INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO
MESTRADO EM ENGENHARIA DE INTELIGÊNCIA ARTIFICIAL

Forecasting Day-Ahead Prices in the Iberian Electricity Market using a Multi-Agent Reinforcement Learning System in Real-Time

DAVID SILVA ARAÚJO
outubro de 2025

P.PORTO

Forecasting Day-Ahead Prices in the Iberian Electricity Market using a Multi-Agent Reinforcement Learning System

David Silva Araújo

1200609

Dissertation to obtain the Master's degree in Artificial Intelligence Engineering

Supervisor: Dr. Zita Vale, Full Professor, Institute of Engineering, Polytechnic of Porto

Co-supervisor: Dr. Gabriel Santos, Junior Researcher, Institute of Engineering, Polytechnic of Porto

Technical Supervisor: MSc Brígida Teixeira, Institute of Engineering, Polytechnic of Porto

Jury

President:

Dr. Luis Manuel da Silva Conceição, Assistant Professor, Institute of Engineering, Polytechnic of Porto

Vocals:

Dr. Pedro Nuno da Silva Faria, Assistant Researcher, Institute of Engineering, Polytechnic of Porto

Dr. Zita Maria Almeida do Vale, Full Professor, Institute of Engineering, Polytechnic of Porto

Porto, October 2025

Declaração de Integridade

Declaro ter conduzido este trabalho académico com integridade. Não plagiei ou apliquei qualquer forma de uso indevido de informações ou falsificação de resultados ao longo do processo que levou à sua elaboração. Declaro que o trabalho apresentado neste documento é original e de minha autoria, não tendo sido utilizado anteriormente para nenhum outro fim. Declaro ainda que tenho pleno conhecimento do Código de Conduta Ética do P. PORTO.

Abstract

The use of renewable energy sources in modern electrical grids introduces significant uncertainty due to the grid's instant response and unexpected fluctuations in generation. While crucial for sustainability, they require the development of optimized energy operations and designs that are complex and robust. Forecasting generation, consumption, or prices is a crucial aspect of an electrical system, being essential to maximise energy efficiency and support strategic planning for both suppliers and consumers. Therefore, this dissertation presents a multi-agent framework for forecasting daily Portuguese prices in the Iberian market to aid bidding proposals. It employs several forecasting models, ranging from statistical to machine learning. A reinforcement learning (RL) methodology is used to select the most appropriate forecasting model for each moment. Forecasting models and RL methodology were subjected to tuning tests to determine hyperparameters. The methodology was tested using real electricity production data and market prices from Portugal and Spain, between 2022 and 2024.

In the application of the RL methodology, the forecasting models were retrained monthly, and the RL methodology was continuously updated with actual prices and forecasts so that the optimal model was selected in each iteration. The forecasting models included in the dissertation are Ridge Regression, Stochastic Gradient Descent, Random Forest, Extreme Gradient Boosting, Deep Neural Networks, and Long Short-Term Memory. These models were integrated into a multi-agent system that manages requests from different users, making the system robust and scalable. It comprises a main agent that receives user requests, organises them, and allocates them to various tasks, creating a "Task Agent" for each task.

All case studies were conducted to evaluate the proposed solution and improve the efficiency of the RL methodology. The final case study simulates real-world conditions and demonstrated a 10% lower performance compared to the others. It highlighted the importance of daily features that are not available in time for bidding proposals. However, the final case study employed context, based on information from each renewable energy production source, which improves the RL model's performance by 3%. Overall, results indicated that the RL methodology surpassed individual forecasting models across all case studies, with gains between 4.66% and 35.71%. As the case studies progressed, it was observed that the average difference between predicted and actual market values was 0.56 €/MWh in the non-reality reflecting case study and 6.13 €/MWh in the (near) real-time scenario. It is further concluded that hydro and wind energy are the primary influences on Portuguese energy prices, with solar energy also playing a notable role. In contrast, nuclear energy in Spain has the least impact.

Keywords: day-ahead market, forecast, iberian electricity market, machine learning, market-clearing price, multi-agent systems, reinforcement learning, statistical models.

Resumo

A utilização de fontes de energia renovável nas redes elétricas modernas introduz uma incerteza significativa devido à resposta instantânea da rede e às flutuações inesperadas na geração. Embora cruciais para a sustentabilidade, exigem o desenvolvimento de operações e projetos energéticos otimizados, que sejam complexos e robustos. A previsão de geração, consumo ou preços é um aspeto crucial de um sistema elétrico, sendo essencial para maximizar a eficiência energética e apoiar o planeamento estratégico para fornecedores e consumidores. Assim sendo, esta dissertação apresenta uma estrutura multiagente para a previsão dos preços diários portugueses no mercado ibérico para auxiliar as propostas de concurso. Emprega vários modelos de previsão, que vão desde técnicas estatísticas até à aprendizagem automática. É utilizada uma metodologia de aprendizagem por reforço (AR) para selecionar o modelo de previsão mais apropriado para cada momento. Os modelos de previsão e a metodologia de AR foram submetidos a testes de ajuste para determinar os hiperparâmetros. A metodologia foi testada utilizando dados reais de produção de eletricidade e de preços de mercado de Portugal e Espanha, entre 2022 e 2024.

Na aplicação da metodologia AR, os modelos de previsão foram retreinados mensalmente, e a metodologia AR foi continuamente atualizada com os preços reais e previsões, de modo que o modelo ótimo fosse selecionado em cada iteração. Os modelos de previsão incluídos na dissertação são o Ridge Regression, Stochastic Gradient Descent, Random Forest, Extreme Gradient Boosting, Deep Neural Networks, e Long Short-Term Memory. Estes modelos foram integrados num sistema multiagente que gere pedidos de diferentes utilizadores, tornando o sistema robusto e escalável. Compreende um agente principal que recebe os pedidos dos utilizadores, organiza-os e aloca-os a várias tarefas, criando um "Task Agent" por cada tarefa.

Todos os casos de estudo foram conduzidos para avaliar a solução proposta e melhorar a eficiência da metodologia AR. O caso de estudo final simula condições do mundo real e demonstrou um desempenho 10% inferior em comparação com os outros. Salientou a importância dos recursos diários que não estão disponíveis a tempo para as propostas de concurso. No entanto, o caso de estudo final empregou contexto, com base em informação de cada fonte de produção de energia renovável, o que melhora o desempenho do modelo AR em 3%. No geral, os resultados indicaram que a metodologia AR superou os modelos de previsão individuais em todos os casos de estudo, com ganhos entre 4,66% e 35,71%. À medida que os casos de estudo foram progredindo, observou-se que a diferença média entre os valores de mercado previstos e reais foi de 0,56 €/MWh no caso de estudo que não reflete a realidade e de 6,13 €/MWh no cenário (quase) em tempo real. Conclui-se ainda que a energia hídrica e eólica são as principais influências nos preços da energia portuguesa, tendo a energia solar também um papel notável. Em contraste, a energia nuclear em Espanha tem o menor impacto.

Palavras-chave: aprendizagem automática, aprendizagem por reforço, mercado diário, mercado ibérico de eletricidade, modelos estatísticos, preço de equilíbrio de mercado, previsão, sistemas multiagentes.

Acknowledgements

I would like to thank everyone who contributed to the completion of this dissertation.

First, I would like to thank my supervisors, Professor Zita Vale, Doctor Gabriel Santos and Eng. Brigida Teixeira, for their guidance and dedication in helping me to complete this work.

Secondly, I would like to thank my family and girlfriend for the encouragement and motivation to achieve this great goal.

Table Of Contents

List of Figures

Lista of Tables

Acronyms and Symbols

List of Acronyms

A3C	Asynchronous Advantage Actor-Critic
AI	Artificial Intelligence
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
ARIMA	Auto Regressive Integrated Moving Average
CB-RL	Controlled Baseline RL
CNN	Convolutional Neural Network
CSV	Comma-Separated Values
DDPG	Deep Deterministic Policy Gradient
DL	Deep Learning
DNN	Deep Neural Network
DP-RL	Dynamic Penalised RL
DPC-RL	Dynamic Penalised Context RL
DQN	Deep Q-Network
DRL	Deep Reinforcement Learning
EM	Electricity Market
ES	Exponential Smoothing
ESS	Energy Storage System
EU	European Union
FIPA	Foundation for Intelligent Physical Agent
FCT	Foundation for Science and Technology
GECAD	Research Group on Intelligent Engineering and Computing for Innovation and Advanced Development

IoT	Internet of Things
JADE	Java Agent Development Framework
KNN	K Nearest Neighbours
LSTM	Long Short-Term Memory
MA	Moving Average
MAPE	Mean Absolute Percentage Error
MAS	Multi-Agent System
MBVE	Model-Based Value Estimation
ML	Machine Learning
MSE	Mean Squared Error
OMIE	Operador de Mercado Ibérico de Eletricidade
PDF	Portable Document Format
PEAK	Python-Based Framework for Heterogeneous Agent Communities
PES	Power and energy systems
PPO	Proximal Policy Optimisation
PRECISE	Power and Energy Cyber-Physical Solutions with Explainable Semantic Learning
QL	Q-Learning
QN	Q-Network
RDF	Resource Description Framework
REE	Red Elétrica Espanhola
REN	Redes Energéticas Nacionais
RES	Renewable energy sources
RF	Random Forest
RL	Reinforcement Learning
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network

RR	Ridge Regression
SAC	Soft Actor-Critic
SES	Simple Exponential Smoothing
SG	Smart Grid
SGD	Stochastic Gradient Descent
SPADE	Smart Python Agent Development Environment
SVM	Support Vector Machine
TLS	Transport Layer Security
TradeRES	New Markets Design & Models for 100% Renewable Power Systems
XGBoost	Extreme Gradient Boosting
XML	Extensible Markup Language
XMPP	Extensible Messaging and Presence Protocol

1 Introduction

This introductory section offers an overview of the dissertation, including the background of the work, the main research questions and objectives, scientific contributions, and the organisation of the document.

1.1 Contextualization

The increasing complexity of modern energy systems, combined with the widespread adoption of renewable energy sources (RES) such as solar and wind, requires intelligent and adaptable energy management, as these are sensitive to meteorological conditions (Kotzur et al. 2021). RES are vital for sustainability, but they present challenges because of their variability and unpredictability, which can obstruct precise energy forecasting and grid stability. Climate change, a global concern affecting natural systems and development, further threatens the availability of renewable based resources like water and wind (Fidalgo, de São José, and Silva 2019). Although Portugal and other European nations are more frequently adopting RES, the European Union (EU) still needs to address the broader issue of climate change. To overcome these challenges, it is necessary to determine optimal energy designs and operations, which is complex and requires robust planning methodologies.

The concept of smart grids (SGs) has emerged as a possible solution to address the variability of the RES (Kataray et al. 2023). SGs apply advanced technologies, including Internet of Things (IoT) devices, automation, and real-time communication, enhancing the efficiency, reliability, and flexibility of power and energy systems (PES). This supports the integration of decentralised RES and allows real-time adjustments to balance supply and demand. The use of SGs enables more efficient and sustainable management of energy systems, potentially reducing CO_2 emissions by up to 72% (Lamnatou, Chemisana, and Cristofari 2022).

SGs can integrate internal market mechanisms to develop platforms that facilitate more efficient electricity markets (EM) and help avoid the necessity of expanding the physical grid (Bagemihl et al. 2018). Recently, the EMs have undergone significant changes driven by the energy transition and decentralisation, including sources such as solar, wind, batteries, electric vehicles, and demand response. This can create vulnerabilities such as price volatility, intermittency of RES, and market regulatory uncertainty (Liu et al. 2025).

A critical component of EMs and SGs is the forecasting of energy-related matters such as energy price, consumption, or generation. Accurate forecasting is essential for the ongoing improvement of energy efficiency and helps suppliers and consumers with their planning (Ghazal 2022). Forecasting models are in demand for their performance and can be broadly classified into statistical and machine learning (ML) models (Schmid et al. 2025). Statistical models are based on deterministic principles and are suitable for linear and stable systems (Qian et al. 2017). However, they often have difficulty dealing with the dynamic and non-linear behaviour of RES. In contrast, ML-based models utilise large data sets to identify hidden patterns and trends, providing greater adaptability and performance in dynamic environments. In this context, the study conducted by Geraldo (Geraldo 2018) highlights the use of forecasting models in an EM to improve decision-making in both day-ahead and intra-day energy markets.

Intelligent systems are becoming increasingly autonomous, and the relevance of ML technologies is fundamental, such as in predicting consumption or energy prices. A study conducted by Ahmad *et al.* (Ahmad et al. 2022) indicates that intelligent automation using ML will result in substantial savings. The research presented by Tschora *et al.* (Tschora et al. 2022) employed various ML models, such as Extreme Gradient Boosting (XGBoost) and Long Short-Term Memory (LSTM), to predict electricity prices, particularly in the day-ahead market. The data utilised covers meteorology, RES, consumption, and prices. The authors conclude that the approach yielded promising results, emphasising the significant role of forecasting AI tools in a more unstable EM. AI has proven to be increasingly effective, with its high accuracy and learning capacity (Lipu et al. 2021).

Although statistical and ML methods are frequently used, selecting the appropriate forecasting model relies on factors like the nature of the problem, the availability of data, computing resources, and implementation needs (Ngo, Nguyen, and Van Nguyen 2023). To optimise the selection of forecasting models in dynamic and uncertain environments, Reinforcement Learning (RL) has emerged as a promising approach. RL allows systems to learn and adapt through interaction with their environment, using feedback mechanisms to improve decision-making over time (Radhamani et al. 2024). RL appears as a promising solution in EM, in price automation, effective decision-making for energy dispatch, and defining market participation strategies, among others (Zhu et al. 2023). For energy systems, RL can be utilised to optimise energy consumption by adapting to user preferences, resulting in cost savings and thereby promoting sustainability (Calegari et al. 2020).

Multi-Agent Systems (MAS) are another crucial technology for managing the high complexity of EMs and SGs (Rando Mazzarino et al. 2023). MAS comprises autonomous agents representing different network entities, such as generators, storage units, or consumers. MAS provides scalability and flexibility, allowing grid operators to manage thousands of distributed energy resources simultaneously. The complexity of EMs and SGs systems arises from integrating various specification areas and collaborating agents. The study conducted by Rando Mazzarino *et al.* (Du et al. 2021) emphasises the importance of coordination and reliable interactions between agents. The authors Samadi *et al.*(Samadi, Badri, and Ebrahimpour 2020), considered that a MAS with RL methodology could be very beneficial when properly developed to optimise energy systems. Another study carried out by Shen *et al.* (Shen et al. 2022) presents a framework, built on MAS and RL technologies, aiming to optimise energy management. The results were promising, showing improvements over traditional methods, including a 43% reduction in unutilized renewable energy and an 8% decrease in energy costs.

1.2 Problem Statement

Forecasting in PES is expanding due to the need for precise forecasting of renewable consumption, generation, and energy prices, driven by the uncertainties associated with renewable generation (Zhang et al. 2022). The accelerated integration of RES, such as solar and wind, and the increase in active users, such as producers and consumers, have made electrical systems more dynamic and complex. This evolution has created some new challenges for the efficient management of resources, stability in electrical networks, and the ability to respond in real-time to variations (Agouzoul and Simeu 2024).

Traditional forecasting methods struggle to handle the volume and velocity of dynamic data, whereas modern electrical systems face significant uncertainty caused by the intermittency of RES and demand fluctuations, among other factors (Ibrahim, Dong, and Yang 2020). Most existing solutions do not utilise a sufficient variety of forecasting models, and this could be limited, as the inclusion of more models is beneficial for better adapting to different problems (Aslam et al. 2021). Another limitation is the absence of forecasting models integrated with an RL model for continuous learning. A limitation in existing solutions is their struggle with scalability and parallelism, which makes it difficult to serve multiple requests simultaneously and operate efficiently in systems requiring high processing speeds (Rane et al. 2024). Future solutions need to be more robust and adaptable.

Implementing RL and MAS remains challenging in current systems, due to difficulties in deploying these methods on a large scale, which requires significant computational power (Bui et al. 2024). Other challenges involve implementing the RL methodology in uncertain environments where data fluctuates, such as EMs. For successful implementation, it is crucial to first make the data as robust and clear as possible, then choose the most appropriate RL approach. MAS-based approaches promote decentralised communication, increasing robustness and scalability. When combined with RL, MAS allows agents to learn cooperative strategies, even in complex environments, improving coordination and adapting to environmental uncertainties (Park and Moon 2022).

1.3 Research Question and Objectives

This dissertation aims to address these gaps through the development of an innovative forecasting MAS. The system utilises advanced statistical models, ML techniques, and RL to generate accurate forecasting in (near) real-time. The approach is based on a main agent responsible for coordinating operations and responding to requests. This main agent can create and delegate specific tasks to other agents, optimising processing and enabling a high level of parallelism. This allows the efficient management of multiple requests simultaneously without compromising accuracy or response speed.

In addition to the multi-agent architecture, the system incorporates various advanced features to enhance flexibility and efficiency. These include the integration of administrative tools for system management and monitoring, the ability to train forecasting models offline to continually improve performance, and an interactive interface that supports engagement with human users. The system allows users to train various models, make forecasts, and add new models to test new methodologies. These characteristics make the system highly adaptable to the needs of energy managers and sufficiently robust to operate in complex scenarios.

This dissertation aims to forecast Iberian market prices to support bid proposals. It provides an accurate response in (near) real-time, essentially in the areas of EMs and energy resource management. To ensure the successful completion of this dissertation, research questions and objectives were outlined to clarify the various processes involved.

Within the scope of this work, a main question and three sub-questions arose, which are addressed in this dissertation. The questions are described below:

- RQ1 - How can multi-agent reinforcement learning improve machine learning or statistical model selection when performing (near) real-time forecasting and/or classifications within electrical power systems?
  - SQ1 - How to combine statistical methods and machine learning algorithms to deal with uncertainties in energy data?
  - SQ2 - What reinforcement learning techniques can be applied to select the best forecasting model?
  - SQ3 - How can multi-agent systems be applied with reinforcement learning to provide forecasting services in (near) real-time within electrical power systems for heterogeneous software agents?

The main objective of this dissertation is to design, realise, test, and validate a solution for robust and effective modelling. To facilitate this main objective and answer all proposed research questions, several specific objectives were defined for this work, as described below:

- O1 - Collection, study, and analysis of the state of the art;
- O2 - Definition of the requirements of the system to be developed;
- O3 - Data collection and preparation;
- O4 - Selection and use of Statistical, Machine Learning, and Deep Learning techniques;

- O5 - Selection and use of reinforcement learning strategies to reward the best model;
- O6 - Design the interactions for each agent, including an administrative interface for system management and offline model training;
- O7 - Development and implementation of unit, integration, and system tests;
- O8 - Design and execution of case studies to apply the solution in real-time;
- O9 - Writing the dissertation document.

In order to assist in comprehending the research questions and the defined objectives, Table 1 shows the correspondence between them.

Table 1 - Research questions and corresponding objectives to support the proposed solution.

Research Questions	Corresponding Objectives
RQ1	O1, O2, O3, O4, O5, O6, O7, O8, O9
SQ1	O1, O2, O3, O4, O7, O8, O9
SQ2	O1, O2, O3, O5, O7, O8, O9
SQ3	O1, O2, O3, O6, O7, O8, O9

1.4 Scientific Contributions

This dissertation was developed within the Research Group on Intelligent Engineering and Computing for Innovation and Advanced Development (GECAD)[1], which focuses on two primary areas: Intelligent Systems and Power and Energy Systems. Its mission is to promote and advance scientific research in the fields of Knowledge and Decision Sciences, supported by Information Technologies. The development of this work was supported by two R&D projects funded by the Portuguese Foundation for Science and Technology (FCT):

- PRECISE[2] - Power and Energy Cyber-Physical Solutions with Explainable Semantic Learning (PTDC/EEI-EEE/6277/2020).
- TradeRES[3] - New Markets Design & Models for 100% Renewable Power Systems.

During the development of this dissertation, scientific work was conducted, resulting in a scientific publication in conference proceedings, one project deliverable, and one journal paper that is currently under preparation:

- [Conference - Published] **Araújo, D.**, Santos, G., Teixeira, B., Faria, P., Vale, Z. (2026). Reinforcement Learning for Real-Time Price Forecasting in Energy Systems. In: Valente de Oliveira, J., Leite, J., Rodrigues, J., Dias, J., Cardoso, P. (eds) Progress in Artificial Intelligence. EPIA 2025. Lecture Notes in Computer Science(), vol 16122. Springer, Cham. https://doi.org/10.1007/978-3-032-05179-0_34

[1] https://www.gecad.isep.ipp.pt/
[2] https://www.gecad.isep.ipp.pt/precise/
[3] https://traderes.eu/

- [public deliverable E1.1 of EnergyP2P[4] project] Luis Gomes, Pedro Rodrigues, Bruna Azevedo, **David Araújo**, Rúben Barreto, Fernando Lezama, "Relatório do estado da arte, identificação dos aspetos diferenciadores e possíveis riscos associados ao projeto"
- [Journal - to be submitted] **David Araújo**, Gabriel Santos, Brígida Teixeira, Pedro Faria and Zita Vale, "Reinforcement Learning for Clearing Price Forecasting in the Iberian Market"

1.5 Document Organisation

This dissertation is divided into six chapters, each carefully written to guide the reader clearly and coherently.

The first chapter (the present chapter) covers the introduction of this dissertation, starting with the context and motivation for this research. It then outlines the research questions and sub-questions and the document structure.

Chapter two examines the current state of the art by investigating, exposing, and describing existing sources of knowledge related to the various domains discussed in this dissertation. These domains include EMs, energy forecasting models, RL, and MAS. This chapter is crucial for understanding previous advancements and identifying gaps at a theoretical level.

The third chapter explains the methods and tools used, as well as how the different processes of the system interact. It also describes the data used in the various case studies of this dissertation, along with considerations about data protection, security, and ethics, ensuring a robust system at every level.

The Experimentation carried out in this dissertation is outlined in chapter four, describing the processes applied for the case studies. The results from the various case studies are presented, analysed, and discussed.

The fifth and final chapter provides a comprehensive conclusion of the various objectives defined, as well as the main findings derived from the results in this dissertation. Finally, it presents limitations and suggestions for potential future developments and improvements.

[4] https://www.gecad.isep.ipp.pt/portfolio/_energyp2p/

2 State of the Art

This chapter presents an overview and discussion of the various themes covered in this dissertation, such as Iberian electricity markets, energy forecasting models, reinforcement learning, and multi-agent systems. In addition to examining these topics individually, it will also explore how these technologies behave within EM, with a particular focus on forecasting. Finally, the chapter concludes by highlighting the gaps in the existing literature on these domains.

2.1 Iberian Electricity Markets

The electricity market is an institutional and economic system that arranges energy transactions between producers, traders, and consumers (Wu et al. 2024). It utilises auction platforms and market tools designed to maximise allocative efficiency and ensure security of supply. The Iberian Electricity Market, which is the Operador del Mercado Ibérico de Eletricidade (OMIE)[5], integrates the electricity markets of Portugal and Spain (OMIE 2020). OMIE includes both daily and intraday markets, which influence the market price, providing a transparent public pricing environment. While Portugal and Spain may share the same hourly price, congestion in the interconnections can cause price differences across zones. OMIE offers various essential data for market analysis, including hourly prices, total energy purchased, supply and demand curves, and technology negotiations (OMIE 2025). It also includes RES such as hydro, nuclear, thermal, wind, and others.

The bidding process is the basic unit of the commercial strategy (Roldan-Fernandez et al. 2024). It specifies the quantity, price, and date details of the energy being offered. Bids are typically submitted for each hour of the following day, with the market applying a uniform price rule. The highest accepted offer establishes the market price for every specific hour. This marginal

[5] https://www.omie.es/pt/

price determines the cost for all energy volumes accepted during that hour. With the bidding strategies, the ongoing integration and increase of RES, and the uncertainty in management due to the non-storability of energy, have a direct impact on the volatility of hourly prices.

OMIE manages the matching platform and publishes market results, working alongside the System Operators, Redes Energéticas Nacionais (REN)[6] in Portugal and Red Elétrica Espanhola (REE)[7] in Spain (OMIE 2025). REN and REE ensure that the transmission network operates in real-time, maintaining a physical balance between generation and consumption, and providing vital network services.

REN is responsible for the overall management of Portugal's electricity system (REN 2025). Its duties include planning, constructing, operating, and maintaining the country's energy transmission networks to ensure a reliable and safe supply. Additionally, REN actively supports the integration of RES, which are crucial for a sustainable environment (REN Data Hub 2025). The information available on REN's website mainly covers electricity generation from sources like natural gas, water, wind, solar, biomass, and others.

REE operates Spain's electricity transmission system, managing the entire high-voltage grid in real-time (REE 2025). Similar to REN, REE ensures supply continuity and security by balancing demand and generation, facilitating RES integration. REE's data collection is comparable to REN's, with the addition of nuclear energy, unique to Spain (Redeia 2025). From REN's and REE's websites, it is possible to monitor verified prices and the forecasts for upcoming periods.

The constant and rapid evolution of various technologies impacts many areas, and the modernisation of grids is no different (Khalid 2024). There is a need to continually modernise all the electrical grids into a sustainable one that incorporates RES and meets the growing demands for flexibility and reliability. SG is a contemporary energy system that combines digital technology with real-time communication. Its main goal is to enhance energy efficiency, boosting sustainability and safety across energy production, distribution, and consumption. SGs offer smart, real-time measurement of multiple electricity parameters, incorporating RES.

SGs employs bidirectional communication between the generator and the consumer and vice versa, helping to create a dynamic and interactive ecosystem (Kataray et al. 2023). This allows for better energy management by supporting the use of renewable energy sources like solar and wind. Consumers can also become producers, encouraging energy sharing and increased user autonomy. Additionally, the systems are automated, capable of predicting failures or overloads in the network and enabling quick responses. Nevertheless, challenges such as cybersecurity, data protection, large-scale data handling, and forecasting remain unresolved issues that lack robust, compact solutions.

A key advantage of SG is its capacity to incorporate different local market mechanisms, enabling the creation of decentralised platforms and supporting EMs (Adak, Bhattacharyya, and Barshilia

[6] https://www.ren.pt/pt-pt
[7] https://www.ree.es/en

2022). This enhances the efficiency of these markets by allowing consumers, producers, and prosumers to trade energy locally in real-time. It is especially advantageous for more remote geographic areas, reducing dependency on national wholesale markets such as OMIE. This approach minimises consumer losses, provides prosumers with quick financial gains, and helps improve grid stability and response times.

Forecasting models are employed to handle the volatility introduced in EMs by analysing fluctuations (Lago et al. 2021). They are vital for the efficient operation of the EM, assisting operators in balancing supply and demand. These models help optimise energy production, preventing overgeneration. Additionally, they provide consumers with tariff recommendations and demand response options. Ensuring that this data is robust and clean is crucial for the forecasting models to deliver more accurate and reliable results.

The adoption of forecasting models in EMs has been shown to be highly effective, resulting in roughly a 20% reduction in electricity consumption and about a 9% improvement in price forecasting accuracy (Yousaf et al. 2021). The recent and impactful use of Artificial Intelligence (AI) models is essential in helping to predict and optimise energy needs in real-time. The upcoming section will cover the importance of forecasting models.

2.2 Energy Forecasting Models

RES are increasingly important, with the use of effective forecasts being essential, increasing the reliability of electrical networks, reducing costs and efficient energy management (Alkhayat and Mehmood 2021). Alkhayat and Mehmood's study, addresses several methods used in RES forecasting: statistical and ML. The special focus in this study is on Deep Learning (DL), demonstrating its growing popularity, due to its ability to identify non-linear relationships. Within DL, Recurrent Neural Network (RNN), and CNN are the most used. However, there is a need to improve the accuracy of forecasts related to RES.

The study of Ahmad *et al.* (Ahmad, Zhang, and Yan 2020) emphasises forecasting at short, medium, and long-term intervals, which is crucial for aggregating energy needs across different time horizons. The authors addressed ML and Artificial Neural Network (ANN) methods in forecasting solar radiation and wind in different regions, such as cities and provinces, among others. For all forecasting methods to deliver accurate forecasting, it's essential to acknowledge the various data types they can process, as demonstrated in Figure 1 (Bhaskar 2023).

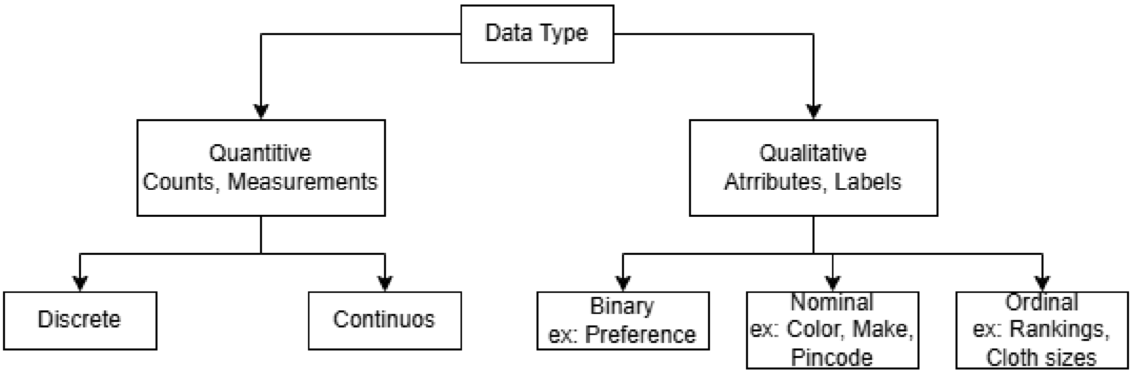

Figure 1 - Possible data type for forecasting models, adapted from (Bhaskar 2023).

Figure 1 represents two types of data: quantitative, being data that can be measured or counted; and qualitative, being categorical or descriptive data. The quantitative is divided into two parts, the discrete, which generally represents whole numbers, and the continuous which represents a continuous scale, such as height or weight. The qualitative is divided into three categories: binary, nominal, and ordinal. Binary means two distinct categories, (e.g., yes or no). The nominal classifies data types, such as colours or service types. The ordinal represents an order or hierarchy but not in numerical form, such as satisfaction levels.

The upcoming subsections will offer a detailed overview of forecasting models, focusing specifically on statistical and ML approaches, and will also review relevant related work.

2.2.1 Statistics Models

Statistical models are characterised by mathematical tools that interpret relationships between variables, based on historical data (Nti et al. 2022). These are methods with applications in regression, time series analysis and classification methods. There are various types of regression, such as the main ones: simple or multiple linear and non-linear and polynomial. Simple linear regression is based on a simple Eq. (1).

$$y = \mathrm{mz} + \mathrm{c} \tag{1}$$

In Eq. (1), the variable y represents what is being predicted, and m denotes the angular coefficient, indicating the extent to which y changes in relation to z. The input variable is z, and c serves as the linear coefficient. Multiple linear regression includes several independent variables, having the ability to relate in a more complex way. Nonlinear and polynomial regression, as the name suggests, considers nonlinear relationships. Time series models deal with data ordered over time, such as the Auto Regressive Integrated Moving Average (ARIMA), which uses data that has trends and seasonality.

Figure 2 represents the mathematical modelling cycle of a statistical model, where it understands, analyses and communicates the results (Turmudi et al. 2017).

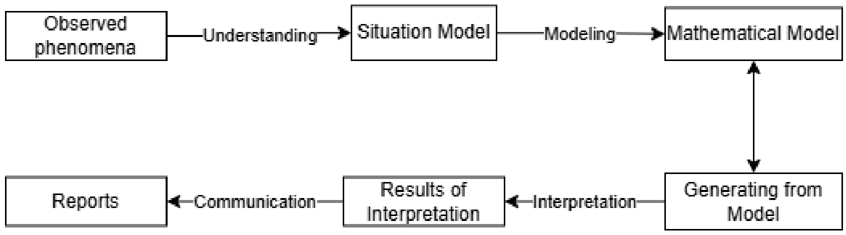

Figure 2 - Mathematical model cycle, adapted from (Turmudi et al. 2017).

Figure 2 demonstrates that the process begins with the understanding of phenomena, leading to the creation of a situational model, which translates into a mathematical model. This can generate results, which are interpreted and compared to real values. If necessary, adjustments are made in an iterative cycle. Finally, the results are communicated.

Other important statistical models for comparing methods include Analysis of Variance (ANOVA), Krustal-Wallis, and post-hoc test methods, which are essential for identifying differences between the performances of various algorithms (Rodriguez, Sued, and Valdora 2025). Hierarchical models, such as test functions, dimensions and instances, return a more structured analysis of variations over time.

One of the advantages of using these statistical models is their generalisation capacity, robustness in the analysis of complex data, and being able to deal with multiple variables (Moroff, Kurt, and Kamphues 2021). These models provide reliable metrics, enabling the evaluation of result significance and facilitating comparisons between methods or configurations. They are able to address a wide range of problems and data types, supporting evidence-based decisions. However, there are disadvantages, such as excessive simplicity, as when the complexity of the data increases, it becomes more challenging to capture it. These models require strong data, demonstrating normality or independence between the variables. When this does not happen, these models are not successful. Another disadvantage is the sensitivity to outliers and noise, which can distort the results.

Statistical models need historical data with well-defined patterns as demonstrated by the methods used to predict energy consumption, in the study conducted by Deb *et al.* (Deb et al. 2017). One of them was time series analysis, which arranges values measured at regular intervals and consists of three main components: the trend over time (excluding seasonality), the seasonal fluctuations, and the residual variability. The other one was Moving Averages (MA), which smooth out fluctuations by averaging previous values. Another method is Exponential Smoothing (ES), which assigns more weight to recent values, making it responsive to changes. Lastly, there is ARIMA, which combines autoregression and integration.

Another article (Verdejo et al. 2017), reveals the various benefits of statistical methods, which are essential for accurate forecasting through identifying models and patterns in historical data. Time series analysis enables the detection of complex data features via decomposition, revealing peak times and seasonal trends. The straightforward application of these models reduces costs for users, as they are simple to implement and require fewer computational resources. Consequently, this facilitates efficient planning and energy optimisation.

2.2.2 Machine learning-based Models

ML is a sub-area of AI focused on creating algorithms that learn from various types of historical data and improve their forecasting over time (Balan, Kumar, and Raj 2025). ML models can automatically make decisions by recognising patterns in datasets, all without human intervention. Using ML is more effective when performing feature selection, tuning and optimising models trained on the data (Sharma and Mishra 2022). ML is an essential tool for dealing with incomplete data, meaning that knowledge on the subject may not be robust. With the rapid growth of various technologies and massive amounts of data, the application of ML continues to expand, bringing new possibilities for real-time accuracy. Several fields have applied this emerging technology, such as bioinformatics, cybersecurity, marketing, EMs and SGs.

