Human-Centred AI Masters
December 2025
THES H6011 Research Thesis, 30 Credits

Economic sustainability of photovoltaic systems

Musa Bakarr

A thesis submitted in the partial fulfilment for the
Masters in Science in Human-Centred Artificial Intelligence

in the
Faculty of Computing, Digital and Data
School of Enterprise Computing and Digital Transformation

Main Supervisor:	PhD, Abdelsalam Busalim, School of Enterprise Computing and Digital Transformation, TU Dublin
Industry Partner:	Ing., Gabriele Piantadosi, PhD student, Sofia Dutto, ENEA, Italy

Declaration

I hereby certify that the material, which I now submit for assessment on the programmes of study leading to the award of Master of Science, is entirely my own work and has not been taken from the work of others except to the extent that such work has been cited and acknowledged within the text of my work. No portion of the work contained in this thesis has been submitted in support of an application for another degree or qualification to this or any other institution.

MB.

Musa Bakarr
September 3, 2025

Abstract

Photovoltaic (PV) systems are pivotal to achieving global sustainability goals. PV systems offer a clean source of energy crucial for the ongoing energy transition, aligning with the United Nations Sustainable Development Goals. A major concern is the fluctuating nature of solar energy, influenced by meteorological factors such as temperature and irradiance. These fluctuations directly impact market participants, including grid operators and prosumers, by introducing uncertainties in energy supply and demand, leading to potential imbalance settlement fees and increased operational costs. This research aims to solve the issue by investigating day-ahead photovoltaic (PV) energy forecasting, with France data from ENTSO-E and Open Meteo chosen as a case study. It investigates the hypothesis which states that by accurately forecasting energy production for energy markets, and simultaneously predicting the potential errors in these forecasts, PV operators can significantly enhance their economic sustainability.

This work employs 6 models (Linear regression, Random forest, MLP, XGBoost, LSTM and CNN-LSTM) which would be compared against each other and the ENTSO-E baseline. Through comprehensive model training, error quantification, and economic analysis, the results demonstrate that the LSTM model outperformed the other models. The proposed model reduces forecasting penalties by 62.7%, corresponding to €136 million in avoided losses, highlighting its potential for substantial financial and operational benefits. Explainable AI (XAI) techniques, SHAP and LIME were applied on the LSTM model. This human centric approach was implemented to provide insights into model behavior and model decision-making process to users and stakeholders, placing energy market participants at the core of the forecasting process.

Keywords

Deep learning, Energy market, Ethical AI, Explainability, Photovoltaic system, Sustainability

Acknowledgement

I would like to express my most sincere gratitude to my academic supervisor Dr. Abdelsalam Busalim, and to my industry supervisors Ing. Gabriele Piantadosi and Sofia Dutto from ENEA, Italy. Their guidance, kind collaboration, and relentless support was fundamental towards this project's success.

Contents

List of Abbreviations

AI- Artificial Intelligence
CNN- Convolution Neural Network
DL - Deep Learning
ENTSO-E - European Network of Transmission System Operators for Electricity
LIME- Local Interpretable Model-agnostic Explanations
LSTM- Long Short Term Memory
ML - Machine Learning
MLP - Multilayer Perceptron
MAE - Mean Absolute Error
MAPE - Mean Absolute Percentage Error
MBE - Mean Bias Error
PV - Photovoltaic
RMSE - Root Mean Square Error
SHAP - SHapley Additive exPlanations
UN SDG- United Nations Sustainable Development Goals

List of Figures

List of Tables

1 Introduction

1.1 Introduction

The transition to renewable energy is reshaping global power systems, with solar photovoltaic (PV) technology playing a pivotal role in achieving sustainability goals. This chapter outlines the significance of PV systems in meeting the energy demands of modern societies while addressing environmental challenges. It explores the critical role of forecasting in optimizing energy markets, particularly within the European Union (EU), where renewable integration is a priority. The motivations for leveraging advanced computational methods are discussed, alongside the studys scope, its alignment with user-focused AI principles, and the research objectives guiding the investigation. The chapter concludes with the thesis structure, providing a roadmap for the exploration of PV forecasting.

1.1.1 Background Information

Photovoltaic systems harness sunlight to generate clean electricity, offering a sustainable alternative to fossil fuels. In the EU, where energy policies aim for carbon neutrality by 2050, PV systems are central to reducing greenhouse gas emissions and diversifying energy sources [1]. PV systems directly align with the United Nations Sustainable Development Goals (SDGs), particularly Goal 7 (Affordable and Clean Energy) and Goal 13 (Climate Action) due to their low carbon emission [2]. The rapid expansion of PV installations, supported by technological advancements and economic incentives, has transformed energy markets, enabling greater flexibility in supply and demand management. However, the intermittent nature of solar energy, driven by weather variability and diurnal cycles, necessitates precise forecasting to ensure grid reliability and economic efficiency.

Forecasting PV output supports energy market operations by reducing operational costs, limiting economic losses, and minimizing reliance on non-renewable backup systems. By predicting energy production, stakeholders can optimize trading strategies, enhance grid stability, and facilitate the integration of renewables into interconnected EU markets.

1.1.2 Motivation

The primary motivation for this research is the pressing need to bridge the gap between the technical capabilities of PV systems and their economic performance in liberalized EU energy markets. While machine learning (ML) and deep learning (DL) models have proven effective in forecasting PV production [3][4], their application must be extended to address the economic dimension of sustainability. This study is driven by the hypothesis that by accurately forecasting energy production for energy markets, and simultaneously predicting the potential errors in these forecasts, PV operators can significantly enhance their economic sustainability [5][6][7]. Furthermore, this research is motivated by the principles of a human-centered approach to AI. For forecasting tools to be adopted, they must be trustworthy and transparent. Therefore, incorporating Explainable AI (XAI) is essential to ensure that stakeholders can understand and rely on the model's outputs [8]. This encourages a symbiotic relationship between technology and human decision-making in the pursuit of a sustainable energy economy.

