the energy markets section of the theoretical framework, it has been explained how the up-regulation prices are, for each hour, equal or higher than the DAM price, while the down-regulation prices are equal or lower. Therefore, in order to be consistent with the prediction of the DAM prices developed in the previous section, it is necessary to ensure the aforementioned behavior for the model developed. For this purpose, it has been decided to use the idea of naive regulation prices predictor described in [51], where the author makes a prediction of the regulation costs. These costs are defined as the difference between regulation prices and the DAM price, and their prediction is made

(a) Histogram of DK2 regulation prices   (b) Imbalance price minus Day-Ahead price
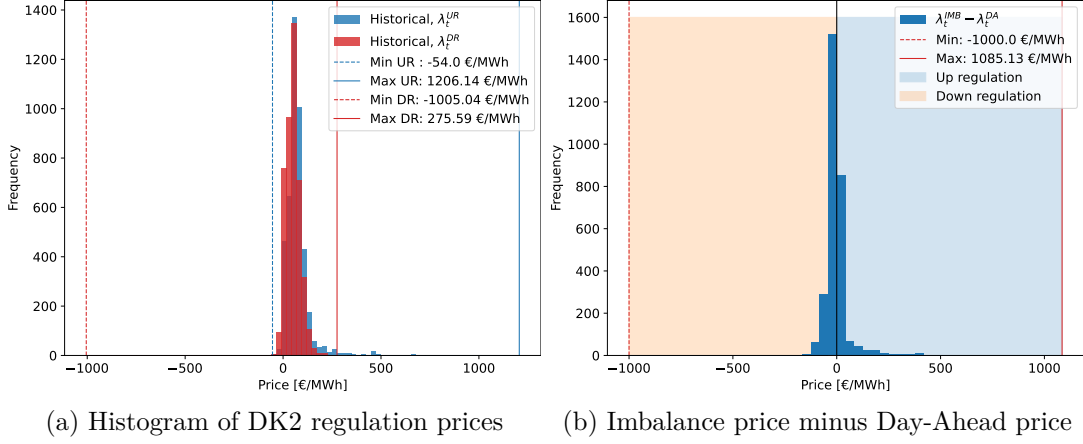
Figure 6.18: Histograms of DK2 regulation prices and difference between imbalance price and day-ahead price. Data from: [50]

by averaging these differences over the decision period. With these predicted costs, the prediction of regulation prices is finally defined as the difference between the actual DAM prices and the regulation costs.

However, the model in [51] is used for trading in the balancing market, so when predicting the regulation prices, the DAM prices for the trading horizon are already known from the day-ahead bid. In this thesis, on the contrary, the DAM prices are unknown when defining the bidding strategy, so the DAM prices used for this predictor are those forecasted as in the previous section. A higher error factor is therefore induced than in the case of [51], since the error introduced by the simplification of the model is added to the error of the prediction of the DAM prices.

$$\hat{C}^{UR} = \frac{1}{N_{\mathcal{T}_{hor}}} \sum_{\tau \in \mathcal{T}_{hor}} \left(\lambda_\tau^{UR} - \lambda_\tau^{DAM} \right) \tag{6.8}$$

$$\hat{C}^{DR} = \frac{1}{N_{\mathcal{T}_{hor}}} \sum_{\tau \in \mathcal{T}_{hor}} \left(\lambda_\tau^{DAM} - \lambda_\tau^{DR} \right) \tag{6.9}$$

Equations (6.8) and (6.9) define the forecasting of up-regulation and down-regulation cost parameters, denoted as $\hat{C}^{UR}$ and $\hat{C}^{DR}$ respectively. In these equations, the time step τ belongs to the selected horizon $\mathcal{T}_{hor}$ of 5 days prior to the bidding strategy decision, from $D-2$ to $D-6$, where $N_{\mathcal{T}_{hor}}$ the number of hours in $\mathcal{T}_{hor}$. It should be noted that the prices used in these equations are the actual historical prices fetched from Energinet API ([33], [50]). Finally, the up-regulation and down-regulation prices are defined in (6.10) and (6.11), where the forecasted prices described in the previous section are used.

$$\hat{\lambda}_t^{UR} = \hat{\lambda}_t^{DAM} + \hat{C}^{UR} \quad \forall t \in \mathcal{T} \tag{6.10}$$

$$\hat{\lambda}_t^{DR} = \hat{\lambda}_t^{DAM} - \hat{C}^{DR} \quad \forall t \in \mathcal{T} \tag{6.11}$$

Figure 6.19 illustrates the historical data used and the results of the forecasting methodology proposed for test day 07-06-2024. The bidding algorithm to obtain the optimal

bidding strategy for 07-06-2024 (day D) is executed at 00:30 a.m of 06-06-2024 ($D-1$) so, as explained above, the regulation prices of $D-1$ are not known. This lack of data is represented by the blank space in Figure 6.19 between the historical data and the results of the forecasting algorithm. The DAM prices of $D-1$ are known, but they are not considered following the proposed method. The same applies to the forecasting of down-regulation prices.

A detailed view of the results for the proposed method can be found in Figure 6.20a and Figure 6.20b for the same test day. The method seems to capture the trend of actual prices, but not the peak values. While this simplified naive predictor is enough for the purpose of this thesis, more sophisticated algorithms should be explored for real application.
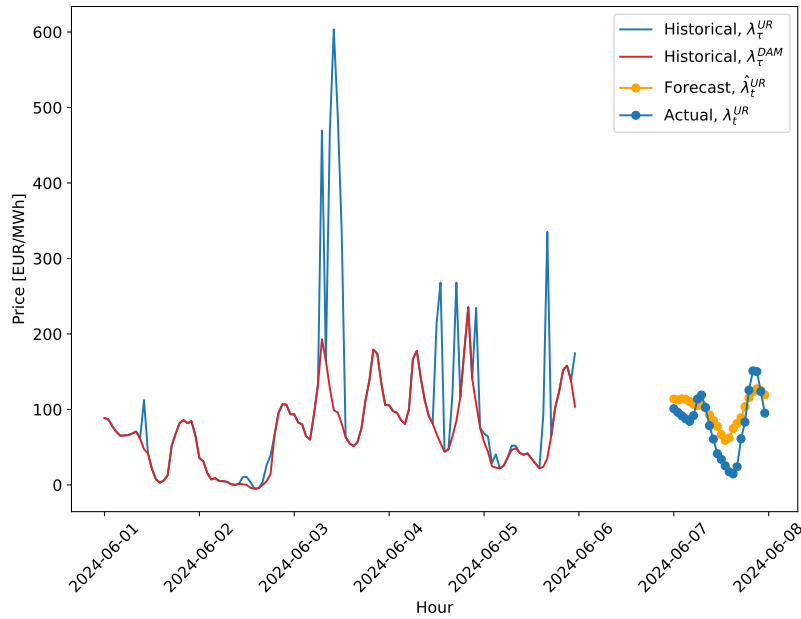
Figure 6.19: Time horizon and results of regulation prices forecasting method
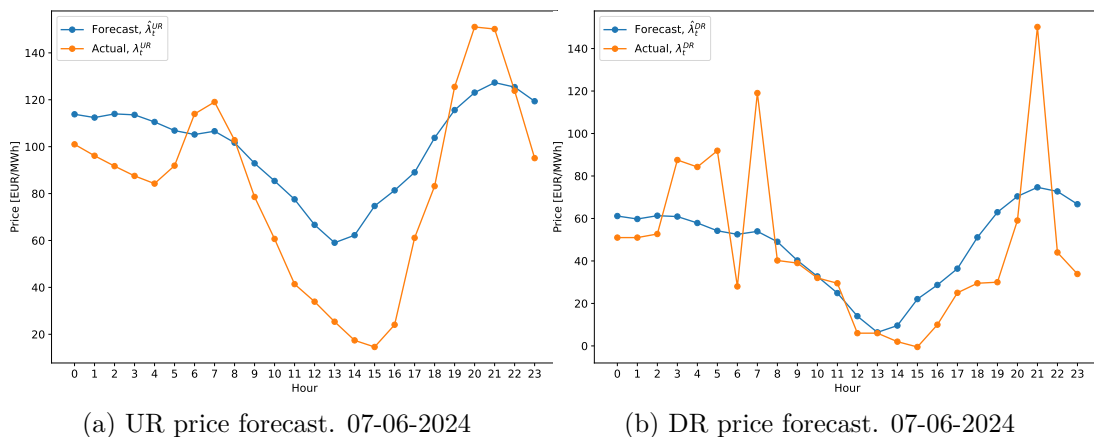
(a) UR price forecast. 07-06-2024 (b) DR price forecast. 07-06-2024

Figure 6.20: Forecasting of up-regulation and down-regulation prices

However, the final regulation and imbalance prices of the market are determined depending on the actual energy activation each hour. Thus, the up-regulation price will be equal to the DAM price if the system is down-regulated and the down-regulation price equals

DAM price if the system is up-regulated. Likewise, the final imbalance price equals the up-regulation or down-regulation prices if the system is up-regulated or down-regulated respectively.

In the previous section it has been explained that the power system need ($S_{t,w}^{Need}$) will be considered as an input to the algorithm. This is done by creating different scenarios for each hour where the system can be in power excess ($S_{t,w}^{Need} = -1$), power deficit ($S_{t,w}^{Need} = 1$) or balanced ($S_{t,w}^{Need} = 0$), where t refers to the set of hours and w refers to the set of scenarios. Therefore, $S_{t,w}^{Need}$ is the power system need at hour t under scenario w. This scenarios are created based on the frequency of activation of each service, already summarized in the overview of historical prices.

Five different scenarios will be considered where, for each hour, the system can be in power excess, deficit or balanced, creating a matrix of scenarios with shape 24x5, where each value is either -1, 0 or 1. The recurrence of each state follows the historical probabilities calculated from the hours of activation seen above. In this line, it is calculated that down-regulation has been activated in 42.2% of hours, while up-regulation has a frequency of activation of 25.7%. Therefore, the imbalance prices and final up-regulation and down-regulation prices used to maximize the revenue in the bidding strategy algorithm will be scenario-indexed and defined as:

$$\lambda_{t,w}^{UR} = \begin{cases} \hat{\lambda}_t^{DAM} & \text{if } S_{t,w}^{Need} = -1 \\ \hat{\lambda}_t^{DAM} & \text{if } S_{t,w}^{Need} = 0 \qquad \forall t \in \mathcal{T}, \forall w \in \mathcal{W} \\ \hat{\lambda}_t^{UR} & \text{if } S_{t,w}^{Need} = 1 \end{cases} \tag{6.12}$$

$$\lambda_{t,w}^{DR} = \begin{cases} \hat{\lambda}_t^{DR} & \text{if } S_{t,w}^{Need} = -1 \\ \hat{\lambda}_t^{DAM} & \text{if } S_{t,w}^{Need} = 0 \qquad \forall t \in \mathcal{T}, \forall w \in \mathcal{W} \\ \hat{\lambda}_t^{DAM} & \text{if } S_{t,w}^{Need} = 1 \end{cases} \tag{6.13}$$

$$\lambda_{t,w}^{IMB} = \begin{cases} \hat{\lambda}_t^{DR} & \text{if } S_{t,w}^{Need} = -1 \\ \hat{\lambda}_t^{DAM} & \text{if } S_{t,w}^{Need} = 0 \qquad \forall t \in \mathcal{T}, \forall w \in \mathcal{W} \\ \hat{\lambda}_t^{UR} & \text{if } S_{t,w}^{Need} = 1 \end{cases} \tag{6.14}$$

Finally, all scenarios are considered to have the same probability of occurrence so π_w, used in the notation to define these probabilities, is 1/5.

6.6 Frequency Reserve Activations

The fraction of FCR-N and FCR-D reserves that are activated in real time to provide balancing service is a function of the frequency measured by the operator and it follows the droop profiles illustrated in Figure 3.3, which can be converted to droop functions. This section explores how to obtain these functions and how to convert them to be used as an input to the bidding strategy algorithm.

The activation functions for FCR-N up-regulation, FCR-N down-regulation, FCR-D up-regulation and FCR-D down-regulation are denoted as $a_t^{N\uparrow}$, $a_t^{N\downarrow}$, $a_t^{D\uparrow}$ and $a_t^{D\downarrow}$ respectively. All functions are defined so that they take positive values, so although down-regulating activation is represented with negative values in the droop profile, for the purpose of this section it is considered as positive. Therefore, 1 is considered full activation and 0 represents no activation for each of the services and direction of activation.