ML is divided into supervised learning, unsupervised learning and RL (Ng et al. 2023). Supervised learning is applicable when the data is clear, connected, and easy to interpret. While the initial training is crucial and relies on robust data, subsequent training sessions of the model may involve incomplete and less reliable data. The different supervised models, such as Support Vector Machine (SVM), Naive Bayes and Logistic Regression, can be directly applied for the classification and forecasting of energy prices (Makkar et al. 2022). Unsupervised learning explores unknown patterns in large volumes of data. For this task to be successful, the ML model uses dimensionality reduction and clustering techniques, which are fundamental in understanding the various data. Unsupervised models such as K-Means or hierarchical clustering are typically used to discover hidden patterns in incomplete data, which is crucial for identifying abnormal patterns in the energy sector. The RL applies learning algorithms to optimise decisions through interaction with the environment, and it will be specifically detailed in the Reinforcement Learning section. Figure 3 (Kumar 2023) illustrates the division of ML.

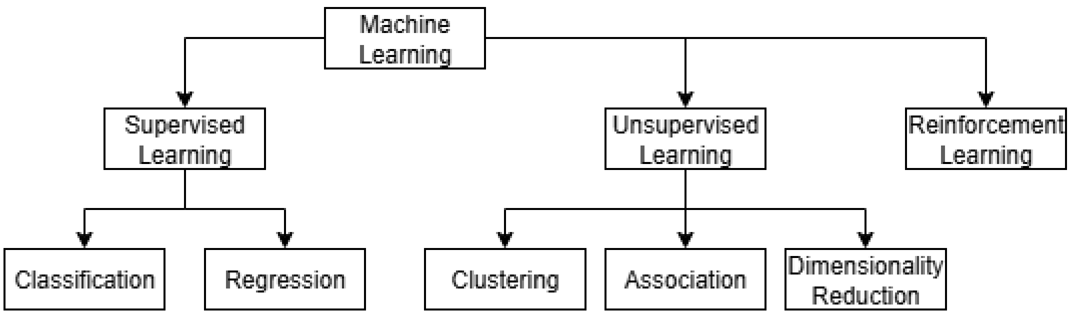

Figure 3 - Supervised and unsupervised learning, adapted from (Kumar 2023).

As shown in Figure 3, supervised learning is applied in regression and classification tasks, whereas unsupervised learning includes clustering, association, and dimensionality reduction. Focusing now on the supervised learning methodology, it is essential to utilise a set of data or features to perform the forecasting, where the output constitutes the target column (Jiang, Gradus, and Rosellini 2020). During training, these types of models learn to map/associate input values with output values, for each specific case.

The most used algorithms dedicated to Supervised Learning are described as follows (Sarker 2021) (Chiang, Rao, and Vannucci 2024):

- Logistic Regression (Rajendra and Latifi 2021): probabilistic model, which uses L1 and L2 regularizations as hyperparameters to prevent overfitting. It is used to estimate the probability of classes;

- SVM (Chhajer, Shah, and Kshirsagar 2022): creates hyperplanes to differentiate high- or infinite-dimensional data, applicable in both classification and regression, especially effective with high-dimensional datasets;

- Decision Tree (Kappelhof et al. 2021): classifies instances using entropy or Gini index while descending the tree. It is applicable in both classification and regression tasks, which makes it easy to implement and suitable for large datasets;

- Random Forest (RF) (Nachouki et al. 2023): uses multiple decision trees at once to enhance accuracy. It can be applied to classification and regression tasks involving continuous and categorical data. By reducing overfitting in single decision trees, it is well-suited for handling large volumes of data;

- XGBoost (Tarwidi et al. 2023): uses gradients to minimise loss functions in ensemble tree models. It reduces overfitting through regularisation, effectively handling large and complex datasets;

- Stochastic Gradient Descent (SGD) (Dagal et al. 2025): is an iterative approach to text classification and natural language processing, applied to large-scale data datasets;

- K Nearest Neighbors (KNN) (Kaneko 2023): is a classification method that predicts new data points based on the proximity of neighbouring points. It performs well with small datasets data;

- Naive Bayes (Oddleifson et al. 2025): this algorithm uses Bayes' theorem to classify data, assuming that features are independent. It needs minimal training data, making it fast and efficient, but it may overly impose the assumption of independence characteristics.

The emergence of DL models revolutionised the entire concept of AI, including the introduction of the Deep Neural Network (DNN) (Morucci 2024). DNN improved the concepts of Supervised Learning, where it minimises the difference between the actual and predicted values, using parameterised functions with weights. Supervised Learning has a central objective, which is to minimise its loss function to improve the model's forecasting. DNNs can adapt to lowdimensional data, even when many parameters are involved in weights. Furthermore, DNNs exploit non-linear and complex structures to obtain greater forecasting capacity.

Unsupervised Learning does not need previously defined and labelled data, operating even with disorganised data (Vladimir 2017). These algorithms identify similarities in data and categorise them into groups called clusters, taking into account the intrinsic properties of the data. Clusters are formed, but there is no prior knowledge about them, which allows unexpected discoveries that may not be recognised in supervised models. Cluster analysis is one of the main references of Unsupervised Learning, being relevant in discovering complex relationships in data. Clusters classify items into groups depending on similarities (Naeem et al. 2023). Clusters can be divided in various ways, such as partitioning, in which each item belongs to only one cluster, using the K-Means method as an example. Hierarchical division starts with a separate cluster and iteratively joins nearby clusters until a desired cluster number is reached. Overlapping occurs when an item can belong to multiple clusters, with the probabilistic aspect dividing clusters based on the likelihood of membership data. Figure 4 illustrates how clustering works (Naeem et al. 2023).

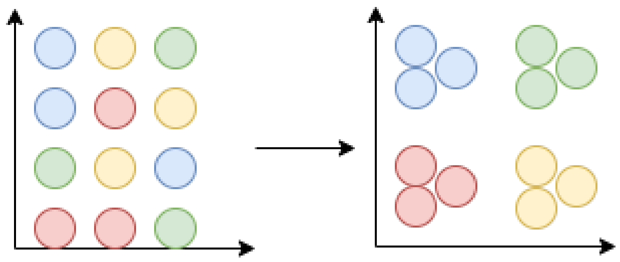

Figure 4 - Example of clustering, adapted from (Naeem et al. 2023).

Figure 4 presents several data points, which represent an item (Naeem et al. 2023). Each item will be assigned to one or more clusters based on their similarities. By analysing their characteristics, clustering algorithms can detect common patterns and form distinct groups. Each group showcases a set of similar properties among its members.

In many dynamic environments, the available data is non-linear or unlabelled, such as in measurements of energy consumption. In this way, Unsupervised Learning plays an important role in discovering hidden patterns (Alanne and Sierla 2022). Clustering and probabilistic models are great for identifying an anomaly in the data, indicating a loss of energy efficiency or even fraud in smart meters. Furthermore, this process can be used to detect intrusions or system security attacks. Clustering is important in grouping consumers, depending on their characteristics, to lead to more personalised forecasting. The dimensionality reduction offered by Unsupervised Learning makes it easier to analyse extensive data generated by sensors, among others.

2.2.3 Related work

This subsection presents several recent works that use various forecasting models, including statistical, ML, and DL. These studies essentially focus on their potential impact in EM.

Statistical models were employed in a study forecasting electricity prices for daily and intraday markets at the Iberian level (Torres, Martínez-Álvarez, and Troncoso 2022). The research analysed historical data from the Iberian market spanning from June 12, 2015, to June 13, 2017. The dataset included variables such as date, period, wind and solar generation, load, among others. This study utilised Gradient Boosting Trees (GBT) and Linear Quantile (LQ) models, one focusing on decision trees and the other on probability forecasting. To evaluate the GBT model, metrics like Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and MAPE were used, yielding results of 3.03 €/MWh, 4.04 €/MWh, and 8.54%, respectively. The LQ model was assessed using the Continuous Ranked Probability Score, which indicated a 1.22% error. In general, lower metric values signify better performance of the tested model.

Another study by Torres *et al.* (Torres et al. 2022) aimed to predict energy consumption in the Spanish electricity market using an LSTM model, which is designed to capture long-term dependencies in time series data. The model shows a strong capacity to manage complex patterns in the electricity market, although training takes considerable time and requires extensive reliable data resources.

A study by Uniejewski *et al.* (Uniejewski 2024) employed the Ridge Regression (RR) model with L2 regularisation as a key hyperparameter for predicting prices in the day-ahead markets of Spain and Germany. The authors found that RR performs well when many variables are interrelated, demonstrating stability and speed. Additionally, it decreased the forecast error relative to actual market values by as much as 10%.

In EM, there has been an emergence and increased importance of using forecasting models such as ML for accurate forecasting, as shown in the study carried out by Mishra *et al.* (Mishra et al. 2024). The authors examined the effects of forecasting electricity prices in the UK market, with real-world applications. The authors utilise ML models such as XGBoost, along with DL models like LSTM and RNN. The dataset included actual energy consumption, gas, and electricity prices, all recorded at 30-minute intervals. Findings demonstrated that DL models outperformed other approaches and highlighted the importance of selecting the appropriate forecasting model based on the data type. Another study conducted by *Kizildağ, Abut, and Akay* (Kızıldağ, Abut, and Akay 2024) used the RF and XGBoost to forecast electricity prices in the Greek market, utilising nine years of data. After feature selection, there was a 5% accuracy improvement. They highlighted the high accuracy of these models, despite the increased computational effort involved.

In a study conducted by *Cardo-Miota, Pérez, and Beltran* (Cardo-Miota, Pérez, and Beltran 2023), various models, such as LSTM and DNN, were tested to assess their accuracy in predicting Spanish prices in the Iberian electricity market. Each model used data on energy production, consumption, temperature, demand, and historical Spanish prices. Results showed that the LSTM model outperformed the others, whereas the DNN was less effective in this context. Another study carried out by *Nascimento, Pinto, and Vale* (Nascimento, Pinto, and Vale 2019), used an ANN model that was dynamically adapted to the selected data, which was applied to daily Iberian day-ahead market price forecasts. This data included day-ahead prices, consumption, generation, and climatology, which underwent a correlation test to identify the most relevant features. The authors highlight the significance of accurate feature selection to minimise noise, resulting in improved model performance.

A study conducted by Ramos *et al.* (Ramos et al. 2021) examined the forecasting of electrical consumption in building rooms to enhance energy management systems and demand response programs. It compared various data architectures and forecasting models, utilising sensor data like consumption and temperature. The methods primarily involved ANN and KNN, focusing on error analysis at different times of the day. Findings showed that customising hyperparameter choices and model configurations to align with daily consumption patterns is essential.

A study by Kumar and Singh (Kumar and Singh 2023) addressed demand and response management, that is, finding a balance between energy generation and consumption in real-time. To address this topic, ML methods were used, specifically Long Short-Term Memory (LSTM) and RNN to perform load forecasting in real-time. The study demonstrated the importance of short, medium, and long-term forecasts, creating a more resilient system and reducing unnecessary costs. LSTM-RNN can address non-linearity in data, being accurate and of medium complexity, and being economical at the computational processing level. The results demonstrated the ability to predict future scenarios.

In the study carried out by Sundaram *et al.* (Sundaram et al. 2024), the LSTM model is used as part of DL which has proven to be very beneficial in practice. LSTM models temporal dependencies and captures complex long-term patterns, crucial in dynamic environments. The authors focus on achieving real-time accuracy and precision in predicting electrical grid stability by utilising data from energy producers and consumers. For comparison, other ML algorithms were tested, and LSTM achieved a higher accuracy of 96%.

In another study, Ribeiro *et al.* (Ribeiro et al. 2024) utilised DL models to forecast energy consumption in SGs. The objective was to predict short-term consumption using the DNN method, tested for various hyperparameters. The model was tested using real-world data, including historical load records, weather data, and others. The authors underlined the importance of DL methods in environments with fluctuating consumption. They also highlight the capability of these models, compared to conventional ones, being more effective in supporting agents' participation in the EM.

2.3 Reinforcement Learning

RL is a subarea of ML, emphasising how agents make sequential decisions in dynamic environments (Ernst, Damien, and Arthur 2024). The main objective of an RL-based model is to maximise rewards over time. RL is crucial in helping decision-making in various systems and various areas, such as EM. RL has the ability to learn independently, without the help of explicit models of the environment, dealing with uncertainties and non-linear values. One of the great benefits is real-time operation in decentralised and interconnected systems. This model can solve problems such as balancing supply and demand, managing renewable resources, optimising cost, efficiency, and predicting energy prices, among others.

There are fundamental principles of RL, such as a decision-making entity, an environment, a state, an action, and a reward (Kanade 2022). Although this entity is referred to as an agent, the RL method can be applied even in the absence of an identified agent. The RL agent is important in making the decision considering all inputs. The environment is where the RL agent executes its actions and learns from them. The State demonstrates the condition of the environment at a given moment. Action means the decision made by the RL agent. The reward is the feedback that the RL agent receives after his decision and then can understand whether its decision was correct. The policy demonstrates the strategy that the RL agent uses based on its previous choices. Figure 5 illustrates how this process works.

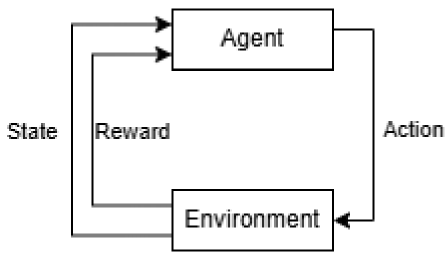

Figure 5 - RL process, adapted from (Kanade 2022).

In RL, there are two types of models, Model-Free and Model-Based (Wang et al. 2025). The Model-Free uses the environment explicitly, so it learns policies or estimates from interactions. They are suitable for systems with high uncertainty and interpretation complexity, being flexible and simple in complicated interaction environments, but they require more interactions with the environment to develop their knowledge and learning. Model-Based already builds a model closely related to the environment, also using it to simulate future interactions. These models are ideal for well-structured systems with a large amount of robust data, as they do not require as many interactions with the environment to build their own. However, it requires more demanding computational terms. Figure 6 (Achiam 2018) presents some of the most common RL models.

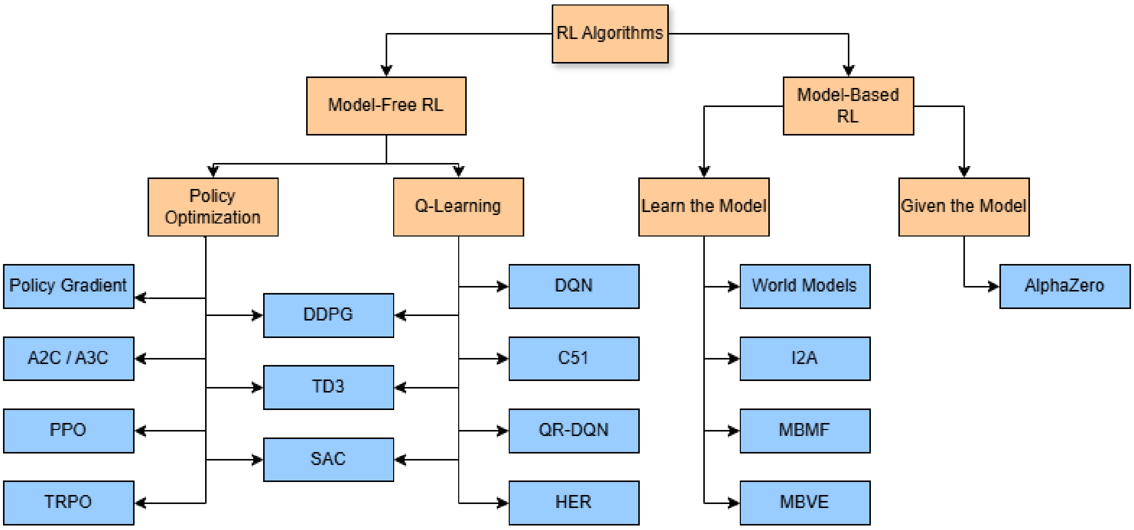

Figure 6 - Taxonomy of RL algorithms, adapted from (Achiam 2018).

There are currently many fundamental and advanced RL algorithms, such as Q-Network (QN) and Deep Q-Network (DQN), Policy Gradient methods and Deep Deterministic Policy Gradient (DDPG), among others (Shakya, Pillai, and Chakrabarty 2023). QN is a table-based algorithm with the aim of solving problems with discrete states, such as the basic control of devices in electrical networks. It is useful in smaller systems and is easily implemented, having a lower computational cost. However, its scalability is limited in large systems. DQN replaces the tables that the QN model uses with DNN, thus managing to deal with more complex problems and offering robustness in continuous problems. Policy Gradient Methods comprise two methods: Proximal Policy Optimisation (PPO) and Asynchronous Advantage Actor-Critic (A3C). These methods directly learn policies to optimise their control, which is most beneficial to use in systems with continuous actions. DDPG is generally used in environments with continuous actions and high complexity, managing to be efficient with lower consumption of computational resources.

One of the challenges in RL focuses on Exploration vs Exploitation (Amit 2024). In the case of exploration, the agent tries new actions in order to receive better rewards. In contrast, in the case of exploitation, the agent uses its knowledge to the fullest to try to obtain greater immediate rewards. However, in exploration, the agent can waste time and overuse resources in an attempt to get better rewards. In the case of exploitation, the agent runs the risk of becoming obsessed with suboptimal options, without ever opening his horizons.

There is a strategy to balance this dilemma, called Epsilon-Greedy, which is one of the approaches in which the agent has a variable based on probability, which is assumed to be 80% (Wei 2020). In other words, 80% of the time, the agent explores, and the remaining 20% the agent takes a random action. The Softmax (Tan et al. 2024) is used to convert numeric values into probabilities within models, making it easier to select classes. It also employs probabilistic methods based on these estimated values, where actions with higher scores are more likely to be chosen.

2.3.1 RL methods

There are several RL models with crucial impact on the dynamic environments. One of them is the RL Bayesian Theorem (Sousa et al. 2014), which selects different forecasting models based on a probabilistic methodology. Each forecasting model is represented by the idea of probabilities, representing success or failure in the range [0,1]. There are two hyperparameters: α and β, representing successes and failures, respectively, which are updated throughout each period of the day. β is used to determine which forecasting model to choose in each period, because if this value is low, it means that the model performs well. This allows a balance between exploration/exploitation, based on the statistical accuracy of each forecast model. Therefore, less-used forecasting methods have a chance of being chosen, avoiding suboptimal convergences.

The Roth-Erev-based RL (Sousa et al. 2014) was developed to explain how RL agents behave within a specific environment, initially in repeated games. The model emphasises propensities, which assign each action an accumulated value, considering previous performances. Each forecasting model is represented by an RL agent, which has a propensity indicating how strongly the environment tends to select it. These propensities are updated every period by the sum of the previous value (with a forgetfulness factor) and the reward received. The propensities are then normalised, complemented by an exploration component that allows for the possibility of choosing different forecast models.

Research by Pinto *et al.* (Pinto et al. 2014) utilised a Q-Learning (QL) methodology, which employs an off-policy algorithm to estimate the value-state-action function. This is the expected value of a future reward when selecting an action in a specific state, then following the optimal policy. Each state may correspond to a period, while actions involve choosing a better approach for each state. Each state-action pair has a value stored in the Q table, which is updated according to the reward function. It has two hyperparameters, which are used in updating the future estimate: α and γ, which respectively induce a learning rate, controlling how quickly the system adapts, while γ is a discount factor that determines the significance of future rewards.

The Double DQN model was designed to select between three forecasting models, applied in the study carried out by Pannakkong *et al.* (Pannakkong et al. 2023). Double DQN reduces the overestimation of Q-values, common in traditional methods such as Q-Learning. It has two layers with 10 and 9 neurons, determined through preliminary tests. The RL agent starts by exploring the models randomly, gradually de-escalating this exploration to more informed choices as it learns. After each action, the agent receives a reward. This aims to favour models with smaller errors and penalise those with larger errors. This approach supports adaptive forecasting and offers robustness against noise, being especially useful in time series data.

Ensemble-based RL (Lee et al. 2021) builds on statistical uncertainty metrics and leverages the ensemble model's generalisation ability through a more grounded exploration mechanism. Each forecasting model used by the RL system is assessed based on the mean and standard deviation of individual rewards collected each day. These two metrics guides the selection of forecasting models to favour those with the highest performance values.

The Actor-Critic RL (Guo et al. 2024) model combines two components in its structure: the actor, which updates the action policy, and the critic, which evaluates the selected action using the state value function. The actor determines a probability distribution over an approach. Meanwhile, the critic assesses the expected value of that state by calculating the temporal error. This error is used to update both the critic's and the actor's parameters, thereby increasing the likelihood of selecting actions that result in high returns. This architecture provides excellent generalisation and efficient online learning.

Deep Reinforcement Learning (DRL), combines RL and DL, typically used in very complex systems, where decisions are made in dynamic environments with a large volume of data (Yu et al. 2021). The DRL agent receives an initial state of the environment, using a neural network to help determine an action. After the action is performed, the environment returns a new reward and a new state. The agent gradually helps the neural network, with learnings carried out over time, regarding its actions (Coraci et al. 2023).

2.3.2 Related work

This subsection examines the application of various RL methodologies, including PPO, QL, DRL, among others. It analyses recent studies, showing the impact of RL primarily in EM, concluding with some in SGs.

The application of RL methodologies in EMs is highly beneficial, as shown in a study conducted by Pinto *et al.* (Pinto et al. 2014). This study explored how bilateral energy contracts are negotiated within local markets. Its goal was to create a conceptual model that helps Iberian electricity market communities enhance their negotiation strategies in competitive settings. A QL model was used to determine the optimal trading strategies. Findings indicated that this learning approach greatly boosted negotiation effectiveness, particularly for small consumers and prosumers.

A study carried out by Teixeira *et al.* (Teixeira et al. 2023), illustrates the impact of forecasting techniques on participation in the wholesale electricity market, with the finding that model selection depends only on accuracy, thus being able to ignore the economic costs associated with forecasting errors. In a simulation of the Iberian market, the authors examined how price forecasting can influence negotiation outcomes. They then proposed an RL method for selecting forecasting models based on both forecasting precision and business value, demonstrating that even minor errors can significantly affect profitability.

A study conducted by Maestre *et al.* (Maestre et al. 2018) employed a Deep Q-Learning (DQL) approach to help define dynamic pricing strategies in markets. The data included demand values and consumer characteristics. When choosing the best RL methodology for their case study, the authors compared the DQL method with the traditional QL method, discovering that the DQL model performed better. Therefore, the results demonstrated that even in environments with volatile user preferences, their approach was practical, identifying pricing policies that enhance profits.

An application developed by Gupta *et al.* (Gupta et al. 2021), is the use of DRL to improve heating control in smart buildings. The system learns to vary heating control in real-time without the need for large volumes of data. This method in this specific case demonstrates that it improved thermal comfort by 15 to 30% and reduced energy costs by 5 to 12%, compared to traditional thermal controllers.

RL algorithms also play a key role in SGs, such as in efficient energy dispatch in decentralised networks, improving stability and avoiding network fluctuations, with the use of PPO (Lee et al. 2024). RL combined with forecasting algorithms has a greater benefit for the system performed by Ruzbahani *et al.* (Ruzbahani 2024), in which the RL algorithm efficiently manages energy components while ML methods such as Convolutional Neural Network (CNN) and LSTM are implemented to detect cyber-attacks. However, there are challenges to successfully applying RL (Zhang, Han, and Deng 2018). Having a large, clear, and robust collection of historical data is crucial for the effectiveness and efficiency of the RL model.

In another study, Ramos *et al.* (Ramos et al. 2022) used an RL approach to dynamically select forecasting methods for electricity consumption forecasting in buildings. This approach utilised the Multi-Armed Bandit algorithm, enabling the methodology to learn and adapt by choosing forecasting models based on their error rates. The dataset included time variables, temperature, humidity, and other factors. The authors found that this RL approach achieved about 71% accuracy in complex buildings, while in buildings with limited data, accuracy dropped to around 61%. They concluded that integrating RL with forecasting models enhances adaptability across different scenarios.

2.4 Multi-agent Systems

An agent is an entity that can perceive its environment and act autonomously to achieve its goals (Ali and Obied 2022). There are different types of agents (Traldi et al. 2022). Some are traditional, distinguished by their internal structures that dictate the agent's behaviour. Reactive architectures respond directly to environmental events with targeted, efficient actions. Deliberative agents utilise symbolic reasoning for more complex actions but require more computational resources. Within this group, a dominant architecture called Belief-Desire-Intention encourages agents to reason about their beliefs, goals, and intentions.

Agents are fundamental components of a MAS, which is made up of multiple agents. These autonomous agents can cooperate or compete to carry out different tasks within a complex environment and to solve one or more problems (Ye, Tao, and Jaques 2025). Each agent performs a microtask independently, intending to complete the total task once all agents have finished. An agent has several characteristics, such as:

- Autonomy: Carry out tasks without the help of others.
- Reactivity: Ability to respond autonomously and intelligently to events and changes in the environment.
- Inferential Ability: Processing information received, and ability to calculate the final decision based on rules and observations.
- Pro-Activity: Can take initiatives, even with the absence of external incentives, in order to achieve the final objectives.
- Social Behaviour: Has the ability to communicate with other agents or systems.

MAS have several capabilities, in which each agent can be programmed to receive a specific task. Each agent's resolution aims to execute their own task, making the process quicker and more efficient due to task distribution (Izmirlioglu et al. 2024). A MAS provides the means to add or remove an agent, as well as reconfigure its roles. Even if there are local failures, the system has the capacity, reliability and robustness to maintain continuous operations. With the implementation of parallelism, data processing and decision making, computational requirements can be much lower. Compared to centralised systems, a MAS is notable for its decentralised applicability, making it more resistant to failures (Kamdar, Paliwal, and Kumar 2018). The diverse range of MAS applications highlights its broad capabilities, including optimisation tasks, software platforms, and real-time simulations. With AI advancements, agents can incorporate techniques such as ML, RL, and more, allowing them to learn simultaneously through ongoing interactions with their environment. Some of these interactions are described in the related work subsection.

2.4.1 Development frameworks

The development of agents and MAS is supported by various frameworks and tools, each offering different features. Choosing the appropriate framework depends on the application's domain and the complexity of the agents that need to be built using a suitable infrastructure. Java Agent Development Framework (JADE) (Eyimaya, Altin, and Nasiri 2024) is one of the development tools that facilitates the creation of MAS. It is an open-source platform implemented in Java, designed to facilitate the creation and management of MAS. JADE implements the Foundation for Intelligent Physical Agents (FIPA) specifications for its communication, which define various standards for agent-to-agent communication (Mequanenit et al. 2025). This communication is based on asynchronous messages, which allows the exchange of independent messages, being fundamental for distributed and scalable systems. JADE has different types of behaviours, thus being configured according to their function. A JADE-based system can be split between multiple machines.

The Smart Python Agent Development Environment (SPADE) (Palanca 2020) and the Python-based Framework for Heterogeneous Agent Communities (PEAK) (Ribeiro et al. 2023) also enables the creation of an MAS. SPADE and PEAK use the Extensible Messaging and Presence Protocol (XMPP), known for exchanging instant real-time messages between entities, standing out for its efficiency and flexibility. Message exchange is carried out using XML structures, making the protocol extensible and human-readable. Provides native support for Transport Layer Security (TLS). It is a decentralised server, where anyone can create an XMPP server. This communication helps with interoperability, consisting of an open protocol with the support of small to large-scale networks.

SPADE (Palanca 2020) uses the Python language for the development of MAS, being inspired in JADE. It provides ways to create, manage and communicate between agents in a distributed, real-time, efficient and scalable way. The benefit of being scalable allows this tool to create more agents, if necessary, quickly and easily. SPADE contributes to the development of distributed systems, as agents can be deployed on different machines or servers. Agents in SPADE, similar to JADE, possess various types of behaviours. Its API is easy and intuitive to use.

PEAK (Ribeiro et al. 2023) was developed to address limitations found in other available Python-based tools. It is based on SPADE, which uses the XMPP protocol, promoting a framework capable of supporting a robust, flexible and scalable MAS. PEAK also has many features inspired by the JADE framework. These features were adapted to Python due to the power of this language, which has been increasingly used in AI algorithms, among others.

2.4.2 Related work

This subsection presents recent work using MAS in the field of EM, as well as in SGs. Finally, it presents several studies on the integration of RL methodology with MAS.

MAS have applications in EM, such as the following two studies. Alberizzi *et al.* (Alberizzi, Di Barba, and Ziel 2025) developed a MAS model to represent the realism of the European electricity market, with agents representing both renewable and thermal producers, as well as aggregated consumers. All agents employed techniques like QL and statistical models to forecast energy production and pricing. The study highlighted that renewable agents tend to correct their behaviours over time and emphasised the importance of differentiating agent profiles. In another study, Morais *et al.* (Morais, Pinto, and Vale 2020) investigated the interaction of various agents across different segments of the Iberian electricity market. They used a MAS that optimises agent behaviour with integrated forecasting models for market perception. The authors concluded that more complex and adaptive strategies lead to greater economic efficiency.

MAS is crucial in the effective management of forecasts in EM, as demonstrated by the study conducted by Santos *et al.* (Santos et al. 2018). This study proposes a MAS to enhance interoperability in the Iberian market as well as in the management of SGs. The authors carried out this study to address the limitations of isolated simulators, enabling more realistic simulations. For this purpose, semantic models related to the various elements within the components of the electrical system were developed. The authors highlighted that this approach contributed to an integrated simulation of EMs and SGs, leading to more comprehensive analyses and decision support.

The integration of RES is increasingly popular, requiring intelligent coordination mechanisms. The study done by Debrah *et al.* (Debrah 2024) proposes a MAS, to facilitate interaction between entities in an energy system, such as producers and consumers. This approach also reinforces the importance of system interoperability, allowing agents to collaborate to manage and optimise the system, as well as supply and demand. MAS allow a dynamic and decentralised response. However, a strong communication capacity is necessary, as demonstrated by the hierarchical and collaborative approach between agents in the study carried out by Amato *et al.* (Amato, Quarto, and Di Lecce 2021), consolidating a balance in the energy network.

An application of MAS for managing renewable energy, aiming to maximise the utilisation of RES and energy storage, results in less dependence on conventional generators (Radhakrishnan and Srinivasan 2016). The results revealed that the solution using MAS was effective, maximising profits when selling energy back to the grid. During climate change, the algorithm remains efficient up to a 20% increase in these changes, with performance degrading beyond this limit.

In the study carried out by Stennikov *et al.* (Stennikov et al. 2022), a MAS is used, in which the agents are mobile, acting autonomously and adapting to needs in real-time. These agents represent consumers, energy sources, and facilities on the electricity grid. The mobility of each agent allows them to navigate the system to perform their tasks, such as collecting relevant data, adjusting changes in the network, among others. With this competence of the agents, the system becomes flexible and decentralised, facilitating scalability, in the quick and easy adoption of new components. In this way, the system can manage energy flows more efficiently, resulting in a 9.94% reduction in energy consumption.

MAS can help in the coordination between devices and autonomous systems, such as sensors, air conditioning systems, solar panels, among others (Yan et al. 2023). Events in agents can optimise communication between devices, because when an event is triggered, a control algorithm will be applied, reducing the frequency of data transmission and saving energy. MAS improves the response to energy demand and facilitates the integration of RES, through communication between different agents with their specific functions (Mahela et al. 2022). MAS facilitates the reaction to topological and operational changes in the network, with cooperation and flexibility between agents, promoting interoperability and scalability between areas of the energy sector, such as: consumption, generation, markets, among others (Shobole and Wadi 2021).

The study conducted by Hernandez *et al.* (Hernandez et al. 2013) demonstrated the use of MAS to support the operations of Virtual Power Plants, which aggregate multiple distributed generation units such as solar panels and wind turbines. The use of MAS enabled improved communication, coordination, and decision-making among different components of energy, leading to better energy management through redirecting excess energy or adjusting consumption during times of scarcity, among other functions. RNN is employed to make consumption or generation forecasting. The study also demonstrated that the system returned an error of 1.5%, which is considered low for this system and thus provides reliable data for MAS.

2.4.2.1 Application of RL in MAS

At MAS, competitiveness and collaboration among different agents are crucial for a system's success. When MAS is combined with RL, it helps in designing effective learning strategies where agents work together towards a shared goal, whether they are competing or cooperating (Wang et al. 2022). The DQN was used in multi-agent environments to investigate the behaviour of two distinct agents with independent DQNs, according to Oroojlooy *et al.* (Oroojlooy and Hajinezhad 2023). The authors concluded that adjusting the value of rewards allows agents to adopt cooperative or competitive behaviours. The autonomous capacity of agents to develop complex strategies depending on the context and rewards was also highlighted.

In MAS it is possible to exchange information even in decentralised systems, as agents can share information in real-time, improving forecasts (Groenewald et al. 2024). One of the advantages of combining RL with MAS is the possibility of prioritising tasks, to deal, for example, with peaks in high demand or extreme weather events. The combination of these two technologies enhances local decision-making and facilitates the coordination of agents to achieve a common goal. However, it is crucial to ensure system stability as these techniques are highly processing-intensive. Since multiple active agents are involved, running tasks at the same time can lead to considerable computational difficulties. Thus, it is crucial to manage the entire MAS efficiently, particularly because environments are dynamic and require the system to be adaptable.

A huge advantage of combining MAS and RL is the multiple processes that can be carried out and controlled in real-time, without exceeding the computational limit, if they are well programmed (Gronauer and Diepold 2022). It is possible to forecast generations and consumption by agents, adjusting and optimising energy distribution, while other agents are sequentially learning the patterns of data received in real-time.

The following authors Jendoubi and Bouffard (Jendoubi and Bouffard 2022) used multiple agents for energy storage, consumption devices and backup generators, in order to collaborate/compete for the success of the system. In addition to these agents, it has RL agents for learning about load data and renewable energy production through training an RL model. It is important to control the forecasting made by each agent, as this can impact the forecasting made by other agents, especially when they are competing to maximise rewards (Harrold, Cao, and Fan 2022). It is crucial to reward successful agents to encourage them to continually improve their forecasting, thereby keeping the system active and searching for better solutions. The study carried out by Harrold *et al.* (Harrold et al. 2022) demonstrated that the combination of MAS and RL improves the control of hybrid energy storage systems, reducing energy costs and maximising the use of renewable energy. The authors also mentioned that when each agent focuses on different parts of a global problem, solving their tasks while coordinating and potentially competing, this approach can be more effective than using MAS for a single overall objective.

2.5 Chapter Conclusions

There is a major modernisation in the energy sector, with the inclusion of automation, monitoring and communication technologies to achieve more efficient energy management. Electricity markets promote efficiency by stimulating investment and integration of renewable energy. However, in the Iberian region, they face problems such as price fluctuations and climate dependence. SGs seeks to reduce unnecessary energy costs and include renewable sources, thus improving environmental quality. However, there are challenges, such as the intermittent nature of renewable sources, achieving scalability, strengthening security and resilience.