1.2 Scope and Limitations of the Research

This thesis focuses on day-ahead PV energy forecasting within the EU, using France energy data from ENTSO-E transparency platform [9] and weather data from Open Meteo [10] as case study due to data availability and transparency. The study employs ML approaches, starting with linear regression as a baseline to establish predictive performance, to DL models like Long Short-Term Memory (LSTM) networks and hybrid models for capturing temporal dependencies. The models leverage ENTSO-E energy data and open meteo meteorological inputs from 2020 to 2025, with a test period starting June 1, 2024. The forecasted PV production of the models is compared with historical ENTSO-E forecasts data, which serves as a baseline. Model performance is evaluated using metrics such as Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). Furthermore, the study incorporates an evaluation of forecast errors, their associated imbalance penalties, and the quantification of net financial impacts in terms of penalties and revenues.

While this research provides insights into improving forecasting accuracy and its economic implications, the analysis is restricted to the day-ahead market. Broader strategies, such as the intra-day and real-time market where forecast adjustments can be made to further reduce imbalance costs, and the integration of energy storage technologies which can mitigate penalties by shifting or smoothing PV output, are acknowledged as complementary but fall outside the scope of this work.

1.3 Problem Description - Relevance to Human-Centric AI

Day-ahead photovoltaic (PV) forecasting faces several real-world challenges that impact its effectiveness in energy markets[5][11]. Variable weather conditions such as frequent cloud cover and low solar irradiance in winter introduce significant uncertainty in predicting PV energy output[12]. This variability complicates grid management, as inaccurate forecasts can lead to energy imbalances, increased reliance on costly fossil fuel reserves, or unnecessary curtailment of renewable energy [5]. A human-centric AI approach, addresses these challenges by placing energy market participants at the core of the forecasting process. The framework begins with identifying real-world needs, which is accurate forecasting to support grid stability and market efficiency. Model development prioritizes both accuracy and interpretability, using glassbox models such as linear regression as a baseline for its transparency and employing XAI techniques, SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) in blackbox models to clarify how inputs influence predictions. These provided interpretable insights, makes the decision making of models transparent thereby building trust and acceptance.

Research Questions

This research is guided by the following pivotal questions:

1. How can advanced machine learning and deep learning models be effectively designed and implemented to achieve superior accuracy in day-ahead energy production forecasting?
2. What methodologies are most effective for quantifying and forecasting the errors associated with PV production predictions, thereby enabling better risk assessment, loss reduction and market participation?
3. To what extent can the integration of these forecasting and error prediction capabilities enhance the overall economic sustainability of PV systems within competitive energy markets?
4. How can Explainable AI (XAI) techniques be successfully applied to deep learning models in PV forecasting to promote transparency, build trust, and support human decision-making in the energy sector?

Research Objectives

To address the above questions, this thesis pursues the following specific objectives:

1. To design, implement, and evaluate a range of forecasting models (from linear regression to advanced deep learning architectures such as LSTM and CNN-LSTM) for day-ahead PV energy forecasting in France.
2. To quantify forecast errors and investigate methods for predicting and mitigating these errors in order to reduce imbalance costs in the day-ahead energy market.
3. To assess the economic sustainability of PV systems by linking forecasting performance to financial impacts, including imbalance penalties and net revenues.
4. To integrate XAI methods (SHAP, LIME) into deep learning models, ensuring model outputs are interpretable, trustworthy, and aligned with human-centered AI principles.

1.4 Structure of the Thesis

This dissertation is structured to systematically address these objectives across six chapters. Chapter 1 lays the groundwork, introducing the research problem and its significance. Chapter 2 provides a critical review of existing literature on PV forecasting, energy markets, ML & DL applications, and the intersection of economic sustainability with human-centered AI. Chapter 3 details the datasets employed and the rigorous data preprocessing steps. Chapter 4 delves into the architectural design and implementation of the proposed forecasting models. Chapter 5 presents a thorough analysis of the experimental results, discussing model performance, economic implications, and the insights gained from XAI. Finally, Chapter 6 synthesizes the key findings, outlines the contributions of this research, and proposes avenues for future exploration.

2 Literature Review

2.1 Introduction to the Literature Review

The literature review begins by establishing the foundational understanding of photovoltaic systems, their global adoption trends, operational characteristics, and their pivotal role in achieving broader sustainability objectives. Subsequently, it delves into the intricacies of European electricity markets, examining their design, various market types (spot, forward, balancing), pricing mechanisms, and the economic implications of forecast errors. A significant portion of this chapter is dedicated to state-of- the-art PV forecasting techniques, exploring different methodologies, forecasting horizons, and the critical aspect of uncertainty quantification and error prediction. Finally, the review addresses the emerging paradigm of human-centered AI in energy forecasting, discussing principles of trustworthiness, transparency, and fairness, and examining the application of Explainable AI (XAI) techniques to enhance the interpretability and adoption of AI-driven solutions in the energy sector. By identifying current research gaps and opportunities, this chapter sets the stage for the novel contributions of this thesis.