Reserves allocated to FCR-N service are fully activated if the frequency exceeds 50.1Hz (down-regulation) or drops below 49.9Hz (up-regulation), while energy activation in between that range follows the linear part of the droop profile. Thus, activation functions for FCR-N can be expressed as in (6.15) and (6.16).

$$a_t^{N\uparrow} = \begin{cases} 1 & \text{if } f_t < 49.9 \text{ Hz} \\ \frac{50 \text{ Hz} - f_t}{0.1 \text{ Hz}}, & \text{if } 49.9 \text{ Hz} \leq f_t \leq 50 \text{ Hz} \\ 0 & \text{if } f_t > 50 \text{ Hz} \end{cases} \tag{6.15}$$

$$a_t^{N\downarrow} = \begin{cases} 0 & \text{if } f_t < 50 \text{ Hz} \\ \frac{f_t - 50 \text{ Hz}}{0.1 \text{ Hz}}, & \text{if } 50 \text{ Hz} \leq f_t \leq 50.1 \text{ Hz} \\ 1 & \text{if } f_t > 50.1 \text{ Hz} \end{cases} \tag{6.16}$$

Regarding to FCR-D service, full activation takes place if the grid frequency is higher than 50.5 Hz for up-regulation and is lower than 49.5 Hz for down-regulation, while no reserve is activated if the frequency is in the range 49.9 Hz - 50.1 Hz. Activations for up-regulation in the range of frequencies 49.5 Hz - 49.9 Hz and down-regulation in the range 50.1 Hz - 50.5 Hz follow the lineal part of the droop profile. Thus, activation functions for FCR-D can be expressed in (6.17) and (6.18).

$$a_t^{D\uparrow} = \begin{cases} 1 & \text{if } f_t < 49.5 \text{ Hz} \\ \frac{49.9 \text{ Hz} - f_t}{0.4 \text{ Hz}}, & \text{if } 49.5 \text{ Hz} \leq f_t \leq 49.9 \text{ Hz} \\ 0 & \text{if } f_t > 49.9 \text{ Hz} \end{cases} \tag{6.17}$$

$$a_t^{D\downarrow} = \begin{cases} 0 & \text{if } f_t < 50.1 \text{ Hz} \\ \frac{f_t - 50.1 \text{ Hz}}{0.4 \text{ Hz}}, & \text{if } 50.1 \text{ Hz} \leq f_t \leq 50.5 \text{ Hz} \\ 0 & \text{if } f_t < 50.1 \text{ Hz} \end{cases} \tag{6.18}$$

For the bidding strategy algorithm it would be ideal to have perfect information at $D-1$ of the frequency deviation events for day of operation D. However, since grid frequency trajectory is unpredictable, the transformation is required to convert the activation functions defined above into the activation parameters $A^{N\uparrow}$, $A^{N\downarrow}$, $A^{D\uparrow}$, $A^{D\downarrow}$ introduced in Chapter 5, keeping these parameters fixed over all hours in the algorithm.

Activation parameters are calculated from a historical dataset of DK grid frequencies with 1-second resolution covering from February to June 2024.

The histogram of frequencies from the dataset is represented in Figure 6.21. In addition to the histogram, vertical lines indicate the maximum and minimum measured frequency values as well as the frequencies at which FCR-N is fully activated and FCR-D starts its activation, 49.9 Hz for up-regulation and 50.1 Hz for down-regulation. The figure shows how frequency deviations that would trigger FCR-D activation rarely happen and how FCR-D was never activated during the data period.
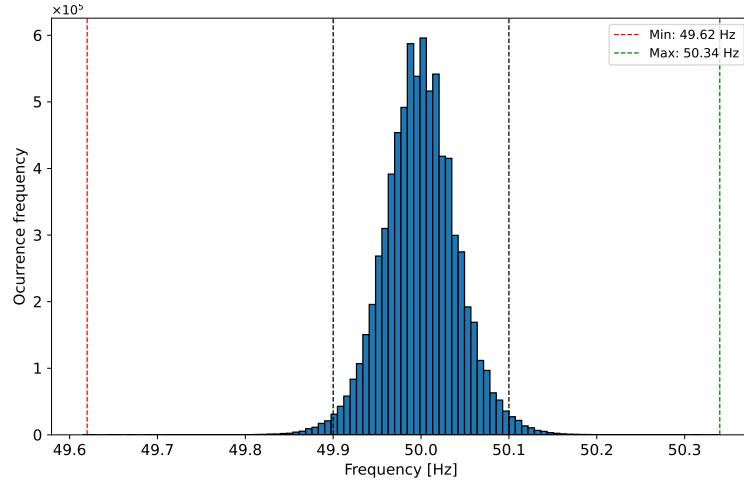
Figure 6.21: Histogram of power system frequencies in 1-second resolution from February 2024 to June 2024.

Since the bidding model has hourly resolution, (6.15) to (6.18) are applied to each of the data points in the dataset and converted into hourly resolution by calculating the average. The distribution of the 1-hour resolution activation variables for the four FCR services is represented in Figure 6.22.

Among the 2133 hours studied, FCR-D up-regulation only has activations above 1% in 29 hours, with a maximum of 6.1%. The activation for FCR-D down-regulation is similar, with only 33 hours with activations higher than 1%, although the maximum is higher than that of up-regulation, being 10.7%. For FCR-N it can be seen that the highest frequency of activation is in the range of 1-10% for down-regulation and 10-20% for up-regulation.
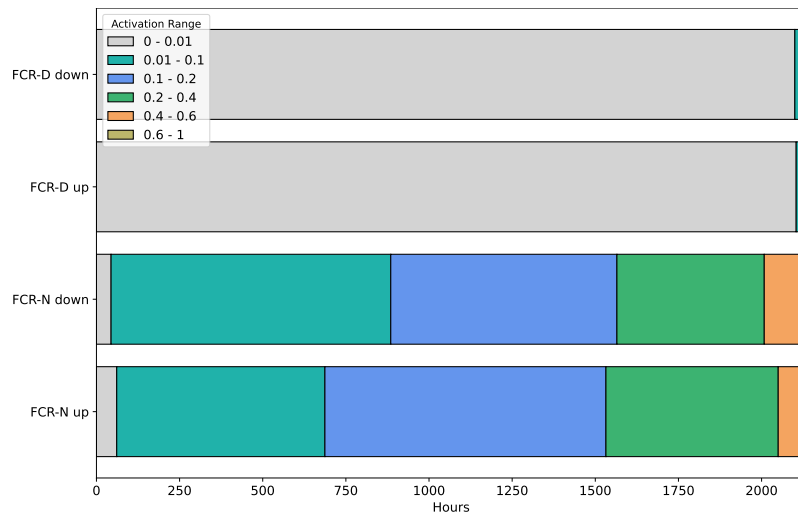
Figure 6.22: FCR activations distribution using 1-hour resolution activation variables.

With the activation variables in hourly resolution, it is now possible to transform them into the activation parameters that are used as an input to the bidding model. However, by calculating the ACF of the hourly activation variables, it is possible to prove that they do not present strong seasonality, so instead of using an activation parameter for each of the bidding optimization hours, it is decided to use the average over all hours for each

service, resulting in the four activation parameters. Thus, these parameters ($A^{N\uparrow}$, $A^{N\downarrow}$, $A^{D\uparrow}$ and $A^{D\downarrow}$) do not have the time sub-index t in the model. An example of the ACF for one of the 1-hour resolution activation variables is shown in Figure 6.23.

The final activation parameters obtained after calculating the average over all the hours are included in Table 6.5.

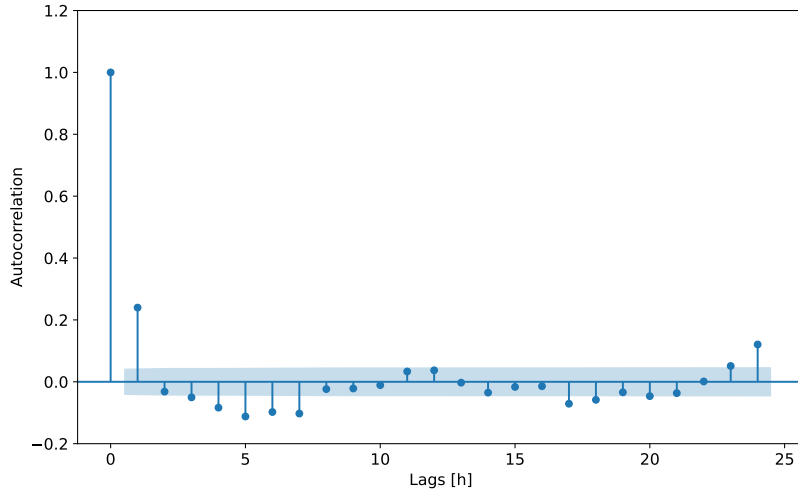Figure 6.23: Autocorrelation factor for 1-hour resolution FCR-N up-regulation activation variable.

Service	$FCR-N\uparrow$	$FCR-N\downarrow$	$FCR-D\uparrow$	$FCR-D\downarrow$
Activation parameter	$A^{N\uparrow}$	$A^{N\downarrow}$	$A^{D\uparrow}$	$A^{D\downarrow}$
Value	0.161	0.157	0.000588	0.000654

Table 6.5: Activation parameters for FCR energy activation.

7 Results

This section analyzes the results obtained after the implementation of the forecasting methods and the bidding strategy algorithm. First, the assumptions, simplifications, and considerations used are presented. Subsequently, an analysis of the optimal bidding strategies obtained from the optimization model is carried out, comparing them with the results obtained in a perfect foresight scenario. Finally, an analysis of the economic profitability of the strategies is conducted.

7.1 Assumptions and considerations

- The WF consists of 12 WTs rated at 2,000KW each, with variable speed and pitch control.

- The wind power forecasting is performed using a LWR model, with wind speed and temperature as features, obtained from external forecasts.

- Due to the poor results obtained from the errors induced by the external forecasts, LWR has been used with real forecasts of the exogenous variables, resulting in predictions with an average MAPE of 7%, representing a more realistic forecast error.

- DAM price is forecasted using SARIMAX.

- Regulation prices and FCR prices are forecasted using naive predictors.

- The APP is assumed to be a price-taker actor.

- The electricity markets considered are DAM, FCR-N and FCR-D.

- Imbalance settlement is also included.

- The bidding strategy is settled at FCR market closing gate (00:30 CET) and is not changed afterwards.

- The bids to DAM and FCR are assumed to be fully accepted.

- Fixed reserve activation parameters based on historical data are used.

- The BESS is constrained to be at 50% of SOC by the end of the day. SOC limits are bounded between 10% and 90%.

- The BESS has a rated capacity of 6MW, with maximum discharge and charge rate of 1C, and discharging and charging efficiency of 98%.

7.2 Bidding strategies

The charts used to illustrate market participation strategies consist of bar charts centered at zero. Above zero, the bars correspond to WF production, BESS discharge or up-regulation reserves. Below zero, the bars correspond to BESS charge or down-regulation reserves.

To analyze the performance against different price configurations, the results of four test days are examined.

Before focusing on each of the days, the optimal strategy for the WF for one of the days is shown in Figure 7.1. It is displayed here because the strategy is the same for all days,

and the only difference from one day to another is the values themselves. Therefore, it is not necessary to show the graph for each day.