Forecasting models enable the anticipation of demand, market prices, energy flows, consumption, and solar energy, among others in the energy sector. These models can be based on probabilities like statistical ones, being simpler and providing robust solutions under pleasant conditions. On the other hand, ML-based models offer high forecasting accuracy in complex situations where data may not be as linear. RL has shown to be one of the greatest advantages in adaptive and real-time decisions, having the ability to learn directly from interactions in the environment, adjusting to different scenarios, whether more linear and coherent or dynamic and uncertain. When combined with MAS, RL enhances its abilities, since MAS, composed of multiple autonomous agents, acquires the distributed and decentralised nature required in both EMs and SGs environments.

However, these technologies face practical challenges, making their applicability on a large scale difficult. Implementing RL requires high computational costs, requiring high and robust volumes of data. For MAS to be efficient, they require interoperability and effective organisation among the various agents. The advantages demonstrated by all these technologies: forecasting models, RL and MAS, are even greater if combined correctly. These technologies revolutionise the fields of EMs and SGs, building a resilient and more sustainable system that enhances environmental quality and offers greater flexibility. Despite numerous challenges, they have become a fundamental component of the global energy sector by adopting AI techniques.

3 Methods and Materials

This chapter explains the system overview, followed by Data Pipeline, Forecasting Models, Reinforcement Learning, Hyperparameter tuning and Sensitivity Analysis, and MAS. A succinct conclusion of the topics covered in this chapter is provided.

3.1 System Overview

This section provides a general overview of the developed system in this dissertation. Figure 7, represents the forecasting system.

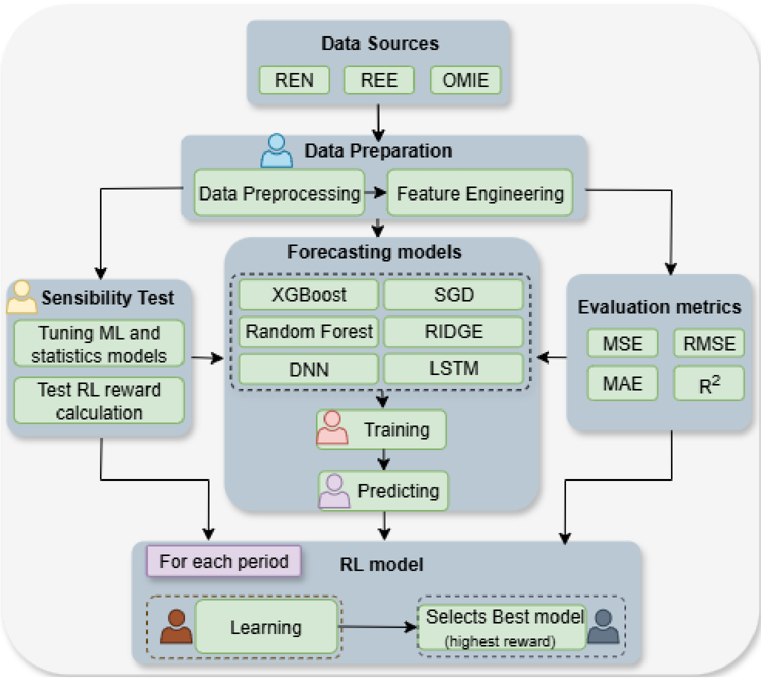

Figure 7 - Forecasting system with data preparation and models evaluated through parameter tuning.

As demonstrated in Figure 7, the data used for this dissertation are from REN, REE, and OMIE. This data undergoes data preparation, which involves data pre-processing tasks such as data acquisition, data correlation, and feature selection. Next, in feature engineering, data standardisation is performed, and finally, categorical variables are transformed using One-hot encoding (Gnat 2021). The following process involves applying various forecasting models, such as statistics, ML, and DL. These methods are subject to tuning, with evaluation from metrics to assess their performance. These models are integrated into an RL approach that iteratively learns from their errors to select the best model at each period. The RL methodology goes through sensitivity testing, which is also evaluated by the metrics. The forecasting system creates a diverse set of robust models that perform well across various scenarios and complement each other.

Multiple libraries were used in the development and data analysis of this system. In data visualisation, analysis, and processing, two libraries are important: Pandas (team 2024) and Numpy (Harris et al. 2020). Pandas is a library that provides high-performance, easy-to-use data structures and data analysis tools in Python. The panda's library is especially suited to handle structured or tabular datasets, a multidimensional table of variables and their respective values. NumPy is a scientific-computing library that introduces a multidimensional array object and other derived structures, such as masked arrays and matrices, along with an enormous variety of efficient operations on arrays, including mathematical, logical, and sorting functions.

3.2 Data Pipeline

Constructing a data pipeline is crucial to ensure the quality and robustness of the data, which is particularly important when using forecasting methods. Therefore, this section will cover data origins, preprocessing and feature engineering, and conclude with data protection, security, and ethics.

3.2.1 Data Sources

The basis for the success of the proposed work focuses on the data that is used for its application. A robust dataset is very important for the success of a forecasting system, especially when it comes to energy data, which is highly variable and uncertain.

To demonstrate the applicability and benefit of this system, one type of dataset was used. This dissertation dataset was constructed using data from REN, REE, and OMIE from 2022, 2023, and 2024. As shown in Figure 8, REN and REE data are accessible on their respective websites through REN Data Hub[8] and REE Todate[9]. Data can be retrieved for different timeframes, such as daily, monthly, or yearly, and downloaded in formats like JSON, PDF, CSV, or Excel. It displays energy sources, including wind, solar, nuclear, hydro, and natural gas. Data is made available

[8] https://datahub.ren.pt/pt/
[9] https://www.ree.es/en/date/todate

every 15 minutes in Portugal, while in Spain it is daily. Both datasets are only accessible the next day at 1 a.m. Central European Time (CET).

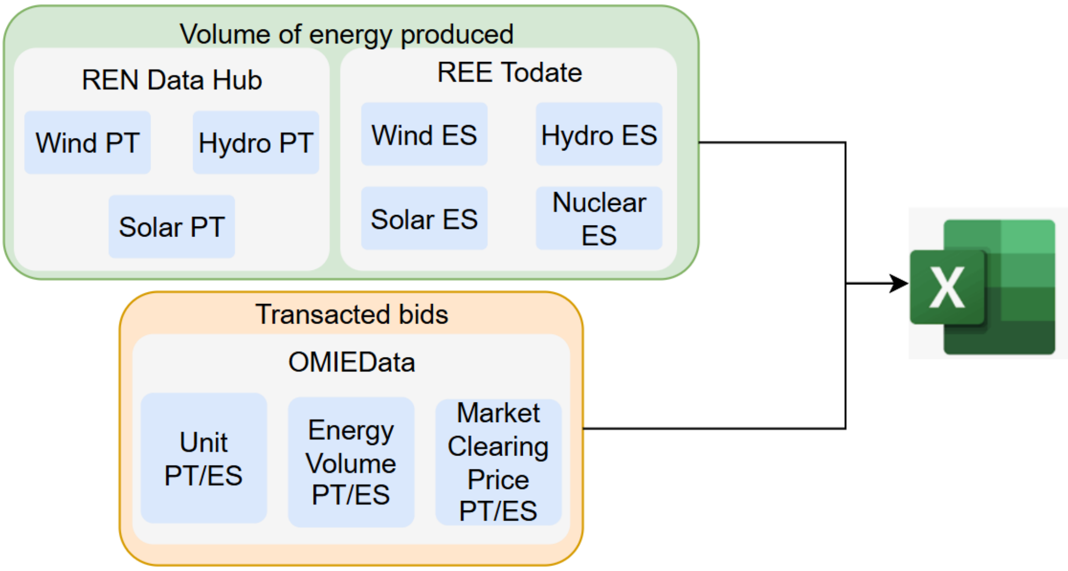

Figure 8 - Data from REN, REE, and OMIE, showing the variables extracted for use in the dataset.

OMIE data is available on his website in OMIEData[10] and allows users to export daily data as CSV files. To access all traded energy volume sources and prices, users should browse the archives and daily market reports. For detailed production data, open the "Total disaggregated power after Day-ahead market" folder, which displays the energy transacted by each energy source for every hour of the day. To view energy prices, visit the prices folder where prices are listed by period for each day. All this data is updated hourly and becomes accessible around 2 p.m. CET on the previous day.

After the clear insights from the data, it is necessary to understand how real-time data events operate in this environment for real-time applications. Table 2 shows the real-time events of the Iberian market and explains how the system implemented in this dissertation should act, regarding the use of models.

[10] https://www.omie.es/pt/market-results/daily/daily-market/day-ahead-price

Table 2 - Example of real-time events from the Iberian market incorporated into the system of this dissertation.

Time PT	Day 1	Day 2
1:00 a.m.	Publication of **final** REN/REE data for the day **before**	Publication of **final** REN/REE data for day **1**
...	Model forecast for day **2**	Model forecast for day **3**
11:00 a.m.	Bid submission for day **2**	Bid submission for day **3**
...		
2:00 p.m.	Publication of **final** OMIE data for day **2**	Publication of **final** OMIE data for day **3**
...	RL model rewards update	RL model rewards update
11:00 p.m.	Publication of **forecasts** REN/REE data for day **2**	Publication of **forecasts** REN/REE data for day **3**

Table 2 illustrates the events in real-time to predict market values and submit bids in time for days 2 and 3. Focusing on the events needed (highlighted in blue) to submit bids in time for day 3, the following occurs. On day 1, at 2:00 p.m., the final market values for day 2 are known, as provided by OMIE. Also on the first day, at 11:00 p.m., the values predicted by REN/REE for day 2 are published. On day 2, at 1 a.m., the final REN/REE value from the previous day (day 1) is made available. All the necessary data is then collected and synchronised chronologically to make forecasting for day 3. Since bids for day 3 must be submitted by 11 a.m. on day 2, an 11-hour window is checked to predict the market value for day 3. After the bids are submitted, the final values for day 3 are known at 2 p.m. on day 2. Finally, the RL model updates its rewards after knowing the actual market values.

3.2.2 Preprocessing and feature engineering

Preprocessing and feature engineering are vital steps in converting data into appropriate information for training models. For the construction of the dataset, a timeline period was used, in which the first features are year, month, day, day of the week and period (hour). The following features are represented in Table 3, where each one identifies the total volume of energy produced by the different sources in Portugal and Spain. The “act” signifies the amount of energy generated from each source. In the dataset, the Portugal data is grouped by hour, but data from Spain is treated as the same daily energy volume for all hours of that day.

Table 3 - Features for the dataset by REN/REE.

Feature Name	Total volume	Country	Source
E-act-wind-pt-(d,h)	Energy produced for wind	Portugal	REN
E-act-wind-es-(d,h)	Energy produced for wind	Spain	REE
E-act-solar-pt-(d,h)	Photovoltaic solar energy produced	Portugal	REN
E-act-solar-es-(d,h)	Photovoltaic solar energy produced	Spain	REE
E-act-hydro-pt-(d,h)	Hydro energy produced	Portugal	REN

Feature Name	Total volume	Country	Source
E-act-hydro-es-(d,h)	Hydro energy produced	Spain	REE
E-act-nuclear-es-(d,h)	Nuclear energy produced	Spain	REE
E-act-load-pt-(d,h)	Energy consumed	Portugal	REN
E-act-load-es-(d,h)	Energy consumed	Spain	REE

OMIE data was also used to complete the required feature set. This data represents the energy traded on the daily market in Portugal and Spain for each energy source. The characteristics are shown in Table 4, where "dam" indicates the day-ahead market, referring to the volume of energy transacted in the market. Ultimately, there is the "mcp" feature, which stands for the Market Clearing Price, representing the actual price value. OMIE data is used as it is, due to its hourly data being provided.

Table 4 – Features for the dataset by OMIE.

Feature Name	Description	Country	Source
E-dam-wind-pt-(d,h)	Energy volume transacted for wind	Portugal	OMIE
E-dam-wind-es-(d,h)	Energy volume transacted for wind	Spain	OMIE
E-dam-solar-pt-(d,h)	Volume of solar photovoltaic energy transacted	Portugal	OMIE
E-dam-solar-es-(d,h)	Volume of solar photovoltaic energy transacted	Spain	OMIE
E-dam-hydro-pt-(d,h)	Volume of hydro energy transacted	Portugal	OMIE
E-dam-hydro-es-(d,h)	Volume of hydro energy transacted	Spain	OMIE
E-dam-nuclear-es-(d,h)	Volume of nuclear energy transacted	Spain	OMIE
E-dam-load-pt-(d,h)	Volume of energy consumed transacted	Portugal	OMIE
E-dam-load-es-(d,h)	Volume of energy consumed transacted	Spain	OMIE
mcp-pt-(d,h)	Market clearing price	Portugal	OMIE
mcp-es-(d,h)	Market clearing price	Spain	OMIE

For each feature of the data represented in Table 3 and Table 4, values from the same period (*h*) for the previous three days (*d*) are added as new features. With the exception of load energy, in which four new characteristics were added: one representing the value in the same period of the previous day (*d-1*, *h*), two more representing the values in periods adjacent to this characteristic (*d-1*, *h+1*; *d-1*, *h-1*) and one representing the value of a week before in the same period (*d-7*, *h*). Also, for the market-clearing price, five features were added: The first four are based on the same logic as the load energy characteristics. The other feature represents the value from two weeks earlier in the same period (*d-14*, *h*). This definition is important because at each moment, previous values are considered, allowing the models to have a better understanding of their decision-making. This data is used in the system, where the target column is "mcp-pt-(d,h)". The forecasting type is regression because the target value represents numerical values (Tan et al. 2021). Finally, the data from REN, REE, and OMIE are integrated into the dissertation dataset, all within the same time period of the day.

With this data set, it was possible to carry out several case studies, aiming to offer the user optimal price suggestions suited to each scenario and requirement.

3.2.3 Data Protection, Security, Ethics

In this dissertation, special attention is given to several issues such as data protection, security and ethics, due to the use of large volumes of historical data, and the efficient management of the system.

Regarding data protection, it is vital to maintain the confidentiality of each user's data, ensuring they only access their own information. The system administrator can view all data and operations but does not share internal details or user information with others.

The integration of AI technologies can introduce potential vulnerabilities to the system, such as biased training data, sensitivity to data noise, among others. Therefore, the administrator will have to continuously monitor its functionality, identifying any failures and acting appropriately to correct them. On the other hand, the system has fault tolerance, allowing it to continue operations even in the event of failures in some active processes, using robustness and redundancy mechanisms.

The values coming from forecasting methods may have ethical implications, however the system ensures that the forecasting provided to the user are reliable, not favouring any user and therefore always acting in accordance with the data that is entered. It is important to highlight that this system indicates a value to the user, resulting from a forecasting, without committing to subsequent results.

The data from OMIE[11], REN[12] and REE[13] that were used for the case studies in this dissertation are public and can be accessed through their official websites. Additionally, the use of this data for research and analysis is permitted due to the terms and conditions set forth by them.

3.3 Forecasting Models

This section outlines the tools required to use the forecasting models, describing each in detail and presenting its characteristics.

For the use of the forecasting methods in this dissertation, tools were necessary. These are Scikit-learn (Wang et al. 2024) and TensorFlow (Abadi et al. 2016), which contain the foundation for programmers to utilise and adapt models within their systems, leveraging the various functions offered by the tools. Additionally, they enable the construction of complex models for classification, forecasting, and big data analysis. Scikit-learn is a Python module that integrates several classic ML algorithms for both supervised and unsupervised problems in the medium-scale dimension. It aims to make access to ML easy for non-experts through a general-purpose high-level language and is a ML library developed focusing on large-scale and diverse

[11] https://www.omie.es/pt/market-results/daily/daily-market/day-ahead-price
[12] https://datahub.ren.pt/pt/
[13] https://www.ree.es/en/date/todate

environments (Konstantinos Roumeliotis and Tselikas 2023). Among the areas which it supports, the most important one is DL models. TensorFlow is an open-source platform, built by Google for running and building ML algorithms on huge distributed systems comprising GPUs. Indeed, it has already been highly employed in both the training and deployment of deep neural networks by a few applications, such as computer vision, robotics, and natural language processing.

Six forecasting models were used, as seen in Figure 7, of which two were statistical models (RR and SGD), two were ML models (RF and XGBoost), and two were DL models (DNN and LSTM). RR is a regularised linear regression technique that subtracts the squared error term by adding a penalty term proportional to the square of the model coefficients (Zhang et al. 2021). This regularisation is important in cases where there is multicollinearity between independent variables, thus ensuring more stable and robust forecasting. In this system, it is important to identify simple relationships in data with many correlated variables. SGD is an iterative model fitting technique for optimising loss functions, such as quadratic error, using random sampling of the data (Saha, Haque, and Sidebottom 2022). It has a specific application in contexts where there is a high volume of data and for which conventional optimisation methods may be computationally inefficient. Since computation time is crucial, providing faster responses to users is better. SGD can quickly refine the forecasting model.

This thesis employs RF as one of the ML models, which is an ensemble learning algorithm that combines multiple decision trees trained on different data subsets to produce more accurate and robust forecasting (Li et al. 2020). It is especially handy in handling complex nonlinear data and finding variable interactions, hence ideal for tasks such as market price forecasting. RF works well with big and small datasets, but its strength lies in large, high-dimensional datasets with many variables. It also does a great job of handling noisy and imbalanced data, avoiding overfitting, which makes it a very versatile tool for a wide range of forecasting modelling tasks. This dissertation also implements XGBoost, which is an optimised version of boosting algorithms that improves model performance incrementally by adding new models to correct the errors of the previous ones (Obiora, Ali, and Hasan 2021). It is fast and efficient and has become widely used in ML competitions due to its superior performance. It works best on complex datasets, especially when fine-tuning and model optimisation can significantly improve forecasting accuracy.

Finally this thesis have DL models, using DNN which is an efficient model that uses multiple layers of artificial neurons (Hossen et al. 2018). This model is used for image recognition, natural language processing, and time series forecasting. These types of models can learn non-linear patterns, as well as analyse relevant characteristics of the data, thus being effective. Its great generalisation capacity is beneficial when dealing with large volumes of data. This model contains fundamental parameterisations that help avoid overfitting, thus increasing the effectiveness of the model. The DNN is very useful in this dissertation, helping to identify variations, abnormalities and outliers in the data, thus leading to ideal forecasting. This thesis also includes LSTM, a type of RNN specifically designed to handle sequential data and capture long-term dependencies (Kartini et al. 2023). Due to their special memory cell architecture and the use of three types of gates (input, forget, and output), LSTMs are able to retain information for a very long time and get rid of irrelevant information, which makes them perfect for tasks such as time series forecasting, natural language processing, and speech recognition, where understanding the context and patterns between sequences is crucial. LSTMs are good at analysing dynamic and time-dependent data and produce good and accurate forecasting and insights in many applications.

3.4 Reinforcement Learning

Throughout the development of the case studies, three RL approaches were used to improve the RL methodology in this dissertation. The RL models focus on identifying the most appropriate forecast model for each time of every day and consist of two components: learning and selecting the best model, as seen in Figure 7.

The initial RL approach, called Controlled Baseline RL (CB-RL), selects the best-known model for each day and period through a transparent and cumulative reward process. The primary basis of the method relies on the MSE measurement metric used in forecasting models. The CB-RL model process begins by calculating the MSE for each forecasting. The MSE measurements of each model are normalised and mapped to a range from 0 to 1, ensuring uniformity in the scales and preventing extreme values. Each instantaneous reward is calculated for each model based on Eq. (2):

$$Immediate\ reward = 1 - \frac{(MSE_i) - \min(MSE_i)}{\max(MSE_i) - \min(MSE_i)} \tag{2}$$

Eq. (2) normalises the reward for all the models and assigns the maximum reward to the model with the lowest MSE. During each period i, the calculated reward is added to the individual rewards of each model. If there is no reward history, the current reward serves as the initial reward. If there is a history, the reward is updated using Eq. (3):

$$\text{Cummulative reward } = \frac{\alpha\ *\ \text{immediate reward}}{(1-\alpha) * \text{previous reward}} \tag{3}$$

In Eq. (3), α is a choice hyperparameter because it allocates importance to instantaneous or historical reward. For instance, if α is 50%, then the terminal reward will have a balance between the instantaneous reward and the historical reward. If the α is more than 50%, then the instantaneous reward has greater priority. However, if α is less than 50%, importance will be given to the historical reward. The CB-RL model determines the reward for each model during each period. Therefore, when the CB-RL model makes a forecasting, it will utilise the model with the highest reward for each period.

Meanwhile, the CB-RL approach was not as robust as desired for this thesis. It aims to be more robust, dynamic, and capable of handling rapid deviations in data, while somehow penalising the forecasting model to avoid bias from historical performance. For this reason, a new RL methodology was developed, designated as Dynamic Penalised RL (DP-RL). Figure 9 represents how the DP-RL model learns.

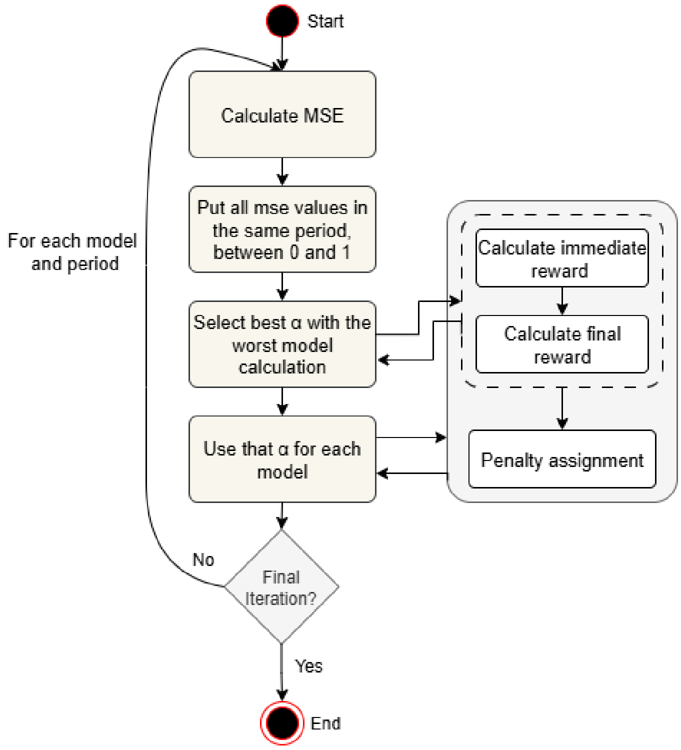

Figure 9 - Reward calculation by DP-RL.

As shown in Figure 9, the DP-RL model selects the most appropriate model for each day by period, through a cumulative reward. The primary basis of the method relies on the MSE measurement metric utilised in forecasting models. The DP-RL model process begins by calculating the MSE for each forecast. Each model's MSE values are normalised and mapped to a range of 0 to 1, ensuring similarity in scales and avoiding extreme values.

The DP-RL model has a hyperparameter called the α list, where the user specifies a list of potential values for α. This hyperparameter is vital, as the best α will be computed per period. This calculation uses the worst forecasting model at each moment, as it exhibits more variance and helps the DP-RL model enhance its knowledge. In assessing this best α, first, the instantaneous reward is calculated for the worst model based on Eq. (2). Then, if there is history for the worst model, the reward is updated according to Eq. (3), where α is a hyperparameter that determines the importance of either the historical or current reward. The value of α that yields the highest reward is then selected as the optimal one for that period.

Once the best value of α is found, it is used to calculate the rewards of all forecasting models for that period. To this end, the calculations for each forecasting model go through Eq. (2) and (3). There is an extra penalty step, though, with the intention of adding greater dynamism and responsiveness to performance changes. Eq. (4) shows the penalty formula, which applies to every forecasting model when the immediate reward is smaller than the previous reward, otherwise, the final reward is not penalised.

$$\text{Final reward} = \text{cummulative reward} * |\text{immediate reward} - \text{previous reward}| \qquad (4)$$

Eq. (4) penalises the final reward of a forecasting model by multiplying this value by the difference between the immediate reward and the previous reward. This applies a proportional penalty, as the resulting value from the absolute difference calculation will be between 0 and 1, since rewards are normalised as previously explained. Therefore, this percentage difference functions as a discount factor on the already calculated accumulated reward. The primary purpose of this penalty is to discourage top-performing models from being chosen repeatedly at the expense of models that currently demonstrate better relative performance. Ultimately, after completing all iterations, the DP-RL model will have a model with a higher reward per period.

To incorporate energy production source contexts into the DP-RL model instead of relying on daily periods, a refined methodology was developed. This final version offers more targeted and detailed information to the model, replicating all the functions of the original DP-RL. Instead of calculating rewards for each forecasting model per period, it uses context numbers. It is called Dynamic Penalised Context RL (DPC-RL).

3.5 Hyperparameter Tuning and Sensitivity Analysis

This section explains how the tuning method is applied to forecasting models to determine the optimal hyperparameters, as well as the sensitivity test used to identify the most effective methodology for the RL model.

The Tuning process is used for all forecasting models, using the Grid Search method (Alibrahim and Ludwig 2021). Grid Search examines several sets of possible parameterisations for a data type and determines the best possible set of hyperparameters. Once the optimal parameterisation of the forecasting models is known, the RL methodology is subjected to a sensitivity test to evaluate its performance. Realising these processes is crucial to determining the best point for each model for a particular data type.

Model evaluation methods were used to determine the quality of each model: MSE, RMSE, Mean Absolute Error (MAE), and Coefficient of determination (R^2). MSE calculates the difference between the actual and predicted value, squared, to penalise larger errors, as illustrated by Eq. (5) (Chugh 2024):

$$\text{MSE} = \frac{1}{n}\sum (\text{y}_\text{i} - \hat{\text{y}}_\text{i})^2 \qquad (5)$$

In Eq. (5), y_i represents the observed value, while $\hat{y}_i$ is the model's predicted value. The difference between them is squared, and ∑ denotes the total of all squared errors. Lastly, (1/n) is used to compute the average of all squared errors. RMSE calculates the square root of the

MSE value because it measures the average deviation of predicted and actual values, demonstrated in Eq. (6) (Chugh 2024) :

$$\mathrm{RMSE} = \sqrt{[\frac{1}{n}\sum (y_i - \hat{y}_i)^2]} \tag{6}$$

MAE estimates the mean absolute error but is less sensitive to outliers, as shown in Eq. (7) (Chugh 2024):

$$\mathrm{MAE} = \frac{1}{n}\sum |y_i - \hat{y}_i| \tag{7}$$

In equation 7, y_i indicates the observed value, whereas $\hat{y}_i$ is the predicted value from the model, with their difference measured as an absolute value. Then, $\sum$ indicates the summed values, and (1/n) is used to calculate the mean. The MSE, RMSE, and MAE measures indicate that the model performs well when their values are closer to zero, and the opposite happens with the R^2 method. The method indicates the percentage of variance in the data explained by the model, using Eq. (8) (Chugh 2024):

$$R^2 = 1 - [\sum \frac{(y_i - \hat{y}_i)^2}{(y_i - \bar{y})^2}] \tag{8}$$

Equation 8 computes a ratio by dividing the sum of errors in the numerator by the total variation of the data relative to the mean in the denominator. Subtracting this ratio from 1 yields the percentage of explained variation, indicating the model's success rate. Without this calculation, the average would instead represent the model's failure percentage.

3.6 Multi-Agent System

The proposed system forecasts recommended prices for bid proposals in the Iberian market. Thus, this is divided into two main interconnected systems: the forecasting system and the MAS, as observed in Figure 10.

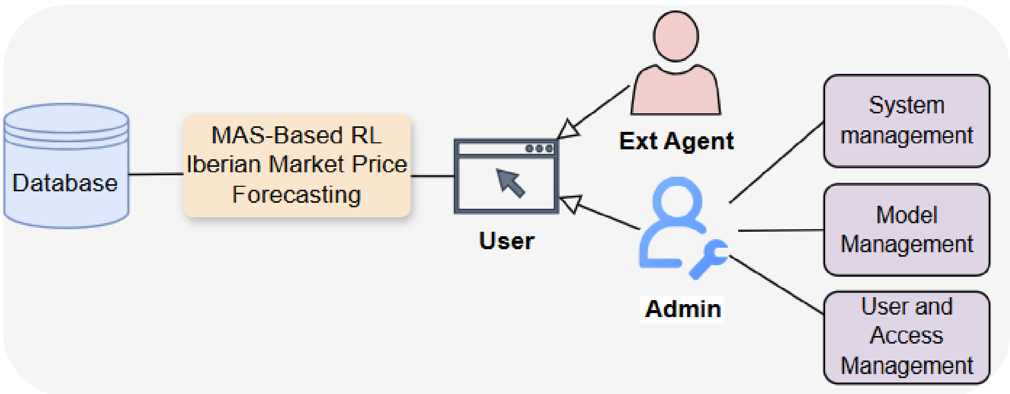

Figure 10 - MAS-Based RL Iberian Market Price Forecasting system, connected to the database and user interface.

The MAS-Based RL Iberian Market Price Forecasting system is connected to a database and aims to store information about forecasting, trained models, and admin configuration. MAS develops a specific agent for each task required by each type of user (administrator, standard user, and external agent), leveraging the MAS to perform multiple tasks simultaneously. External Agent and Administrator expand user functionalities. The administrator plays a vital role in managing the system, templates, and user access.

The MAS-Based RL Iberian Market Price Forecasting designed for this dissertation intends to use the SPADE framework, because it uses native communication via XMPP, allowing the construction of distributed agents even in complex environments, as already discussed in section 2.4 Multi-agent Systems. MAS-Based RL Iberian Market Price Forecasting presents an architecture with autonomous and cooperative agents to execute different tasks related to data forecasting in the energy sector, as shown in Figure 11.

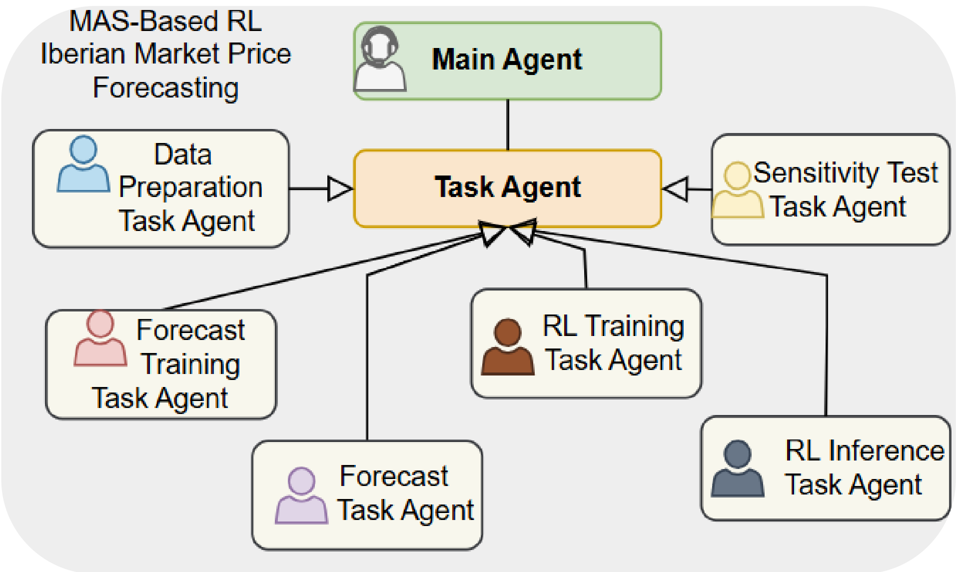

Figure 11 - Main and task agents, illustrating their different actions.

As illustrated in Figure 11, the Main Agent is cyclical as well as active. This is intended to wait for user requests and respond to them later. Therefore, the Main Agent is very important as it creates the Task Agents needed to perform the different actions required. Task agents are single-run and terminate after the task completes. For each forecasting model used, the tasks available for use by the respective agents are as follows:

- data preparation (independent of forecasting model use);
- tuning/sensitivity testing of all types of models;
- training forecasting models (statistical and ML);
- training RL model;
- execution of forecasting models or RL.

3.6.1 Database infrastructure

A PostgreSQL database is implemented to store system configurations, the users, their trainings and forecasting, which includes key features of each model and results, enhancing the system's scalability and robustness. As mentioned before, both the Administrator and the External Agent extend the user's functions, as illustrated in Figure 10. The External Agent, due to the configuration required for communication with the system agents, can only make forecasting with the models already present. On the other hand, the administrator has a connection to the "Configuration" table, as shown in Figure 12, because it has several important system management and support functions. One of them involves configuring each system action, whether training, retraining or predicting each existing model. In addition, it provides models already created for the Iberian market price. However, the administrator chooses models of interest for user applicability. This way, it does not overload the system with all the models.

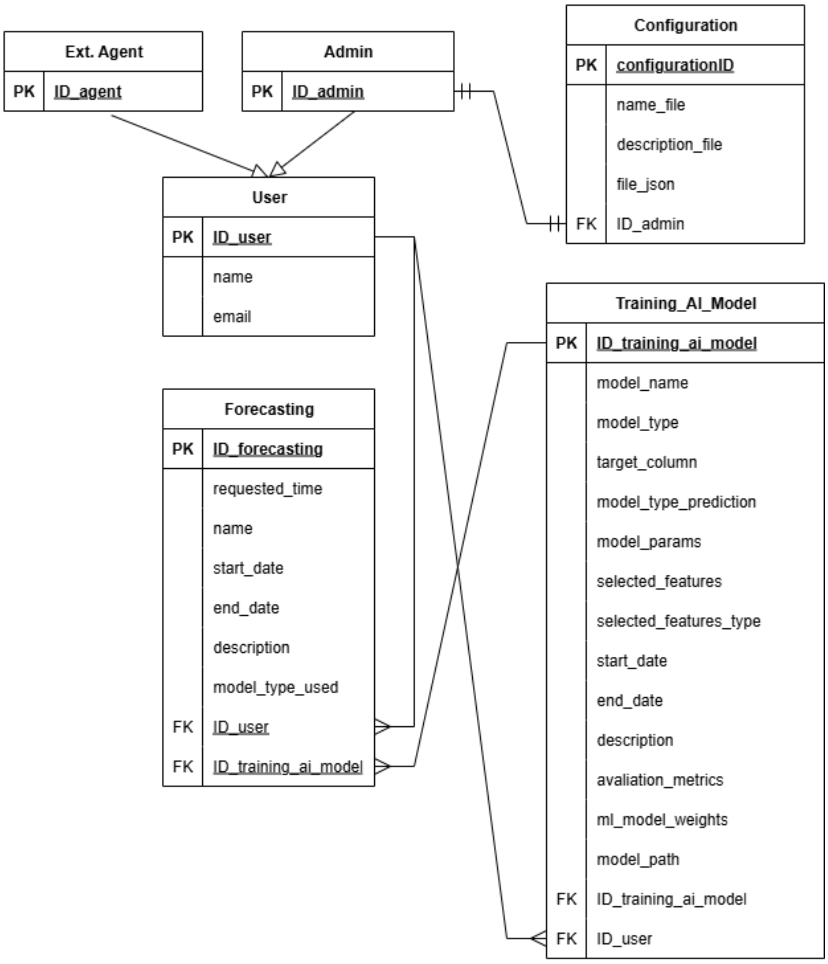

Figure 12 - Relational model, showing tables and their relationships.