2.2 Photovoltaic Systems and Energy Transition

2.2.1 PV system structure & operation

A PV system is an assembly of components designed to convert sunlight directly into electricity. At its core, a PV system comprises one or more solar panels, commonly known as PV panels, which contain photovoltaic cells. These cells absorb photons from sunlight and convert solar energy into direct current (DC) electricity through the photovoltaic effect [13], [14]. The basic structure of a typical grid-tied PV system includes several key components:

- Solar Panels (PV Modules): These are the primary components, consisting of multiple PV cells wired together and encapsulated to form a single unit. [13].
- Inverter: This instrument converts DC electricity produced by the solar panels into to AC. It is a crucial component for grid integration [15].
- Mounting System: PV arrays must be securely mounted on stable structures, typically rooftops or ground-mounted racks, designed to support the panels and withstand various weather conditions over decades [13].
- Balance of System (BOS) Components: This category includes all other necessary electrical and mechanical hardware, such as wiring, circuit breakers, disconnects, and metering equipment, which ensure the safe and efficient operation of the system [16].

2.2.2 Global PV adoption trend

The global adoption of solar PV technology has witnessed unprecedented growth over the past decade. In 2023, power generation from solar PV increased by a record 320 TWh, marking a 2% increase from 2022 and accounting for 5.4% of total global electricity generation [17]. For the first time, PV energy production exceeded 10% of the world's electricity consumption [18].

By the end of 2022, the global cumulative installed capacity of solar PV systems reached 1,177 GW, a significant jump from 938 GW in the previous year [SolarPowerEurope]. Projections indicate continued growth, with the global photovoltaics market expected to be valued at €524,14 billion in 2025 and reach €827,19 billion by 2030, demonstrating a Compound Annual Growth Rate (CAGR) of 9.6% [19].

2.2.3 Sustainability contribution

The UN-SDGs are a global framework that serves to balance the main objectives of sustainability [20].The primary contribution of PV technology lies in its ability to generate electricity with minimal environmental impact, especially concerning greenhouse gas emissions. This directly aligns with SDG 13. The widespread adoption of PV technology strengthens SDG 9 (Industry, Innovation, and Infrastructure) through the development of modern, sustainable energy systems. In addition, the expansion of solar energy contributes directly to SDG 7 (Affordable and Clean Energy) by providing accessible, sustainable electricity. Accurate forecasting empowers stakeholders to plan more effectively, reducing inefficiencies and supporting SDG 12 (Responsible Consumption and Production). Collectively, these advancements contribute to SDG 11 (Sustainable Cities and Communities), helping to create resilient, sustainable urban environments powered by clean energy[2].

Figure 2.1: Sustainable Development Goals. Source: United Nations.

2.3 Energy Markets in the European Union

2.3.1 Overview of the EU electricity market

The EU electricity market design is characterized by a phased liberalization process that began in the late 1990s, aiming to move away from national monopolies towards an integrated and competitive market [21]. The structure of the EU electricity market is built upon a framework of common energy market rules. Key elements include:

- Unbundling: Separation of generation, transmission, distribution, and supply activities to prevent anti-competitive practices and ensure fair access to the grid for all market participants.
- Harmonization of Rules: Development of common regulations and network codes to facilitate seamless cross-border electricity flows and market integration.
- Market Coupling: A mechanism that implicitly allocates cross-border transmission capacity by matching bids and offers across different bidding zones, leading to a more efficient use of interconnectors and convergence of prices [22][23].

2.3.2 Market types

The EU electricity market operates through a series of interconnected market segments with each serving distinct functions and timeframes. These market types allow for the continuous trading of electricity, from long-term commitments to real-time balancing. The interplay of these markets, spot (day-ahead and intraday) market, forward market, and balancing markets, form the cornerstone of modern electricity trading and grid reliability.

Market Type	**Details**	**Definition**
Spot Market	Day-Ahead	Participants submit their bids and offers for the delivery of electricity during each hour of the following day before a certain cutoff time.
	Intraday	This operates between the closure of the day-ahead market and the real-time delivery; typically in 15-minute or 30-minute intervals
Balancing Market	Real time operation	Operates after the day-ahead and intra day market is closed, ensures an immediate equilibrium between electricity supply and demand.
Forward Market	Forward & future contracts	These markets enable participants to buy and sell electricity for future delivery, typically months or years in advance.

Table 2.1: Market types

2.4 Economic Sustainability of PV Systems

Economic sustainability in PV systems refers to the ability of solar energy producers to maintain long-term financial viability while contributing to a decarbonized energy system. Beyond the environmental benefits, PV operators must ensure that their investments remain profitable in increasingly competitive electricity markets. This requires maximizing revenues from energy sales while minimizing costs associated with imbalance penalties and operational inefficiencies. In the context of European liberalized markets, day-ahead forecasts directly influence market participation and profitability. Inaccurate forecasts expose producers to imbalance settlement fees and reliance on costly balancing energy, which erodes financial sustainability.

2.4.1 Economic implications of forecast errors

Forecast errors, particularly in the context of variable renewable energy sources like PV, have significant economic implications for market participants and the overall electricity system. The core issue arises from the discrepancy between forecasted and actual electricity generation or consumption, leading to imbalances that must be resolved to maintain grid stability.

European electricity markets rely on a sophisticated system of penalties to maintain grid stability, ensure market participants adhere to their commitments, and recover costs associated with managing imbalances. These penalties are designed to be effective, dissuasive, and non-discriminatory, as mandated by EU regulations[24][25]. The EU's Electricity Balancing Guideline (Regulation (EU) 2017/2195) provides a framework for harmonizing imbalance settlement rules across Member States, emphasizing cost-reflectivity, transparency, and non-discrimination. This guideline allows for either a single imbalance price or a dual price system, where different prices are applied for positive and negative imbalances. National TSOs then develop specific methodologies for calculating these prices, subject to approval by their respective regulatory authorities (Table 2.2). These pricing mechanisms, whether single or dual, aim to incentivize Balance Responsible Parties (BRPs) to accurately forecast and manage their positions, thereby minimizing deviations and contributing to overall system balance [26][27][28]. In instances of energy shortages or overproduction, the TSO would have to add or withdraw power depending on the transmission grid's status. The cost of this stabilization measure as a result leads to a penalty for the suppliers responsible for the imbalance. In dual imbalance pricing systems, penalties are classified as either falling penalties, for underproduction, or rising penalties, for overproduction [29]. Penalties are only imposed when a TSO intervention has taken place. Beyond imbalance charges, penalties for market manipulation or non-compliance with market rules are also enforced by national regulatory authorities and bodies like ACER, which can result in significant fines [5][30].