The results show that for the WF, the optimum strategy is always to offer its most accurate power forecast in the DAM and FCR-D down-regulation. This outcome is due to the fact that if the WF wanted to bid in FCR-D up-regulation or FCR-N, it would have to operate under curtailment to leave enough capacity to provide up-regulation reserve. Given this situation and the current price levels, it is more optimal for the turbine to bid its capacity for down-regulation, as this allows it to operate at its realized power and only curtail if necessary to provide the service.
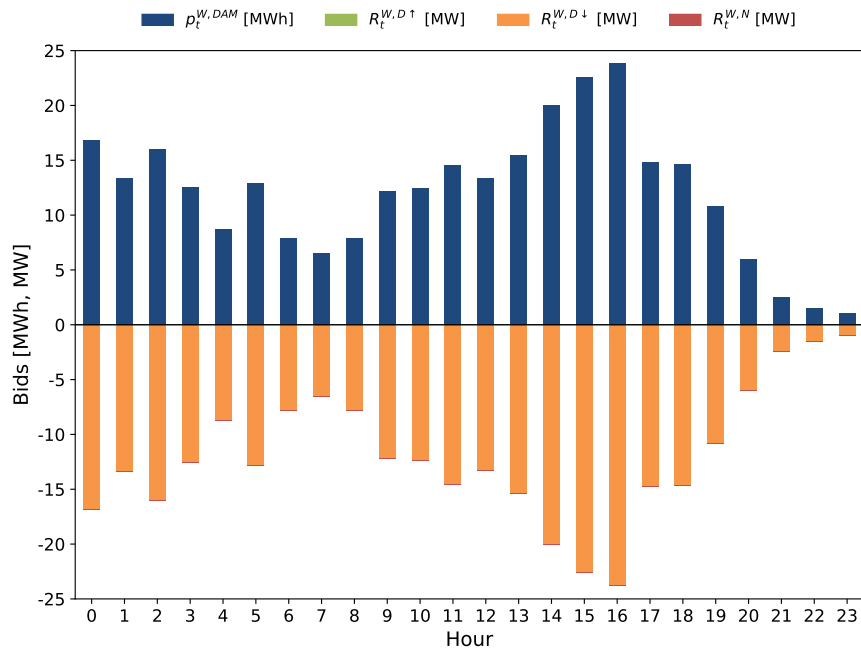

Figure 7.1: Optimal WF bids into the markets, test day 1

7.2.1 Test Day 1

The strategy for the BESS is illustrated in Figure 7.2. It should be noted that although it is rated at 6 MW, there are bars that can reach 12 MW. This is because the up-regulation and down-regulation reserves can be up to twice the rated power of the battery. For example, in hour 13, the battery is operating with a baseline of charging 6 MW. For that hour, the optimal result is to offer 12 MW of up-regulation reserve, which accounts for reducing the baseline of 6 MW of charge plus discharging 6 MW. The same happens in hours 0 and 21, where the BESS has a baseline of around 4 MW of discharge and offers a FCR-D up-regulation reserve of about 9 MW, accounting for the reduction of the 4 MW baseline and 5 MW of actual reserve provision. In these two hours, the BESS also saves some capacity to bid in FCR-N.

Since FCR-N is a symmetrical service, it is doubly represented in the graph. Using as an example the hours 1 to 12, the BESS offers 6MW of FCR-N reserve, although 12MW are represented since the activation can be in both directions.

The evolution of SOC is also depicted in Figure 7.2 with a black line, where it can be appreciated that the BESS starts and ends the day with a SOC of 50%.
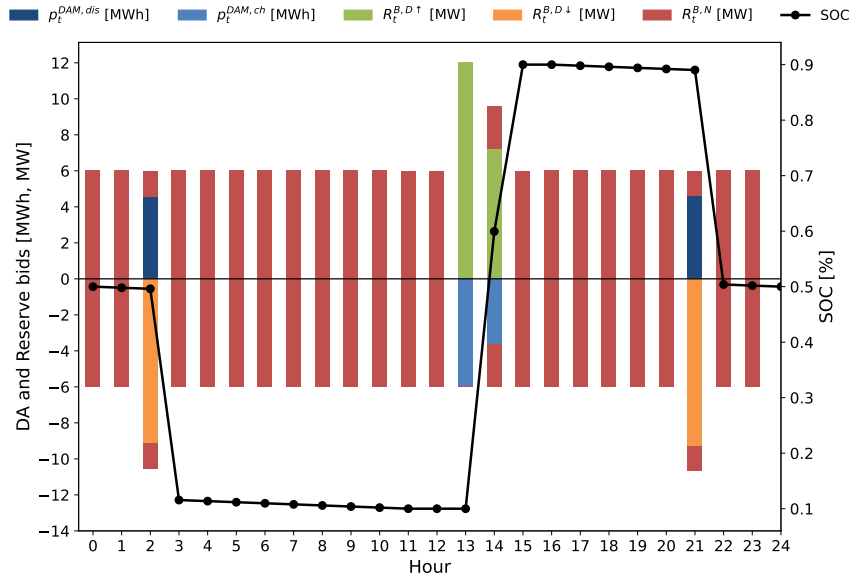
Figure 7.2: Optimal BESS bids, test day 1

To analyze the bidding strategy, the predicted prices for this day are shown in Figure 7.3.

It can be observed that the DAM prices are higher than the reserve prices for all but three hours. Given this situation, the optimal strategy for the battery is to take advantage of the high DAM prices during the first hours of the day to discharge in the third hour, which also allows the BESS to bid down-regulation reserve in FCR-D and save some capacity for FCR-N. Once discharged, the BESS takes advantage of the fact that FCR-N prices are higher than those of FCR-D to make full capacity bids in this market. This is maintained until hours 13 and 14, and coinciding with a low DAM price, an offer is placed for charging, which also allows bidding capacity in FCR-D up-regulation. When DAM prices are again higher than FCR-N prices, the battery continues to bid only in FCR-N. Additionally, with the BESS charging during the hours of lower DAM prices, it is possible to discharge when prices are higher, which occurs in hour 21.
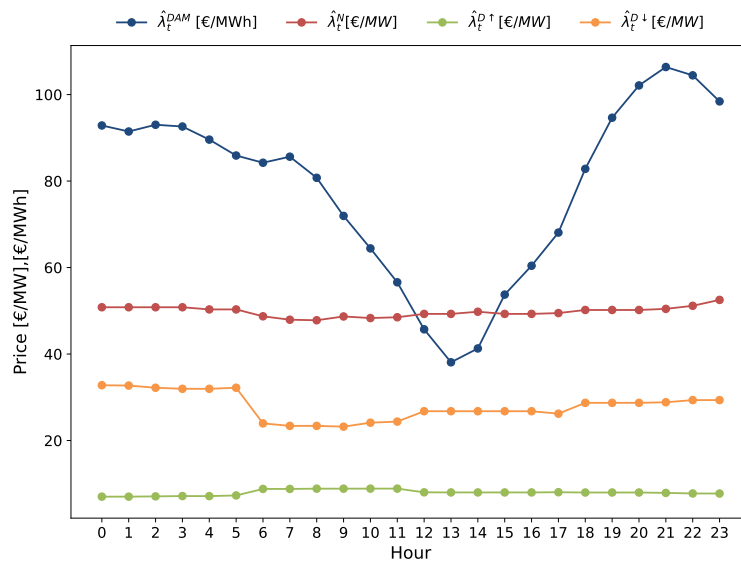

Figure 7.3: Forecasted prices for test day 1

7.2.2 Test Day 2

The next test day has been chosen because the FCR-N and FCR-D-down prices are higher than the DAM price, presenting a different price configuration as the one seen before. The results are illustrated in Figure 7.4.

During the first hours of the day, the FCR-D down price is the highest, so the BESS seeks to participate in this market. To do so, it charges in the first hour of the day. During the next five hours, when the FCR-D down price remains the highest, the BESS can also place a bid for discharging in the DAM, which allows it to bid even more reserve. Once the BESS's SOC is at its minimum, it supplements the FCR-D down bid with FCR-D up bids. During the remaining hours, the BESS seeks to participate in FCR-N, as it is the service with the highest prices, with the exception of a charging baseline in hours 11 and 12, when the DAM price is lower, and discharging in hours 20 and 21, coinciding with the peak DAM prices. In the last hour, a bid for charging is placed to reach 50% SOC.
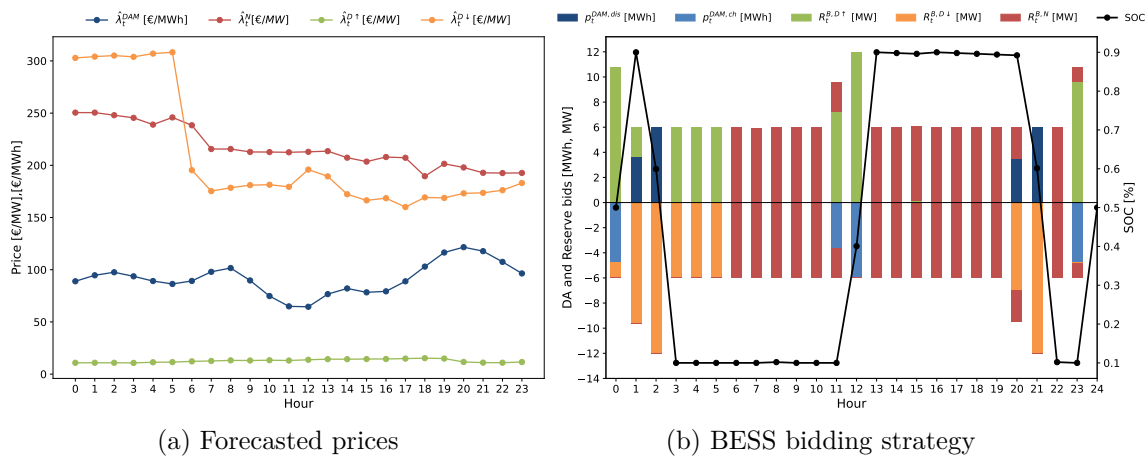

| (a) Forecasted prices | (b) BESS bidding strategy |

Figure 7.4: Bidding strategy results, test day 2

7.2.3 Test Day 3

The third day presents a new price configuration, as the price of FCR-D-down is the highest during the whole day, followed by FCR-N and DAM, which does not exceed the previous ones at any time.

From the results included in Figure 7.5 it can be observed how the BESS always seeks to participate in FCR-D-down. Moreover, as in this case the FCR-N price is always lower, it does not seek to participate in it throughout the day. In addition, a greater arbitrage is observed with the DAM bids, seeking to charge at local minima such as hours 6, 7, 12, 13 and discharge at local maxima such as hour 9.

There are also other charging bids placed for the DAM in hours 16, 18 and 19. The results seem to indicate that they are placed to take advantage of the DAM price rise in hour 17 and the peak in hours 20 and 21, since these discharging baselines also allow to make bids for FCR-D-down reserve, taking advantage of the high price.
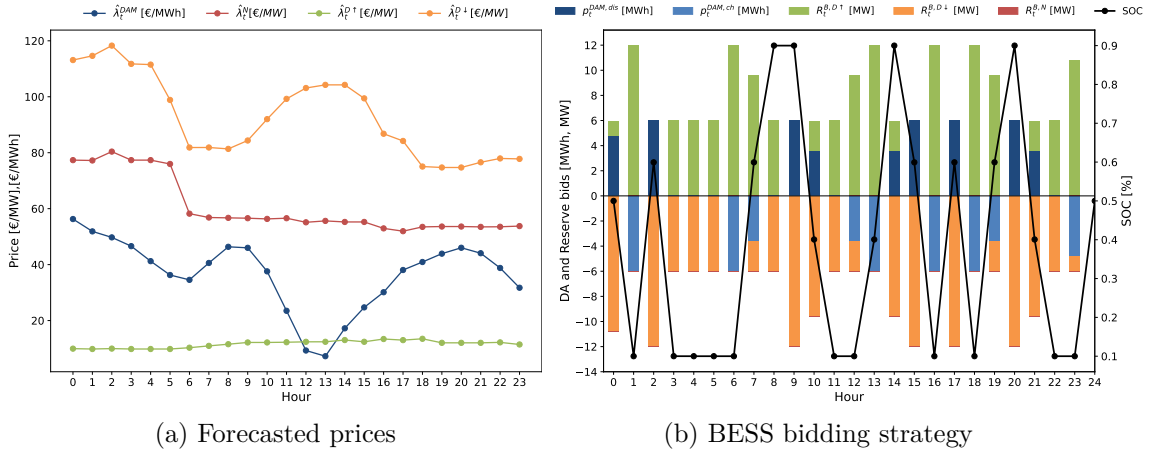
(a) Forecasted prices

(b) BESS bidding strategy

Figure 7.5: Bidding strategy results, test day 3

7.2.4 Test Day 4

The last test day studied presents a more variable price configuration, with the reserve prices crossing the DAM prices in several hours. In this case, a combination of strategies similar to those seen on the previous test days is presented. When FCR-N prices are higher than FCR-D, the BESS bids on the former unless the DAM prices have a local maximum, as is the case in hour 2. However, it can be seen that if FCR-N prices are not sufficiently higher than FCR-D, as is the case in the middle hours of the day, it is more optimal to bid on FCR-D, which allows for simultaneous FCR-D bids in the other reserve direction or in the DAM. The DAM peak events are again used for discharging in hours 20 and 21.