Figure 12 shows the relational model of the implemented System. The basis of the dissertation is AI models: statistical, ML, or RL models. All these models are saved in the same table called

"Training_AI_Model," which allows for covering any type of model in an abstracted way. This table stores several important pieces of information:

- the name of the model and what source it is from (statistical, ML or RL);
- the features and type of features that this model supports;
- the type of forecasting of the model;
- the hyperparameters of the model;
- the path to where the zip model is stored, as it takes up less space;
- description of the model. Ex: data you use among other relevant information;
- evaluation of metrics.

The "Forecasting" table represents a request made by any user. For each request, various details are recorded: at what time the request was made, the user's name, the data range considered, and the description of the request. Finally, it specifies the type of model (statistical, ML, or RL) to be used, selecting the specific model.

3.6.2 Workflow

MAS-Based RL Iberian Market Price Forecasting's operations and communications follow a consistent pattern. Initially, as can be seen in Figure 13, the MAS is requested by the user to perform data preparation.

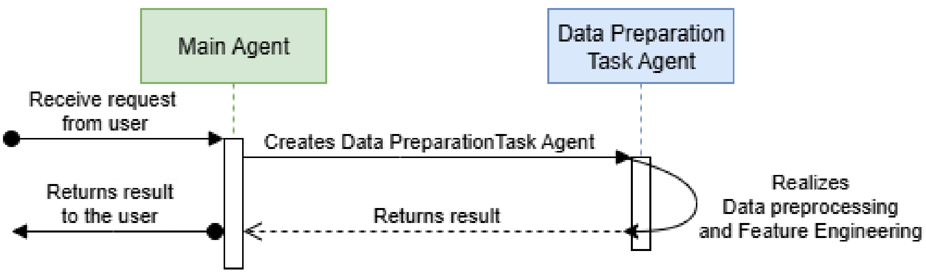

Figure 13 - Sequence diagram - Interaction between agents for data Preparation.

In Figure 13, it is evident that the Main Agent receives the user's request. This Agent creates a specific Task Agent for data preparation. This data preparation involves initial steps such as data acquisition, identifying the most relevant features through correlation testing, and selecting the most effective ones. Finally, in feature engineering, the transformation of categorical variables is performed using One-Hot Encoding.

After the data is well prepared, the user asks the System to perform a tuning for the various forecasting models, to find the best parameterisation for each one, as shown in Figure 14.

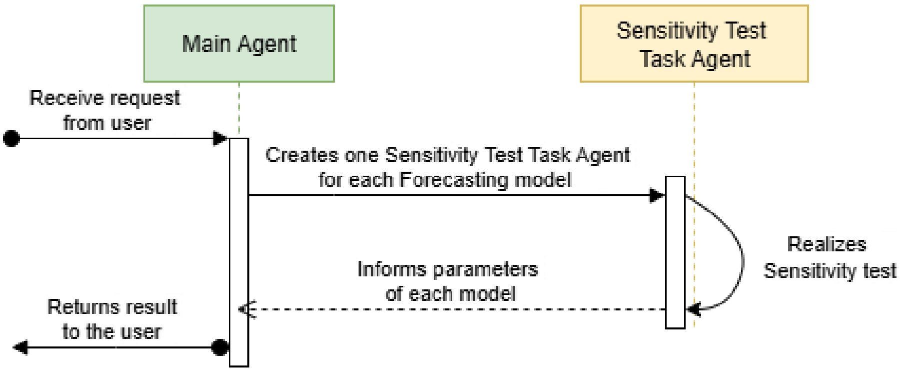

Figure 14 - Sequence diagram - Interaction between agents for tuning forecasting models.

Therefore, in Figure 14, the Main Agent receives the request and instantiates six specialised task agents, one for each forecasting model: RF, XGBoost, LSTM, SGD, DNN, and RR. Each of these agents individually adjusts its respective model. After completing this task, the agents return their parameterisations to the Main Agent. This, in turn, indicates to the user the best parameterisations for each model.

Once the ideal parameterisations for each forecast model have been found, it is possible to perform the sensitivity test on the RL model, as illustrated in Figure 15.

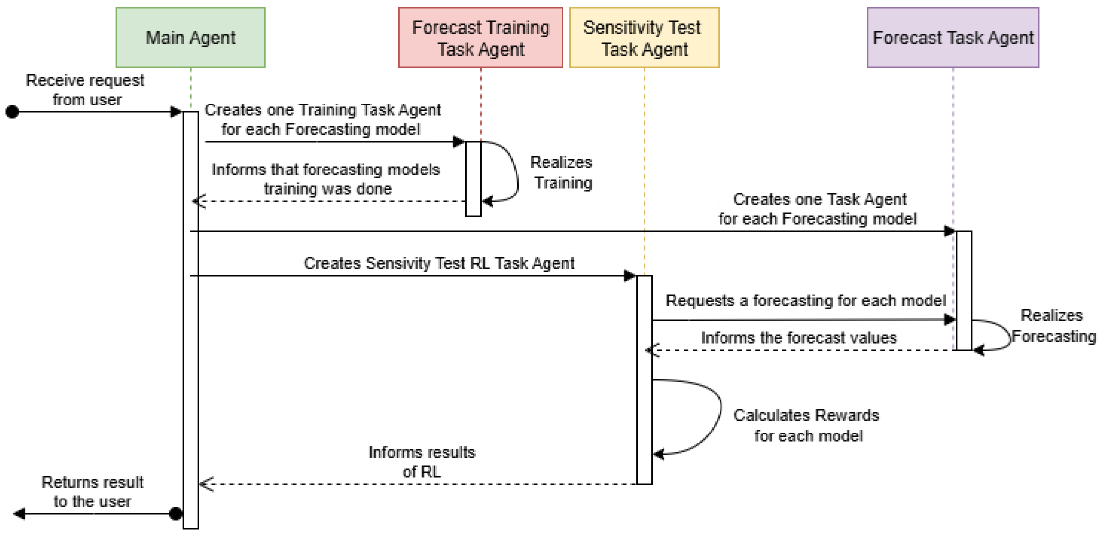

Figure 15 - Sequence diagram - Interaction between agents for RL sensitivity test.

In Figure 15, the Main Agent receives the original user request to test the sensitivity of the RL model. To do so, the user must provide the parameterisations for the forecasting models and the RL sensitivity test itself. The Main Agent initiates the process by generating six task agents, each responsible for training the forecasting models according to their respective configurations. At the end of the training, six task agents are created, one for each forecasting model, along with an additional task agent specifically assigned to conduct sensitivity testing of the RL model. The RL model agent then starts a task cycle by sending forecasting requests to the forecasting model agents, which provide their results. It then calculates each model's reward based on the responses received. After the cycle is completed, the agent responsible for the RL sensitivity test informs the Main Agent that the task has been finished, also sending the results obtained. Finally, the Main Agent forwards these results to the user.

After the ideal parameterisation for the RL sensitivity test is found, it is possible to perform the validation test, represented in Figure 16.

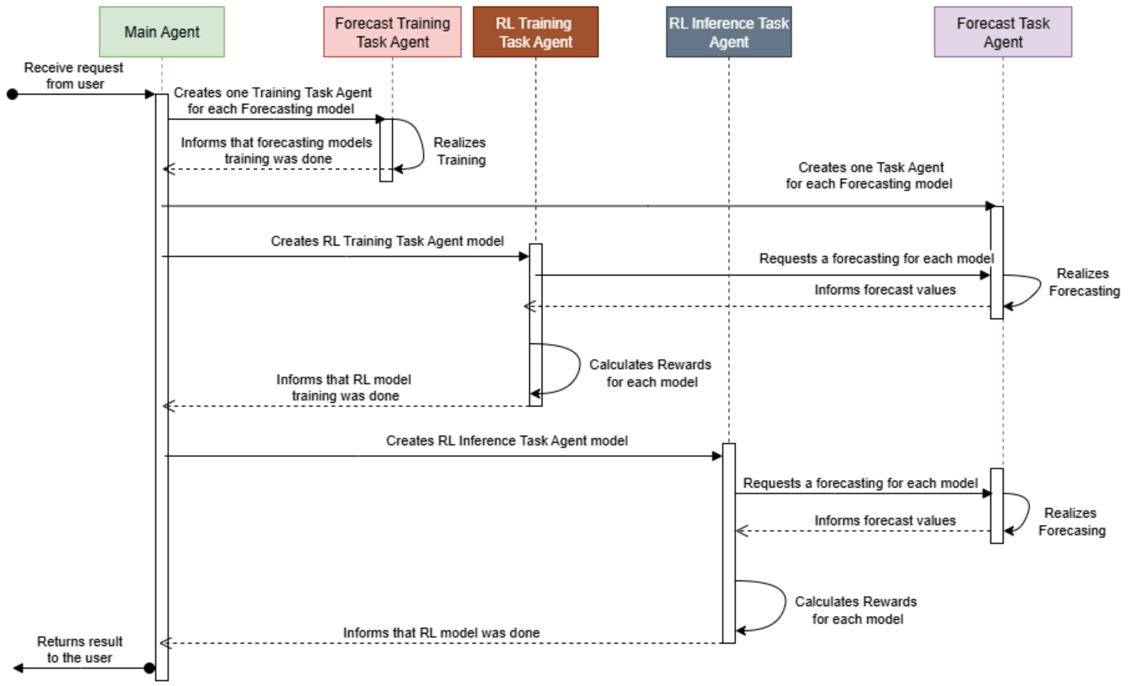

Figure 16 - Sequence diagram - Interaction between agents for validation test.

As shown in Figure 16, the Main Agent is tasked to perform the validation test. The user needs to input the respective hyperparameters for each model. Consequently, the Main Agent creates six task agents to train each forecasting model. After training, each task agent reports to the Main Agent that training has been completed. The Main Agent can then create seven task agents: six forecasting task agents for each forecasting model and one for RL training. The RL training agent model will initiate a cyclic task: it requests forecasting from each forecasting model and calculates the reward for each one. When the cycle is complete, the Training RL Agent informs the Main Agent that the RL model has been trained. For the final task, the Main Agent creates a task agent to predict the RL model, which will begin the same training process. In the end, it informs the Main Agent of the RL test results.

3.7 Chapter conclusions

This chapter covers essential topics for this dissertation. It begins by describing how each system module is connected and its respective processes. Following the description of the data used for the case studies, providing details, as well as issues related to data protection, security, and ethics. This highlights the system’s robustness, fault tolerance, and commitment to user data privacy.

The significance of selecting and implementing various forecasting models, including the RL model, is highlighted. The RL model has evolved through several versions to adapt to changing needs and incorporate innovations for better performance and robustness. These two models form the core of a MAS, including details about its design and agents. To safeguard various and crucial information, the MAS uses a PostgreSQL database. The system is highly adaptable, handling different requests and maximising the RL model’s capabilities.

4 Experimentation

This chapter conducted case studies, evaluating each model used. The objective of the studies is to predict prices to be proposed in a bid for the Iberian market. Finally, the findings describe the performance of the proposed methodology and compare the results of each case study.

Data preparation is conducted before the realisation of the case studies. The tuning/sensitivity tests are executed for each case study, evaluating the performance of both the forecasting models and the RL methodology, using data from 2022 and 2023. Figure 17 shows the timeline of the tuning conducted for forecasting models, and Figure 18 illustrates the sensitivity test of the RL.

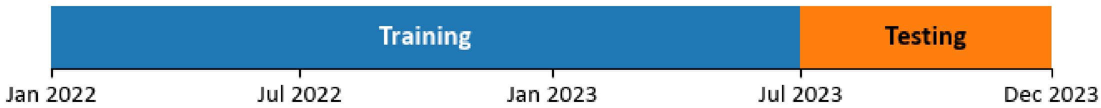

Figure 17 – Dataset used for tuning the forecasting models, considering data from January 2022 to June 2023 for training and data from July 2023 to December 2023 for testing.

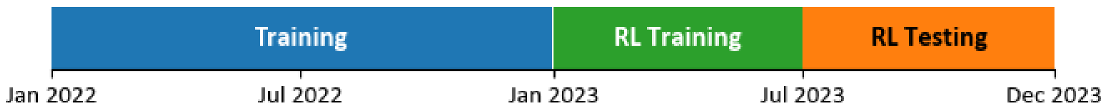

Figure 18 – Dataset used for sensitivity test of the RL methodology, considering data from 2022 for training forecasting models, data from January 2023 to June 2023 for training RL, and data from July 2023 to December 2023 for RL testing.

Figure 17 illustrates that the forecasting models were trained from January 2022 to June 2023. Figure 18, in the same period, shows the training for forecasting models and RL. Both figures display the forecasting model and RL testing, respectively, from July to December 2023.

These two tests allow each model to be better adjusted to the type of data. Finally, the validation test is carried out following the chronology shown in Figure 19.

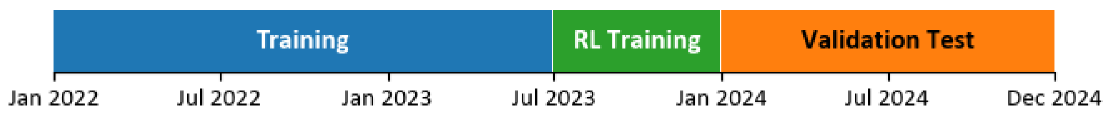

Figure 19 - Dataset used for the proposed solution validation, considering data from January 2022 to June 2023 for training forecasting models, data from July 2023 to December 2023 for training RL, and data from 2024 for validation test.

As referenced in Figure 19, the forecasting models are trained from January 2022 to June 2023. The RL methodology is trained from July to December 2023, with validation test being carried out for the full year of 2024. During the RL training and testing phases, the forecasting models are cyclically retrained at the end of each month.

The following subsections present the four case studies that validate the proposed system. Case studies 1-3 validate and confirm the methodology's effectiveness, while case study 4 tests it under conditions similar to the real context. Table 5 presents the four case studies, each with its own sub-case studies.

Table 5 - Description of case studies, including their objectives, whether feature selection was applied, the RL methods used, and the simulated environment.

Case Study	Sub Case studies	Objective	Considers feature selection?	RL methodology used	Simulates a real-time environment?
CS1	CS1.1	Analyses a one-month forecast	No	CB-RL	No
	CS1.2	Analyses a one-year forecast	No	CB-RL	No
CS2	CS2.1	Compares CB-RL, DP-RL, and other approaches to improve reward calculation	Yes	CB-RL, DP-RL, and other approaches	No
	CS2.2	Compares DP-RL with existing methods to identify gaps and differences	Yes	DP-RL and different RL methodologies	No
CS3	-	Compares the performance of each energy source to the Iberian market price	Yes, but only for the basic features	DP-RL	No
CS4	CS4.1	Runs the DP-RL model	Yes	DP-RL	Yes

Case Study	Sub Case studies	Objective	Considers feature selection?	RL methodology used	Simulates a real-time environment?
	CS4.2	Runs the DPC-RL model with context from energy source, production, and demand	Yes	DPC-RL	Yes

The first case study serves as a baseline, using monthly and yearly forecasting with no feature selection, along with a CB-RL model. The second case study is more comprehensive: the first sub-case study compares the CB-RL methodology with other approaches to improve RL reward calculation, leading to the development of the DP-RL model. The second sub-case evaluates the DP-RL approach against different RL methodologies (Bayesian Theorem, Roth-Erev model, Q-Learning, Double DQN, Ensemble-based RL, DRL, and Actor-Critic RL), to assess how the DP-RL method aligns with current techniques. The third case study compares various feature usages for each energy production source, starting from a base to identify the best performance. In the fourth case study, real-world conditions are reflected across two scenarios: one using the DP-RL model and the other employing more robust contexts, using the DPC-RL methodology.

All case studies were carried out on a Windows 11 Pro computer equipped with an Intel Core i5 (13th Gen) processor, 16 GB of RAM, and integrated Intel Iris Xe Graphics. A market price proposal session occurs daily, during which the system makes 7 forecasting per period. There are 24 periods each day, and each session takes about 45 seconds to complete.

4.1 Baseline

This case study is split into two sub-studies: the monthly forecasting analysis and the annual forecasting study. The first used reduced data, while the second used complete data already presented in the Data Pipeline section. The objective of this case study is to predict market prices, particularly in the Iberian daily market, providing users with more accurate predicted values of the market-clearing price. Both case studies utilise all the columns of the dataset as features, being 91 without the target column.

4.1.1 Monthly forecasting

Setup

In this study, the various models were trained and tested over a shorter period. Following basic data preparation, the forecasting methods were tuned using data from January to March 2024. Table 6 presents the hyperparameters selected for each forecasting model after the tuning process.

Table 6 - Case Study 4.1.1, hyperparameters of forecasting models in the monthly forecasting study.

Model	Hyperparameters
RF	***'n_estimators'***: 200; ***'n_jobs'***: -1; ***'random_stat'***: 42
XGBoost	***'learning_rate'***: 0.1; ***'n_estimators'***: 200; ***'random_state'***: 42
RR	***'α'***: 1.0; ***'copy_X'***: True; ***'fit_intercept'***: True; ***'max_iter'***: 2000 ; ***'positive'***: True; ***'random_state'***: 42; ***'solver'***: 'auto'; ***'tol'***: 0.0001
SGD	***'loss'***: 'squared_error'; ***'max_iter'***: 1000; ***'penalty'***: 'l2'; ***'tol'***: 0.0001 ; ***'warm_start'***: False
LSTM	**'n_celss_with_returning_sequences'**: 64; **'n_celss_with_not_returning_sequences'**: 64; **'dropout'**: 0.2; **'dense'**: 128
DNN	**'activation'**: 'elu', **'optimizer'**: 'adam', **'neuron'**: 128, **'dropout'**: 0.2

After the tuning, the forecast models were applied in sensitivity tests conducted on the CB-RL methodology to determine the optimal hyperparameter. This involved training the forecasting models in January and the CB-RL approach in February 2024, with the CB-RL forecasting conducted in March 2024. As illustrated in Table 7, three sensitivity tests were done, employing α percentages of 20%, 50%, and 80%.

Table 7 - Case Study 4.1.1, comparison of the CB-RL model performance for each hyperparameter in the monthly forecasting study.

Sensitivity Test	MSE (€/MWh)	RMSE (€/MWh)	MAE (€/MWh)	R^2
α = 20%	56.8956	7.6878	3.1758	0.9598
α = 50%	61.2736	7.9898	3.6847	0.9545
α = 80%	73.8495	8.3764	3.7985	0.9512

In the process of analysing these sensitivity tests, four evaluation metrics are used, as shown in Table 7. The CB-RL model consistently demonstrated superior performance to the other forecasting methods, achieving the best results with α = 20%.

Results and Discussion

With the hyperparameters established for each forecasting method, it is safe to proceed to the execution of the validation test. It was performed using the forecast models trained from January to March 2024. These were used both in the training of the CB-RL methodology from April to May 2024, and in the testing of the CB-RL model from June 2024. After this validation test, Table 8 presents the performance of each method throughout June of 2024.

Table 8 - Case Study 4.1.1, evaluation and comparison of CB-RL with forecasting models in the monthly forecasting study.

Model	MSE (€/MWh)	RMSE (€/MWh)	MAE (€/MWh)	R^2
CB-RL	50.8164	7.1286	3.1479	0.9755
XGBoost	54.5114	7.3832	3.5470	0.9630
RF	73.9264	8.5980	5.2631	0.9498
RR	76.2892	8.7344	5.0942	0.9482
LSTM	218.535	14.9075	12.1127	0.9032

Model	MSE (€/MWh)	RMSE (€/MWh)	MAE (€/MWh)	R^2
DNN	255.926	15.4391	19.7845	0.8970
SGD	356.997	16.8374	20.1983	0.8613

Table 8 illustrates that the CB-RL approach outperformed the forecasting models based on the evaluation metrics, recording the lowest error values for MSE, RMSE, and MAE, as well as achieving the highest R^2 value. It also highlights that the CB-RL forecasts came very close to the actual values throughout June, with an average variance of around 3.15 €/MWh (see the MAE column).

To assess the CB-RL model's performance, the RMSE metric is more effective at identifying the worst and best-performing days, when errors are larger. Figure 20 displays the RMSE values for June 2024.

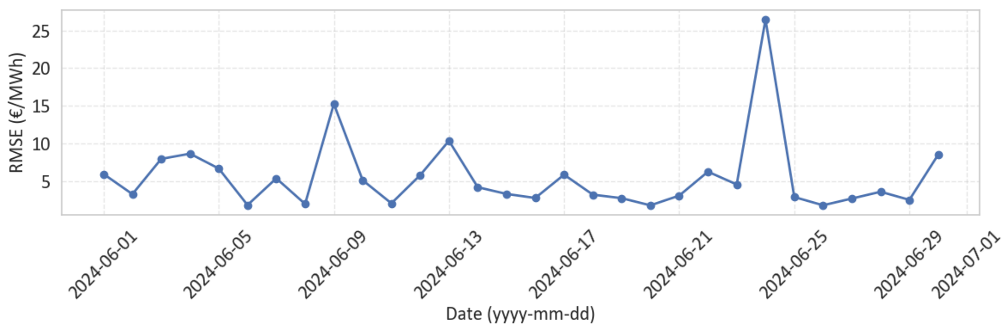

Figure 20 - Case Study 4.1.1, RMSE through June 2024 in the monthly forecasting study.

As illustrated in Figure 20, RMSE values remained low throughout the month of June, with most days recording values below 10 €/MWh. This indicates that the forecasting models chosen by the CB-RL achieved good accuracy over time, demonstrating consistent performance. The day with the worst performance was June 24th, which had an RMSE of around 26.45 €/MWh, while the best day was June 20th, with an RMSE of approximately 1.77 €/MWh.

The forecasting models chosen by the CB-RL are shown in Figure 21.

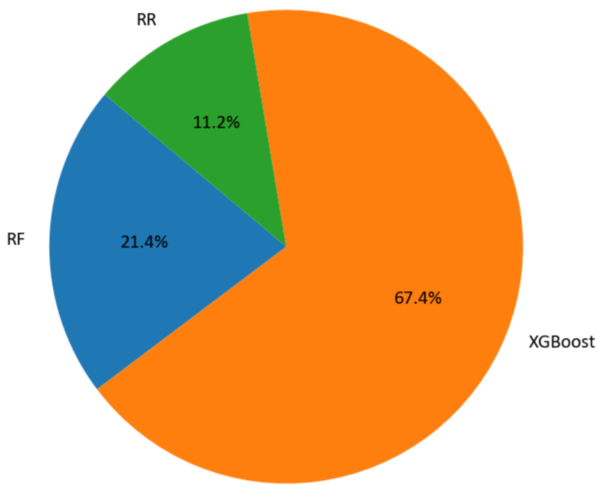

Figure 21 - Case Study 4.1.1, CB-RL Selection Distribution in the monthly forecasting study.

As shown in Figure 21, the forecasting model most frequently chosen by the CB-RL methodology was XGBoost, selected 67.4% of the time, followed by RF, which was chosen 21.4% of the time, and RR came in third place with 11.2%. The LSTM, DNN, and SGD were not selected during June 2024.

Figure 22 compares the performance of the forecasting models for June 24th, 2024.

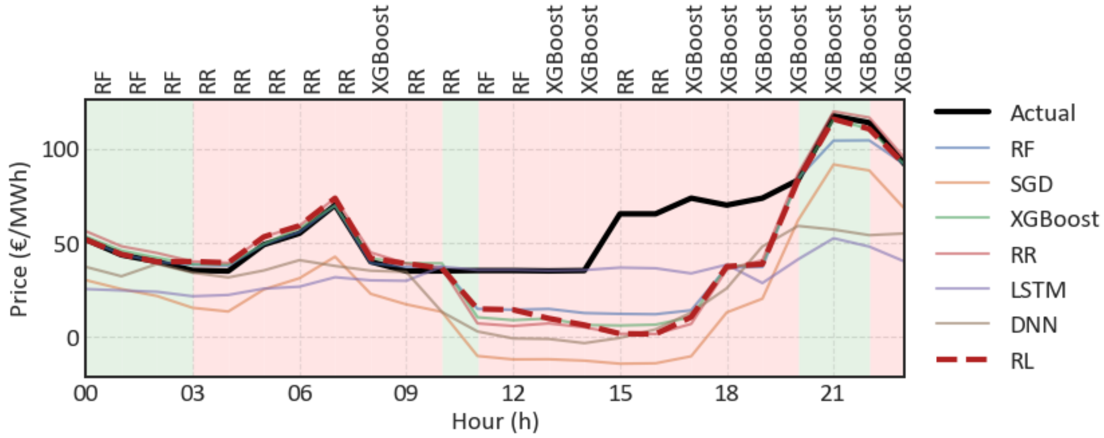

Figure 22 - Case Study 4.1.1, Forecasts vs Actual on the worst day (June 24th) in the monthly forecasting study.

The analysis of Figure 22 reveals that the XGBoost was the most chosen in 41.7% of the periods, followed by RR with 37.5% and RF with 20.8%. Although the CB-RL did not always choose the best-performing model, it often selected one with lower forecasting errors. Furthermore, since accumulated rewards drive CB-RL decisions, it tends to favour models with strong historical performance. This trend is evident in the periods 10 to 20, where price forecasts showed greater variation, leading the CB-RL methodology to select a model that did not adapt as well to the specific conditions of that day.

Figure 23 illustrates how competitive the selected forecast models are for this day.

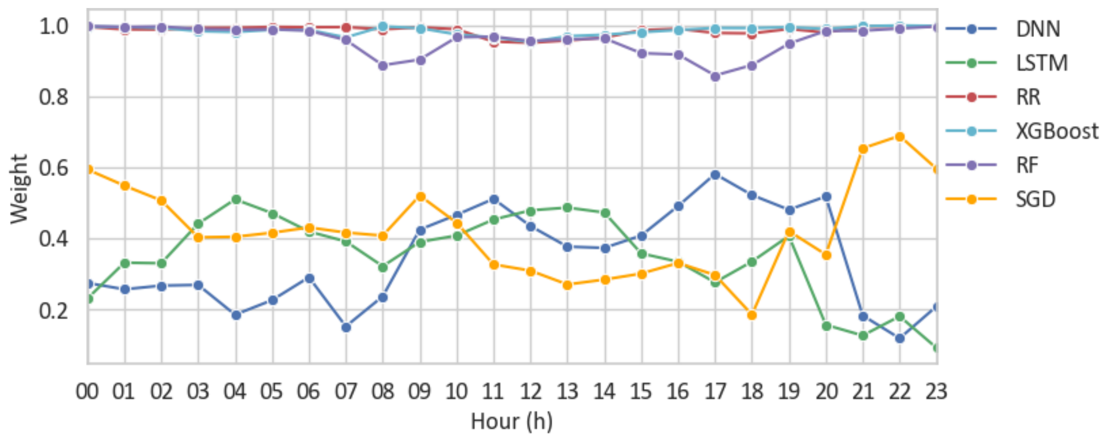

Figure 23 - Case Study 4.1.1, evolution of rewards for the worst day (June 24th) in the monthly forecasting study.

As depicted in Figure 23, the RR, RF, and XGBoost maintained weights consistently near the maximum value throughout the day. Between periods 10 and 20, LSTM was identified as the most suitable model based on Figure 22, even though its weight never went beyond 0.5. This is explained by the fact that the reward calculation for forecasting models relies on cumulative rewards.

Figure 24 shows the comparison between the predicted values of each model and the actual market-clearing price on June 20th, 2024. On that day, XGBoost was the most chosen by the CB-RL agent, being selected in 66.7% of the instances, while RF was chosen 25.0% of the times, and RR only 8.3%. The results indicate a strong preference for XGBoost, especially in the second half of the day, when it was selected almost exclusively.

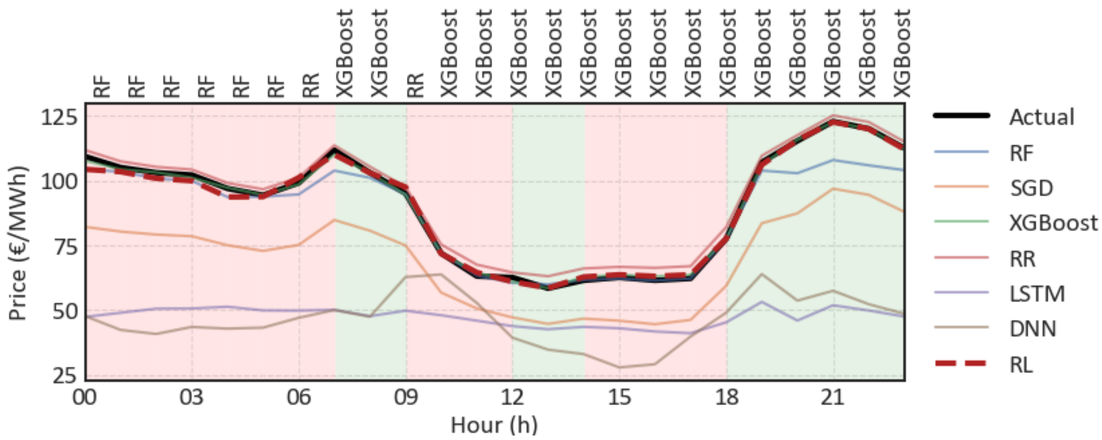

Figure 24 - Case Study 4.1.1, Forecasts vs Actual on the best day (June 20th) in the monthly forecasting study.

The CB-RL forecasting closely matched the actual values, as shown in Figure 24, suggesting that the CB-RL methodology effectively adjusted its selection strategy to maintain forecasting accuracy throughout the day. The consistent performance across different periods indicates that the CB-RL has successfully identified and enforced the use of forecasting models with more reliable historical performance.

The competitiveness among the models can be analysed by observing the evolution of the reward values assigned to each forecasting model, as shown in Figure 25.

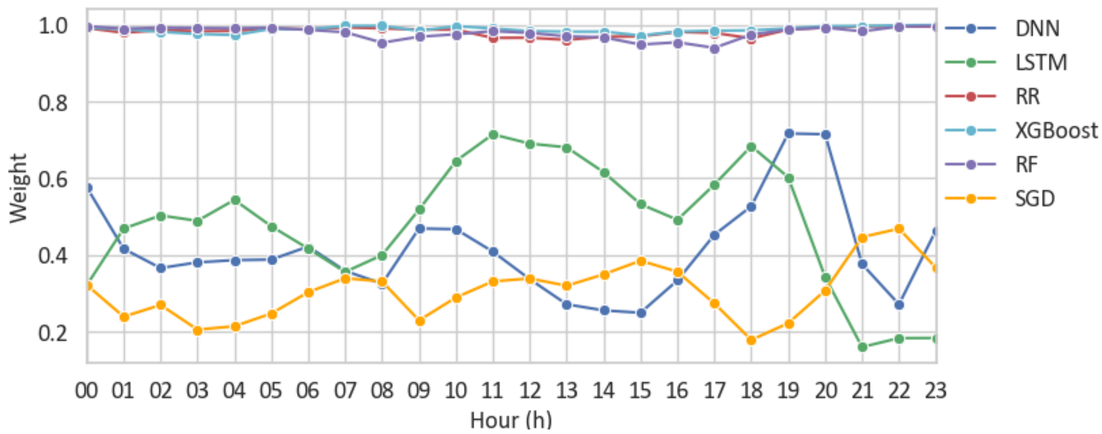

Figure 25 - Case Study 4.1.1, evolution of rewards for the best day (June 20th) in the monthly forecasting study.

As illustrated in Figure 25, the reward values for XGBoost, RF, and RR are similar, with one of these models being consistently selected by the CB-RL. In contrast, SGD, LSTM, and DNN are never selected by the CB-RL since their rewards are consistently ranked lower at each iteration.

4.1.2 Annual forecasting

Setup

First, the forecasting models undergo tuning, followed by sensitivity tests for the CB-RL model. The hyperparameters for the forecasting models are displayed in Table 9:

Table 9 - Case Study 4.1.2, hyperparameters of forecasting models in the annual forecasting study.

Model	Hyperparameters
RF	***'n_estimators'***: 50; ***'n_jobs'***: -1; ***'random_stat'***: 42
XGBoost	***'learning_rate'***: 0.2; ***'n_estimators'***: 200; ***'random_state'***: 42
RR	***'α'***: 1.0; ***'copy_X'***: True; ***'fit_intercept'***: True; ***'max_iter'***: 1000 ; ***'positive'***: True; ***'random_state'***: 42; ***'solver'***: 'auto'; ***'tol'***: 0.0001
SGD	***'loss'***: 'squared_error'; ***'max_iter'***: 2000; ***'penalty'***: 'l2'; ***'tol'***: 0.0001 ; ***'warm_start'***: False
LSTM	**'n_celss_with_returning_sequences'**: 64; **'n_celss_with_not_returning_sequences'**: 64; **'dropout'**: 0.3; **'dense'**: 128
DNN	**'activation'**: 'elu', **'optimizer'**: 'adam', **'neuron'**: 128, **'dropout'**: 0.3

The hyperparameters resulting from the tuning process are used in each forecasting model employed in the sensitivity tests conducted on the CB-RL approach. In Table 10, three sensitivity tests were performed to identify the best hyperparameter, with percentages of α: 20%, 50%, and 80%.

Table 10 - Case Study 4.1.2, comparison of the CB-RL model performance for each hyperparameter in the annual forecasting study.

Sensitivity Test	MSE (€/MWh)	RMSE (€/MWh)	MAE (€/MWh)	R^2
α = 20%	42.3608	6.5085	2.2503	0.9680
α = 50%	44.1121	6.6417	2.3301	0.9667
α = 80%	51.6461	7.1865	2.5623	0.9610

As seen in Table 10, the CB-RL stood out as the best in relation to the other forecasting models (see Position column), having the best performance with α = 20%.

Results and Discussion

With the hyperparameters defined, the validation test is executed, and it is possible to view the performance of each model throughout 2024 in Table 11.

Table 11 - Case Study 4.1.2, evaluation and comparison of CB-RL with forecasting in the annual forecasting study.