2.5 Existing PV Forecasting Techniques

Forecasting methods for PV generation can be broadly divided into physical, statistical, and data-driven approaches:

Physical methods:

PV forecasting has traditionally relied on physical models that simulate the underlying processes governing solar energy generation. Physical models rely on engineering principles and system characteristics to simulate PV output. These models are grounded in the physics of solar irradiance, module characteristics, and environmental conditions.

A widely adopted approach is the Nominal Operating Cell Temperature (NOCT) model, which estimates the PV cell temperature as a function of ambient temperature, solar irradiance, and wind speed. Since module temperature directly affects conversion

Country	Electricity Market Operator(s)	Transmission System Operator (TSO)
Austria	EPEX SPOT, EXAA	Austrian Power Grid (APG)
Belgium	EPEX SPOT	Elia Transmission Belgium
Bulgaria	IBEX (Independent Bulgarian Power Exchange)	ESO (Electroenergien Sistemen Operator)
Croatia	CROPEX	HOPS (Croatian TSO)
Czechia	OTE (Market Operator)	ČEPS
Denmark	Nord Pool, EPEX SPOT	Energinet
Estonia	Nord Pool, EPEX SPOT	Elering
Finland	Nord Pool, EPEX SPOT	Fingrid
France	EPEX SPOT	RTE (Réseau de Transport d'Électricité)
Germany	EPEX SPOT, EEX	TenneT, Amprion, TransnetBW, 50Hertz
Greece	Designated NEMO(s)	IPTO / ADMIE
Hungary	PXE, HUPX	MAVIR
Ireland	SEMO	EirGrid
Italy	GME (IPEX)	Terna
Latvia	Nord Pool, EPEX SPOT	AST (Augstsprieguma tīklS)
Lithuania	Nord Pool, EPEX SPOT	Litgrid
Luxembourg	EPEX SPOT, Nord Pool	Creos Luxembourg
Netherlands	EPEX SPOT, Nord Pool	TenneT
Poland	POLPX, Nord Pool, EPEX SPOT	PSE (Polskie Sieci Elektroenergetyczne)
Portugal	OMIE	REN (Redes Energéticas Nacionais)
Romania	OPCOM	Transelectrica
Slovakia	PXE	SEPS
Slovenia	BSP (BSP Energetska Borza)	ELES
Spain	OMIE, OMIP, MEFF, EEX	REE (Red Eléctrica de España)
Sweden	Nord Pool, EPEX SPOT	Svenska Kraftnät

Table 2.2: EU Countries: Electricity Market Operators and TSOs

efficiency, NOCT-based models provide a simplified but effective way to forecast PV output under varying environmental conditions [31]. The Sandia PV Array Performance Model extends physical modeling by incorporating empirical coefficients derived from laboratory and field testing of PV modules. This approach enables more accurate characterization of module behavior under diverse irradiance and temperature conditions, and has been widely adopted for both research and utility-scale forecasting applications [32]. Similarly, the Fuentes thermal model provides a refined estimation of cell temperature by explicitly considering thermal exchanges between the PV module, the surrounding air, and the mounting structure. This is particularly relevant in forecasting, as temperature sensitivity remains one of the largest sources of uncertainty in PV performance modeling [33].

While physical models have the advantage of interpretability and do not require large volumes of historical training data, they are sensitive to the accuracy of meteorological inputs (irradiance, wind, temperature) and system-specific parameters. Moreover, their deterministic nature means they may struggle to capture nonlinearities or site-specific anomalies present in real-world PV performance. For this reason, there has been a growing trend toward hybrid approaches that combine the interpretability of physical models with the adaptability of ML methods, aiming to achieve higher forecasting accuracy while retaining transparency [4].

Statistical methods:

The incorporation of historical data in PV forecasting introduced classical statistical models. These approaches aimed to identify statistical relationships and patterns using these historical data. Common statistical methods included:

- Time series models: such as Autoregressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), and Exponential Smoothing, which model the temporal dependencies in PV power data [34].

- Regression models: Including linear regression, multiple linear regression, and support vector regression (SVR), which establish relationships between PV output and input variables like weather data [35].

These models proved effective in capturing linear dependencies and predictable patterns however, their performance was often limited by their inability to adequately capture the complex non-linear relationships.

Machine learning methods:

ML models offer greater flexibility and can capture more complex, non-linear patterns than traditional statistical methods. Popular ML techniques for PV forecasting include:

- Ensemble Methods: The mid-2010s saw increasing interest in ensemble learning techniques, such as Random Forests and Gradient Boosting Machines (GBMs). These methods proved particularly effective in handling complex interactions between input features and were less sensitive to noise compared to single models. Random Forest models have been applied on day ahead predictions and achieved a 92% accuracy [36][37][38].

- Support Vector Machines (SVMs): They are effective for both regression and classification tasks and are often used for their generalization capabilities[39]. SVMs operate by finding an optimal hyperplane that best separates data points. These models generalizes well on unseen data, are robust to outliers, and effective in capturing non-linear relationships through various kernel functions. [40] achieved 98% accuracy, specificity, and sensitivity by combining KNN and SVM for improved structural and data diversity in solar power forecasting.