(a) Forecasted prices

(b) BESS bidding strategy
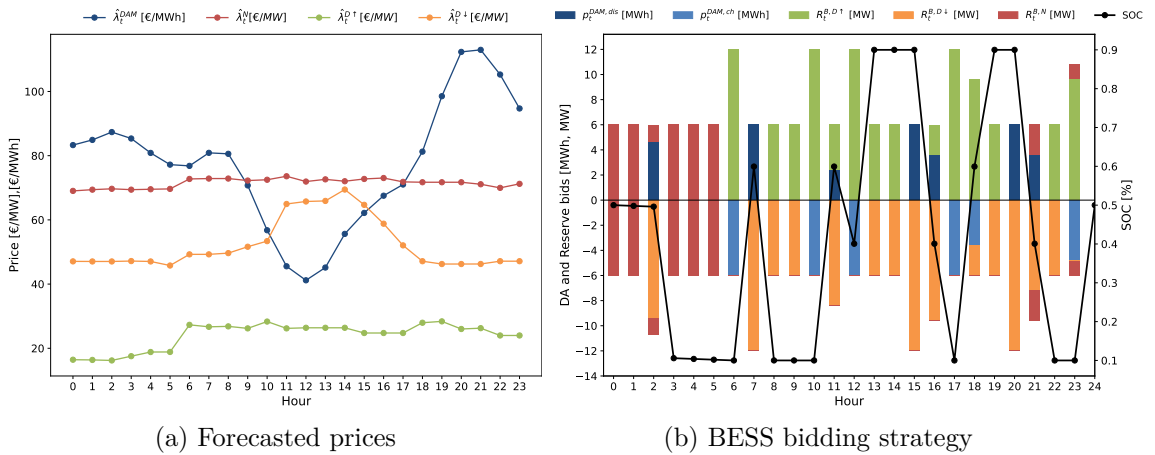
Figure 7.6: Bidding strategy results, test day 4

7.3 Profitability

In the previous section, the strategies adopted by the APP in terms of the expected electricity prices have been analyzed. However, these prices are actually forecasts, so the impact of the wrong prediction on the results has not been analyzed. In this section, the expected revenue when planning the bidding strategy will be compared with the actual perceived revenue, as well as the revenue that would have been obtained if perfect prices and wind power information had been available at the time of calculating the optimal bids.

7.3.1 Expected vs. Realized revenue

The revenue expected by the APP is the result of the objective function defined in (5.6). On the other hand, the realized revenue is the actual revenue that the APP will receive based on the actual wind power realization and actual electricity prices.

The realized revenue is calculated as the sum of the revenues from the energy markets (DAM + Imbalance settlement) and the FCR-D and FCR-N services, following (5.2), (5.4) and (5.5) respectively. However, in this case, the actual market prices and the bidding strategy resulting from the optimization model are used. Furthermore, the realization of the WF is also considered to calculate the imbalance settlement for deviations from the DAM schedule. Finally, actual frequency measurements from different days transformed to hourly frequency data are used to calculate reserve activations due to frequency deviations. Although the frequency measurements corresponding to the days studied are not available, the use of actual data corresponding to other dates is representative of a typical daily situation. It should be noted that for calculating the imbalance settlement and FCR-N activated energy payments it is no longer necessary to perform scenario indexing, since the system state and final imbalance price is known when calculating the realized revenue.

The expected and realized revenues over the four testing days are listed in Table 7.1. It can be seen that although the revenues are similar for test day 1 and test day 4, the optimal bidding strategy for test day 2 represents a very optimistic scenario, while for test day 3 it represents a pessimistic scenario.

	Test Day 1	Test Day 2	Test Day 3	Test Day 4
Expected revenue (€)	37,648	43,157	37,535	36,777
Realized revenue (€)	34,258	14,058	52,634	31,343

Table 7.1: Expected vs. realized revenue

Before analyzing these deviations in depth, the realized revenue per hour over the four days (96 hours) is illustrated in Figure 7.7, where the final height of each stacked bar reflects the total revenue for that hour. FCR-N revenues include both reservation and activation payments.
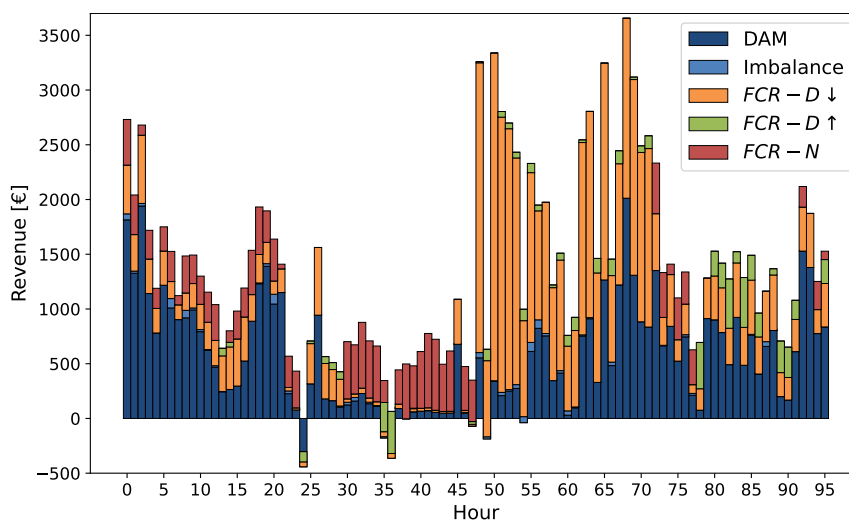

Figure 7.7: Hourly realized revenue per market for the APP over four testing dates

The graph shows variability in terms of the most influential market on total revenue. In general, the DAM has a significant influence, especially on the first day. However, on the second day, FCR-N acquires more relevance, as does FCR-D down-regulation on the third day. The total contribution of each market for each day is shown in Table 7.2.

	Day 1		Day 2		Day 3		Day 4	
	Expected	Realized	Expected	Realized	Expected	Realized	Expected	Realized
DAM	22,026	21,321	2,813	2,645	6,706	14,496	15,158	17,097
Imbalance	0	400	0	36	0	488	0	32
FCR-D up	154	112	638	996	1,671	1,447	2,740	3,347
FCR-D down	8,828	6,269	20,564	2,769	29,157	36,204	15,943	8,385
FCR-N	6,224	5,478	18,860	7,174	0	0	2,852	2,421
FCR-N activations	415	677	285	440	0	0	84	63

Table 7.2: Expected vs. realized revenue in € per market for four test days

Several conclusions can be drawn from the above table. On the one hand, the difference between expected revenue and realized revenue from the DAM is within an acceptable range except for day 3. On the other hand, for the days studied, the final imbalance settlement ends up resulting in an extra profit beyond what was expected. As seen, the WF offers its most accurate forecast in the DAM, making the expected imbalance null. The extra profit comes from the imbalance created because of the realization of the WF, which has the net effect of helping with the total imbalance of the system.

As for the reserves, the largest differences in revenue are created from FCR-D down-regulation and FCR-N participation, being the main factors that explain the largest difference between expected and realized revenue for days 2, 3 and 4. The revenue for FCR-N activations also varies considerably in relative terms when using real frequency measures instead of the fixed parameters used in this thesis. However, its contribution to total APP revenue is minor, so it does not explain the total difference.

It has been seen that the biggest differences between expected revenue and realized revenue come from the DAM and FCR markets. Obviously, the revenue of these markets is sensitive to the forecasts used for prices and wind power, so the differences between these forecasts and actual values are presented below.

Figure 7.8 shows the difference between expected and realized revenue in the DAM, wind power, and DAM price. It can be clearly seen how the difference in revenue can be explained based on prediction errors, with the third day serving as an example of an incorrect forecast, which is precisely the day that experiences the largest deviation in terms of DAM revenues.

Although significant wind power deviations are not present in this case, the contribution to price discrepancies can still be observed. For instance, between hours 5 and 10, the actual price is considerably higher than the forecasted price. However, since the WF produces 2 MW less than expected, the realized revenue does not exceed expectations by a large margin.

The next market where significant differences are observed in terms of expected and realized revenue is FCR-D down-regulation. For this market, three of the tested days show optimistic revenue estimates, particularly exaggerated for days 2 and 4. Conversely, day 3 proves to be pessimistic, with actual revenues exceeding expectations. To analyze the four days on an hourly basis, Figure 7.9 is included, following the same format as Figure 7.8 but showing revenues and price forecasts for FCR-D down-regulation.
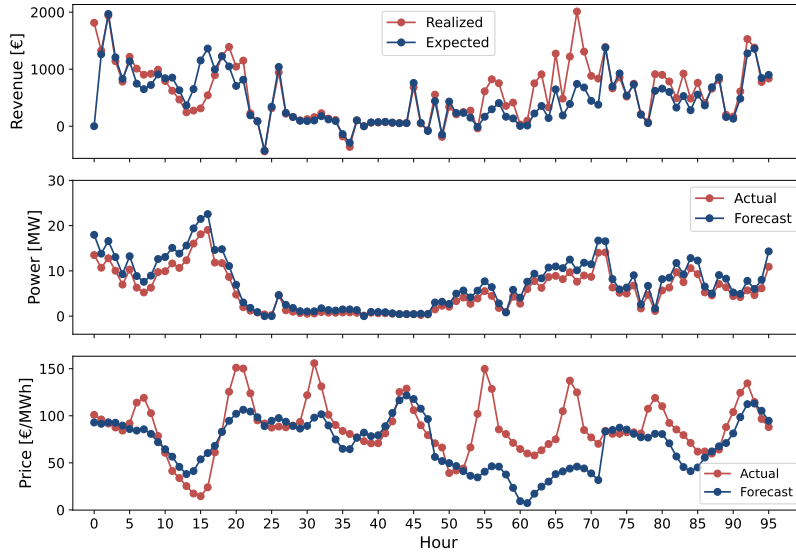
Figure 7.8: Hourly DAM realization. (From top to bottom: Realized vs Expected profit, Wind Power Imbalance, DAM price forecast)
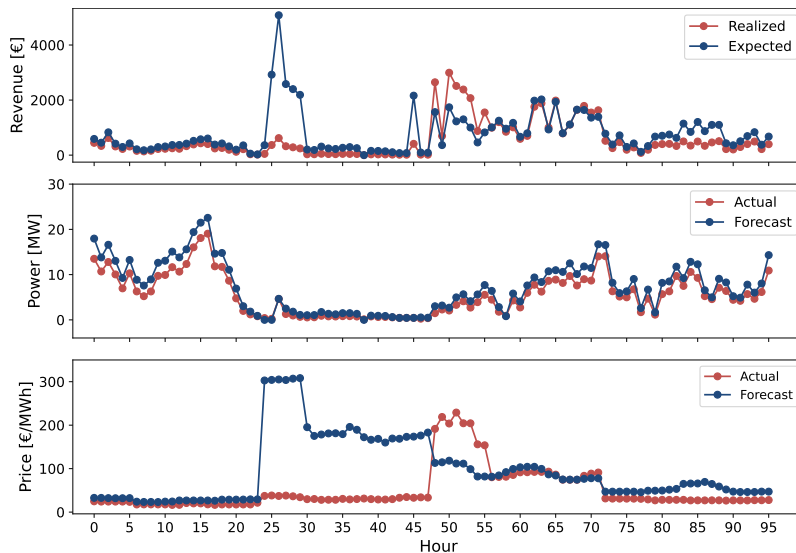

Figure 7.9: Hourly FCR-D down-regulation realization. (From top to bottom: Realized vs Expected profit, Realized vs Expected wind power, FCR-D price forecast)

As mentioned, days 2 and 4 result in higher-than-expected revenues in FCR-D down-regulation. It can be clearly seen that this is due to a significantly deviated price estimate on day 2 and a combination of both price and wind power deviations on day 4. Conversely, day 3 results in higher revenue than expected due to a price deviation in the early hours of the day (hours 48 to 57).

It is worth noting that the reserve price prediction was based on a naive predictor, so the deviations were to be expected.