Model	MSE (€/MWh)	RMSE (€/MWh)	MAE (€/MWh)	R^2
CB-RL	30.3543	5.5095	2.0933	0.9854
RF	31.3066	5.5952	2.1143	0.9849
XGBoost	42.2686	6.5014	2.5011	0.9796
RR	58.7581	7.6654	4.7361	0.9717
LSTM	162.8445	12.7611	8.4785	0.9216
DNN	174.3808	13.2053	9.8035	0.9160
SGD	228.8464	15.1277	10.3181	0.8898

The CB-RL methodology presented the best performance among the others, as shown in Table 11, although the other methods performed well.

In Figure 26, the evolution of the MAE per month throughout the year 2024 is visible. This graph illustrates the absolute average difference between the value predicted by the model chosen by CB-RL and the real value.

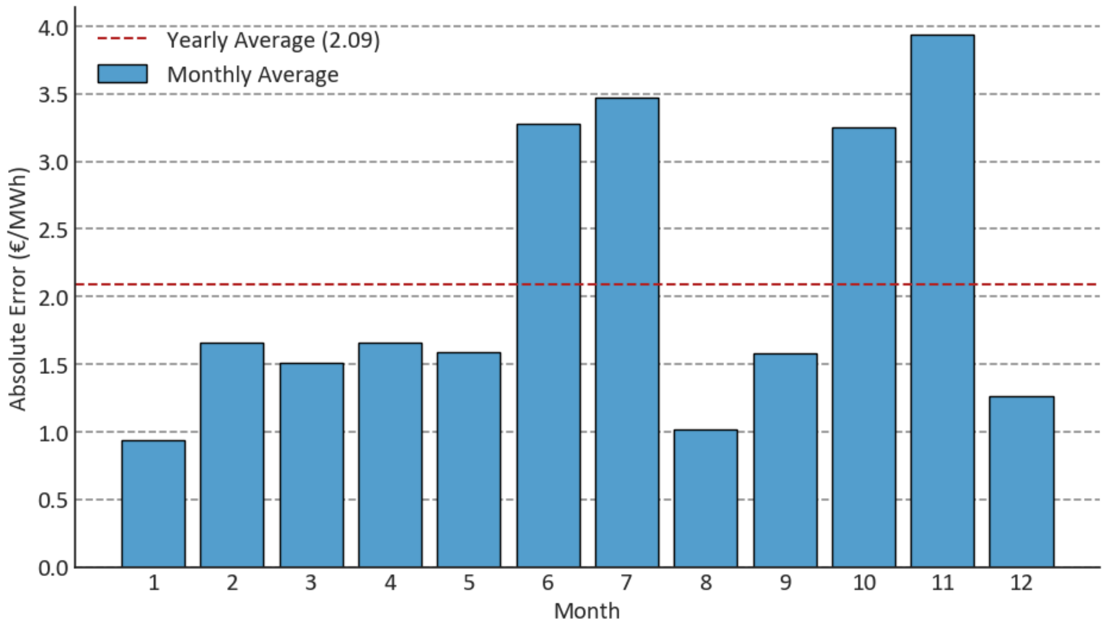

Figure 26 - Case Study 4.1.2, Monthly vs Yearly MAE in the annual forecasting study.

Figure 26 demonstrates how close the CB-RL was to the real value throughout the year, with just a 2.09 €/MWh difference. The worst month was November, with an MAE value of 3.94, and the best month was January, with an MAE value of 0.94. This highlights the CB-RL perception by selecting the top forecasting models for each period. In Figure 27, the distribution of forecasting model selection by the CB-RL is represented.

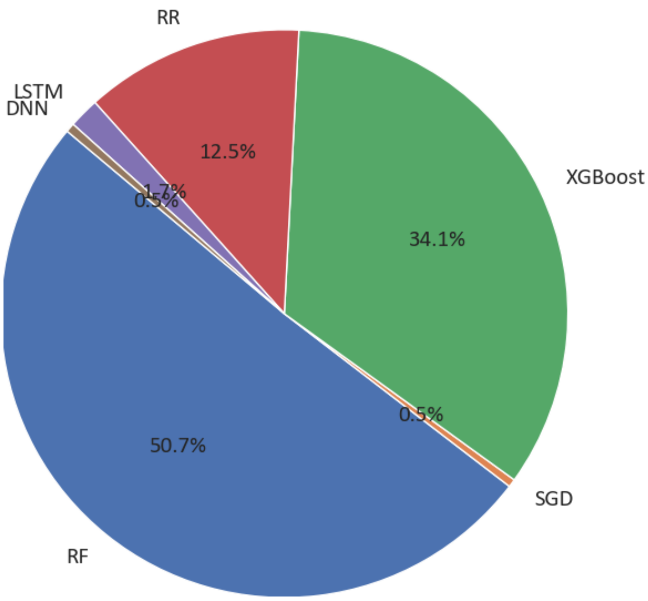

Figure 27 - Case Study 4.1.2, CB-RL Selection Distribution in the annual forecasting study.

It is possible to see in Figure 27 that the CB-RL mainly chose the RF with 50.7%, followed by the XGBoost with 34.1% and the RR with 12.5%. Finally, the last ones chosen were the LSTM, SGD, and DDN, with 1.7%, 0.5%, and 0.5%, respectively. The selection of forecasting models perfectly reflects the order of performance represented in Table 11.

Figure 28, shows how effective the selection of the CB-RL approach is.

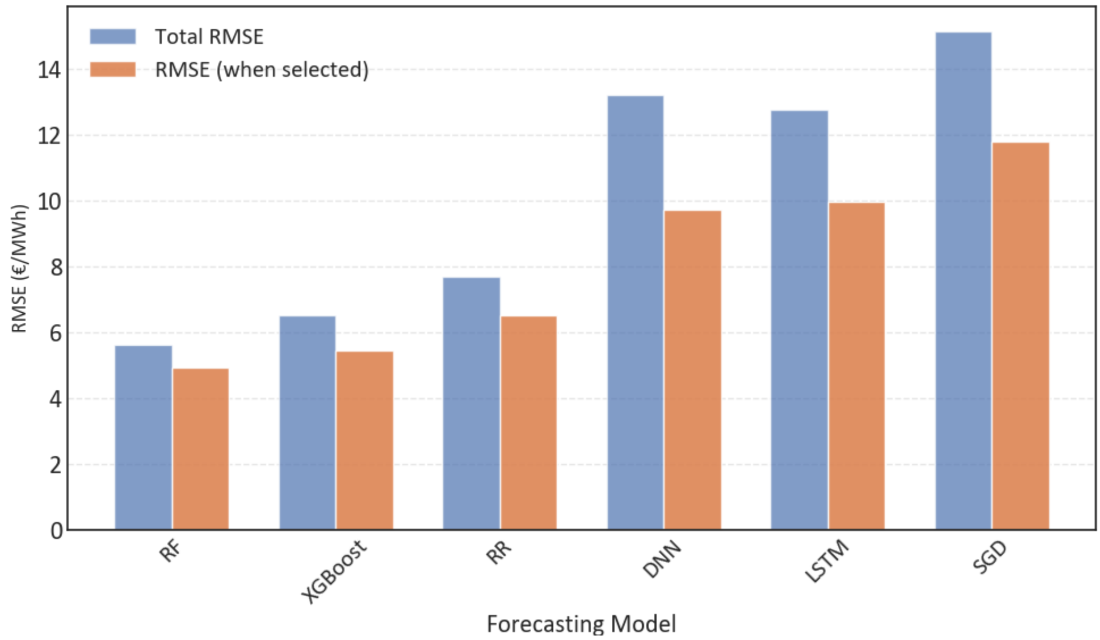

Figure 28 - Case Study 4.1.2, RMSE per Model - Total vs When selected by CB-RL in the annual forecasting study.

In Figure 28, to evaluate the performance of the CB-RL approach, the RMSE metric is used, because it can better identify when there are large differences in errors. This metric is applied to all forecasting models throughout the year, including the value of the same metric at the time the CB-RL chose them. CB-RL effectively selects the forecast models, as it chooses them when they have the lowest RMSE. Still, the total RMSE of the forecasting models was low.

The same metric was used to evaluate the worst and best day of the year, for the worst and best month identified in Figure 26. For the worst month, the worst day was November 24th, with an RMSE value of 21.71 €/MWh. On the other hand, in the best month, the best day was January 21st, with an RMSE value of 0.26 €/MWh. Figure 29 represents the worst day of the year, with the x-axis indicating the time of day, while the y-axis shows the actual and predicted values.

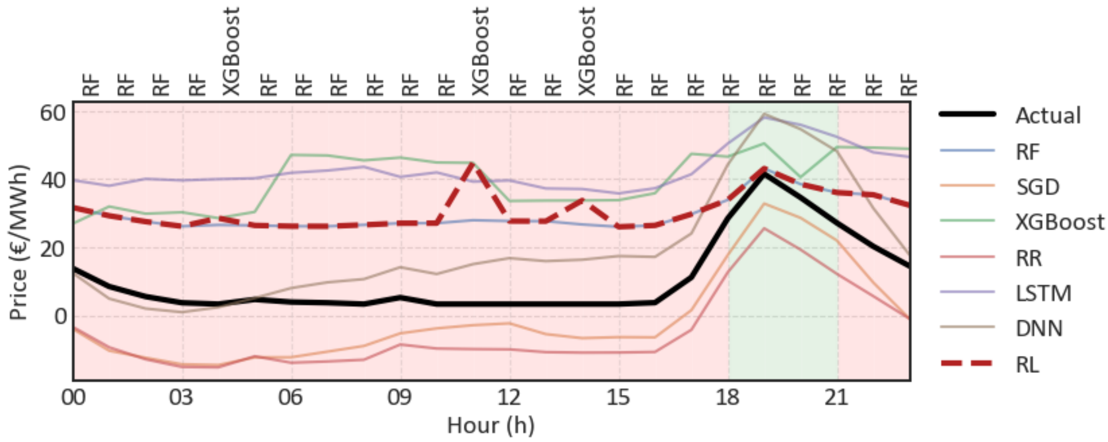

Figure 29 - Case Study 4.1.2, Forecasts vs Actual on the worst day (November 24th) in the annual forecasting study.

As shown in Figure 29, only in 12.50% of the day, the CB-RL choose the model that was closest to the real value. Although in period 11 the CB-RL chose the worst one (XGBoost), in the other periods in which it did not choose the best, the CB-RL managed to choose intermediate methods (RF and XGBoost). Thus, it indicates that the choice of the CB-RL was not perfect but ultimately proved optimal. Figure 30 shows the evolution of rewards on this day.

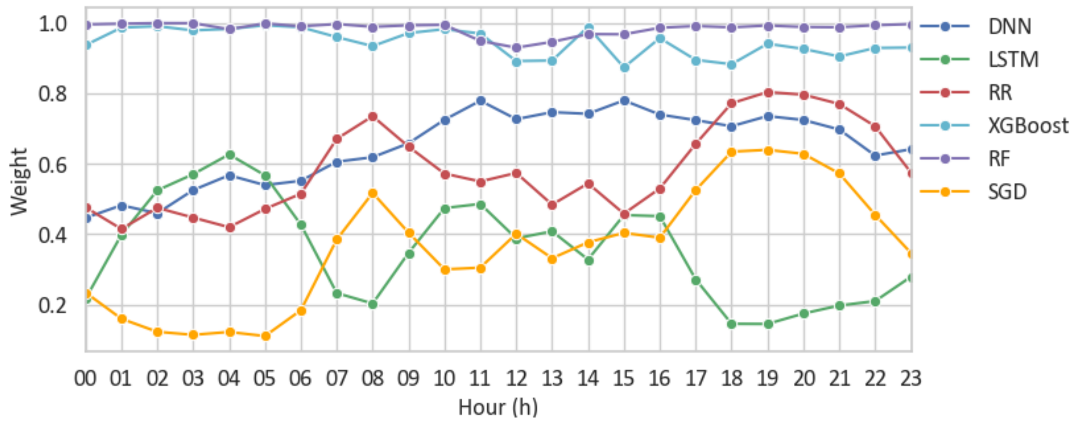

Figure 30 - Case Study 4.1.2, evolution of rewards for the worst day (November 24th) in the annual forecasting study.

As shown in Figure 30, only the XGBoost and RF achieved rewards above 0.8, indicating the poor historical performance of the other models. Despite their better historical results, RF and XGBoost were not the top choices on this particular day.

Figure 31 is similar to Figure 29 but illustrates the best day of the year.

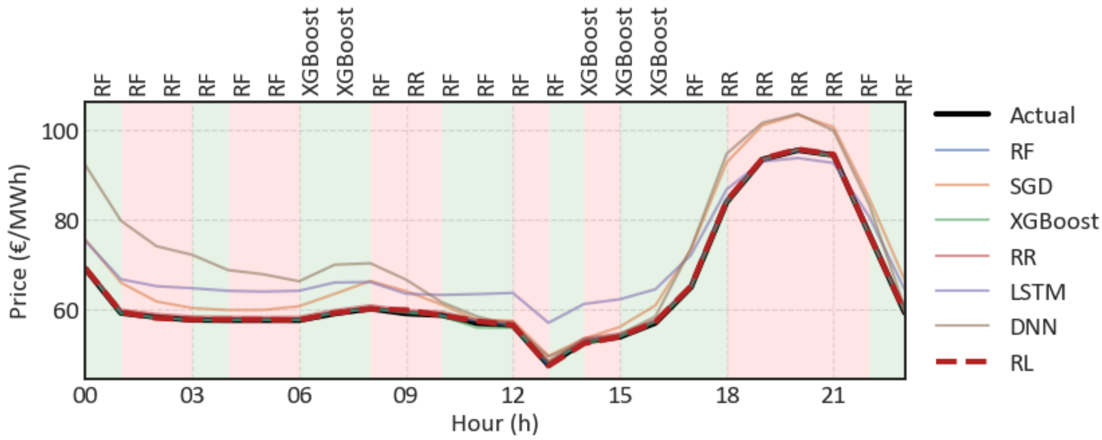

Figure 31 - Case Study 4.1.2, Forecasts vs Actual on the best day (January 21st) in the annual forecasting study.

In Figure 31, it is noticeable that the forecasting models selected by CB-RL consistently show very minimal differences compared to the actual values. On this day, despite the forecasting models closely matching the real values, the CB-RL achieved an accuracy of 45.83% in its selections for each period. In the periods when the CB-RL did not select the best forecasting model, the one chosen by CB-RL was practically similar to the best one.

Competitiveness of methods is demonstrated in Figure 32, which shows the evolution of the rewards throughout the day.

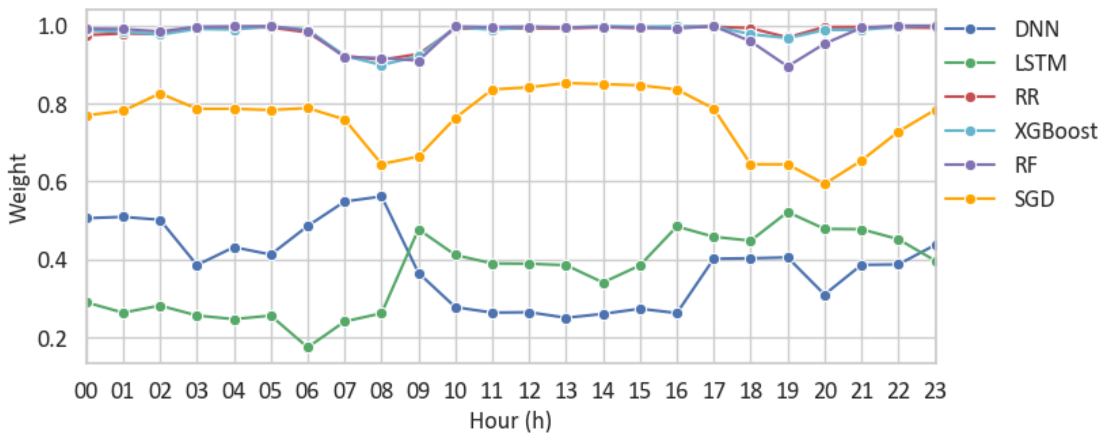

Figure 32 - Case Study 4.1.2, evolution of rewards for the best day (January 21st) in the annual forecasting study.

As shown in Figure 32, the RF, XGBoost, and RR were always very close to being chosen by CB-RL. The three models were also very close to 1, meaning that they were, at the time, the best ones to be selected.

4.2 Comparison of RL methodologies

This study is essential for improving the RL approach used in this thesis, aiming to make it more robust, automated, and improve forecasting accuracy. Additionally, comparing the RL methodology with existing approaches is crucial for identifying gaps and differences, as well as assessing how this thesis's RL method aligns with current techniques.

This case study is divided into two sub-studies: one on the most effective RL approach and the other comparing various RL methodologies. The first focuses on developing the CB-RL approach, as addressed in the Baseline case studies. On the other hand, the second sub-study aims to compare the best RL approach, identified in the first sub-study, with existing RL methodologies from the literature.

General Setup

The data used in this case study undergoes a process of selecting the best features to use in the various forecasting models. This process is beneficial as it provides models with clear data information that is most closely correlated to the target column, which is: "mcp-pt-(d,h)". The column "mcp-es-(d,h)", which represents the Spanish market price on the same day, is not considered a resource because it is not possible to obtain this information when forecasting. Figure 33 shows the correlation matrix.

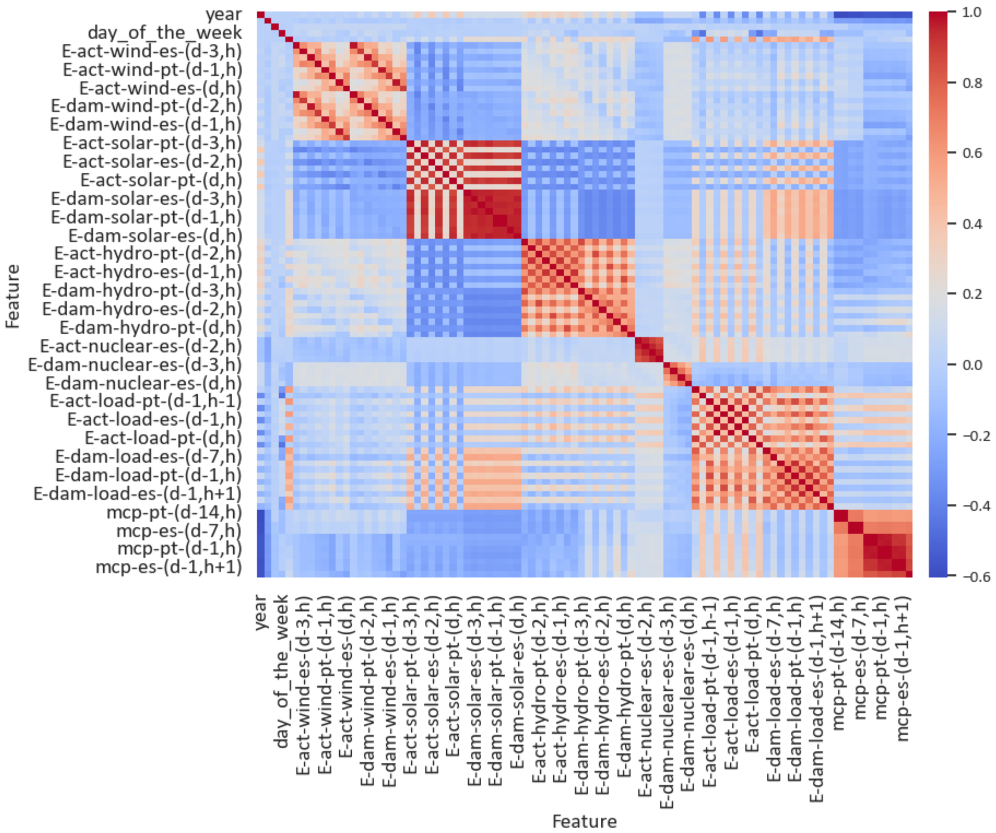

Figure 33 - Case Study 4.2, correlation matrix of the dataset for comparing RL methodologies.

Figure 33 highlights the connections between the columns of the dataset. However, not all features are visible, as the dataset is large. Nevertheless, a correlation can be observed between features in the dataset and the target column. The figure shows values ranging from -0.6 (bluish colour) to 1 (reddest colour), with the highest correlation near 1 and the lowest near -0.6. It is visible that the strongest correlation of mcp-pt-(d,h) is the mcp of the previous days, followed by some energy sources. However, it is not clear which ones, and for this reason, it is necessary to use the F-score technique for better selection of features. The F-score measures the importance of each feature related to the target column by calculating the variance between the different classes. The higher the F-score, the greater the correlation with the target column, offering better conditions for model forecasting. Since the dataset is very large, a minimum F-score of 100 was set for all features, filtering only the relevant ones. Figure 34 illustrates the ranking of selected features.

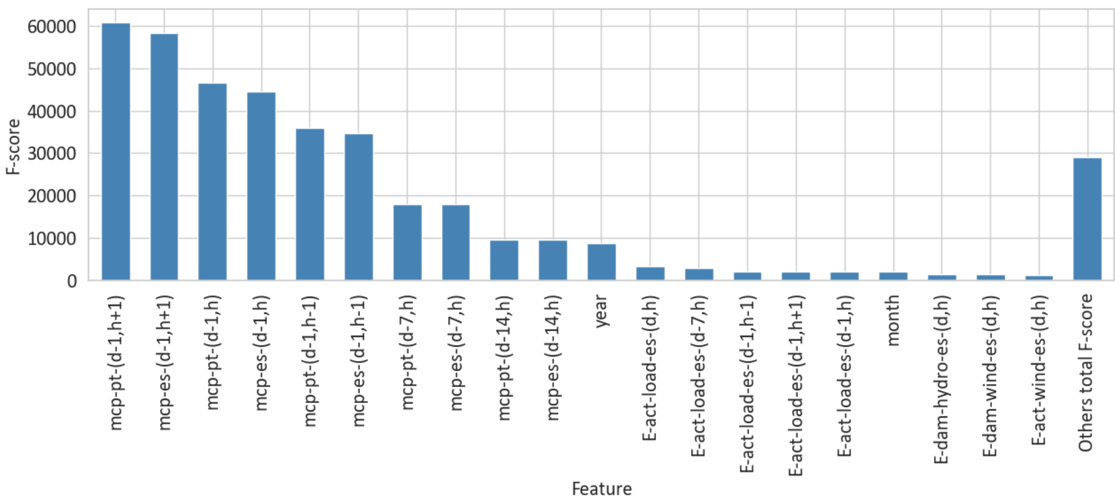

Figure 34 - Case Study 4.2, ranking of selected features of the dataset for comparing RL methodologies.

In Figure 34, it is possible to see 20 variables along with 61 in the “Others total F-score” column on the x-axis. The most important are the MCP variables for Portugal and Spain, from the previous days, with greater emphasis on Portuguese market prices. Other variables with good correlation with the target column are the year, the month, the energy load of Spain on the same day and previous days. To improve the graph’s visual clarity, a single column titled “Others total F-score” has been added, showing the combined significance of all other variables. This column includes data on solar, wind, and hydroelectric energy from Portugal and Spain, along with nuclear energy from Spain. All features represented on the x-axis in Figure 34 are used as a new dataset, incorporated into the forecasting models in the following two subsections.

After data preparation, the process of tuning the forecasting models begins, and the resulting hyperparameters are shown in Table 12:

Table 12 - Case Study 4.2, hyperparameters of forecasting models for comparing RL methodologies.

Model	Hyperparameters
RF	***'n_estimators'***: 200; ***'n_jobs'***: -1; ***'random_stat'***: 42
XGBoost	***'learning_rate'***: 0.2; ***'n_estimators'***: 100; ***'random_state'***: 42
RR	***'α'***: 1.0; ***'copy_X'***: True; ***'fit_intercept'***: True; ***'max_iter'***: 1000; ***'positive'***: True; ***'random_state'***: 42; ***'solver'***: 'auto'; ***'tol'***: 0.0001
SGD	***'loss'***: 'squared_error'; ***'max_iter'***: 1000; ***'penalty'***: 'l2'; ***'tol'***: 0.0001; ***'warm_start'***: False
LSTM	**'n_celss_with_returning_sequences'**: 64; **'n_celss_with_not_returning_sequences'**: 64; **'dropout'**: 0.2; **'dense'**: 128
DNN	**'activation'**: 'elu', **'optimizer'**: 'adam', **'neuron'**: 256, **'dropout'**: 0.2

The hyperparameters for the forecasting models are used in the RL sensitivity test, as well as in the validation test carried out in the following two subsections.

4.2.1 Improvement of the RL reward calculation

Setup

To improve the RL methodology, eight different approaches were tested and compared with the CB-RL. To calculate the reward, all strategies follow Eq. (2) and (3), but with variances in the calculation of α. The approaches are outlined in the list below:

1) **Same α per period, using the worst approach:** The RL utilised the reward from the worst forecasting model at that moment to determine the best α for each period.
2) **Same α per period, using the best approach:** The logic follows the 1) attempt, but uses the best forecasting model in each period for decision α.
3) **Different α per period and model:** The RL employed the reward from each forecasting model to identify the best α for each period.
4) **Use of K-Means:** This aims to identify contextual patterns, creating a specific number of groups. The number of defined groups will replace the number of periods for each day, becoming the new temporal segmentation. A methodology called Silhouette Score was applied to determine the optimal number of groups based on the features. To this end, the validation test was conducted using only the target column and then with the target plus additional features.
    a. **Target column:** two groups of contexts were identified, with a score of 0.76.
    b. **Target column + 81 features:** three groups of contexts were identified, with a score of 0.84.
        i. **With static α at 20%:** to calculate the reward of each forecasting model, the Baseline case study is used as a reference for the α value.
        ii. **Same α per period, using the worst approach:** uses the 1) attempt approach to calculate the reward for each forecasting model.

    The attempts a. and b. are combined into separate tests with approaches i and ii.
5) **DP-RL**

To test all these approaches, a sensitivity test was performed. The CB-RL was also tested, being used as a reference to verify whether the proposed methodologies surpass its performance. The results are presented in Table 13, with the approaches ordered from best to worst.

Table 13 - Case Study 4.2.1, comparison of RL approaches for improving the calculation of the RL reward.

RL sensitivity test	Better than baseline?	RL rank	MSE (€/MWh)	RMSE (€/MWh)	MAE (€/MWh)	R^2
5)	Yes	1º	33.3465	5.2743	2.0007	0.9712
CB-RL	-	1º	47.9868	6.9869	2.3536	0.9673
1)	No	1º	50.0910	7.0775	2.4621	0.9622
3)	No	1º	61.6461	8.1865	3.5623	0.9510
2)	No	1º	83.3547	9.3480	3.8096	0.9350
4) b. i.	No	3º	99.1634	12.9581	6.0865	0.9251
4) a. i.	No	3º	126.4698	14.6534	8.9384	0.9074
4) b. ii.	No	3º	189.9573	15.9489	10.8578	0.8845
4) a. ii.	No	3º	205.2384	17.4564	13.9357	0.8687

After all the approaches were carried out, only one stood out as the best in relation to the methodology used in the baseline study (CB-RL), as demonstrated in Table 13. The best one was the 5) approach, the DP-RL. Only approach 1) came very close to the CB-RL methodology, while methodologies 3) and 2) performed slightly worse. On the other hand, none of the four approaches that used K-Means were successful, since the methodology did not outperform the forecasting models used in the sensitivity tests. Intending to find an automatic and dynamic RL methodology, the DP-RL proved to be the most appropriate.

Results and Discussion

With the hyperparameters defined for the forecasting models and a better approach found as the DP-RL, the validation test is executed. Table 14 shows the performance evaluation of each method.

Table 14 - Case Study 4.2.1, evaluation and comparison of DP-RL (the best performing approach in this sub-case study) with forecasting models.

Model	MSE (€/MWh)	RMSE (€/MWh)	MAE (€/MWh)	R^2
DP-RL	5.4006	2.3239	0.5627	0.9974
RF	6.0602	2.4617	0.5927	0.9969
XGBoost	6.8816	2.6233	0.6131	0.9967
RR	19.5186	4.4180	1.9692	0.9906
LSTM	20.9876	4.5812	1.6899	0.9899
DNN	29.1243	5.3967	3.4880	0.9860
SGD	52.1734	7.2231	5.3390	0.9749

As shown in Table 14, the DP-RL obtained good metrics, being the best among the others despite the satisfactory performance by the rest.

The evolution of the monthly MAE throughout 2024 is illustrated in Figure 35, which demonstrates the absolute average difference between the predicted value by DP-RL and the actual value.

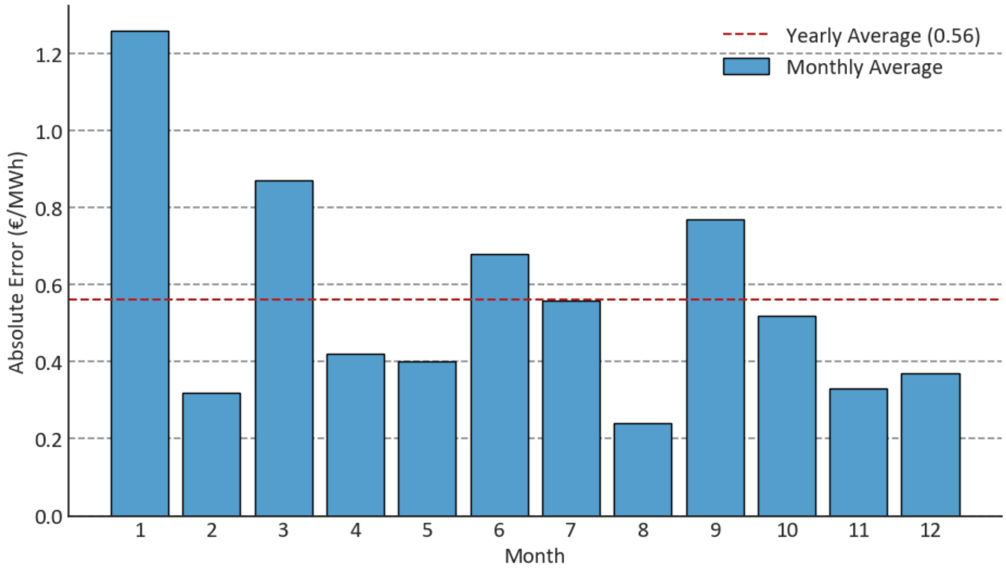

Figure 35 - Case Study 4.2.1, Monthly vs Yearly MAE using DP-RL.

The difference between the value predicted by the DP-RL compared to the actual value throughout the year was very small, at 0,56 €/MWh, as shown in Figure 35. The worst month was January, with the highest MAE value at 1.26 €/MWh. However, the best month was August, with the lowest MAE value at 0.24 €/MWh.

Figure 36 represents the selection number of each forecasting model by the DP-RL. Note that the year 2024 has 8784 hours, due to a leap year.

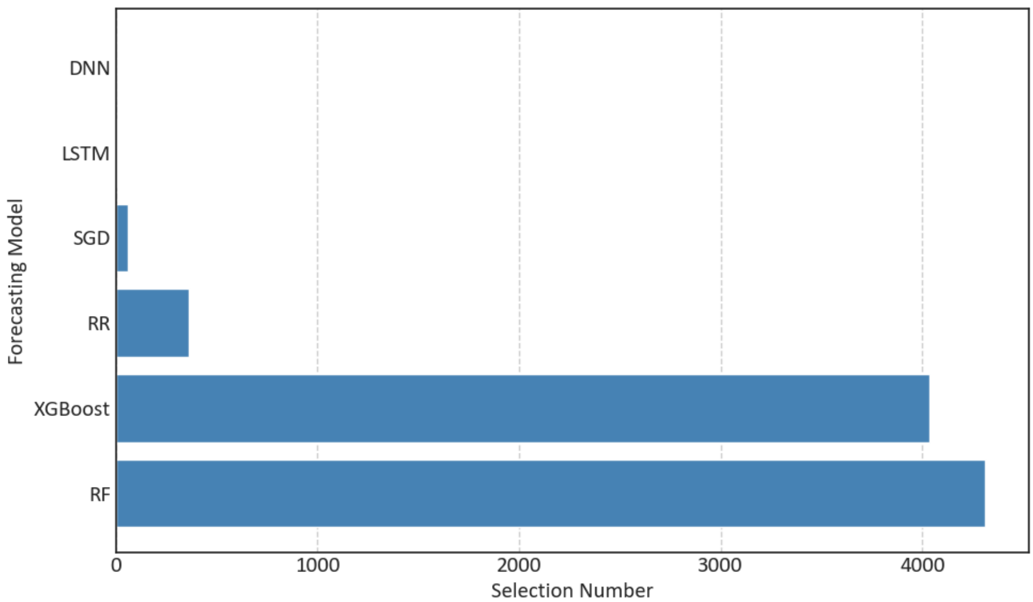

Figure 36 - Case Study 4.2.1, DP-RL Selection Distribution.

In Figure 36, it is noticeable that the DP-RL primarily preferred the RF, which was selected in 4313 periods, followed by XGBoost, chosen in 4038 periods. Conversely, the least selected were RR, SGD, LSTM, and DNN, which were chosen in 361, 62, 6, and 4 periods, respectively. This selection of forecasting models almost aligns with the performance order shown in Table 14. However, the SGD, which performed the worst, was the most selected compared to the LSTM and DNN.

The RMSE metric measures how the DP-RL performs with deviations from the actual values. In Figure 37, this metric is applied to all forecasting models throughout the year, including their value at the time the DP-RL selected them.

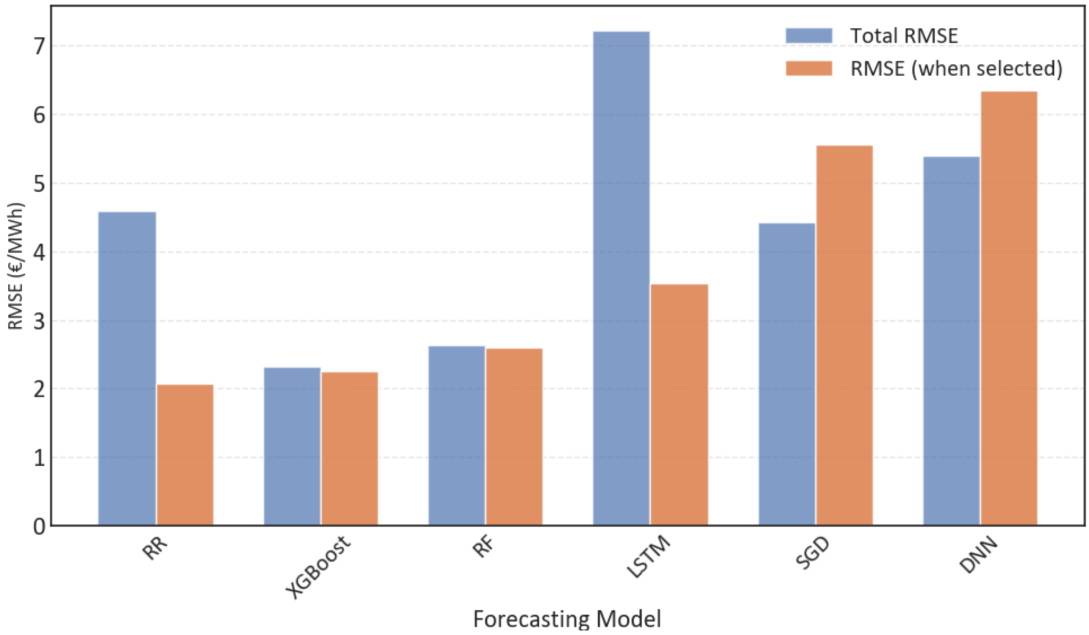

Figure 37 - Case Study 4.2.1, RMSE per Model - Total vs When selected by DP-RL.