- Artificial Neural Networks (ANNs): ANNs, particularly Multilayer Perceptrons (MLPs), gained popularity due to their universal approximation capabilities, allowing them to model complex non-linear relationships between inputs and outputs without explicit functional forms [41]. Radicioni et al (2021) customized an ANN on the basis of the particular season of the year [42]. The aim of this work was to furnish accurate predictions even in the case of strong irregularities of solar irradiance, providing accurate results in rapidly changing scenarios. Nicoletti and Bevilacqua (2024) proposed using numerical weather prediction (NWP) data instead of solar radiation on feedforward neural networks (FFNNs) [43]. Their model achieved an R^2 of 0.879 and RMSE of 10.5%.

References	Title	Results	Method
[44]	Probabilistic day-ahead prediction of PV generation. A comparative analysis of forecasting methodologies and of the factors influencing accuracy	Various	Statistical approach
[45]	Random forest machine learning algorithm based seasonal multi-step ahead short-term solar photovoltaic power output forecasting	Accuracy - 50%	Random Forest
[37]	Industry Experience of Developing Day-Ahead Photovoltaic Plant Forecasting System Based on Machine Learning	Accuracy - 92%.	Random Forest, Gradient Boosting Regressor, Linear Regression, Decision Trees regression
[46]	Machine Learning-Based Forecasting of Temperature and Solar Irradiance for Photovoltaic Systems	Accuracy	Random Forest, Decision Trees, Support Vector Machines, XGBoost
[36]	Machine Learning Based Solar Photovoltaic Power Forecasting: A Review and Comparison	Skill score - 37.33%	Gradient Boost, XGBoost, SVR, Random Forest, Lasso Regression, Ada Boost, MLP, Neural Networks
[42]	Power Forecasting of a Photovoltaic Plant Located in ENEA Casaccia Research Center	Result dependent on weather	Artificial Neural Network
[47]	Forecasting of Photovoltaic Solar Power Production Using LSTM Approach	MAPE - 8.93%	Long Short-Term Memory (LSTM)
[48]	Solar Photovoltaic Forecasting of Power Output Using LSTM Networks	RMSE - 0.11368	Long Short-Term Memory (LSTM)
[49]	Deep learning approach for one-hour ahead forecasting of energy production in a solar-PV plant	Accuracy - 98%	Long Short-Term Memory (LSTM)
[50]	Short-term self consumption PV plant power production forecasts based on hybrid CNN-LSTM, ConvLSTM models	Various	CNN-LSTM, ConvLSTM
[51]	Short-term self consumption PV plant power production forecasts based on hybrid CNN-LSTM, ConvLSTM models	MAE - 14.6079	CNN-BiLSTM

Table 2.3: State-of-the-art models

Deep learning methods:

- Long Short-Term Memory (LSTMs): With the ability to capture and learn from long-term dependencies in sequential data, LSTMs are well-suited for tasks involving time series data, where the order and temporal relationships between data points are crucial. Konstantinou acheived an RMSE of 0.11368 in a work focused on the impact of time scale on forecasting errors using LSTM networks [48]. There have been other groundbreaking results from the incorporation of LSTM in PV forecasting. These include Ozbek et al. (2022) achieving an accuracy of 0.98 in a work that highlights the importance of accurate predictions to mitigate random fluctuations [49], and a MAPE of 8.93 achieved by Harrou in a work that used LSTM networks for forecasting PV solar power production[47].

- Hybrid Deep Learning Models: Hybrid deep learning models are a more novel models that aims to combine two or more different neural network architectures.

> The goal is to achieve superior performance by addressing limitations that a single model might have, especially when dealing with complex, heterogeneous data or requiring features like interpretability. Salman et al. (2024) focused on hybrid models for time series forecasting of solar power [52], Agga worked on short-term self-consumption PV plant power production forecasts using CNN-LSTM and ConvLSTM hybrid models[50], and Liu and Mao (2024) used CNN-BiLSTM with attention mechanisms to improve the stability of the model against adverse weather conditions [51].

These models achieve very high accuracy, with reported R^2 values up to 0.98 and RMSE as low as 0.113, demonstrating their potential for operational use. However, they are computationally intensive and require substantial datasets for training and their "black-box" nature limits interpretability, which may hinder adoption in operational decision-making contexts. Previous works involving state-of-the-art models can be found in Table 2.3.

Overall, a critical comparison of these methods highlights a clear performance hierarchy: physical and statistical models provide moderate accuracy suitable for baseline forecasts; ML approaches improve predictive capability through modeling nonlinear interactions; and deep learning, particularly hybrid architectures, consistently delivers the highest forecasting accuracy, capturing both temporal and nonlinear complexities inherent in PV generation.

2.6 Human-Centered AI in Energy Forecasting

The core principles of Human-Centered AI are fundamental to its successful and ethical deployment in the energy sector. To operationalize these principles, this thesis applies an ethical matrix framework, which systematically maps key ethical dimensions to the primary stakeholders in energy systems (see Table 2.4). The ethical matrix provides a structured way to evaluate how AI forecasts can support or potentially conflict with the interests and responsibilities of each stakeholder group. By integrating this ethical matrix into the design and evaluation of AI forecasting models, this thesis ensures that technical performance is aligned with human-centered principles. Subsequent sections on model implementation and XAI will refer back to these ethical considerations, demonstrating how the proposed methodologies not only improve predictive accuracy but also uphold the key human-centred values across all stakeholders.

2.6.1 Explainable AI

Unlike ML models, deep learning models are complex and require the use of XAI techniques to bridge the gap between model complexity and human interpretability. SHAP is a widely applied XAI technique used for explainability [53]. It is based on cooperative game theory where SHAP values quantify the contribution of each feature to a prediction by considering all possible combinations of features. This provides a consistent and locally accurate explanation for individual predictions, as well as into feature importance. Another widely used XAI technique, LIME, explains the predictions of any black-box machine learning model by approximating it locally with an interpretable model (e.g., linear model). It generates explanations by perturbing the input data and observing how the model's predictions change [54].