Finally, the same analysis is carried out for FCR-N, as it shows a significant deviation in revenue on day 2. In Figure 7.10, it can be clearly seen that this difference is due to an extremely optimistic price forecast. It is also observed that the forecast was highly

deviated during the early hours of the third day, but since it was not considered optimal to participate in this market, it has no impact on the revenue.
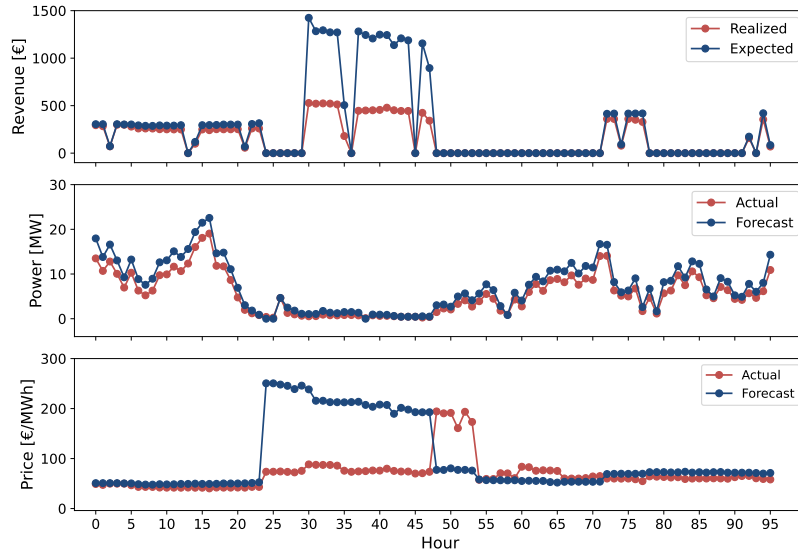

Figure 7.10: Hourly FCR-N realization. (From top to bottom: Realized vs Expected profit, Realized vs Expected wind power, FCR-N price forecast)

7.3.2 Forecasts vs. perfect information

With the previous result analyses, it has been demonstrated how the expected revenue of the plant, based on the bidding strategy, is sensitive to the forecasts used. Therefore, this section analyzes the outcomes of the model in an ideal situation where perfect forecasts are available for planning the strategy. To achieve this, the same model described in Chapter 5 is executed, but using actual wind power and market price data as inputs.

It is worth noting that even if perfect forecasts of balancing energy prices (up-regulation and down-regulation) are used, it is still considered that a reliable forecast of the system state for each hour cannot be made. Therefore, scenario-based indexing is still maintained. The same applies to the activation of reserves; the method for calculating them does not change.

Firstly, the expected vs. realized values are listed in Table 7.3. As expected, when using perfect information forecasts, the volatility between the expected and realized values is much lower than when using imperfect forecasts as the ones studied in this thesis. The difference between both revenues is precisely due to the uncertainty introduced by the system state and the activation of reserves. If they were known exactly, the revenues should be the same.

	Test Day 1	Test Day 2	Test Day 3	Test Day 4
Expected revenue (€)	32,484	14,736	52,615	31,846
Realized revenue (€)	33,408	14,882	52,615	32,265

Table 7.3: Realized vs. expected revenue using perfect information forecasts

Additionally, the realized revenues for each market are analyzed for the case with perfect information and real forecasts. The results are presented in Table 7.4. Firstly, it is observed that the payment for imbalance disappears, as the wind power is known exactly

and offered in the DAM market. Again, the remaining differences between both cases are due to the uncertainty of reserve activation and the final imbalance price. Generally, the perceived revenue from the DAM, FCR-N market, and reserve activation tends to increase, while the income from FCR-D decreases.

	Day 1		Day 2		Day 3		Day 4	
	Perfect info	Forecasts	Perfect info	Forecasts	Perfect info	Forecasts	Perfect info	Forecasts
DAM	20,760	21,321	2,707	2,645	15,314	14,496	17,491	17,097
Imbalance	0	400	0	36	0	488	0	32
FCR-D up	165	112	718	996	1,452	1,447	2,951	3,347
FCR-D down	6,225	6,269	1410	2,769	35,849	36,204	7,834	8,385
FCR-N	5,393	5,478	9403	7,174	0	0	3,565	2,421
FCR-N activations	865	677	644	440	0	0	424	63
Total [€]	33,408	34,258	14,882	14,058	52,615	52,634	32,265	31,343

Table 7.4: Comparison of realized revenue using perfect and imperfect forecasts

When using perfect information forecasts, the BESS participation strategies change slightly throughout the day. However, since the explanation of these follows the same principles as those already explained with real forecasts, they have not been included.

All comparative graphs between perfect information forecasts and real forecasts, including strategies for the WF and BESS, can be found in Appendix A, as well as the considered system state scenarios.

8 Discussion

In Chapter 4, different wind power prediction models have been tested and analyzed, and in Chapter 6, the remaining inputs used for the bid optimization model developed in Chapter 5 have been defined. These inputs include DAM forecasts, for which different models have been analyzed, and reserve and regulation price forecasts, for which naive predictors have been employed. Other analyzed inputs include the scenarios of power need for the electrical system and historical frequency values to define the activation parameters. The outcome of all these analyses are the results presented in Chapter 5, where the different bidding strategies adopted by the APP and their economic performance have been analyzed.

Based on all the analysis of results done so far, this chapter elaborates on them, summarizing the main findings and developing insights.

8.1 Wind power forecasts

A correct estimation of wind power has proven important in defining optimal market participation strategies, as it has been shown that a significant contributor to the APP's final revenue has been the participation of the WF in the DAM on most of the days studied. This does not mean that it always provided the highest income, as depending on the price configuration, there were days when bids in reserve markets were more lucrative. However, the revenue from participation in the DAM was never neglectable.

The analysis of the WF's SCADA data available for this thesis has identified wind speed as the main factor for wind power prediction, as expected due to their cubic relationship. The strong correlation presented led to the exploration of regression-based models. Thus, a simple linear regression model was tested first, followed by an autoregressive model in which past values of wind power were considered as features for the model. Both methods had the drawback of predicting particularly poorly during periods of low wind power, although the use of autoregressive variables proved more efficient in accurately predicting peak power values. Finally, given the cubic relationship between wind speed and power, a model based on local weighted regression was implemented, which allowed for a smoother fit curve and achieved the best prediction results.

However, the use of external variables requires their own forecasts. Although external forecasts obtained from DMI have been used, the difficulty of adapting these forecasts to the particular conditions of the WF has resulted in predictions with considerable errors, highlighting the importance of having proprietary forecasts for the exogenous variables to be considered.

8.2 Electricity prices forecasts

Price forecasting has been the most important factor both for planning market participation strategies and for the final revenue obtained on the studied days, with deviations between expected revenue and realized revenue coinciding with deviations in price forecasts.

Unlike with wind power, from the analysis of the day-ahead market price time series, it was decided to evaluate time series forecasting methods of varying complexity. The first model tested was Prophet, which resulted in worse outcomes than a SARIMA model that

considered the seasonality of the price time series. For this model, it was more accurate to manually tune it by selecting parameters based on the ACF and PACF functions rather than automatically with predefined functions. Finally, the possibility of incorporating exogenous variables into the model was explored, resulting in a SARIMAX model, which proved to be the best both in absolute and relative terms.

As with the wind power forecasts, external forecasts for the exogenous variables were used, in this case from the ENTSO-E platform. These forecasts were useful for improving the model, although there was still significant variability in results depending on the day tested. Therefore, for a real application, more complex models should be studied.

Regarding the FCR reserve market prices, it has been analyzed that during the first half of 2024, FCR-N prices have generally been higher than FCR-D prices, with up-regulation prices being lower than down-regulation prices. This price profile may indicate that participation in FCR-N could be more lucrative than in FCR-D. However, since the objective of this thesis is not to conduct an economic study on the integration of the APP into the markets, and given the need to simplify the methodology due to time constraints, the decision was made to opt for naive predictors. While these predictors have resulted in significant error, they have allowed this price trend to be roughly maintained.

Although the price trend has been maintained with these simplifications and has allowed for the identification of potential bidding strategies for the BESS, they have produced a significant deviation between the expected revenue and the final perceived revenue. This has highlighted the need for accurate predictions.

The same applies to the prediction of balancing energy prices, where the simplification of methods has induced a lot of error in price prediction. However, these errors have not significantly influenced the final strategies or the final revenue, as they were used to quantify the benefit from reserve activations, which have turned out to be a minimal fraction of the final revenue.

8.3   Other modelling considerations

Other areas where assumptions have been made when implementing the optimization algorithm include the creation of scenarios to model the system's state at each hour, as well as the creation of reserve activation parameters.

The first aspect has been implemented to define the final regulation and imbalance prices of the system, as these depend on the net state of the system at each hour (power deficit, excess, or balanced). Due to the difficulty in predicting the above, considering various scenarios and calculating the average revenue across all of them allows for an adequate modeling of these conditions. These scenarios have been created based on historical data regarding the system's state for the first half of 2024, aiming to reflect this frequency in the set of scenarios. The implementation of this technique allows for a more conservative bidding strategy and reduces the potential error between expected and realized revenue. Additionally, when evaluating expected and realized revenue in the case of exact price forecasts, where scenario indexing has been maintained, the difference between them has not been significant, so this technique has proven adequate. This is also related to the fact that, as mentioned, the revenue from FCR-N reserve activations constitutes a small fraction of the total profits.

On the other hand, real frequency data measured in the system with a 1-second resolution have been analyzed. Using the FCR-N and FCR-D activation curves, activation signals have been calculated and, after averaging them to obtain hourly resolution, the mean has

been used as the reserve activation parameter. When calculating the realized revenue, the deviation in revenue related to FCR-N activations is significant, but since the contribution to the APP's total revenue is small, this procedure has been sufficient for the modelling of the reserve algorithm.

8.4   Bidding strategies

The bidding strategies based on prices have been explained in the previous chapter for each of the four days explored. However, although the price configuration for each day is different, many strategies are repeated, so they are discussed in this section.

Firstly, the strategy for the WF is repeated for all the study days. It offers its most accurate prediction in the DAM and the same amount as downward reserve in the FCR-D market. With the current price level, this occurs because a WT would have to operate under curtailment if it offered in FCR-D up-regulation or FCR-N to have sufficient capacity to increase power if a sudden reserve activation occurs. In this situation, the WF prefers to offer in the DAM and gain extra benefit from the FCR-D down-regulation reserve, as it does not require curtailment in real time unless activation.

Regarding the BESS, repetitive patterns are also observed throughout the days. On the one hand, it is observed that the battery prefers to offer in reserve markets rather than in the DAM. However, during periods when it is more optimal to offer in FCR, it takes advantage of local maxima in day-ahead prices to discharge and have more reserve to offer for up-regulation, or local minima to charge and have more reserve to offer for down-regulation.

It is also observed that the expected revenues from reserve activations in FCR-N are not significant enough to determine the optimal strategy, and it ends up choosing the reserve service with the highest price per capacity. Thus, whenever FCR-D prices are higher than FCR-N prices, it opts for the former, even if the latter allows to earn money from activated energy. In fact, the BESS opts for FCR-D even when the FCR-N price is slightly higher. The reason is that since FCR-D is an asymmetric service, it can combine offering in the DAM with offering in FCR-D to maximize profit. For example, the BESS can offer a baseline discharge at maximum discharge power in the DAM and a downward reserve of up to twice its capacity (half by stopping discharge and the other half by charging capacity), earning money from both the DAM and FCR-D for the same hour. However, these cases may encounter limitations not studied in this thesis, such as leading to impossible real-time operation due to SOC limits. This situation could be avoided if the SOC for each hour is limited by considering that full reserve activations might occur, but this would lead to overly conservative BESS operation and would not allow for the analysis of a wide variety of strategies. In addition, it would be interesting to study the price difference at which the BESS opts for FCR-D or FCR-N.

Regarding the profitability of the results, the expected revenues have been compared with the realized revenues. While for two of the four days studied the differences are not very large, the same cannot be said for the other two days. After analyzing the revenues by market, it has been observed that the largest differences come from erroneous forecasts of reserve prices. As already mentioned, these prices come from naive forecasts, so the induced error explains the revenue deviations. For that reason, the expected revenue has also been analyzed against the realized revenue using perfect price forecasts, where it has been shown that the volatility between expected and realized revenues is reduced, highlighting once again the need for accurate forecasts to obtain realistic expected revenues.

9 Conclusions

This thesis has focused on developing a bidding strategy algorithm for the participation and management of an APP, consisting of a WF and BESS, in electricity markets and ancillary services. By achieving its specific objectives, this research has addressed the integration of data-driven models for both electricity prices and wind power forecasting with optimization models to derive optimal bidding strategies in the joint markets of DAM, FCR-D, and FCR-N.