In Figure 37, it is possible to see how low the total RMSE values were for each forecasting model. With this, the DP-RL was able to select the forecasting models effectively. It chose better the RR and LSTM, as evidenced by the significant difference in average RMSE compared to when they were selected. However, it selected the SGD and DNN when they had a slightly higher RMSE value.

For a more in-depth analysis, RMSE was used to evaluate the worst and best day of the year, for the worst and best month, respectively. For the worst month, the worst day was January 6th, with an RMSE value of 13.60. On the other hand, in the best month, the best day was August 13th, with an RMSE value of 0.10.

The worst day of the year is represented in Figure 38. The x-axis represents the time of day, and the y-axis indicates both the actual and predicted prices by each model.

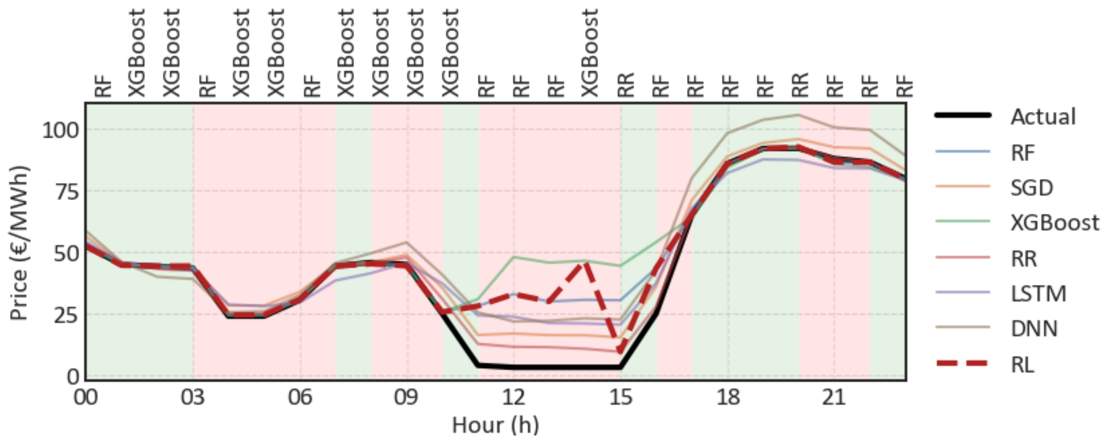

Figure 38 - Case Study 4.2.1, Forecasts vs Actual on the worst day (January 6th) using DP-RL.

The DP-RL chose the forecasting model that was closest to the real value in 45.83% of the 24 periods of the day, as illustrated by Figure 38. Between periods 10 and 15, the DP-RL moved considerably away from the real value, as well as choosing the worst model (XGBoost) in period 14. In the remaining periods, the DP-RL managed to follow the real value closely, thus indicating that the choice of the DP-RL was not always perfect, but it was considerably optimal.

To understand the DP-RL choices a little better, Figure 39 illustrates the evolution of the rewards for this day.

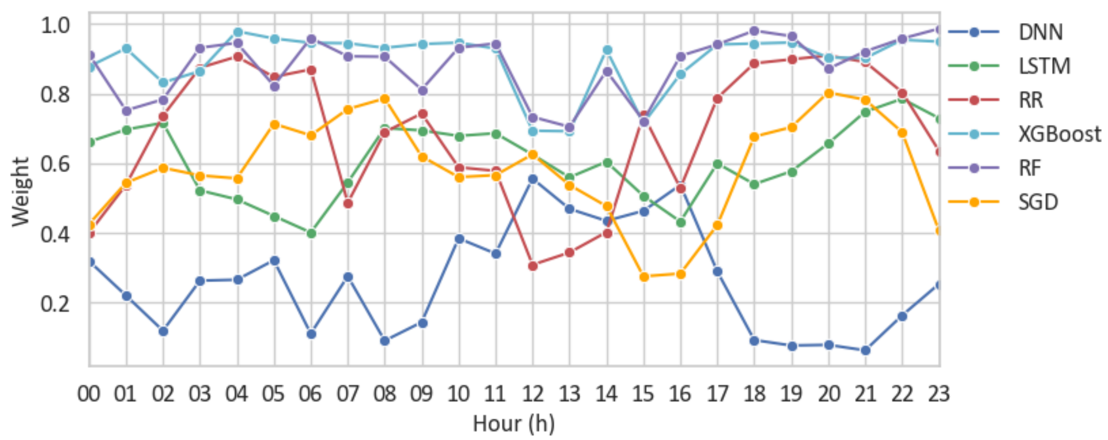

Figure 39 - Case Study 4.2.1, evolution of rewards for the worst day (January 6th) using DP-RL.

In Figure 39 it is possible to see that between the periods (0 to 9 and 16 to 23), the highest reward value in each period never dropped below 0.8. However, in 3 periods (12, 13 and 15), where the DP-RL deviates most from the real value, the rewards were below 0.8, indicating low accuracy of the forecasting models in these periods. However, in periods 10, 11 and 14, the rewards were higher than 0.8, indicating that they were performing well in these periods in the previous days, but on January 6th they were not successful. Note that these rewards are updated at the end of the day, resulting in a one-day disadvantage.

Figure 40 is similar to Figure 38 but illustrates the best day of the year.

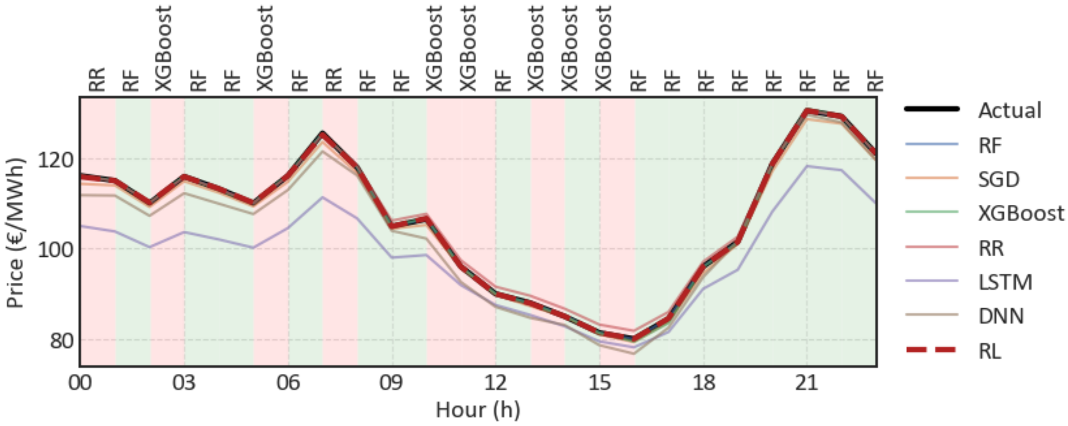

Figure 40 - Case Study 4.2.1, Forecasts vs Actual on the best day (August 13th) using DP-RL.

The forecasting models that the DP-RL chose always show a very minimal difference from the actual value, as shown in Figure 40. On this day, the DP-RL had an accuracy of 66.67% in choosing the best forecasting model in each period.

This day demonstrates the competitiveness and quality of each forecasting model, as represented by the rewards evolution in Figure 41.

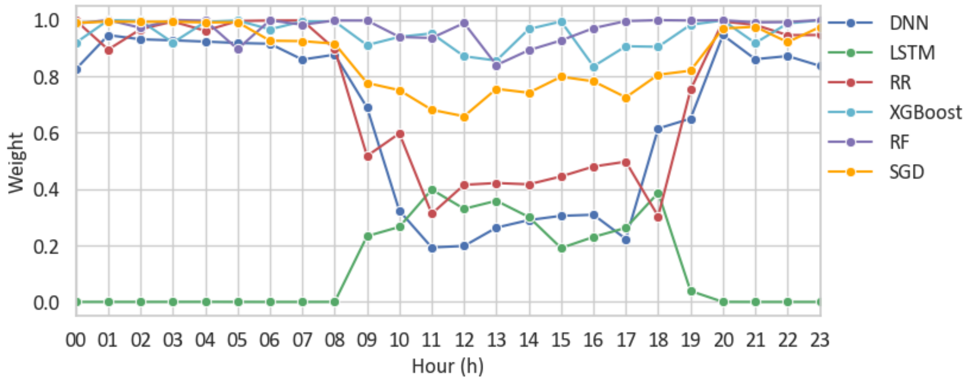

Figure 41 - Case Study 4.2.1, evolution of rewards for the best day (August 13th) using DP-RL.

As observed in Figure 41, in the initial and final periods of the day, all forecasting models presented rewards greater than 0.8, with the possibility of being selected, with the exception of the LSTM. However, during periods 10 to 18, only two (RF and XGBoost) demonstrated viability of being chosen, since their rewards were above 0.8.

4.2.2 Comparison with different RL methodologies

In addition to the DP-RL, seven RL methodologies were used for comparison. Among which three older models and four newer ones.

Setup

The three oldest RL methodologies are the Bayesian Theorem, Roth-Erev model, and Q-Learning. The most recent ones are Double DQN, Ensemble-based RL, DRL, and Actor-Critic RL. These RL approaches were adapted to accomplish the objectives of this thesis by calculating their rewards for each forecasting model during each daily period. Among these, only the Bayesian Theorem, the Roth-Erev model, and Ensemble-based RL did not require significant modifications to achieve the goals of this dissertation. For Q-Learning, each state represented a specific time period during the day, with the action being the selection of a forecasting model for that period.

The Double DQN was originally designed to select the best-suited one among three forecasting models for each month, as used in the study by Pannakkong *et al.* . However, in this dissertation, it is employed daily to choose the best among six different forecasting models. The DRL integrates multiple forecasting models, each with a state vector containing current context information such as forecasts and recent errors. Every daily period, a neural network predicts the best action by choosing a forecasting model and updates the reward according to its accuracy. Actor-Critic RL considers each time period in a day as a separate state. The actor then generates a probability distribution over the available forecasting models.

Results and Discussion

Sensitivity tests were conducted, with results being displayed in Table 15, ordered by the best-performing RLs.

Table 15 - Case Study 4.2.2, comparison of different RL methodologies with DP-RL.

Model	RL rank	MSE (€/MWh)	RMSE (€/MWh)	MAE (€/MWh)	R^2
DP-RL	1º	33.3465	5.7745	2.0007	0.9712
Bayesian	1º	38.3602	6.3404	2.3254	0.9682
Double Q-Learning	1º	42.9549	6.5968	2.4320	0.9660
Q-Learning	1º	45.2548	6.7264	2.4635	0.9646
Roth-Erev	1º	46.9072	6.8489	2.4856	0.9640
DRL	3º	62.8543	7.9281	3.2995	0.9525
Actor-Critic RL	4º	181.2730	12.8301	7.1645	0.8756
Ensemble-Based RL	4º	220.2959	14.8424	8.3336	0.8436

Table 15 presents the results of the sensitivity test for various RL methodologies, where none of the other RL methods managed to surpass the DP-RL. This is partly because external RL methodology covers different areas, while the one designed in this dissertation is focused and tested on Iberian market prices. However, three RL methods came relatively close to the DP-RL, namely: Bayesian, Double Q-Learning, Roth-Erev. They scored highest in their respective sensitivity tests compared to the individual forecast models (see RL first? column). The DRL also closely resembled the DP-RL and performed well, but only ranked as the third-best one in its sensitivity test. The RL models that performed the worst were the Actor-Critic RL and the Ensemble-Based RL, which ranked fourth in their respective sensitivity tests. Therefore, it can be concluded that the DP-RL model beats the others. This highlights its robustness and effectiveness, making it suitable for the context of this dissertation.

The DP-RL closely resembles the Actor-Critic RL, since both incorporate an accumulated reward and a penalty component. However, the Actor-Critic RL was significantly worse than the DP-RL. The Q-Learning and Roth-Erev share some similarities with the DP-RL, mainly in their use of a cumulative reward. Other ones differ entirely, as the DP-RL method does not use neural networks like DRL, Bayesian uncertainty, or variance-based exploration strategies, such as in Ensemble-Based RL.

4.3 Best-performing energy production sources

This case study evaluates the performance of each energy production source in relation to the actual Iberian market price. However, it does not utilise a general selection of resources, as in case study 4.2. Instead, it uses a fundamental set of resources, allowing each source of energy production to contribute individually, making its impacts more evident.

Setup

To carry out this case study, it is necessary to define a feature base. To this end, a correlation test was applied to determine the best ones, considering the following features: Year, Month, Day, Day of the Week, Period, plus all features involving the MCP (PT/ES). Figure 42 shows the correlation map of all features considered as the base.

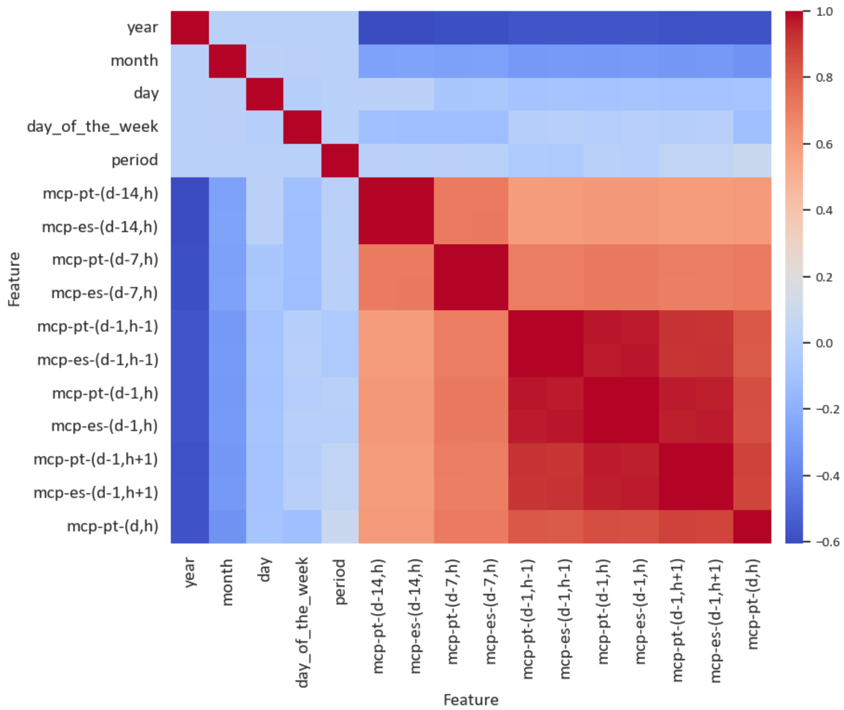

Figure 42 - Case Study 4.3, correlation matrix of base features, with chronological and market-clearing price features.

In Figure 42, the features of the previous mcp values significantly influence the results, while the impact of the features related to time chronology is less important. However, to make the importance of each feature quantitatively perceptible, the F-Score is used. Figure 43 is representative of the features that exceeded the minimum F-Score value, which is 100.

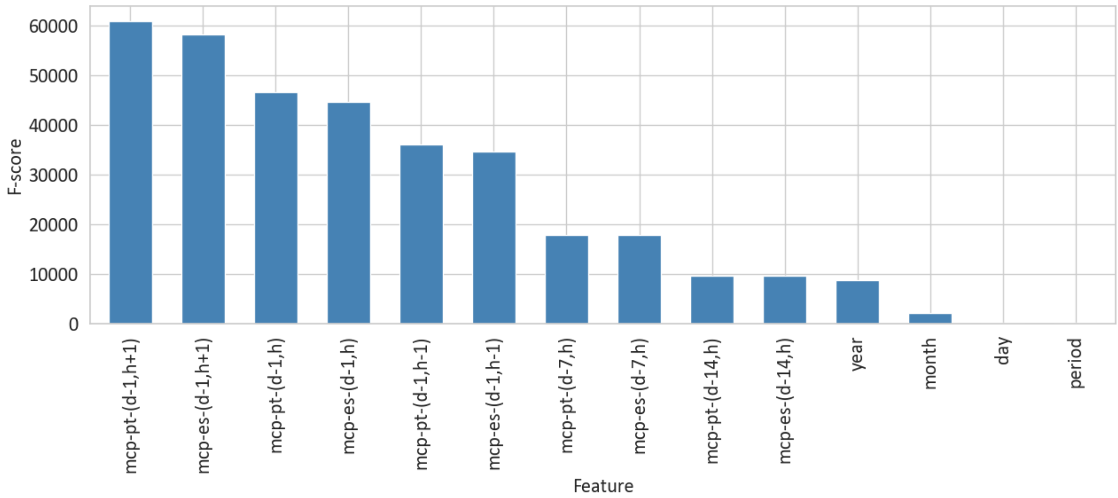

Figure 43 - Case Study 4.3, feature selection ranking considering the base features, with chronological and market-clearing price features.

Figure 43 shows that the most influential variable is "mcp-pt-(d-1,h+1)", closely followed by the other features that incorporate the mcp values from previous days for Portugal and Spain. Among the time-related features, only the "day_of_week" feature fell below the required threshold. As a result, a total of 14 features were chosen as the basis, in addition to the target column.

Having defined the feature base, it is necessary to define the other features for each energy production source: firstly, with the individual addition of each production source (there are four, since LOAD is not a source of energy production) and finally, with all production sources combined. For each source, the energy produced is considered, as well as the volume of energy traded in the day-ahead market ("act" and "dam" features). For each of these datasets, it is necessary to perform the tuning for each forecasting model to determine the best parameterisation. The hyperparameters for the base dataset are represented in Table 16.

Table 16 - Case Study 4.3, parameterisation of forecasting models for the base dataset, without any energy production sources data.

Model	Hyperparameters
RF	***'n_estimators'***: 100; ***'n_jobs'***: -1; ***'random_stat'***: 42
XGBoost	***'learning_rate'***: 0.2; ***'n_estimators'***: 100; ***'random_state'***: 42
RR	***'α'***: 0.1; ***'copy_X'***: True; ***'fit_intercept'***: True; ***'max_iter'***: 1000; ***'positive'***: True; ***'random_state'***: 42; ***'solver'***: 'auto'; ***'tol'***: 0.0001
SGD	***'loss'***: 'squared_error'; ***'max_iter'***: 1000; ***'penalty'***: 'l2'; ***'tol'***: 0.0001; ***'warm_start'***: False

Model	Hyperparameters
LSTM	**'n_celss_with_returning_sequences'**: 64; **'n_celss_with_not_returning_sequences'**: 64; **'dropout'**: 0.3; **'dense'**: 128
DNN	**'activation'**: 'elu', **'optimizer'**: 'adam', **'neuron'**: 128, **'dropout'**: 0.3

In Table 17, the hyperparameters for the base dataset plus the wind production source are shown.

Table 17 - Case Study 4.3, parameterisation of forecasting models for the base + wind energy production sources data.

Model	Hyperparameters
RF	***'n_estimators'***: 200; ***'n_jobs'***: -1; ***'random_stat'***: 42
XGBoost	***'learning_rate'***: 0.1; ***'n_estimators'***: 200; ***'random_state'***: 42
RR	***'α'***: 0.1; ***'copy_X'***: True; ***'fit_intercept'***: True; ***'max_iter'***: 1000; ***'positive'***: True; ***'random_state'***: 42; ***'solver'***: 'auto'; ***'tol'***: 0.0001
SGD	***'loss'***: 'squared_error'; ***'max_iter'***: 2000; ***'penalty'***: 'l2'; ***'tol'***: 0.0001; ***'warm_start'***: False
LSTM	**'n_celss_with_returning_sequences'**: 64; **'n_celss_with_not_returning_sequences'**: 64; **'dropout'**: 0.3; **'dense'**: 128
DNN	**'activation'**: 'relu', **'optimizer'**: 'adam', **'neuron'**: 256, **'dropout'**: 0.3

Table 18 illustrates the hyperparameters for the base dataset plus the solar production source.

Table 18 - Case Study 4.3, parameterisation of forecasting models for the base + solar energy production sources data.

Model	Hyperparameters
RF	***'n_estimators'***: 200; ***'n_jobs'***: -1; ***'random_stat'***: 42
XGBoost	***'learning_rate'***: 0.1; ***'n_estimators'***: 300; ***'random_state'***: 42
RR	***'α'***: 0.1; ***'copy_X'***: True; ***'fit_intercept'***: True; ***'max_iter'***: 1000; ***'positive'***: True; ***'random_state'***: 42; ***'solver'***: 'auto'; ***'tol'***: 0.001
SGD	***'loss'***: 'squared_error'; ***'max_iter'***: 1000; ***'penalty'***: 'l2'; ***'tol'***: 0.0001; ***'warm_start'***: False
LSTM	**'n_celss_with_returning_sequences'**: 64; **'n_celss_with_not_returning_sequences'**: 64; **'dropout'**: 0.3; **'dense'**: 128
DNN	**'activation'**: 'relu', **'optimizer'**: 'rmsprop', **'neuron'**: 256, **'dropout'**: 0.3

Table 19 shows the hyperparameters used in the forecast models for the base dataset plus the hydro production source.

Table 19 - Case Study 4.3, parameterisation of forecasting models for the base + hydro energy production sources data.

Model	Hyperparameters
RF	***'n_estimators'***: 200; ***'n_jobs'***: -1; ***'random_stat'***: 42
XGBoost	***'learning_rate'***: 0.1; ***'n_estimators'***: 300; ***'random_state'***: 42

Model	Hyperparameters
RR	***'α'***: 0.1; ***'copy_X'***: True; ***'fit_intercept'***: True; ***'max_iter'***: 1000; ***'positive'***: True; ***'random_state'***: 42; ***'solver'***: 'auto'; ***'tol'***: 0.0001
SGD	***'loss'***: 'squared_error'; ***'max_iter'***: 1000; ***'penalty'***: 'l2'; ***'tol'***: 0.0001; ***'warm_start'***: False
LSTM	**'n_celss_with_returning_sequences'**: 64; **'n_celss_with_not_returning_sequences'**: 64; **'dropout'**: 0.3; **'dense'**: 128
DNN	**'activation'**: 'relu', **'optimizer'**: 'rmsprop', **'neuron'**: 64, **'dropout'**: 0.3

Table 20 shows the hyperparameters of each forecasting model for the base dataset plus the nuclear production source.

Table 20 - Case Study 4.3, parameterisation of forecasting models for the base + nuclear energy production sources data.

Model	Hyperparameters
RF	***'n_estimators'***: 200; ***'n_jobs'***: -1; ***'random_stat'***: 42
XGBoost	***'learning_rate'***: 0.2; ***'n_estimators'***: 100; ***'random_state'***: 42
RR	***'α'***: 0.1; ***'copy_X'***: True; ***'fit_intercept'***: True; ***'max_iter'***: 1000; ***'positive'***: True; ***'random_state'***: 42; ***'solver'***: 'auto'; ***'tol'***: 0.0001
SGD	***'loss'***: 'squared_error'; ***'max_iter'***: 1000; ***'penalty'***: 'l2'; ***'tol'***: 0.0001; ***'warm_start'***: True
LSTM	**'n_celss_with_returning_sequences'**: 64; **'n_celss_with_not_returning_sequences'**: 64; **'dropout'**: 0.3; **'dense'**: 128
DNN	**'activation'**: 'relu', **'optimizer'**: 'rmsprop', **'neuron'**: 64, **'dropout'**: 0.2

Table 21 presents the hyperparameters of the forecast models for the base data set plus all production sources.

Table 21 - Case Study 4.3, parameterisation of forecasting models for the base + all energy production sources.

Model	Hyperparameters
RF	***'n_estimators'***: 200; ***'n_jobs'***: -1; ***'random_stat'***: 42
XGBoost	***'learning_rate'***: 0.2; ***'n_estimators'***: 100; ***'random_state'***: 42
RR	***'α'***: 0.1; ***'copy_X'***: True; ***'fit_intercept'***: True; ***'max_iter'***: 1000; ***'positive'***: True; ***'random_state'***: 42; ***'solver'***: 'auto'; ***'tol'***: 0.0001
SGD	***'loss'***: 'squared_error'; ***'max_iter'***: 2000; ***'penalty'***: 'l2'; ***'tol'***: 0.0001; ***'warm_start'***: False
LSTM	**'n_celss_with_returning_sequences'**: 64; **'n_celss_with_not_returning_sequences'**: 64; **'dropout'**: 0.2; **'dense'**: 128
DNN	**'activation'**: 'elu', **'optimizer'**: 'adam, **'neuron'**: 64, **'dropout'**: 0.4

Results and Discussion

With the parameterisations of the forecasting models defined for each dataset, it is possible to realise the validation test for each of them. There is no sensitivity test for the DP-RL model, as it is automatic and dynamic. Table 22 shows the results ordered by the best performance of each data set used.

Table 22 - Case Study 4.3, evaluation and comparison of DP-RL with each energy production source data, to identify the most performing one.

Model	MSE (€/MWh)	RMSE (€/MWh)	MAE (€/MWh)	R^2
Base + Hydro	25.7914	4.9760	1.5646	0.9875
Base + Wind	27.8874	5.2809	1.6853	0.9866
Base + Solar	28.9591	5.3814	1.7804	0.9861
Base	29.5089	5.4449	1.8292	0.9858
Base + Nuclear	31.1975	6.1035	1.9281	0.9849
Base + ALL	38.9513	7.1728	2.1101	0.9838

Based on the analysis in Table 22, the hydro energy source provides more accurate forecasting of the actual market value. The next ones were the wind and solar, while the feature base had a lower performance than the first three. In turn, the source of nuclear energy production is the one that contributes the least relevance. Interestingly, the lowest performance occurs when all energy production sources are combined with the base features, rather than including each energy source separately with the base features.

4.4 Context-based RL methodology mirroring real-time conditions

This case study aims to understand whether incorporating contexts, based on energy production sources, into models offers greater benefits compared to using daily periods. This case study describes two scenarios conducted in a controlled simulation environment, replicating real-time conditions. The first scenario focuses on applying the DP-RL model, compared to a second scenario that uses DPC-RL, which considers contexts based on energy production sources, volume of production, and demand for energy sources.

General Setup

With the conceptualising of the events in real-time from Table 2, it was found that data on the volume of energy production (REN/REE) and the volume of energy transacted by each energy source (OMIE) for the target day are not available. Consequently, forecasting were made to estimate these features based on the previous day's energy volume data, including production and demand.

4.4.1 Application of the DP-RL model

Setup

Focusing on the first scenario, it is necessary to realise feature selection. A correlation scenario was conducted, followed by an F-Score evaluation, mirroring the process already used in case studies 4.2 and 4.3. Figure 44 illustrates the visualisation of the most important features with an F-score above 100.

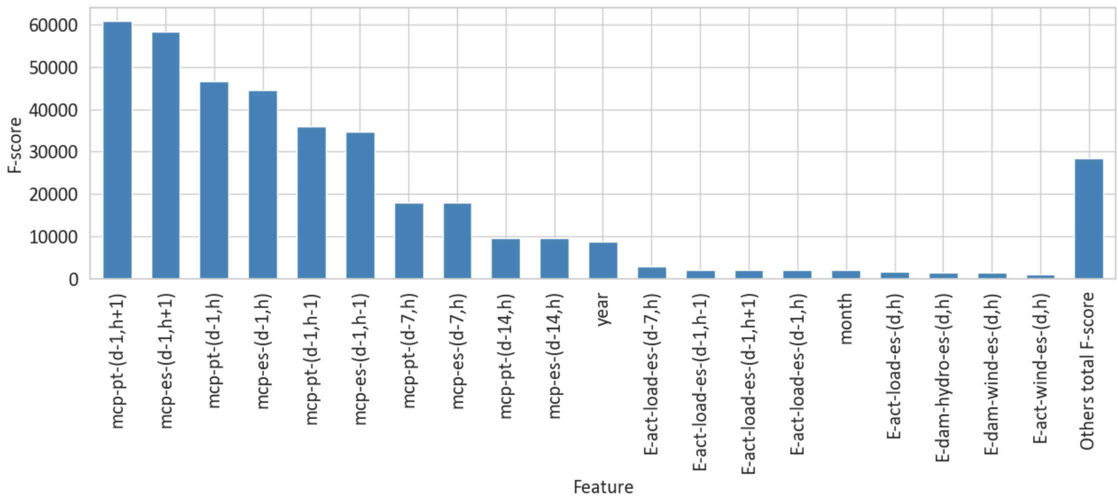

Figure 44 - Case Study 4.4.1, ranking of the feature selection for the first scenario, which uses DP-RL.

Figure 44 shows the 20 plus 61 features listed in the "Others total F-Score" column, which are chosen for the first scenario. The most important features include the previous mcp, the year, and the Spanish load for both the forecast day and previous days. Additionally, hydro and wind data for Spain on the forecast day are considered. Other features include various act and dam data for Portugal and Spain, covering all energy sources for both forecast days and previous days, along with indicators for the day, period, and day of the week.

With the features identified, the tuning of the various forecasting models for the first scenario is carried out, and the results are presented in Table 23.

Table 23 - Case Study 4.4.1, parameterisation of forecasting models for the first scenario, which uses DP-RL.

Model	Hyperparameters
RF	***'n_estimators'***: 100; ***'n_jobs'***: -1; ***'random_stat'***: 42
XGBoost	***'learning_rate'***: 0.1; ***'n_estimators'***: 200; ***'random_state'***: 42
RR	***'α'***: 10.0; ***'copy_X'***: True; ***'fit_intercept'***: True; ***'max_iter'***: 1000; ***'positive'***: False; ***'random_state'***: 42; ***'solver'***: 'auto; ***'tol'***: 0.001
SGD	***'loss'***: 'squared_error'; ***'max_iter'***: 2000; ***'penalty'***: 'l2'; ***'tol'***: 0.001; ***'warm_start'***: True
LSTM	**'n_celss_with_returning_sequences'**: 64; **'n_celss_with_not_returning_sequences'**: 64; **'dropout'**: 0.3; **'dense'**: 128
DNN	**'activation'**: 'relu', **'optimizer'**: 'rmsprop', **'neuron'**: 64, **'dropout'**: 0.2

After conducting the tuning of each forecast method in the first scenario, it is possible to proceed with the validation test. However, there is no sensitivity test for the DP-RL methodology, as it was already done in Case Study 4.2.

Results and discussion

Table 24 presents the results ordered by the best performance for the first scenario.

Table 24 - Case Study 4.4.1, evaluation and comparison of DP-RL with forecasting models for the first scenario.

Model	MSE (€/MWh)	RMSE (€/MWh)	MAE (€/MWh)	R^2
DP-RL	144.3924	12.6047	8.7539	0.9083
DNN	156.8579	14.0910	9.4657	0.9031
LSTM	168.8756	15.7586	10.2850	0.8959
RF	173.5633	16.1245	10.4589	0.8946
XGBoost	192.7583	19.4739	13.1750	0.8716
RR	215.4670	21.4728	14.3875	0.8567
SGD	229.7584	28.5887	17.8598	0.8209

Table 24 shows that the DP-RL outperformed the others. However, all models have an R^2 greater than 0.82. A notable finding was the DL's superior performance compared to ML and statistical models, demonstrating its greater adaptability to a case study that simulates real-world conditions. Compared to case study 4.2.1, there was a general decline in the performance of several methods, notably in the average difference between the DP-RL predicted value and the actual market price (MAE column), which is 8.7539 €/MWh. This decline is understandable because this study simulates real-time operations, incorporating forecasts for daily act and dam features for each energy production source, being 18 features.

This case study is essential for understanding real-time events, as well as the values predicted by the DP-RL for user bid submissions. To accomplish this, it is necessary to analyse the percentage difference between the predicted value by DP-RL and the actual market price throughout 2024. Figure 45 displays the average predicted values alongside the actual market values for each month. It also shows a light orange area when the predicted value is above the actual value, and a light blue area when the opposite occurs.

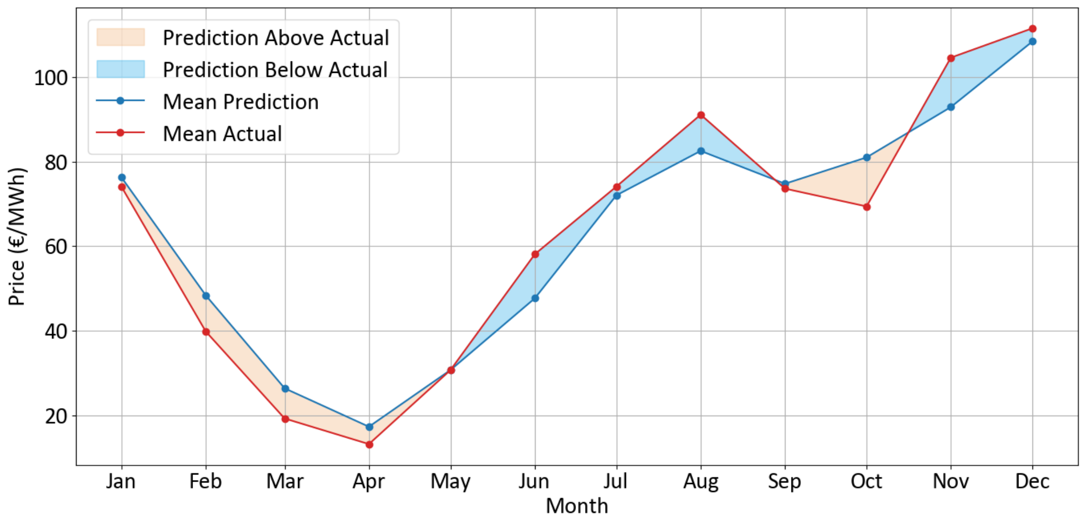

Figure 45 - Case Study 4.4.1, comparison between the predicted and actual prices for the first scenario, which uses DP-RL.

Figure 45 shows that from January to May and again from September to mid-October, predicted prices are generally slightly lower than the actual market. In other months, the trend reverses, but the difference is more significant. Therefore, throughout the year, forecasts were on average 4.96 €/MWh higher than the actual market value and 7.17 €/MWh. Furthermore, the number of times forecasting are over or under is nearly balanced, occurring over 51.1% of the time and under 48.9% of the time.