Despite the clear benefits of XAI implementation, it still faces some limitations. These include the computational overhead of some XAI methods, the complexity of explaining highly non-linear deep learning models, and the need for domain-specific expertise to interpret XAI outputs effectively. However, ongoing research is focused on

Principles	Producers	Consumers	Regulators & Policy Makers	AI Developers & Researchers
Trust	Reliable forecasts optimize bidding and reduce penalties.	Confidence that energy supply is reliable and stable, which in turn avoids unnecessary cost volatility.	Ensures market integrity by minimizing risks from inaccurate forecasts.	Build models that are efficient, validated, and secure against adversarial attacks.
Transparency	Clear explanations of forecast outputs support better operational decisions.	Understand how energy availability may affect household bills.	Ability to audit models for compliance with fairness and competition rules.	Use interpretable ML/Explainable AI to clarify model reasoning.
Fairness	Equal access to accurate forecasts across large and small producers.	Prevents bias in allocation of energy resources or discriminatory pricing.	Upholds fairness across market participants and prevents systemic disadvantages.	Ensure datasets and models do not embed biases.
Accountability	Producers remain accountable for decisions based on AI forecasts.	Consumers can raise issues if forecasts indirectly cause unfair energy costs.	Regulators can assign responsibility when forecasts cause imbalance.	Developers document design choices and limitations of models.
Human Oversight	Operators can intervene when forecasts diverge significantly from reality.	Consumers retain rights to question automated decisions affecting them.	Regulators oversee the use of AI in energy markets to ensure compliance.	AI tools are designed with "human-in-the-loop" mechanisms.
Privacy	Sensitive operational data from plants is safeguarded during data sharing.	Consumer energy usage data is protected from misuse.	Compliance with GDPR and EU energy data regulations.	Adopt privacy-preserving ML methods (e.g., anonymization, federated learning).

Table 2.4: Ethical Matrix for Human-Centered AI in Energy Forecasting

developing more efficient and user-friendly XAI tools, paving the way for more intelligent and human-aligned energy systems[55].

2.7 Research Gaps and Opportunities

While PV forecasting has achieved high levels of accuracy in predicting power output, there is a notable gap in directly integrating economic sustainability metrics into the forecasting models themselves. Most existing models focus on minimizing technical errors (e.g., RMSE, MAE) rather than optimizing economic outcomes (e.g., maximizing revenue, minimizing imbalance costs, improving ROI).

Few studies explicitly incorporate economic or market-based metrics as part of the model design or evaluation [5]][7]. For instance, while a model may achieve high accuracy in predicting PV output, it may not minimize the financial penalties associated with forecast deviations in day-ahead electricity markets. This gap highlights an opportunity to develop forecasting frameworks that integrate market prices, imbalance penalties, and revenue optimization directly into model training and evaluation to improve PV system economic sustainability.

3 Data Collection and Preparation

3.1 Data Collection

Data were sourced from the ENTSO-E Transparency Platform and Open Meteo weather API, selected for their comprehensive energy and meteorological variables critical for PV forecasting. The ENTSO-E dataset was accessed via API, providing hourly records for France from 2020 to 2025. The Open Meteo dataset, also accessed via API, provides hourly meteorological data, averaged across 11 different locations in France, to support solar angle calculations. Both datasets were processed in CSV format using Python (pandas) for temporal alignment and integration.

3.2 Data Description

The ENTSO-E dataset for France comprises 47,472 hourly records with variables: energy forecasted (MWh), energy consumed (MWh), & energy price (EUR/MWh). Since France uses a dual imbalance pricing system, two variables (+ imbalance (EUR/MWh), and - imbalance (EUR/MWh) were collected to indicate hourly imbalance prices. The following variables were not directly obtained from the ENTSO-E platform but were generated from existing variables. The variable forecast error (MWh) is the difference between the energy consumed and energy forecasted, and total imbalance price (EUR/MWh) is the absolute value of the summation of the negative and positive imbalance prices.

The open meteo dataset contained 8 variables across 11 locations in France. The averaged hour data contained 48792 rows. The variables present in this dataset were: temperature, pressure (msl), cloud cover, wind speed (10m), shortwave radiation, diffuse radiation, direct normal irradiance, & relative humidity (2m).

3.3 Data quality and preprocessing

Data quality assessment revealed minimal missing values ($< 1\%$) in both datasets, addressed via linear interpolation [5]. Outliers in PV production and energy price were capped using z-scores (threshold: ± 3) [7]. Preprocessing steps included:

- Temporal Alignment: Standardized timestamps to UTC.
- Lagged Features: Created ghi_lag24, ghi_lag48, pv_lag24 for temporal dependencies.
- Solar Angles: Computed zenith and azimuth for Paris (48.86, 2.35) using pvlib.
- Seasonal Dummies: Added monthly binary variables.
- Normalization: Scaled features to [0,1].

3.4 Exploratory data analysis

3.4.1 Normality test

A Shapiro Wilk test was conducted for all key variables in the ENTSO-E dataset. None of the investigated variables followed a normal distribution ($p < 0.001$ in all cases). As such, a non-parametric method, Spearman's rank correlation, was used for subsequent bivariate analysis.

3.4.2 Correlation analysis

- A strong positive correlation between energy forecasted and energy consumed: $p(23{,}416) = 0.94$, $p < 0.001$.
- Energy price vs. consumption showed a weak inverse relationship: $p = -0.16$, $p<0.001$.
- Very strong positive correlations between energy price and both + imbalance and – imbalance: $p = 0.84$ for each.
- Forecast error correlations with imbalance prices were minimal ($p = 0.01$), barely reaching significance for + imbalance ($p = 0.036$).
- The correlation between + imbalance and – imbalance was nearly perfect: $p = 1.00$, $p < 0.001$, indicating near-identical pricing movements in both imbalance directions.