The initial step involved the refinement and filtration of SCADA data collected from multiple WTs, ensuring proper data quality for model training. After selecting the explanatory variables to be considered, forecasting techniques were developed to predict wind power generation and electricity prices, focusing on regressive methods for the first and time series models for the second. However, forecasts for reserve and balancing energy prices have been done with naive predictors. These forecasts served as critical inputs for the formulation of a MILP optimization model designed to determine optimal bidding schedules for APP participation in the aforementioned markets.

For wind power forecasts, the explanatory variables chosen from data analysis were incorporated into the model using external forecasts. However, the difficulty in adjusting these forecasts to the specific characteristics of the WF (such as location, hub height, etc.) resulted in poor forecast accuracy. In contrast, including external forecasts for predicting DAM prices did not distort the results as significantly.

Given that the objective of this work was not to conduct an economic viability study of the markets, the developed model was tested only over four days with different price configurations to analyze various market participation strategies. This approach allowed the study of optimal participation strategies for the WF, revealing that the best strategy involves bidding its most accurate forecast into the DAM and FCR-D down-regulation reserve. Additionally, the strategy for the BESS was analyzed, showing a preference for reserve markets over the DAM. It was also found that, given the assumptions made, FCR-N reserve activations generally do not constitute the fundamental criteria for choosing FCR-N over FCR-D.

The developed strategies were analyzed by comparing expected revenues with realized revenues. This analysis included scenarios where perfect information forecasts were used, in which actual electricity prices and wind power were known at the time of planning the bid offering strategy. This comparison highlighted the necessity for more precise forecasts to reduce the gap between expected and realized revenues. Achieving more accurate forecasts could enable the creation of realistic long-term business scenarios, enhancing the economic viability and strategic planning of APPs in actual market conditions.

9.1 Future Works

Although this thesis has successfully demonstrated the integration of data-driven models and optimization techniques to define bidding strategies, numerous simplifications and assumptions were made during the development process. Therefore, this section outlines potential areas for future work.

9.1.1 Wind power forecasts
- **Advanced data filtering techniques:** In this thesis, the initial step for wind power forecasts involved data filtering using the normal operation ranges provided by

the manufacturer. However, if such operational ranges are not available, it would be necessary to develop advanced models for data filtering. Examples of such techniques have been discussed in the literature review.

- **Exploration of more complex models:** Future research should explore more complex models, such as time series models, to improve the accuracy of wind power forecasts. Additionally, it would be advantageous to develop forecasts for the exogenous features considered.

- **Consideration of additional inputs:** Future work could include other aspects to improve the final forecast of the WF, such as the layout and orographic obstacles. These factors were not considered in this thesis due to their complexity, but their inclusion could improve the accuracy of the forecasts.

9.1.2 Electricity prices forecasts

- **Forecast of reserve and balancing prices:** Besides improving the DAM price forecasts with more complex models, future work should focus on developing more sophisticated models for predicting reserve and balancing energy prices.

9.1.3 Bidding strategy algorithm

- **Dynamic bidding strategy**: This thesis assumes that the bidding strategy is defined at the closure of the FCR market and not modified afterwards. However, in practice, DAM offers can be modified until the DAM market closure, which occurs almost 12 hours later. This allows the use of forecasts with a shorter horizon, potentially increasing accuracy. Future work should consider this dynamic adjustment capability to refine the bidding strategy. Furthermore, the active participation of the BESS in the balancing market to reduce deviations could be analysed.

- **FCR-D and FCR-N late auctions:** Future research could also explore the potential benefits of participating in late auctions for FCR-D and FCR-N.

- **Real-time control model:** This thesis did not consider a real-time control model for the plant. Consequently, situations where the BESS might fail to activate reserves due to state-of-charge (SOC) constraints were not studied. Future work should incorporate a real-time control model to address these scenarios, ensuring more reliable and resilient operation of the APP.

- **Stochastic vs deterministic models:** The current model is deterministic, relying on point forecasts. Future work should focus on developing a stochastic model that incorporates probabilistic forecasts and scenario indexing, similar to the approach used for regulation prices. This would enable the inclusion of risk aversion metrics in the objective function.

Bibliography

[1] Wind Europe. *Wind energy in Europe. 2023 Statistics and the outlook for 2024 - 2030.* February 2024. URL: https://windeurope.org/intelligence-platform/product/wind-energy-in-europe-2023-statistics-and-the-outlook-for-2024-2030/.

[2] McKinnon. C et al. "Comparison of novel SCADA data cleaning technique for wind turbine electric pitch system". In: *Journal of Physics: Conference Series* Vol. 2151.No. 1 (2022), p. 0120051. DOI: https://iopscience.iop.org/article/10.1088/1742-6596/2151/1/012005.

[3] Llombart. A et al. "Robust data filtering in wind power systems". In: *European wind energy conference EWEC, Athens* Vol. 2 (2006, February), p. 149–54.

[4] Fahim. M et al. "Machine learning-based digital twin for predictive modeling in wind turbines". In: *IEEE Access,* 10 (2022), pp. 14184–14194.

[5] Tuton. E, Ma. X, and Dethlefs. N. "Intelligent digital twin–machine learning system for real-time wind turbine wind speed and power generation forecasting". In: *E3S Web of Conferences* 10 (2023), p.01008.

[6] Conejo. A. J. et al. "Forecasting electricity prices for a day-ahead pool-based electric energy market." In: *International journal of forecasting,* 21(3) (2005), pp. 435–462.

[7] Tschora. L. et al. "Electricity price forecasting on the day-ahead market using machine learning." In: *Applied Energy,* 313,118752 (2022).

[8] Hameed. Z, Træholt. C, and Hashemi. S. "Investigating the participation of battery energy storage systems in the Nordic ancillary services markets from a business perspective". In: *Journal of Energy Storage,* 58 (2023), p. 106464.

[9] Zhan. S et al. "Co-optimized trading of hybrid wind power plant with retired EV batteries in energy and reserve markets under uncertainties." In: *International journal of electrical power & energy systems* 117,105631 (2020).

[10] Alavijeh N. M et al. "Optimal Scheduling of Battery Storage Systems in the Swedish Multi-FCR Market Incorporating Battery Degradation and Technical Requirements." In: *arXiv preprint arXiv:2406.07301.* (2024).

[11] Thingvad A. et al. "Economic value of multi-market bidding in Nordic frequency markets." In: *2022 International Conference on Renewable Energies and Smart Technologies (REST)* 1, (2022), pp. 1–5.

[12] Glas J., Semerow A., and Luther M. "General analysis of frequency containment and restoration reserves of wind power plants in power systems." In: *Renewable Energy and Power Quality Journal* 1(14), (2016), pp. 96–101.

[13] Zhao X., Xue Y., and Zhang X. P. "Fast frequency support from wind turbine systems by arresting frequency nadir close to settling frequency." In: *IEEE Open Access Journal of Power and Energy* 7, (2020), pp. 191–202.

[14] Hansen M. *Aerodynamics of wind turbines.* Routledge, 2015.

[15] Hansen A. D. *Introduction to wind power models for frequency control studies.* DTU Wind Energy E, 2016. ISBN: 978-87-93278-60-8.

[16] Nordic Balancing Model NBM. *Single price and Single position – implementation in the Nordics, Common Market Design description.* [Online]. URL: https://nordicbalancingmodel.net/wp-content/uploads/2020/10/Single-Price_Common-Market-Design.pdf.

[17] Energinet. *Ancillary service to be delivered in Denmark - Tender conditions.* February 2024. URL: https://en.energinet.dk/media/gmci1x5h/ancillary-services-to-be-delivered-in-denmark-tender-conditions-1-2-2024.pdf.

[18] Energinet. *Prequalification of Units and aggregated portfolios.* September 2023. URL: https://en.energinet.dk/media/r0rbaoqb/prequalification-of-units-and-aggregated-portfolios.pdf.

[19] Astolfi D. et al. "Multivariate Data-Driven Models for Wind Turbine Power Curves including Sub-Component Temperatures". In: *Energies* 16(1) (2022), p. 165.

[20] Huang H. et al. "Feature selection and hyper parameters optimization for short-term wind power forecast". In: *Applied Intelligence* (2021), pp. 1–19.

[21] Restrepo. L. F and Gonzalez J. "De pearson a spearman". In: *Revista Colombiana de Ciencias Pecuarias* 20(2) (2007), pp. 183–192.

[22] Ke. Z and Zhang. Z. "Testing autocorrelation and partial autocorrelation: Asymptotic methods versus resampling techniques". In: *British Journal of Mathematical and Statistical Psychology* 71(1) (2018), pp. 96–116.

[23] Herlau T., Schmidt. M. N., and Mørup M. *Introduction to machine learning and data mining.* Technical University of Denmark(DTU), 2016.

[24] Sen S., Das M. N., and Chatterjee R. "A weighted kNN approach to estimate missing values". In: *3rd International Conference on Signal Processing and Integrated Networks (SPIN)* IEEE (2016), pp. 210–214.

[25] DMI. *DMI Open Data.* [Online]. URL: https://opendatadocs.dmi.govcloud.dk/DMIOpenData.

[26] DMI. *Weather Model (HARMONIE) for DINI and IG.* [Online]. URL: https://opendatadocs.dmi.govcloud.dk/Data/Forecast_Data_Weather_Model_HARMONIE_DINI_IG.

[27] Irwin J. S. "A theoretical variation of the wind profile power-law exponent as a function of surface roughness and stability". In: *Atmospheric Environment (1967)* 13(1) (1979), pp. 191–194.

[28] Kubik M. L., Coker P. J., and Hunt C. "Using meteorological wind data to estimate turbine generation output: a sensitivity analysis". In: *World Renewable energy congress.* Linköping University Electronic Press, Linköping; Sweden (2011), pp. 4074–4081.

[29] Montgomery D. C., Peck E. A., and G. G. Vining. *Introduction to linear regression analysis.* John Wiley and Sons, 2021.

[30] Cheng L. et al. "An output-only ARX model-based sensor fusion framework on structural dynamic measurements using distributed optical fiber sensors and fiber Bragg grating sensors". In: *Mechanical Systems and Signal Processing* 152 (2021), p. 107439.

[31] Atkenson C.G., Moore A.W., and Schaal S. "Locally weighted regression for control". In: *Artificial Intelligence Review* 11 (1997), pp. 75–113.

[32] Fan M. et al. "Coupling the K-nearest neighbors and locally weighted linear regression with ensemble Kalman filter for data-driven data assimilation". In: *Open Geosciences* 13(1) (2021), pp. 1395–1413.

[33] Energinet. *Elspot Prices.* [Online]. URL: https://www.energidataservice.dk/tso-electricity/Elspotprices.

[34] Shohan M. J. A., Faruque M. O, and Foo S. Y. "Forecasting of electric load using a hybrid LSTM-neural prophet model". In: *Energies* 15(6) (2022), p. 2158.

[35] Taylor S. J. and Letham B. "Forecasting at scale". In: *The American Statistician* 72(1) (2018), pp. 37–45.

[36] Meta Open Source. *Quick Start | Prophet.* [Online]. URL: https://facebook.github.io/prophet/docs/quick_start.html#python-api.

[37] Dubey A. K. et al. "Study and analysis of SARIMA and LSTM in forecasting time series data." In: *Sustainable Energy Technologies and Assessments* 47 (2021), p. 101474.

[38] Box G. E. et al. *Time series analysis: forecasting and control.* John Wiley and Sons, 2015.

[39] Ariyanti V. P. and Yusnitasari T. "Comparison of Arima and Sarima for forecasting crude oil prices". In: *RESTI (Rekayasa Sistem dan Teknologi Informasi)* 7(2) (2023), pp. 405–413.

[40] Zhao Z. et al. "Improving short-term electricity price forecasting using day-ahead LMP with ARIMA models". In: *IEEE Power and Energy Society General Meeting* (2017), pp. 1–5.

[41] *pmdarima.* https://pypi.org/project/pmdarima/. Retrieved April 2024.

[42] *pmdarima.arima.auto$_a$rima.* https://alkaline-ml.com/pmdarima/modules/generated/pmdarima.arima.auto_arima.html. Retrieved April 2024.

[43] Wu. L and Shahidehpour. M. "A hybrid model for day-ahead price forecasting". In: *IEEE Transactions on Power Systems* 25(3) (2010), pp. 1519–1530.