4.4.2 Application of the DPC-RL model

Setup

In scenario 2, before executing all the processes and analyses as in the first scenario, it's important to understand how the contexts are formed. There are three main factors for defining contexts: the source of production, the volume of production by source, and the volume of demand by energy source. The energy production sources include wind, solar, hydro, and nuclear, from Portugal and Spain.

Once the requirements are clearly specified, the next step is to create datasets for each production source and their combined data. A moving average of up to one month was calculated for the variables: E-act-source-country-(d,h) and E-dam-source-country-(d,h) across 2022, 2023, and 2024, helping to reveal the trends of each variable. It's important to note that for the aggregated dataset by production source context, these two variables will be combined based on each energy source.

The K-Means algorithm is used because it can detect clusters in the data, with the Silhouette Score guiding the optimal choice of k. The result was K=2 for both, but this performed poorly, so k was manually adjusted to 3 for both act and dam. Using the two new moving average columns, K-Means forms three independent groups based on the act and dam means. After clustering, the centroids are ordered from the lowest to the highest mean. A mapping then assigns group 0 to the lowest, group 1 to the intermediate, and group 2 to the highest. With three groups per variable, a context column is created to summarise the K-Means clusters. Since there are nine possible combinations (three groups for act and dam), each combination is assigned a specific context. Table 25 represents the meaning of each numerical context.

Table 25 - Case Study 4.4.2, description of K-Means contexts for the second scenario.

Context number	Production (Act)	Demand (Dam)
C0	Low	Low
C1	Low	Medium
C2	Low	High
C3	Medium	Low
C4	Medium	Medium
C5	Medium	High
C6	High	Low
C7	High	Medium
C8	High	High

Table 25 provides a description of each context number, which ranges from zero to eight. Zero represents both low production and demand, while eight represents the opposite.

Heat maps are generated for each energy production source, as well as their aggregate, covering the entire year of analysis (2024). These graphs are important for understanding where the contexts are most concentrated on average during each period of the day and per month. Figure 46 presents the heatmap graph of wind energy for the year 2024.

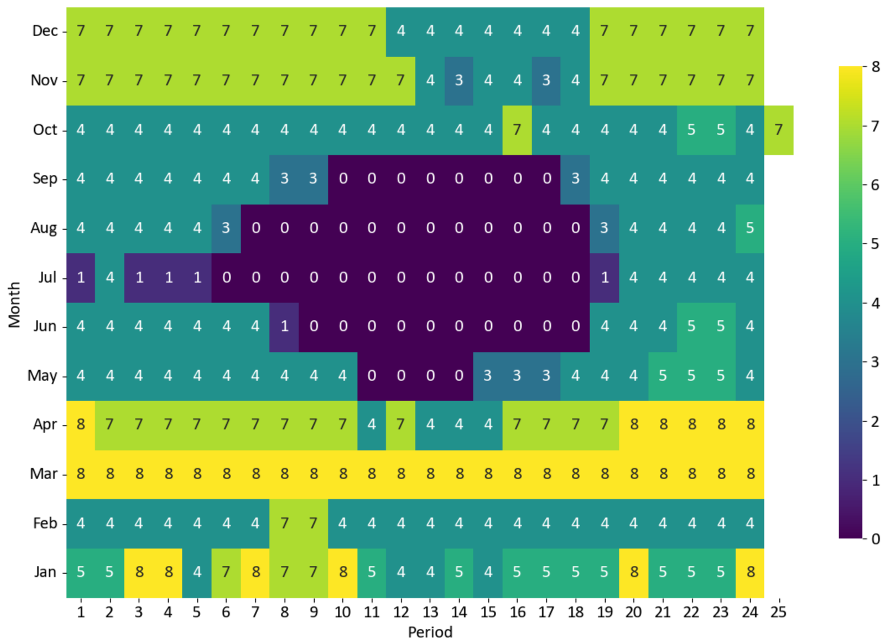

Figure 46 - Case Study 4.4.2, context heatmap for wind energy production.

In Figure 46, it is noticeable that in months and periods of heat, low production is expected as wind demand, through the zero context. In turn, production and demand remain average for the remaining months, as represented by context four. High production and demand are observed in March, April, November, and December, primarily in contexts 7 and 8.

Figure 47 illustrates the heat of solar energy for the year 2024.

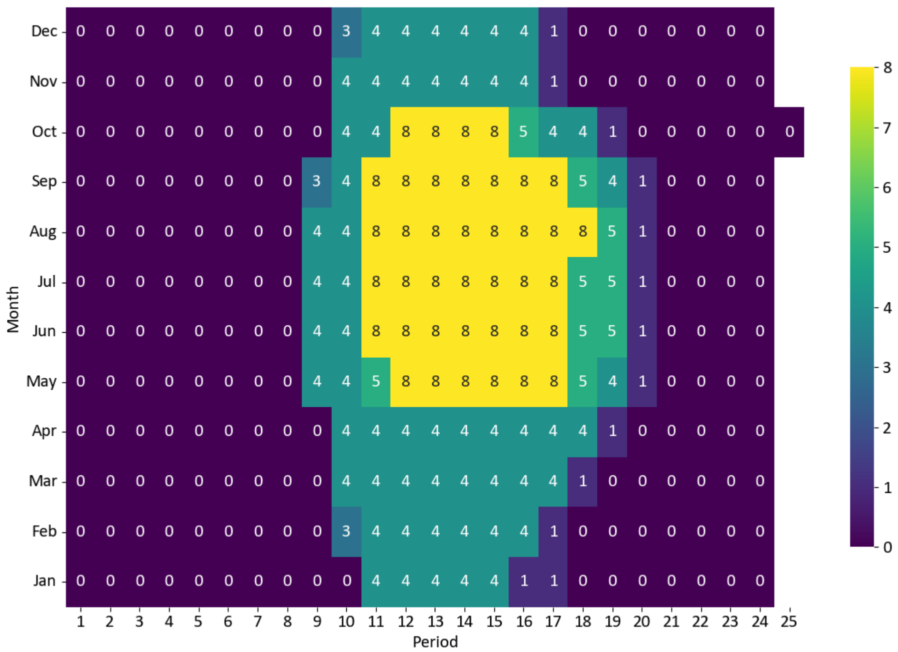

Figure 47 - Case Study 4.4.2, context heatmap for solar energy production.

In Figure 47, it is evident that production and demand are low from period 1 to 8 and from 20 to 24 each day, reflecting nighttime hours. The other periods show average to high levels of both production and demand, with these levels increasing during summer months and peak heat hours, particularly between periods 11 and 17.

Figure 48 shows the 2024 hydro source heatmap.

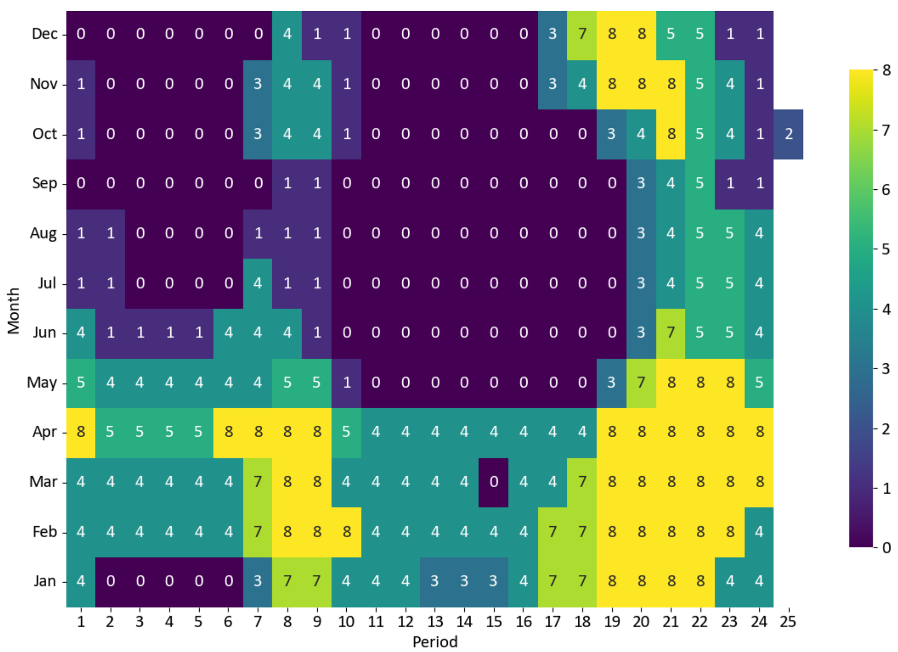

Figure 48 - Case Study 4.4.2, context heatmap for hydro energy production.

In Figure 48, no clear pattern is observable in the contexts. Still, it is evident that during the first half of the year, there is an intermediate to high level of production and demand for hydropower, which contrasts with the second half of the year, when production and demand are expected to be lower. However, in general, the final period of the day, both production and demand should be higher.

Figure 49 presents the heatmap of Spanish nuclear energy production throughout 2024.

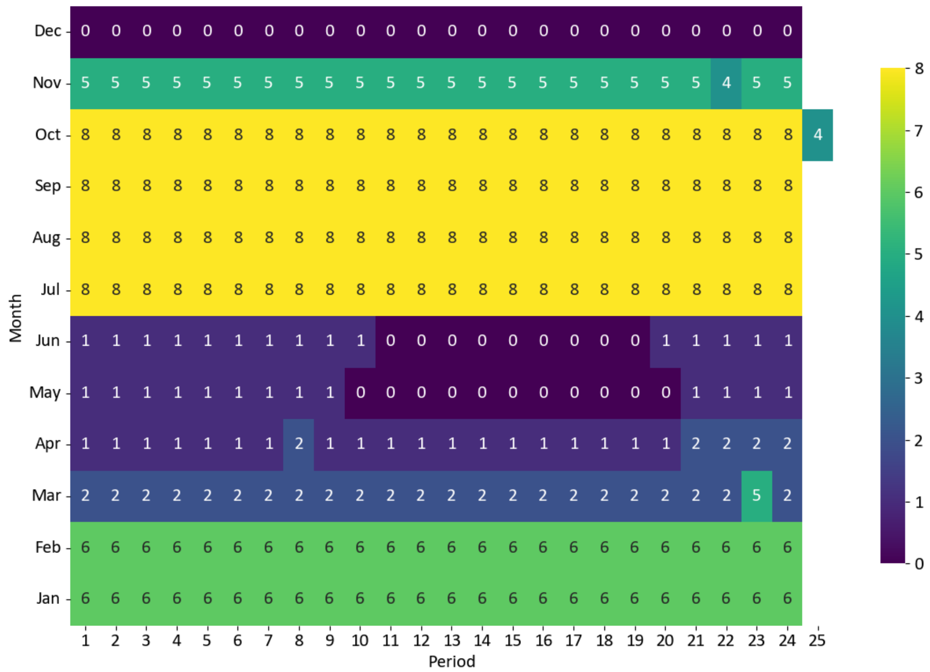

Figure 49 - Case Study 4.4.2, context heatmap for nuclear energy production.

Figure 49 shows that in January and February, there will be high nuclear energy production alongside low demand, contrasting with March, where the opposite is expected. From April to June, low production and medium-low demand are forecasted for this energy. During July to October, high production and demand are anticipated, while in November, intermediate production is expected alongside high demand. The year concludes with low production and demand for Spanish nuclear energy.

Finally, Figure 50 presents the heat map of the combined contexts of each energy source.

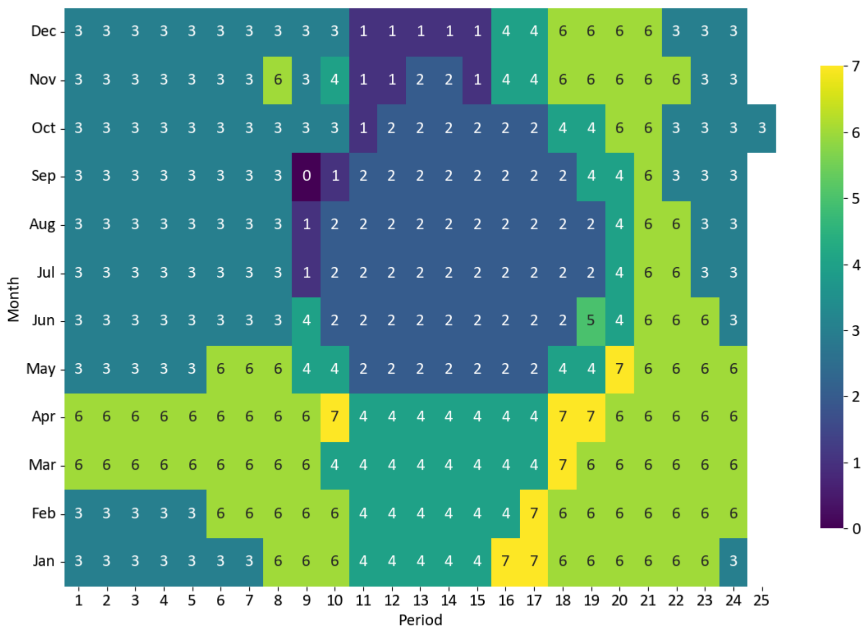

Figure 50 - Case Study 4.4.2, context heatmap for the aggregate of contexts for each energy production source.

Figure 50 shows from zero to seven contexts and not the number eight, revealing that in an aggregate of all energy sources, there is not a combination of high production and demand. It indicates that in all months, between periods 1 to 8 and from 20 to 24, intermediate production and low to high demand are expected from all energy sources. However, in the middle periods, there is a notable decline in production, while demand is expected to be medium to higher.

Figure 51 illustrates for the second scenario, the visualisation of the most important features with an F-score above 100.

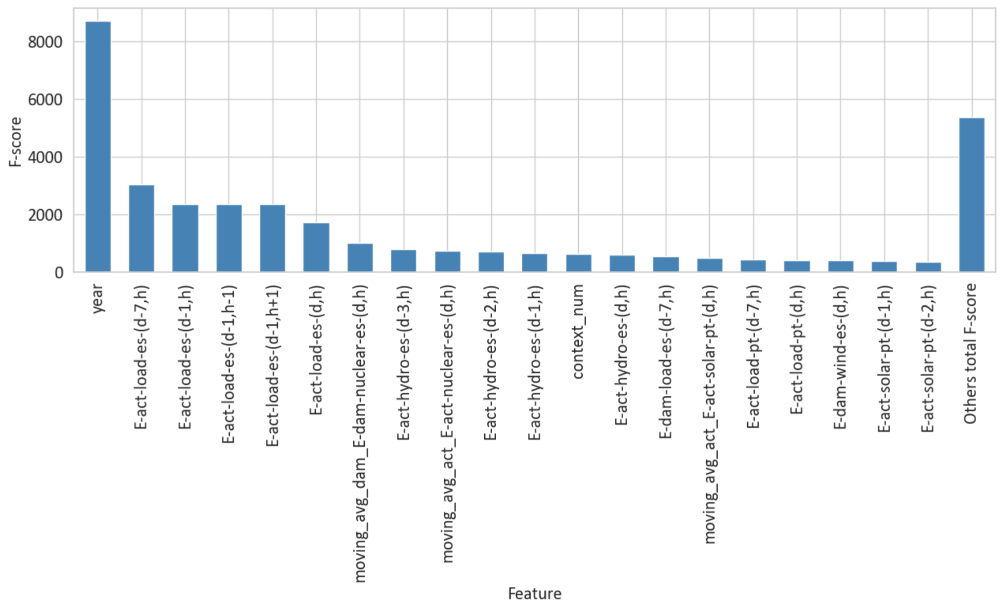

Figure 51 - Case Study 4.4.2, ranking of the feature selection for the second scenario, which uses DPC-RL.

Figure 51 shows the 50 most important features for this scenario. The key features include the year, the Spanish and Portuguese loads for the forecast day and previous days, and the moving averages of each energy production source for the forecast day. Additional resources consist of diverse act and dam data for Portugal and Spain, covering all energy sources for both forecast and prior days, along with indicators for the day, period, and weekday. With these features identified, the tuning for forecasting models can now be conducted for the second scenario.

Table 26 shows the hyperparameters identified for the forecasting models in the second scenario.

Table 26 - Case Study 4.4.2, parameterisation of forecasting models for the second scenario, which uses DPC-RL.

Model	Hyperparameters
RF	***'n_estimators'***: 50; ***'n_jobs'***: -1; ***'random_stat'***: 42
XGBoost	***'learning_rate'***: 0.01; ***'n_estimators'***: 300; ***'random_state'***: 42
RR	***'α'***: 10.0; ***'copy_X'***: True; ***'fit_intercept'***: True; ***'max_iter'***: 1000; ***'positive'***: False; ***'random_state'***: 42; ***'solver'***: 'lsqr'; ***'tol'***: 0.001
SGD	***'loss'***: 'squared_error'; ***'max_iter'***: 2000; ***'penalty'***: 'l1'; ***'tol'***: 0.001; ***'warm_start'***: False
LSTM	**'n_celss_with_returning_sequences'**: 64; **'n_celss_with_not_returning_sequences'**: 64; **'dropout'**: 0.3; **'dense'**: 128
DNN	**'activation'**: 'relu', **'optimizer'**: 'rmsprop', **'neuron'**: 64, **'dropout'**: 0.4

Results and discussion

After realising the tuning, the validation test can be performed for the second scenario. Table 27 presents the results, sorted by the best performance.

Table 27 - Case Study 4.4.2, evaluation and comparison of DPC-RL with forecasting models for the second scenario.

Model	MSE (€/MWh)	RMSE (€/MWh)	MAE (€/MWh)	R^2
DPC-RL	102.4824	13.0423	6.1294	0.9291
LSTM	106.8579	14.0910	6.4657	0.9214
DNN	110.8756	15.1586	7.2850	0.9159
SGD	111.5633	15.3245	7.4589	0.9146
RF	127.7583	17.4739	9.1750	0.8816
XGBoost	134.4670	18.4728	10.3875	0.8767
RR	160.7584	21.5887	13.8598	0.8509

Table 27 shows that the DPC-RL surpassed the others. All models performed well, with R^2 greater than 0.85. DL methods (LSTM and SGD) in conjunction with a statistical one (SGD) were better than ML and other statistical models. In this scenario, the average difference between the predicted values by the DPC-RL and the current market price is 6.1294 €/MWh (MAE column). Therefore, this scenario produced better results than the first one.

Figure 52 illustrates the monthly average difference between predicted values by the DPC-RL and actual market prices throughout 2024. In the figure, the light orange shading indicates months where forecasting exceeded actual prices, and the light blue shading indicates months where forecasting were below.

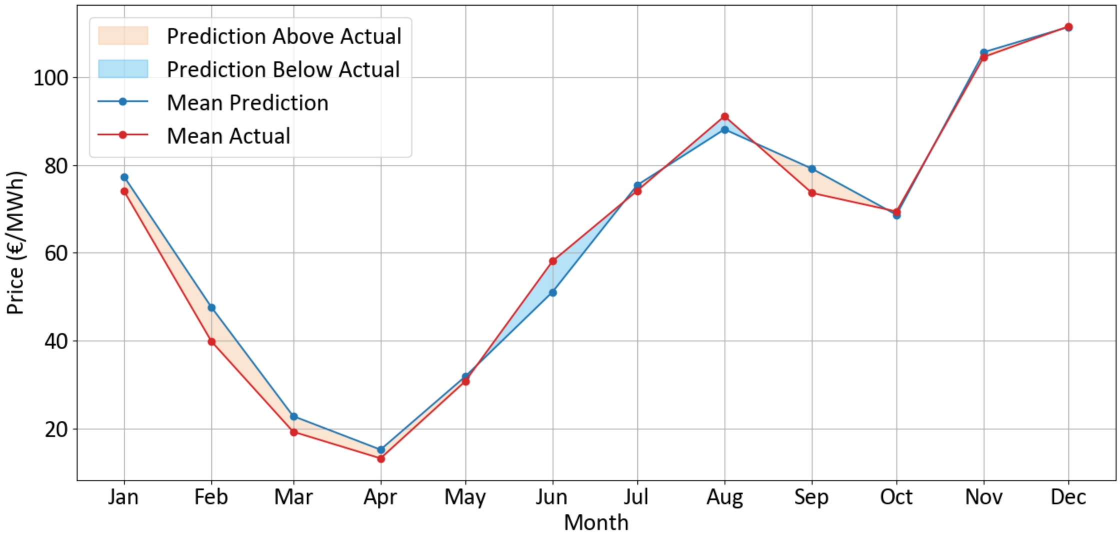

Figure 52 - Case Study 4.4.2, comparison between the predicted and actual prices for the second scenario, which uses DPC-RL.

Figure 52 shows that from January to May and again from mid-August to the end of the year, predicted prices tend to be slightly below the actual market values. Between May and mid-August, the trend shifts, except in July. Overall, forecasts were, on average, 3.20 €/MWh higher than actual market values and 2.73 €/MWh lower. Furthermore, it predicts above 54.4% of the time and below 45.6% of the time. Compared to the first scenario (see Figure 45), this second scenario has lower average forecasting for both above and below cases, making it more advantageous, with a marginally higher frequency of forecasts above.

Since the second scenario proved to be a more beneficial approach, it's essential to analyse how each context and each energy production source affects the actual market value. This analysis is shown in Table 28.

Table 28 - Case Study 4.4.2, impact of aggregated energy sources on the actual market prices.

Context number	Average Wind (act) (MWh)	Average Solar (act) (MWh)	Average Hydro (act) (MWh)	Average Nuclear (act) (MWh)	Production (Act) (MWh)	Demand (Dam) (MWh)	Actual Average price (€/MWh)
0	5652.47	2649.20	3594.31	167.87	Low	Low	104.20
1	6088.43	5044.98	3303.25	142.27	Low	Medium	75.44
2	4418.49	7147.74	2476.89	151.78	Low	High	42.63

Context number	Average Wind (act) (MWh)	Average Solar (act) (MWh)	Average Hydro (act) (MWh)	Average Nuclear (act) (MWh)	Production (Act) (MWh)	Demand (Dam) (MWh)	Actual Average price (€/MWh)
3	7027.56	276.86	6539.61	148.76	Medium	Low	84.41
4	6727.59	3959.38	8506.39	136.05	Medium	Medium	42.27
5	5365.27	3757.12	4814.06	140.59	Medium	High	34.69
6	7506.47	321.83	16047.73	134.37	High	Low	61.77
7	7478.63	2436.68	12645.12	129.23	High	Medium	30.14

Table 28 shows that as demand increases, the actual average market price tends to fall, decreasing even further as overall production increases. Hydro and wind sources have the most significant impact, generating the highest energy volumes across all contexts, exceeding 2,476.89 MWh. Solar energy plays a supporting role when overall energy production is low, as seen in context 2, reaching a peak output of 7,147.74 MWh. Nuclear energy remains stable, with a low energy output ranging from 129.23 MWh to 167.87 MWh, and has minimal impact on market prices. Based on this information and Figure 50, it can be concluded that the highest market prices are likely to occur at the beginning and end of the day, whereas the lowest prices tend to happen during the middle hours.

4.5 Final Remarks

This chapter covered four case studies: the first was a baseline; the second examined the most effective RL approach and compared it with the different RL methodologies existing in the literature; the third analysed how each energy source performed better for the real market price; and the fourth case study involved a controlled environment that simulated real-time events. The fourth study also compared DP-RL and DPC-RL to determine whether context-based RL is a more effective approach.

In the first case study, there was no adequate data preparation. However, the results obtained were satisfactory for the two sub-case studies presented, in which the absolute difference between the values predicted by the CB-RL and the current value throughout 2024 was approximately 3.15 €/MWh and 2.09 €/MWh, respectively.

In the second case study with general feature selection, both the forecasting models and the DP-RL showed excellent performances, with these results being superior to those demonstrated in the first case study. The absolute difference over the year 2024 was 0.56 €/MWh in the first sub-study. This improvement reflects the impact of an automatic and dynamic RL methodology, as well as the feature selection process, on the performance of each forecasting model. The DP-RL was compared in a second sub-study with different existing RL methodologies, where it remained superior to the others.

The third case study demonstrated that hydro had the best performance. However, the performance of the DP-RL with the hydro energy production source dataset was slightly lower than that demonstrated in case study 4.2. This statement is justified because there was only a selection of features in the base dataset in case study 4.3.

In the fourth case study, two scenarios were executed in a controlled environment, simulating real-time events. The first scenario utilised the DP-RL methodology, which showed a performance decline of approximately 10%. This decline occurred because no data were available at the time of the forecast for the same day, leading to forecasts being generated for 18 variables related to that day. The first scenario was roughly 3% less effective than the second scenario, which employed a DPC-RL based on contexts such as energy production sources, production volume, and demand volume. Additionally, the second scenario showed a more balanced forecast, with forecasting values above and below the actual market value at 3.20 €/MWh and 2.73 €/MWh, respectively. In conclusion, the analysis shows that hydro and wind energy are the main sources influencing the overall Portuguese price in the Iberian market, with solar energy also having a significant impact. In contrast to Spain's nuclear energy, which has the smallest influence. This internal analysis reveals the effects of real-time energy production, consistent with the findings in non-real-time case study 4.3.

Finally, these findings highlight the advantages of using DPC-RL, which outperforms all other forecasting models and the RL methodologies compared across all the case studies. Furthermore, these studies are valuable because they provide effective analyses and satisfactory results for a bid proposal in the Iberian market.

5 Conclusions

This chapter summarises the main conclusions of this dissertation, including the objectives achieved. It also discusses limitations and suggests new functionalities for future integration to enhance the system.

5.1 Main Conclusions

At the end of this dissertation, with the final completion of the research, investigation, and development, it is possible to indicate that all objectives were achieved.

The first objective involves conducting a study and analysis of the state of the art in relation to the various areas and topics presented in this dissertation, highlighting the promising advances proposed in the fields of AI, EM, and SGs, and filling existing gaps. Therefore, O1 was successfully fulfilled and partially answered RQ1, defined in the Research Question and Objectives section.

The second objective was to outline the different requirements for implementing the various AI technologies. This O2 was successfully achieved, deploying a flexible solution with multiple AI methods. The third objective covers data analysis and collection. This O3 is highly important, as data analysis and preparation are essential, being the main basis for AI models to be effective. These two objectives partially addressed RQ1, since they were both completed.

Regarding the fourth objective, which consists of the selection and implementation of forecasting models, specifically statistical and ML, these are essential for predicting energy market prices. This objective was successfully achieved, addressing SQ1. In the fifth objective, there is the selection and implementation of an RL methodology. This objective is a differentiator, as it combines the knowledge of all available forecasting models, with RL choosing the most suitable model for a given situation and moment. O5 was completed successfully, responding to SQ2.

Objective six focuses on designing the actions of various agents, an administrative interface for system management, and offline model development, addressing RQ1. The last three objectives, essential for the development and testing of the system, with different case studies, were all successfully fulfilled, thus answering RQ1. O7 covers the development and implementation of unit and integration tests, as well as system tests. O8 involves designing and executing various case studies to assess the system's versatility. In O9, the dissertation document was written in a clear, succinct, and objective manner, revealing new applications of AI technologies and innovating the energy sector.

Ultimately, the main objective of this dissertation was to explore how an RL-based MAS improves forecasting by selecting the most appropriate model, within the scope of MCP forecasting in the Iberian electricity market. The initially proposed RL methodology — CB-RL — started by using a static α and MSE-based cumulative rewards to select the best model. To increase dynamism, DP-RL introduced a flexible α and performance penalties for forecasting models' responsiveness. Finally, DPC-RL incorporated contextual information from RES volume, creating a robust methodology that adaptively selects the most suitable forecast model for each period and context. The proposed MAS highlights the relevance of distributed and autonomous decision-making in complex environments, providing a foundation for future developments by enabling scalability, parallel processing, and coordination among agents.

The results demonstrate and confirm the ability of RL to adaptively select the most appropriate forecasting model and to achieve higher accuracy than the individual models tested. CB-RL achieved absolute errors of approximately 2.09 €/MWh, while DP-RL reduced these errors to 0.56 €/MWh. However, when DP-RL was tested in a simulated environment reflecting the participation in the day-ahead market, its forecasts deviated significantly from the actual MCP — due to the absence of the 18 features resulting from the execution of the final Iberian market program —, with average errors of 4.96 €/MWh above and 7.17 €/MWh below. These results motivated the application of DPC-RL in this environment, which improved performance, achieving balanced deviations of approximately 3.20 €/MWh above and 2.73 €/MWh below the actual MCP. Overall, DPC-RL demonstrates the effectiveness of context-aware RL in achieving the highest accuracy and robust forecast performance.

5.2 Limitations and Future Work

The system described in this dissertation offers innovative features, such as the ability to perform various trainings and forecasting using datasets from different periods. It can easily incorporate additional forecasting or RL models, thereby enhancing the system's versatility. Designed to be scalable and robust, it supports price forecasting in the Iberian market for bid proposals. Nonetheless, this section recognises limitations and suggests potential improvements for the future.

Regarding the simulation of a real-time scenario, as seen in case study 4.4, data were not available for the same period. Therefore, access to energy production data from REN and REE, along with OMIE data, was limited to the day before the forecast. To address this, a forecasting method was developed to fill the missing data. This is a significant limitation because incomplete data reduces the accuracy of forecasts. Therefore, a proposed improvement is to enhance the forecasts for the missing data.

An upcoming feature is the administrative interface, through which the administrator will be able to execute and manage actions necessary for the system's effective operation. This functionality will also enable the administrator to train forecasting and RL models, which will be made available to users. An additional feature involves developing a user interface where users can submit requests, train or predict with any available method, view results, and provide feedback. Users will have the option to select existing models or add new ones. The administrator will have access to this user interface, but the admin interface itself will be restricted solely to administrators.

For the final future implementation, it would be beneficial to establish a login system for authorised users. In other words, unauthorised users could only make forecasts using the methods provided by the administrator. Meanwhile, authorised users would have access to all features permitted to them and have their own created models. This implementation is essential to restrict user access to processes that require prior knowledge.

References

Abadi, Martín, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mane, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viegas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. 2016. 'TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems'.

Achiam, Josh. 2018. 'Spinningup/Docs/Spinningup/Rl_intro2.Rst at Master · Openai/Spinningup'. https://github.com/openai/spinningup/blob/master/docs/spinningup/rl_intro2.rst.

Adak, Deepanjana, Raghunath Bhattacharyya, and Harish Barshilia. 2022. 'A State-of-the-Art Review on the Multifunctional Self-Cleaning Nanostructured Coatings for PV Panels, CSP Mirrors and Related Solar Devices'. *Renewable and Sustainable Energy Reviews* 159:112145. doi:10.1016/j.rser.2022.112145.

Agouzoul, Abdelali, and Emmanuel Simeu. 2024. 'Predictive Control Method for Comfort and Thermal Energy Enhancement in Buildings'. Pp. 1–6 in *2024 International Conference on Control, Automation and Diagnosis (ICCAD)*.

Ahmad, Tanveer, Rafal Madonski, Dongdong Zhang, Chao Huang, and Asad Mujeeb. 2022. 'Data-Driven Probabilistic Machine Learning in Sustainable Smart Energy/Smart Energy Systems: Key Developments, Challenges, and Future Research Opportunities in the Context of Smart Grid Paradigm'. *Renewable and Sustainable Energy Reviews* 160:112128. doi:10.1016/j.rser.2022.112128.

Ahmad, Tanveer, Hongcai Zhang, and Biao Yan. 2020. 'A Review on Renewable Energy and Electricity Requirement Forecasting Models for Smart Grid and Buildings'. *Sustainable Cities and Society* 55:102052. doi:10.1016/j.scs.2020.102052.

Alanne, Kari, and Seppo Sierla. 2022. 'An Overview of Machine Learning Applications for Smart Buildings'. *Sustainable Cities and Society* 76:103445. doi:10.1016/j.scs.2021.103445.

Alberizzi, Andrea, Paolo Di Barba, and Florian Ziel. 2025. 'Agent Based Modeling for Intraday Electricity Markets'. *OPSEARCH* 62(1):178–97. doi:10.1007/s12597-024-00805-w.

Ali, Mohammed, and Ali Obied. 2022. 'Interactive Situated Autonomic Multi-Agents System-Comprehensive Survey'. *Journal of Al-Qadisiyah for Computer Science and Mathematics* 14(3):Page 22-32. doi:10.29304/jqcm.2022.14.3.983.

Alibrahim, Hussain, and Simone A. Ludwig. 2021. 'Hyperparameter Optimization: Comparing Genetic Algorithm against Grid Search and Bayesian Optimization'. Pp. 1551–59 in *2021 IEEE Congress on Evolutionary Computation (CEC)*.

Alkhayat, Ghadah, and Rashid Mehmood. 2021. 'A Review and Taxonomy of Wind and Solar Energy Forecasting Methods Based on Deep Learning'. *Energy and AI* 4:100060. doi:10.1016/j.egyai.2021.100060.

Amato, Alberto, Alessandro Quarto, and Vincenzo Di Lecce. 2021. 'An Application of Cyber-Physical System and Multi-Agent Technology to Demand-Side Management Systems'. *Pattern Recognition Letters* 141:23–31. doi:10.1016/j.patrec.2020.09.004.

Amit, Hey. 2024. 'Model-Free vs. Model-Based Reinforcement Learning | by Hey Amit | Nov, 2024 | Medium'. https://medium.com/@heyamit10/model-free-vs-model-based-reinforcement-learning-0019cb219884.

Aslam, Sheraz, Herodotos Herodotou, Syed Muhammad Mohsin, Nadeem Javaid, Nouman Ashraf, and Shahzad Aslam. 2021. 'A Survey on Deep Learning Methods for Power Load and Renewable Energy Forecasting in Smart Microgrids'. *Renewable and Sustainable Energy Reviews* 144:110992. doi:10.1016/j.rser.2021.110992.

Bagemihl, Joachim, Frank Boesner, Jens Riesinger, Michael Künzli, Gwendolin Wilke, Gabriela Binder, Holger Wache, Daniel Laager, Jürgen Breit, Michael Wurzinger, Juliana Zapata, Silvia Ulli-Beer, Vincent Layec, Thomas Stadler, and Franz Stabauer. 2018. 'A Market-Based Smart Grid Approach to Increasing Power Grid Capacity without Physical Grid Expansion'. *Computer Science - Research and Development* 33(1):177–83. doi:10.1007/s00450-017-0356-5.

Balan, Ganapathy Sakthi, Santhosh Kumar, and Aravind Raj. 2025. 'Machine Learning and Artificial Intelligence Methods and Applications for Post-Crisis Supply Chain Resiliency and Recovery'. *Supply Chain Analytics* 10:100121. doi:10.1016/j.sca.2025.100121.

Bhaskar, Yash. 2023. 'Introduction to Statistical Methods in AI — Overview'. https://medium.com/@yash9439/introduction-to-statistical-methods-in-ai-overview-9bc981ba91d0.