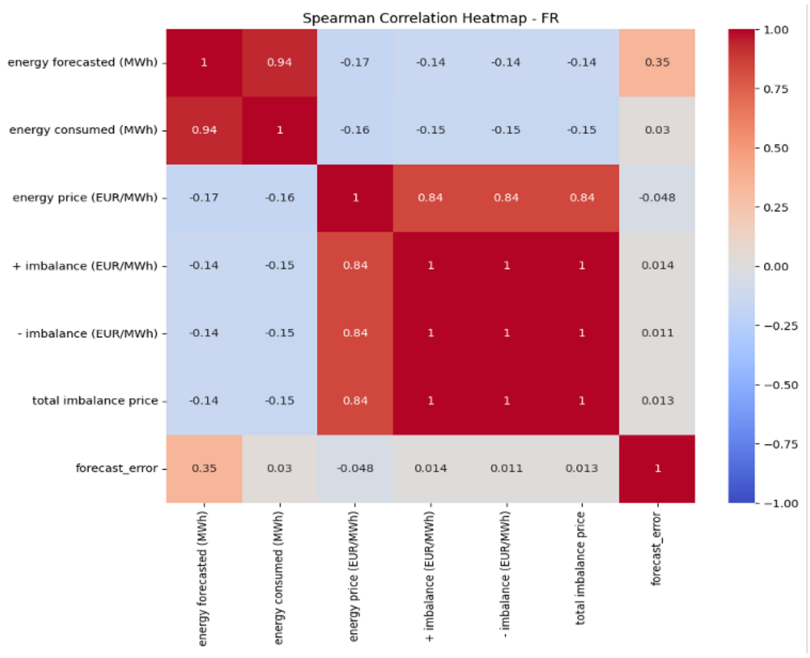

Figure 3.1: Correlation plot of France ENTSO-E variables.

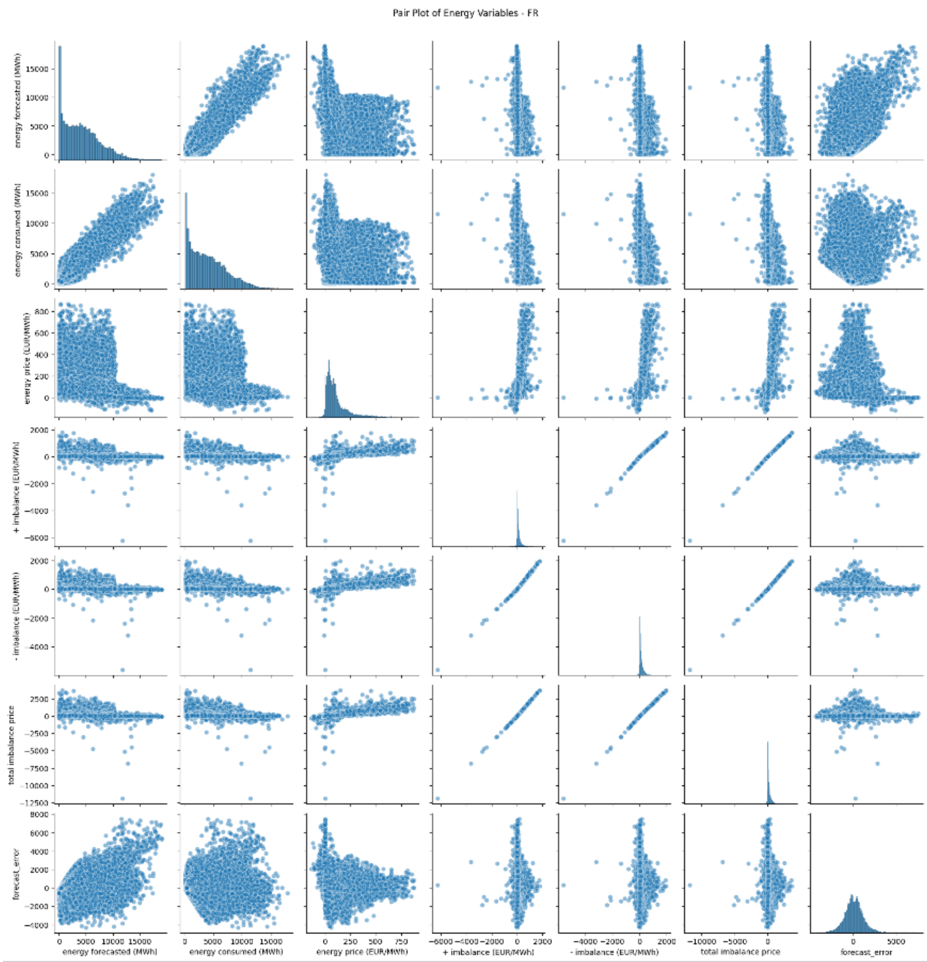

Figure 3.2: Pairwise plot of France ENTSO-E variables.

4 AI Modelling

4.1 Model Selection and Development

Model selection:

Six models were chosen for the study: Linear Regression, Random Forest, Multi-Layer Perceptron (MLP), XGBoost, Long Short-Term Memory (LSTM), and Convolutional Neural Network-LSTM (CNN-LSTM). Linear Regression and Random Forest were selected for their interpretability, MLP and XGBoost chosen for their ability to model complex non-linear relationships and LSTM and CNN-LSTM were included to capture temporal and spatial-temporal dependencies in time-series data.