[44] ENTSO-E. *Transparency Platform.* [Online]. URL: https://transparency.entsoe.eu/.

[45] Usoro A. E. "Some basic properties of cross-correlation functions of n-dimensional vector time series". In: *Journal of Statistical and Econometric Methods* 4(1) (2015), pp. 63–71.

[46] ENTSO-E Transparency Platform. *Total Load - Day Ahead / Actual.* [Online]. URL: https://transparency.entsoe.eu/load-domain/r2/totalLoadR2/show.

[47] ENTSO-E Transparency Platform. *Actual Generation per Production Type.* [Online]. URL: https://transparency.entsoe.eu/generation/r2/actualGenerationPerProductionType/show.

[48] ENTSO-E Transparency Platform. *Generation Forecasts for Wind and Solar.* [Online]. URL: https://transparency.entsoe.eu/generation/r2/dayAheadGenerationForecastWindAndSolar/show.

[49] Energinet. *FCR N and D, Frequency Containment Reserves, DK2.* [Online]. URL: https://www.energidataservice.dk/tso-electricity/FcrNdDK2.

[50] Energinet. *Regulating and Balance Power, Overall Data.* [Online]. URL: https://www.energidataservice.dk/tso-electricity/RegulatingBalancePowerdata.

[51] Andersen P.H.D. "Optimal trading strategies for a wind-storage power system under market conditions". Master's thesis. DK-2800 Kgs. Lyngby, Denmark: Technical University of Denmark, DTU, 2009.

A   Sustainable Development Goals

A positive outcome of this project for a real application would imply the feasibility of developing aggregated asset management models that improve the energy efficiency of these assets and create a business case for their participation in electricity markets and ancillary services. On the one hand, economic viability would motivate generation companies to continue installing wind farms and consider hybridization with energy storage assets. On the other hand, ancillary services have traditionally been performed by non-conventional plants, so a correct performance of aggregated renewable assets would facilitate the displacement of these non-conventional plants and their replacement by renewables. Therefore, this project has a clear alignment with Sustainable Development Goals (SDGs) 7, Affordable and Clean Energy, and 13, Climate Action .

A   Appendix

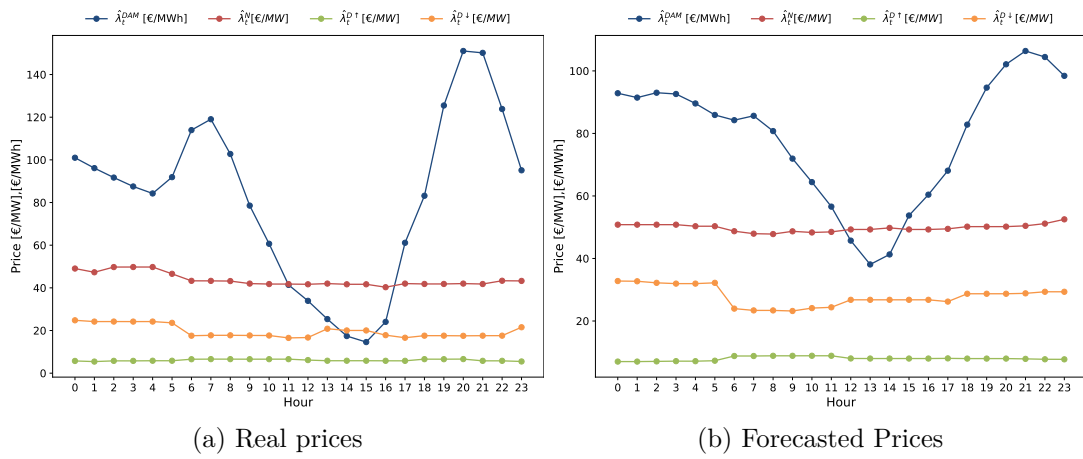## A.1   Real vs. perfect information forecasts

A.1.1   Test day 1

(a) Real prices

(b) Forecasted Prices

Figure A.1: Real vs. Forecasted prices for test day 1

(a) Perfect information
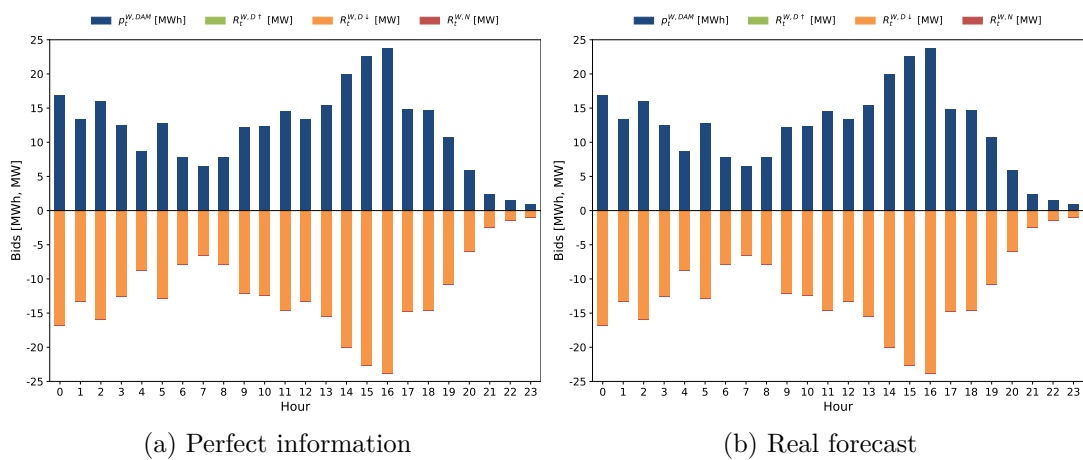
(b) Real forecast

Figure A.2: WF bidding with real forecasts and perfect information for test day 1

(a) Perfect information   (b) Real forecast
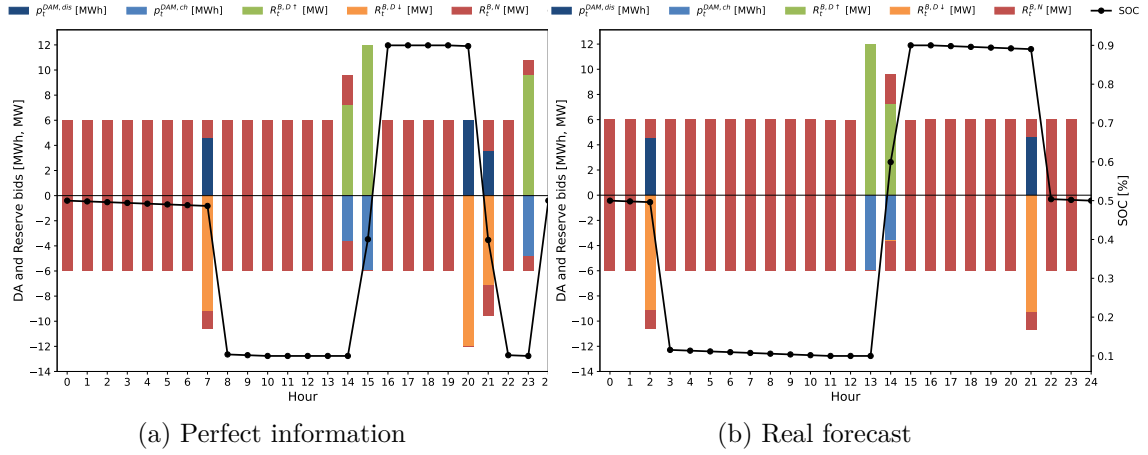
Figure A.3: BESS bidding with real forecasts and perfect information for test day 1
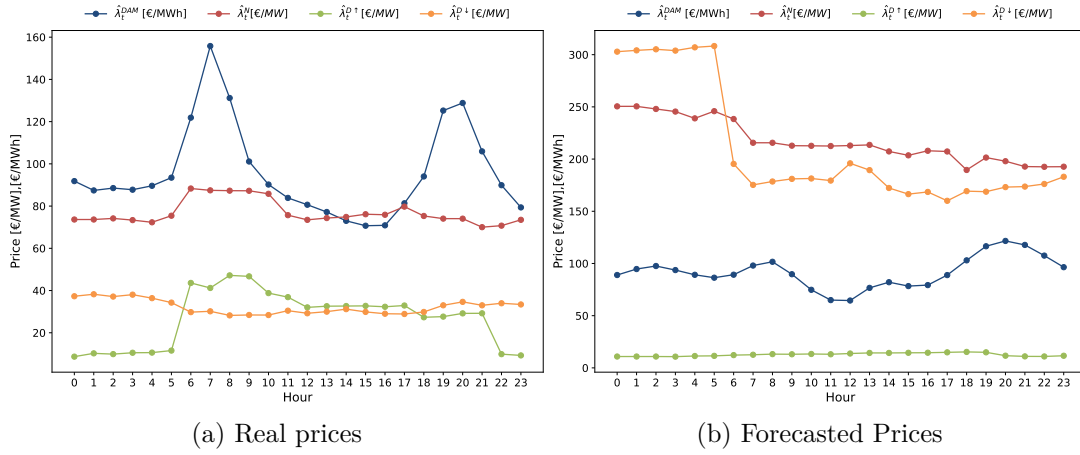
A.1.2   Test day 2
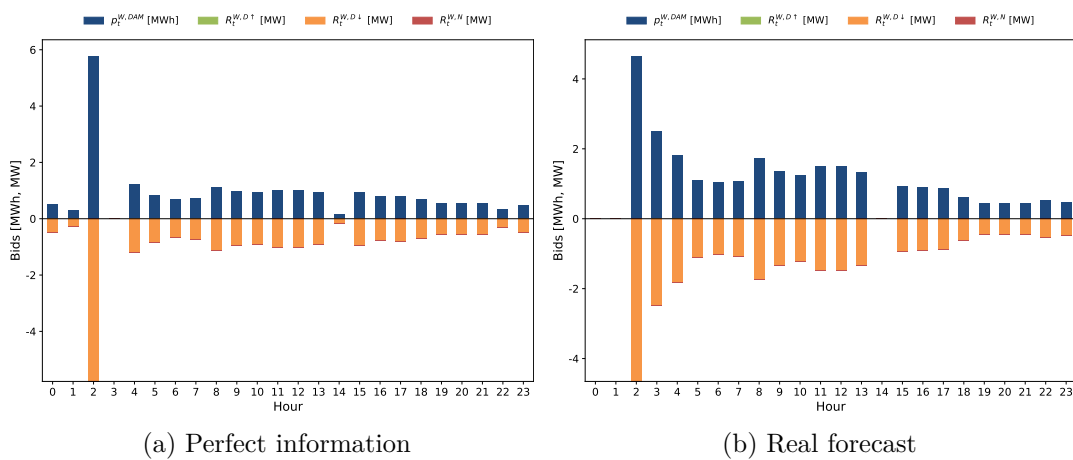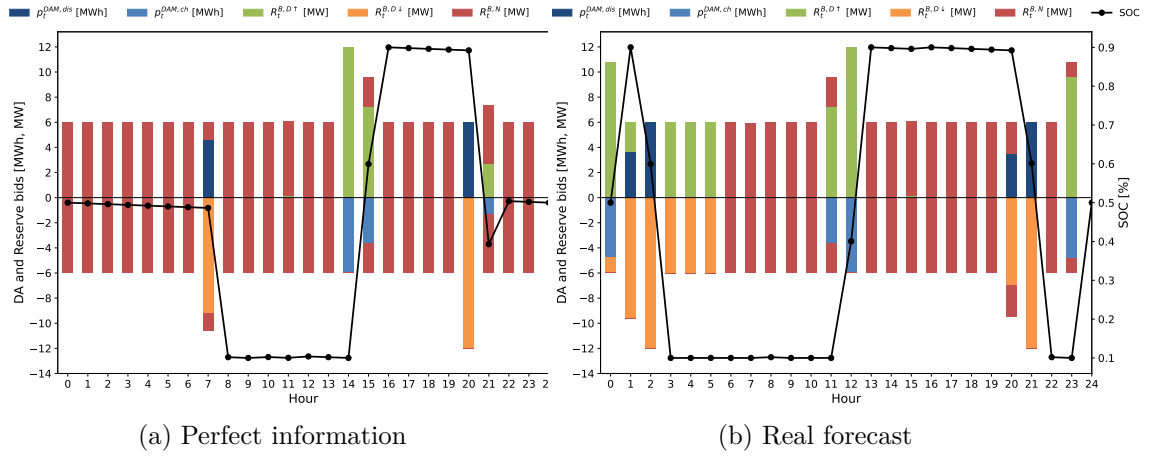