Bui, Van-Hai, Srijita Das, Akhtar Hussain, Guilherme Vieira Hollweg, and Wencong Su. 2024. 'A Critical Review of Safe Reinforcement Learning Techniques in Smart Grid Applications'.

Calegari, Roberta, Giovanni Ciatto, Viviana Mascardi, and Andrea Omicini. 2020. 'Logic-Based Technologies for Multi-Agent Systems: A Systematic Literature Review'. *Autonomous Agents and Multi-Agent Systems* 35(1):1. doi:10.1007/s10458-020-09478-3.

Cardo-Miota, Javier, Emilio Pérez, and Hector Beltran. 2023. 'Deep Learning-Based Forecasting of the Automatic Frequency Reserve Restoration Band Price in the Iberian Electricity Market'. *Sustainable Energy, Grids and Networks* 35:101110. doi:10.1016/j.segan.2023.101110.

Chhajer, Parshv, Manan Shah, and Ameya Kshirsagar. 2022. 'The Applications of Artificial Neural Networks, Support Vector Machines, and Long–Short Term Memory for Stock Market Prediction'. *Decision Analytics Journal* 2:100015. doi:10.1016/j.dajour.2021.100015.

Chiang, Sharon, Vikram Rao, and Marina Vannucci. 2024. *Statistical Methods in Epilepsy*. CRC Press.

Chugh, Akshita. 2024. 'MAE, MSE, RMSE, Coefficient of Determination, Adjusted R Squared — Which Metric Is Better?' https://medium.com/analytics-vidhya/mae-mse-rmse-coefficient-of-determination-adjusted-r-squared-which-metric-is-better-cd0326a5697e.

Coraci, Davide, Silvio Brandi, Tianzhen Hong, and Alfonso Capozzoli. 2023. 'Online Transfer Learning Strategy for Enhancing the Scalability and Deployment of Deep Reinforcement Learning Control in Smart Buildings'. *Applied Energy* 333:120598. doi:10.1016/j.apenergy.2022.120598.

Dagal, Idriss, Kürşat Tanriöven, Ahmet Nayir, and Burak Akın. 2025. 'Adaptive Stochastic Gradient Descent (SGD) for Erratic Datasets'. *Future Generation Computer Systems* 166:107682. doi:10.1016/j.future.2024.107682.

Deb, Chirag, Fan Zhang, Junjing Yang, Siew Eang Lee, and Kwok Wei Shah. 2017. 'A Review on Time Series Forecasting Techniques for Building Energy Consumption'. *Renewable and Sustainable Energy Reviews* 74:902–24. doi:10.1016/j.rser.2017.02.085.

Debrah, Kwame. 2024. 'A Decentralised Peer-to-Peer Energy Trading Platform for Residential Homes'. https://salford-repository.worktribe.com/output/2683138.

Du, Mengyao, Miao Zhang, Yue Hu, Jie Yang, Yiping Yao, and Quanjun Yin. 2021. 'Blockchain for Distributed Consistency: A Cliquebased Framework for Multi-Agent Systems'. Pp. 421–27 in *2021 7th International Conference on Big Data and Information Analytics (BigDIA)*.

Ernst, Damien, and Louette Arthur. 2024. 'Introduction to Reinforcement Learning'. *Introduction to Reinforcement Learning* 67.

Eyimaya, Süleyman Emre, Necmi Altin, and Adel Nasiri. 2024. 'A Review on Multi-Agent Systems and JADE Applications in Microgrids'. Pp. 623–28 in *2024 12th International Conference on Smart Grid (icSmartGrid)*.

Fidalgo, J. Nuno, Débora de São José, and Carlos Silva. 2019. 'Impact of Climate Changes on the Portuguese Energy Generation Mix'. Pp. 1–6 in *2019 16th International Conference on the European Energy Market (EEM)*.

Geraldo, Osório. 2018. 'Hybrid Forecasting Model for Short-Term Electricity Market Prices with Renewable Integration'. https://www.mdpi.com/2071-1050/11/1/57.

Ghazal, T. M. 2022. 'Energy Demand Forecasting Using Fused Machine Learning Approaches'. *Intelligent Automation & Soft Computing* 31(1):539–53. doi:10.32604/iasc.2022.019658.

Gnat, Sebastian. 2021. 'Impact of Categorical Variables Encoding on Property Mass Valuation'. *Procedia Computer Science* 192:3542–50. doi:10.1016/j.procs.2021.09.127.

Groenewald, Coenrad Adolph, Gonesh Chandra Saha, Garima Mann, Bharat Bhushan, Eric Howard, and Elma Sibonghanoy Groenewald. 2024. 'Multi-Agent Systems in Robotics: Coordination and Communication Using Machine Learning.' *NATURALISTA CAMPANO* 28(1):882–97.

Gronauer, Sven, and Klaus Diepold. 2022. 'Multi-Agent Deep Reinforcement Learning: A Survey'. *Artificial Intelligence Review* 55(2):895–943. doi:10.1007/s10462-021-09996-w.

Guo, Hongshu, Yining Ma, Zeyuan Ma, Jiacheng Chen, Xinglin Zhang, Zhiguang Cao, Jun Zhang, and Yue-Jiao Gong. 2024. 'Deep Reinforcement Learning for Dynamic Algorithm Selection: A Proof-of-Principle Study on Differential Evolution'. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 54(7):4247–59. doi:10.1109/TSMC.2024.3374889.

Gupta, Anchal, Youakim Badr, Ashkan Negahban, and Robin G. Qiu. 2021. 'Energy-Efficient Heating Control for Smart Buildings with Deep Reinforcement Learning'. *Journal of Building Engineering* 34:101739. doi:10.1016/j.jobe.2020.101739.

Harris, Charles, Jarrod Millman, Stéfan van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. 2020. 'Array Programming with NumPy'. *Nature* 585(7825):357–62. doi:10.1038/s41586-020-2649-2.

Harrold, Daniel, Jun Cao, and Zhong Fan. 2022. 'Renewable Energy Integration and Microgrid Energy Trading Using Multi-Agent Deep Reinforcement Learning'. *Applied Energy* 318:119151. doi:10.1016/j.apenergy.2022.119151.

Hernandez, Luis, Carlos Baladron, Javier M. Aguiar, Belen Carro, Antonio Sanchez-Esguevillas, Jaime Lloret, David Chinarro, Jorge J. Gomez-Sanz, and Diane Cook. 2013. 'A Multi-Agent System Architecture for Smart Grid Management and Forecasting of Energy Demand in Virtual Power Plants'. *IEEE Communications Magazine* 51(1):106–13. doi:10.1109/MCOM.2013.6400446.

Hossen, Tareq, Arun Sukumaran Nair, Radhakrishnan Angamuthu Chinnathambi, and Prakash Ranganathan. 2018. 'Residential Load Forecasting Using Deep Neural Networks (DNN)'. Pp. 1–5 in *2018 North American Power Symposium (NAPS)*.

Ibrahim, Muhammad Sohail, Wei Dong, and Qiang Yang. 2020. 'Machine Learning Driven Smart Electric Power Systems: Current Trends and New Perspectives'. *Applied Energy* 272:115237. doi:10.1016/j.apenergy.2020.115237.

Izmirlioglu, Yusuf, Loc Pham, Tran Cao Son, and Enrico Pontelli. 2024. 'A Survey of Multi-Agent Systems for Smartgrids'. *Energies* 17(15):3620. doi:10.3390/en17153620.

Jendoubi, Imen, and François Bouffard. 2022. 'Data-Driven Sustainable Distributed Energy Resources' Control Based on Multi-Agent Deep Reinforcement Learning'. *Sustainable Energy, Grids and Networks* 32:100919. doi:10.1016/j.segan.2022.100919.

Jiang, Tammy, Jaimie L. Gradus, and Anthony J. Rosellini. 2020. 'Supervised Machine Learning: A Brief Primer'. *Behavior Therapy* 51(5):675–87. doi:10.1016/j.beth.2020.05.002.

Kamdar, Renuka, Priyanka Paliwal, and Yogendra Kumar. 2018. 'A State of Art Review on Various Aspects of Multi-Agent System'. *Journal of Circuits, Systems and Computers*. doi:10.1142/S0218126618300064.

Kanade, Vijay. 2022. 'Everything You Should Know About Reinforcement Learning'. https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-reinforcement-learning/.

Kaneko, Hiromasa. 2023. 'Local Interpretation of Nonlinear Regression Model with K-Nearest Neighbors'. *Digital Chemical Engineering* 6:100078. doi:10.1016/j.dche.2022.100078.

Kappelhof, Ramos, Kappelhof, van Os, Chalos, van Kranendonk, Kruyt, Roos, van Zwam, van der Schaaf, M. A. A. van Walderveen, Wermer, van Oostenbrugge, Hester Lingsma, Diederik Dippel, Majoie, and Marquering. 2021. 'Evolutionary Algorithms and Decision Trees for Predicting Poor Outcome after Endovascular Treatment for Acute Ischemic Stroke'. *Computers in Biology and Medicine* 133:104414. doi:10.1016/j.compbiomed.2021.104414.

Kartini, Unit Three, Hariyati, Priyo Heru Adiwibowo, Sayyidul Aulia Alamsyah, Arrahmad Dwi Budiarto, and Khoirul Fadli. 2023. 'Long Short-Term Memory Decomposition Model for Forecasting Solar Irradiance Photovoltaic Household Scale on Grid'. Pp. 124–27 in *2023 IEEE 7th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*.

Kataray, Tarun, Nitesh, Bharath Yarram, Sanyukta Sinha, Erdem Cuce, Saboor Shaik, Pethurajan Vigneshwaran, and Abin Roy. 2023. 'Integration of Smart Grid with Renewable Energy Sources: Opportunities and Challenges – A Comprehensive Review'. *Sustainable Energy Technologies and Assessments* 58:103363. doi:10.1016/j.seta.2023.103363.

Khalid, Muhammad. 2024. 'Smart Grids and Renewable Energy Systems: Perspectives and Grid Integration Challenges'. *Energy Strategy Reviews* 51:101299. doi:10.1016/j.esr.2024.101299.

Kızıldağ, Mehmet, Fatih Abut, and Mehmet Fatih Akay. 2024. 'Development of New Electricity System Marginal Price Forecasting Models Using Statistical and Artificial Intelligence Methods'. *Applied Sciences* 14(21):10011. doi:10.3390/app142110011.

Konstantinos Roumeliotis, and Nikolaos Tselikas. 2023. 'A Machine Learning Python-Based Search Engine Optimization Audit Software'. *Informatics* 10(3):68. doi:10.3390/informatics10030068.

Kotzur, Leander, Lars Nolting, Maximilian Hoffmann, Theresa Groß, Andreas Smolenko, Jan Priesmann, Henrik Büsing, Robin Beer, Felix Kullmann, Bismark Singh, Aaron Praktiknjo, Detlef Stolten, and Martin Robinius. 2021. 'A Modeler's Guide to Handle Complexity in Energy Systems Optimization'. *Advances in Applied Energy* 4:100063. doi:10.1016/j.adapen.2021.100063.

Kumar, Ajitesh. 2023. 'Supervised & Unsupervised Learning Difference'. https://vitalflux.com/dummies-notes-supervised-vs-unsupervised-learning/.

Kumar, Jatinder, and Ashutosh Kumar Singh. 2023. 'A Demand and Response Management Model Using Load Forecasting Technique for Smart Grid'. Pp. 1–6 in *2023 5th International Conference on Energy, Power and Environment: Towards Flexible Green Energy Technologies (ICEPE)*.

Lago, Jesus, Grzegorz Marcjasz, Bart De Schutter, and Rafał Weron. 2021. 'Forecasting Day-Ahead Electricity Prices: A Review of State-of-the-Art Algorithms, Best Practices and an Open-Access Benchmark'. *Applied Energy* 293:116983. doi:10.1016/j.apenergy.2021.116983.

Lamnatou, Chemisana, and Cristofari. 2022. 'Smart Grids and Smart Technologies in Relation to Photovoltaics, Storage Systems, Buildings and the Environment'. *Renewable Energy* 185:1376–91. doi:10.1016/j.renene.2021.11.019.

Lee, Kimin, Michael Laskin, Aravind Srinivas, and Pieter Abbeel. 2021. 'SUNRISE: A Simple Unified Framework for Ensemble Learning in Deep Reinforcement Learning'.

Lee, Yongsun, Hoyun Choi, Laurent Pagnier, Cook Hyun Kim, Jongshin Lee, Bukyoung Jhun, Heetae Kim, Juergen Kurths, and Kahng. 2024. 'Reinforcement Learning Optimizes Power Dispatch in Decentralized Power Grid'. *Chaos, Solitons & Fractals* 186:115293. doi:10.1016/j.chaos.2024.115293.

Li, Haoxiang, Qi Zhou, Jing Tian, and Xiaoyu Lin. 2020. 'Energy Demand Forecasting for an Office Building Based on Random Forests'. Pp. 29–32 in *2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2)*.

Lipu, Hossain, Sazal Miah, Hannan, Aini Hussain, Mahidur Sarker, Afida Ayob, Mohamad Hanif Md Saad, and Sultan Mahmud. 2021. 'Artificial Intelligence Based Hybrid Forecasting Approaches for Wind Power Generation: Progress, Challenges and Prospects'. *IEEE Access* 9:102460–89. doi:10.1109/ACCESS.2021.3097102.

Liu, Zhenhua, Yushu Wang, Xinting Yuan, Zhihua Ding, and Qiang Ji. 2025. 'Geopolitical Risk and Vulnerability of Energy Markets'. *Energy Economics* 141:108055. doi:10.1016/j.eneco.2024.108055.

Maestre, Roberto, Juan Duque, Alberto Rubio, and Juan Arévalo. 2018. 'Reinforcement Learning for Fair Dynamic Pricing'.

Mahela, Om Prakash, Mahdi Khosravy, Neeraj Gupta, Baseem Khan, Hassan Haes Alhelou, Rajendra Mahla, Nilesh Patel, and Pierluigi Siano. 2022. 'Comprehensive Overview of Multi-Agent Systems for Controlling Smart Grids'. *CSEE Journal of Power and Energy Systems* 8(1):115–31. doi:10.17775/CSEEJPES.2020.03390.

Makkar, Kartika, Pardeep Kumar, Monika Poriye, and Shalini Aggarwal. 2022. 'A Comparative Study of Supervised and Unsupervised Machine Learning Algorithms on Consumer Reviews'. Pp. 598–603 in *2022 IEEE World Conference on Applied Intelligence and Computing (AIC)*.

Mequanenit, Azanu Mirolgn, Eyerusalem Alebachew Nibret, Pilar Herrero-Martín, María S. García-González, and Rodrigo Martínez-Béjar. 2025. 'A Multi-Agent Deep

Reinforcement Learning System for Governmental Interoperability'. *Applied Sciences* 15(6):3146. doi:10.3390/app15063146.

Mishra, Bhupesh Kumar, Vjosa Preniqi, Dhavalkumar Thakker, and Erich Feigl. 2024. 'Machine Learning and Deep Learning Prediction Models for Time-Series: A Comparative Analytical Study for the Use Case of the UK Short-Term Electricity Price Prediction'. *Discover Internet of Things* 4(1):24. doi:10.1007/s43926-024-00075-4.

Morais, Hugo, Tiago Pinto, and Zita Vale. 2020. 'Adjacent Markets Influence Over Electricity Trading—Iberian Benchmark Study'. *Energies* 13(11):2808. doi:10.3390/en13112808.

Moroff, Nikolas Ulrich, Ersin Kurt, and Josef Kamphues. 2021. 'Machine Learning and Statistics: A Study for Assessing Innovative Demand Forecasting Models'. *Procedia Computer Science* 180:40–49. doi:10.1016/j.procs.2021.01.127.

Morucci, Marco. 2024. 'Model Complexity for Supervised Learning: Why Simple Models Almost Always Work Best, And Why It Matters for Applied Research'.

Nachouki, Mirna, Elfadil Mohamed, Riyadh Mehdi, and Mahmoud Abou Naaj. 2023. 'Student Course Grade Prediction Using the Random Forest Algorithm: Analysis of Predictors' Importance'. *Trends in Neuroscience and Education* 33:100214. doi:10.1016/j.tine.2023.100214.

Naeem, Samreen, Aqib Ali, Sania Anam, and Munawar Ahmed. 2023. 'An Unsupervised Machine Learning Algorithms: Comprehensive Review'. https://www.researchgate.net/publication/368983958_An_Unsupervised_Machine_Learning_Algorithms_Comprehensive_Review.

Nascimento, João, Tiago Pinto, and Zita Vale. 2019. 'Day-Ahead Electricity Market Price Forecasting Using Artificial Neural Network with Spearman Data Correlation'. Pp. 1–6 in *2019 IEEE Milan PowerTech*.

Ng, Sandra, Sara Masarone, David Watson, and Michael R. Barnes. 2023. 'The Benefits and Pitfalls of Machine Learning for Biomarker Discovery'. *Cell and Tissue Research* 394(1):17–31. doi:10.1007/s00441-023-03816-z.

Ngo, Vu Minh, Huan Huu Nguyen, and Phuc Van Nguyen. 2023. 'Does Reinforcement Learning Outperform Deep Learning and Traditional Portfolio Optimization Models in Frontier and Developed Financial Markets?' *Research in International Business and Finance* 65:101936. doi:10.1016/j.ribaf.2023.101936.

Nti, Isaac Kofi, Juanita Ahia Quarcoo, Justice Aning, and Godfred Kusi Fosu. 2022. 'A Mini-Review of Machine Learning in Big Data Analytics: Applications, Challenges, and Prospects'. *Big Data Mining and Analytics* 5(2):81–97. doi:10.26599/BDMA.2021.9020028.

Obiora, Chibuzor, Ahmed Ali, and Ali Hasan. 2021. 'Implementing Extreme Gradient Boosting (XGBoost) Algorithm in Predicting Solar Irradiance'. Pp. 1–5 in *2021 IEEE PES/IAS PowerAfrica*.

Oddleifson, Carly, Stephen Kilgus, David Klingbeil, Alexander Latham, Jessica Kim, and Ishan Vengurlekar. 2025. 'Using a Naive Bayesian Approach to Identify Academic Risk Based on Multiple Sources: A Conceptual Replication'. *Journal of School Psychology* 108:101397. doi:10.1016/j.jsp.2024.101397.

OMIE. 2020. *Day-Ahead Market Operation*. https://www.omie.es/sites/default/files/inline-files/day_ahead_market.pdf.

OMIE. 2025. 'File Access | OMIE'. https://www.omie.es/en/file-access-list.

Oroojlooy, Afshin, and Davood Hajinezhad. 2023. 'A Review of Cooperative Multi-Agent Deep Reinforcement Learning'. *Applied Intelligence* 53(11):13677–722. doi:10.1007/s10489-022-04105-y.

Palanca, Javi. 2020. 'SPADE — SPADE 4.0.1 Documentation'. https://spade-mas.readthedocs.io/en/latest/readme.html.

Pannakkong, Warut, Vu Thanh Vinh, Nguyen Ngoc Minh Tuyen, and Jirachai Buddhakulsomsiri. 2023. 'A Reinforcement Learning Approach for Ensemble Machine Learning Models in Peak Electricity Forecasting'. *Energies* 16(13):5099. doi:10.3390/en16135099.

Park, Keonwoo, and Ilkyeong Moon. 2022. 'Multi-Agent Deep Reinforcement Learning Approach for EV Charging Scheduling in a Smart Grid - ScienceDirect'. https://www.sciencedirect.com/science/article/pii/S030626192201368X?via%3Dihub.

Pinto, Tiago, Zita Vale, Tiago Sousa, Isabel Praça, Gabriel Santos, and Hugo Morais. 2014. 'Adaptive Learning in Agents Behaviour: A Framework for Electricity Markets Simulation'. *Integrated Computer-Aided Engineering* 21(4):399–415. doi:10.3233/ICA-140477.

Qian, Yulin, Yang Zhang, Yuanhe Tang, and Pengfei Ye. 2017. 'Mathematical Modeling and Control Optimization for Energy Internet'. Pp. 1–6 in *2017 IEEE Conference on Energy Internet and Energy System Integration (EI2)*.

Radhakrishnan, Bharat Menon, and Dipti Srinivasan. 2016. 'A Multi-Agent Based Distributed Energy Management Scheme for Smart Grid Applications'. *Energy* 103:192–204. doi:10.1016/j.energy.2016.02.117.

Radhamani, Karthick, Kishore Kumar, and Gokulraj. 2024. 'Deployment of an IoT-Integrated Home Energy Management System Employing Deep Reinforcement Learning'. Pp. 1–4 in *2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)*.

Rajendra, Priyanka, and Shahram Latifi. 2021. 'Prediction of Diabetes Using Logistic Regression and Ensemble Techniques'. *Computer Methods and Programs in Biomedicine Update* 1:100032. doi:10.1016/j.cmpbup.2021.100032.

Ramos, Daniel, Pedro Faria, Luis Gomes, Pedro Campos, and Zita Vale. 2022. 'Selection of Features in Reinforcement Learning Applied to Energy Consumption Forecast in Buildings According to Different Contexts'. *Energy Reports* 8:423–29. doi:10.1016/j.egyr.2022.01.047.

Ramos, Daniel, Mahsa Khorram, Pedro Faria, and Zita Vale. 2021. 'Load Forecasting in an Office Building with Different Data Structure and Learning Parameters'. *Forecasting* 3(1):242–55. doi:10.3390/forecast3010015.

Rando Mazzarino, Pietro, Alberto Macii, Lorenzo Bottaccioli, and Edoardo Patti. 2023. 'A Multi-Agent Framework for Smart Grid Simulations: Strategies for Power-to-Heat Flexibility Management in Residential Context'. *Sustainable Energy, Grids and Networks* 34:101072. doi:10.1016/j.segan.2023.101072.

Rane, Nitin Liladhar, Jayesh Rane, Suraj Kumar Mallick, and Ömer Kaya. 2024. *Scalable and Adaptive Deep Learning Algorithms for Large-Scale Machine Learning Systems*. Deep Science Publishing.

Redeia. 2025. 'Todate'. https://www.ree.es/en/date/todate.

REE. 2025. 'Electricity Business in Spain | Red Eléctrica'. https://www.ree.es/en.

REN. 2025. 'Atividade REN | Energia e Infraestruturas'. https://www.ren.pt/pt-pt/atividade/o-que-fazemos.

REN Data Hub. 2025. 'REN Data Hub'. https://datahub.ren.pt/.

Ribeiro, Bruno, Helder Pereira, Luis Gomes, and Zita Vale. 2023. 'Python-Based Ecosystem for Agent Communities Simulation'. Pp. 62–71 in *17th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2022)*, edited by P. García Bringas, H. Pérez García, F. J. Martinez-de-Pison, J. R. Villar Flecha, A. Troncoso Lora, E. A. de la Cal, Á. Herrero, F. Martínez Álvarez, G. Psaila, H. Quintián, and E. S. Corchado Rodriguez. Cham: Springer Nature Switzerland.

Ribeiro, Bruno, Rafael Silva, Luis Gomes, and Zita Vale. 2024. 'Detailed Analysis of a Deep Learning Energy Forecast Model Considering Different Input Units and Dimensions'. *IFAC-PapersOnLine* 58(13):308–13. doi:10.1016/j.ifacol.2024.07.500.

Rodriguez, Daniela, Mariela Sued, and Marina Valdora. 2025. 'A Kruskal-Wallis Type Test for Functional Data'. *Communications in Statistics - Simulation and Computation* 0(0):1–15. doi:10.1080/03610918.2025.2455418.

Roldan-Fernandez, Juan Manuel, Javier Serrano-Gonzalez, Angel Gaspar Gonzalez-Rodriguez, Manuel Burgos-Payan, and Jesus Manuel Riquelme-Santos. 2024. 'The Effect of Hydropower Bidding Strategy on the Iberian Day-Ahead Electricity Market'. *Energy Strategy Reviews* 55:101517. doi:10.1016/j.esr.2024.101517.

Ruzbahani, Ali Mohammadi. 2024. 'Enhancing Smart Grids with Internet of Energy: Deep Reinforcement Learning and Convolutional Neural Network'.

Saha, Sajal, Anwar Haque, and Greg Sidebottom. 2022. 'Towards an Ensemble Regressor Model for ISP Traffic Prediction with Anomaly Detection and Mitigation'. Pp. 1–6 in *2022 International Symposium on Networks, Computers and Communications (ISNCC)*.

Samadi, Esmat, Ali Badri, and Reza Ebrahimpour. 2020. ‘Decentralized Multi-Agent Based Energy Management of Microgrid Using Reinforcement Learning’. *International Journal of Electrical Power & Energy Systems* 122:106211. doi:10.1016/j.ijepes.2020.106211.

Santos, Gabriel, Tiago Pinto, Isabel Praça, and Zita Vale. 2018. ‘Iberian Electricity Market Ontology to Enable Smart Grid Market Simulation’. *Energy Informatics* 1(1):13. doi:10.1186/s42162-018-0018-2.

Sarker, Iqbal H. 2021. ‘Machine Learning: Algorithms, Real-World Applications and Research Directions’. *SN Computer Science* 2(3):160. doi:10.1007/s42979-021-00592-x.

Schmid, Lena, Moritz Roidl, Alice Kirchheim, and Markus Pauly. 2025. ‘Comparing Statistical and Machine Learning Methods for Time Series Forecasting in Data-Driven Logistics—A Simulation Study’. *Entropy* 27(1):25. doi:10.3390/e27010025.

Shakya, Ashish Kumar, Gopinatha Pillai, and Sohom Chakrabarty. 2023. ‘Reinforcement Learning Algorithms: A Brief Survey’. *Expert Systems with Applications* 231:120495. doi:10.1016/j.eswa.2023.120495.

Sharma, Ajay, and Pramod Kumar Mishra. 2022. ‘Performance Analysis of Machine Learning Based Optimized Feature Selection Approaches for Breast Cancer Diagnosis’. *International Journal of Information Technology* 14(4):1949–60. doi:10.1007/s41870-021-00671-5.

Shen, Rendong, Shengyuan Zhong, Xin Wen, Qingsong An, Ruifan Zheng, Yang Li, and Jun Zhao. 2022. ‘Multi-Agent Deep Reinforcement Learning Optimization Framework for Building Energy System with Renewable Energy’. *Applied Energy* 312:118724. doi:10.1016/j.apenergy.2022.118724.

Shobole, Abdulfetah Abdela, and Mohammed Wadi. 2021. ‘Multiagent Systems Application for the Smart Grid Protection’. *Renewable and Sustainable Energy Reviews* 149:111352. doi:10.1016/j.rser.2021.111352.

Sousa, Tiago, Tiago Pinto, Isabel Praça, Zita Vale, and Hugo Morais. 2014. ‘Reinforcement Learning Based on the Bayesian Theorem for Electricity Markets Decision Support’. Pp. 141–48 in *Distributed Computing and Artificial Intelligence, 11th International Conference*, edited by S. Omatu, H. Bersini, J. Corchado, S. Rodríguez, P. Pawlewski, and E. Bucciarelli. Cham: Springer International Publishing.

Stennikov, Valery, Evgeny Barakhtenko, Gleb Mayorov, Dmitry Sokolov, and Bin Zhou. 2022. ‘Coordinated Management of Centralized and Distributed Generation in an Integrated Energy System Using a Multi-Agent Approach’. *Applied Energy* 309:118487. doi:10.1016/j.apenergy.2021.118487.

Sundaram, Karthic, Akash, Dhanush, Dheenadayalan, and Naveen Jagannathan. 2024. ‘A Deep Learning Approach to Smart Grid Stability Prediction with LSTM Network’. Pp. 265–70 in *2024 2nd International Conference on Advancement in Computation & Computer Technologies (InCACCT)*.

Tan, Chang Wei, Christoph Bergmeir, Francois Petitjean, and Geoffrey I. Webb. 2021. ‘Time Series Extrinsic Regression’.

Tan, Xiaoyu, Yongfeng Zhang, Hui Zhang, and Tongxing Li. 2024. 'Exploration of Hybrid Teaching of Higher Mathematics Based on Artificial Intelligence'. *Open Access Library Journal* 11(12):1–8. doi:10.4236/oalib.1112709.

Tarwidi, Dede, Sri Redjeki Pudjaprasetya, Didit Adytia, and Mochamad Apri. 2023. 'An Optimized XGBoost-Based Machine Learning Method for Predicting Wave Run-up on a Sloping Beach'. *MethodsX* 10:102119. doi:10.1016/j.mex.2023.102119.

team, The pandas development. 2024. 'Pandas-Dev/Pandas: Pandas'.

Teixeira, Brígida, Ricardo Faia, Tiago Pinto, and Zita Vale. 2023. 'Study of Forecasting Methods' Impact in Wholesale Electricity Market Participation'. Pp. 267–76 in *Distributed Computing and Artificial Intelligence, Special Sessions I, 20th International Conference*, edited by R. Mehmood, V. Alves, I. Praça, J. Wikarek, J. Parra-Domínguez, R. Loukanova, I. de Miguel, T. Pinto, R. Nunes, and M. Ricca. Cham: Springer Nature Switzerland.

Torres, Martínez-Álvarez, and Troncoso. 2022. 'A Deep LSTM Network for the Spanish Electricity Consumption Forecasting'. *Neural Computing and Applications* 34(13):10533–45. doi:10.1007/s00521-021-06773-2.

Traldi, Andrea, Francesco Bruschetti, Marco Robol, Marco Roveri, and Paolo Giorgini. 2022. 'Real-Time BDI Agents: A Model and Its Implementation'.

Tschora, Léonard, Erwan Pierre, Marc Plantevit, and Céline Robardet. 2022. 'Electricity Price Forecasting on the Day-Ahead Market Using Machine Learning'. *Applied Energy* 313:118752. doi:10.1016/j.apenergy.2022.118752.

Turmudi, Setya Utari, Suprih Widodo, Ratnaningsih, Turmudi, Setya Utari, Suprih Widodo, and Ratnaningsih. 2017. 'Mathematics Instruction Based on Science Using Didactical Phenomenology Approach in Junior Secondary School in Indonesia'. in *Science Education - Research and New Technologies*. IntechOpen.

Uniejewski, Bartosz. 2024. 'Regularization for Electricity Price Forecasting'. *Operations Research and Decisions* 34(3). doi:10.37190/ord240314.

Verdejo, Humberto, Almendra Awerkin, Cristhian Becker, and Gabriel Olguin. 2017. 'Statistic Linear Parametric Techniques for Residential Electric Energy Demand Forecasting. A Review and an Implementation to Chile'. *Renewable and Sustainable Energy Reviews* 74:512–21. doi:10.1016/j.rser.2017.01.110.

Vladimir, Nasteski. 2017. 'An Overview of the Supervised Machine Learning Methods'. *ResearchGate* 11.

Wang, Feng, Shicheng Li, Shanshui Niu, Haoran Yang, Xiaodong Li, and Xiaotie Deng. 2025. 'A Survey on Recent Advances in Reinforcement Learning for Intelligent Investment Decision-Making Optimization'. *Expert Systems with Applications* 282:127540. doi:10.1016/j.eswa.2025.127540.

Wang, Jianrui, Yitian Hong, Jiali Wang, Jiapeng Xu, Yang Tang, Qing-Long Han, and Jürgen Kurths. 2022. 'Cooperative and Competitive Multi-Agent Systems: From Optimization to

Games'. *IEEE/CAA Journal of Automatica Sinica* 9(5):763–83. doi:10.1109/JAS.2022.105506.

Wang, Zhaobin, Ke Liu, Jian Li, Ying Zhu, and Yaonan Zhang. 2024. 'Various Frameworks and Libraries of Machine Learning and Deep Learning: A Survey'. *Archives of Computational Methods in Engineering* 31(1):1–24. doi:10.1007/s11831-018-09312-w.

Wei, Pengyuan. 2020. 'Exploration-Exploitation Strategies in Deep Q-Networks Applied to Route-Finding Problems'. *Journal of Physics: Conference Series* 1684(1):012073. doi:10.1088/1742-6596/1684/1/012073.

Wu, Haoke, Tao Huang, Stefania Conti, and Ettore Bompard. 2024. 'A Framework for Assessing Electricity Market Performance under Different Bidding Zone Configurations'. *Energies* 17(11):2743. doi:10.3390/en17112743.

Yan, Ziwei, Liang Han, Xiaoduo Li, Xiwang Dong, Qingdong Li, and Zhang Ren. 2023. 'Event-Triggered Formation Control for Time-Delayed Discrete-Time Multi-Agent System Applied to Multi-UAV Formation Flying'. *Journal of the Franklin Institute* 360(5):3677–99. doi:10.1016/j.jfranklin.2023.01.036.

Ye, Eric, Ren Tao, and Natasha Jaques. 2025. 'An Efficient Open World Environment for Multi-Agent Social Learning'.

Yousaf, Adnan, Rao Muhammad Asif, Mustafa Shakir, Ateeq Ur Rehman, Fawaz Alassery, Habib Hamam, and Omar Cheikhrouhou. 2021. 'A Novel Machine Learning-Based Price Forecasting for Energy Management Systems'. *Sustainability* 13(22):12693. doi:10.3390/su132212693.

Yu, Liang, Shuqi Qin, Meng Zhang, Chao Shen, Tao Jiang, and Xiaohong Guan. 2021. 'A Review of Deep Reinforcement Learning for Smart Building Energy Management'. *IEEE Internet of Things Journal* 8(15):12046–63. doi:10.1109/JIOT.2021.3078462.

Zhang, Dongxia, Xiaoqing Han, and Chunyu Deng. 2018. 'Review on the Research and Practice of Deep Learning and Reinforcement Learning in Smart Grids'. *CSEE Journal of Power and Energy Systems* 4(3):362–70. doi:10.17775/CSEEJPES.2018.00520.

Zhang, Nan, Xiaojie Lin, Wei Zhong, and Fanqi Kong. 2022. 'A Study on Integrated Energy System Long-Term Planning Optimization Considering Complementarity and Economy'. Pp. 3129–34 in *2022 IEEE 6th Conference on Energy Internet and Energy System Integration (EI2)*.

Zhang, Shihan, Wei Nai, Yujie Qiu, Wei Xu, Zan Yang, Dan Li, and Yidan Xing. 2021. 'Ridge Regression Based on Glowworm Swarm Optimization Algorithm with T-Distribution Parameters'. Pp. 1–4 in *2021 IEEE 11th International Conference on Electronics Information and Emergency Communication (ICEIEC)2021 IEEE 11th International Conference on Electronics Information and Emergency Communication (ICEIEC)*.

Zhu, Ziqing, Ze Hu, Ka Wing Chan, Siqi Bu, Bin Zhou, and Shiwei Xia. 2023. 'Reinforcement Learning in Deregulated Energy Market: A Comprehensive Review'. *Applied Energy* 329:120212. doi:10.1016/j.apenergy.2022.120212.