(a) Real prices   (b) Forecasted Prices

Figure A.4: Real vs. Forecasted prices for test day 2

(a) Perfect information   (b) Real forecast

Figure A.5: WF bidding with real forecasts and perfect information for test day 2

(a) Perfect information

(b) Real forecast

Figure A.6: BESS bidding with real forecasts and perfect information for test day 2
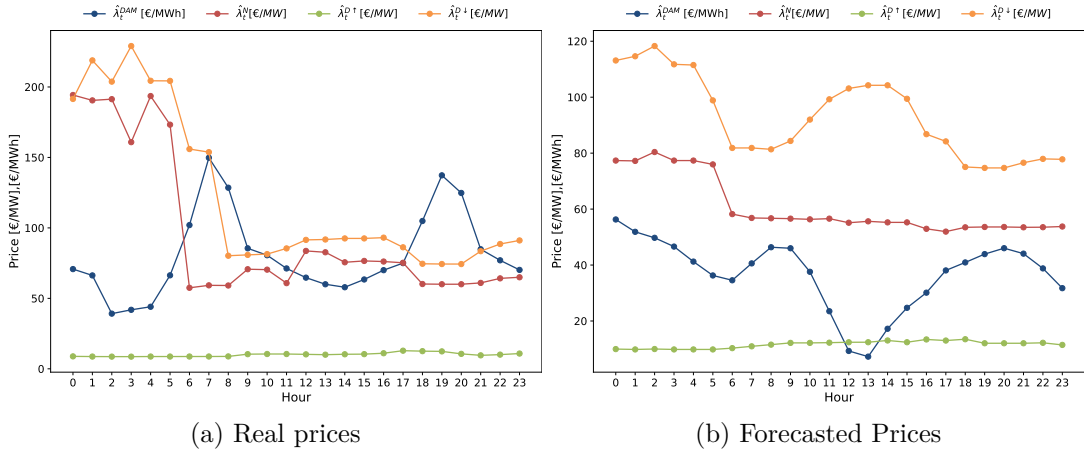
A.1.3   Test day 3
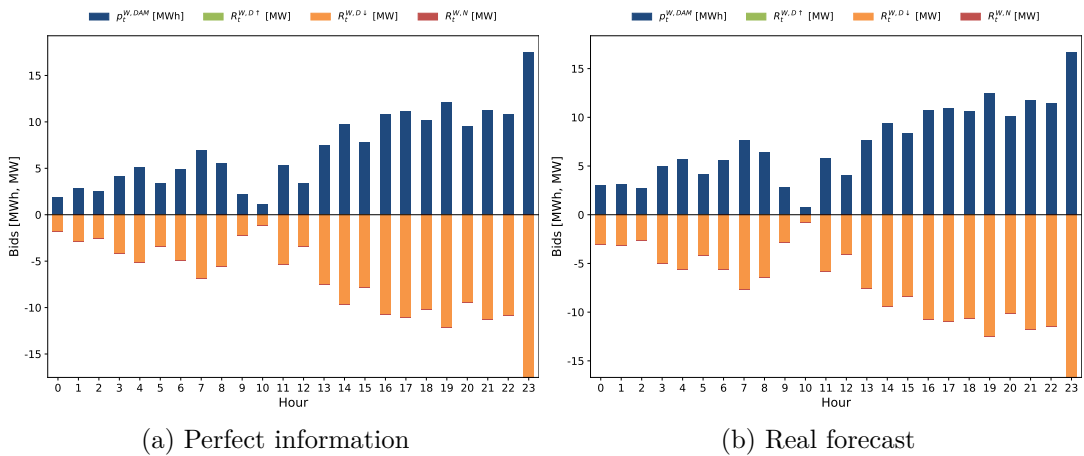

(a) Real prices

(b) Forecasted Prices

Figure A.7: Real vs. Forecasted prices for test day 3

(a) Perfect information

(b) Real forecast

Figure A.8: WF bidding with real forecasts and perfect information for test day 3
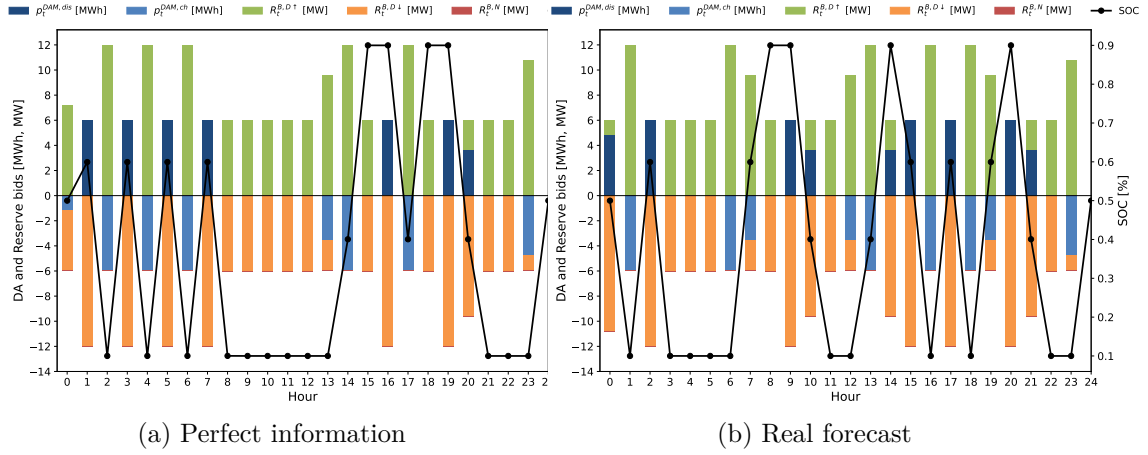
(a) Perfect information

(b) Real forecast

Figure A.9: BESS bidding with real forecasts and perfect information for test day 3
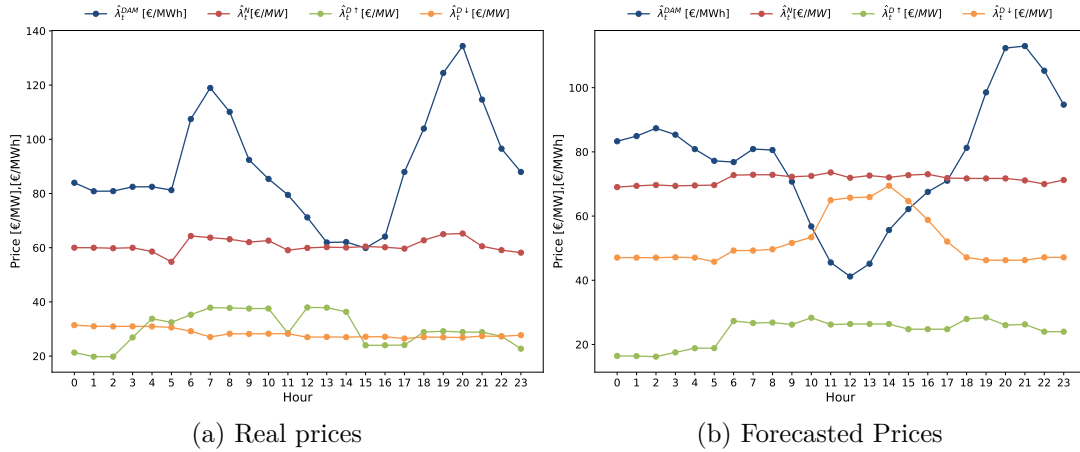
A.1.4 Test day 4
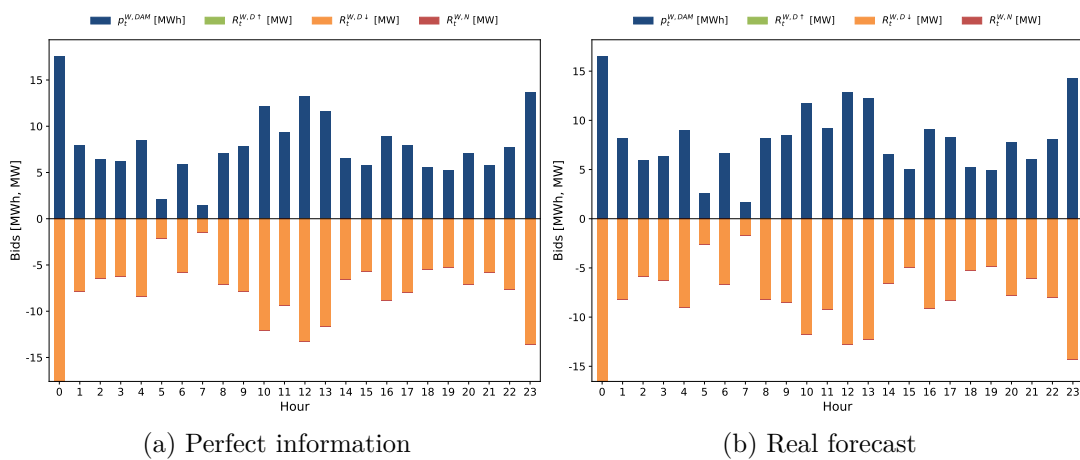

(a) Real prices

(b) Forecasted Prices

Figure A.10: Real vs. Forecasted prices for test day 4

(a) Perfect information

(b) Real forecast

Figure A.11: WF bidding with real forecasts and perfect information for test day 4
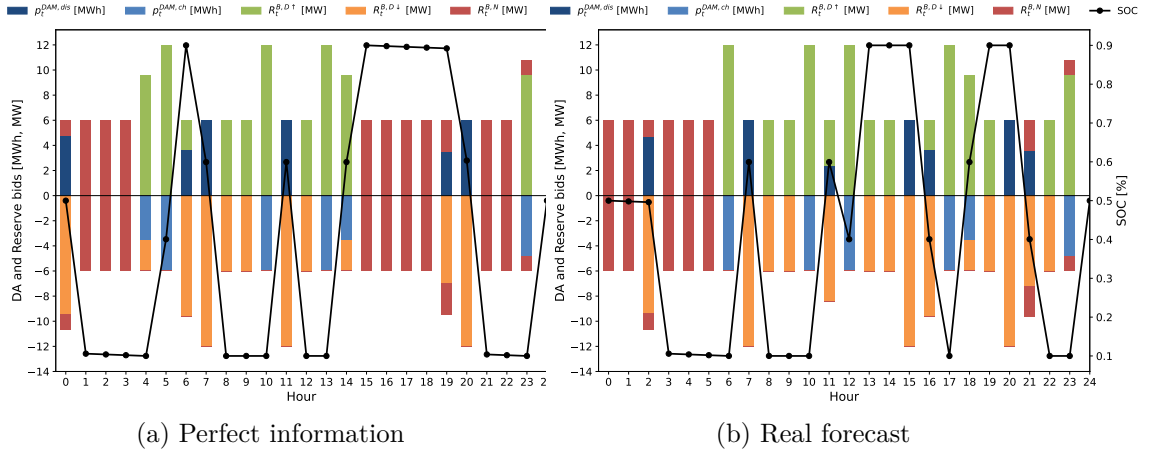
(a) Perfect information

(b) Real forecast

Figure A.12: BESS bidding with real forecasts and perfect information for test day 4

A.2 Scenario matrix of power system states

Hour	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
0	0	1	1	-1	0
1	0	0	0	0	1
2	1	1	1	-1	1
3	1	0	-1	0	1
4	-1	1	1	-1	0
5	1	0	-1	0	-1
6	1	0	-1	1	0
7	1	1	1	-1	1
8	0	0	1	1	-1
9	0	-1	-1	-1	-1
10	-1	1	-1	1	1
11	-1	-1	-1	-1	1
12	0	1	1	-1	1
13	-1	1	-1	1	1
14	-1	-1	-1	-1	1
15	0	-1	1	-1	1
16	1	1	1	-1	0
17	0	-1	1	1	1
18	1	-1	0	1	-1
19	0	0	1	1	0
20	-1	-1	1	1	1
21	-1	-1	-1	0	-1
22	1	0	-1	-1	-1
23	-1	1	1	-1	1

Table A.1: Power system state per scenario per hour (1 : Power deficit, 0: Balanced, -1: Power excess)