Data processing:

The datasets were merged on a common datetime index (UTC), with missing values removed, and outliers eliminated using the interquartile range (IQR) method. Normalized production data (via StandardScaler) was split using an 80:20 train-test ratio, with the test set starting June 1, 2024 for all models.

Architecture:

Linear Regression used a single layer with L2 regularization and 100 decision trees were employed for the Random Forest model. MLP consisted of two hidden layers (100, 50 neurons, ReLU activation. XGBoost utilized a gradient boosting framework. LSTM and CNN-LSTM were implemented using the TensorFlow framework with randomly selected layers, refined through iterative tuning. LSTM included one LSTM layer, dropout, and a dense output layer for 24-hour sequences and CNN-LSTM combined a 1D convolutional layer (Conv1D), max pooling, an LSTM layer, dropout, and a dense output layer to extract spatial-temporal features.

Training:

For LSTM and CNN-LSTM, epoch loss training and validation curves were monitored, using a 10% validation split. ReLU activation functions and Adam optimizers were employed for neural models. Early stopping (LSTM patience=10, CNN-LSTM patience = 15) prevented overfitting in neural models, while Linear Regression, Random Forest, and XGBoost relied on scikit-learn's default convergence criteria.

4.2 Hyper Parameter Tuning

A 70-15-15 train-validation-test split was used, with 5-fold cross-validation for non-neural models and a 10% validation split for neural models. Grid search analysis was applied to all the models (except random forest) for the best model. Table 4.1 summarizes the hyperparameter search spaces.

Hyperparameter	**Model**	**Search Space**
Hidden Layer Sizes	MLP	{ 500, 100, 50 }
Learning Rate		{ 0.001}
Batch Size		{ auto}
Maximum Iterations		{ 1000 } (fixed)
Validation Fraction		{ 0.1 } (fixed, early stopping)
N estimators	XGBoost	{ 50, 75, 100, 200 }
Learning Rate		{ 0.001, 0.01, 0.1 }
Max Depth		{ 3, 6 }
Units	LSTM	{ 50, 100 }
Dropout Rate		{ 0.2, 0.3 }
Learning Rate		{ 0.001, 0.01 }
Batch Size		{ 32, 64 }
Epochs		{ 50 }
Validation Split		{ 0.1 } (early stopping, patience=10)
Units	CNN-LSTM	{ 50, 100 }
Dropout Rate		{ 0.2, 0.3 }
Learning Rate		{ 0.001, 0.01 }
Batch Size		{ 32, 64 }
Filters		{ 32, 64 }
Kernel Size		{ 3, 5 }
Epochs		{ 50 }
Validation Split		{ 0.1 } (early stopping, patience=10)

Table 4.1: Hyperparameter search spaces used by models

4.3 Model Evaluation

The performance of the models was evaluated using four metrics: Mean Absolute Error (MAE), Mean Bias Error (MBE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE). These metrics assess the accuracy and bias of day-ahead PV energy forecasts, comparing model predictions against the ENTSO-E baseline. MAE quantifies the average absolute difference between actual and predicted values. MBE measures the average bias, with positive values indicating overestimation and negative values indicating underestimation. MAPE provides the average percentage error, suitable for relative performance assessment. RMSE, derived from squared error differences, is expressed in the same units as the target (MWh) and is sensitive to outliers. The metrics are defined as follows:

$$\text{MBE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i) \tag{4.1}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{4.2}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \tag{4.3}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4.4}$$

$$\tag{4.5}$$

where:

- n is the number of observations,
- y_i is the actual value for the i-th observation,
- $\hat{y}_i$ is the predicted value for the i-th observation.

5 Results and Discussion

The performance of the six models was evaluated against the ENTSO-E baseline for day-ahead PV energy forecasting in France. The evaluation used the test set covered the period 01/06/2024 to 31/05/2025. Four metrics, MBE, MAE, MAPE, & RMSE were computed to assess model performance. These metrics provide a comprehensive view of model accuracy, bias, and error distribution. The results are summarized in Table 5.1, with MAE and MAPE distributions visualized in Figures 5.1 and 5.2

5.1 Results

Among all tested models, the LSTM network delivers the best overall balance with the lowest absolute and proportional errors (MAE = 340.83 MWh and MAPE 30.19%), outperforming the other 5 models. This indicates superior accuracy of the model in capturing PV production patterns, likely due to its ability to model temporal dependencies in the sequential data. The CNN-LSTM model follows closely, with an MAE of 363.32 MWh and MAPE of 30.88%, benefiting from its convolutional layers that extract spatial and temporal features from meteorological inputs. Random Forest and XGBoost also perform well, with MAEs of 351.62 MWh and 351.35 MWh, respectively, and MAPEs around 31%, outperforming Linear Regression (MAE = 420.74 MWh, MAPE = 33.90%) and MLP (MAE = 365.20 MWh, MAPE = 33.94%). The Linear Regression model exhibits the highest MAE and RMSE (782.30 MWh), indicating it struggles to capture non-linear relationships in the data.

Model	**MBE**	**MAE (MWh)**	**MAPE (%)**	**RMSE (MWh)**
ENTSO-E	-48.56	371.30	34.06	698.04
Linear Regression	-230.84	420.74	33.90	782.30
Random Forest	-249.00	351.62	31.62	714.02
MLP	-212.55	365.20	33.94	734.23
XGBoost	-256.59	351.35	31.24	707.58
LSTM	-234.38	340.83	30.19	700.96
CNN-LSTM	-278.04	363.32	30.88	742.51

Table 5.1: Overall Performance Metrics for ENTSO-E Baseline and Forecasting Models (2024–2025)

.

5.2 Explainable AI

To better understand the contribution of individual input variables to the best forecasting model (LSTM), SHAP and LIME analysis were conducted. Figures 5.3 and 5.4

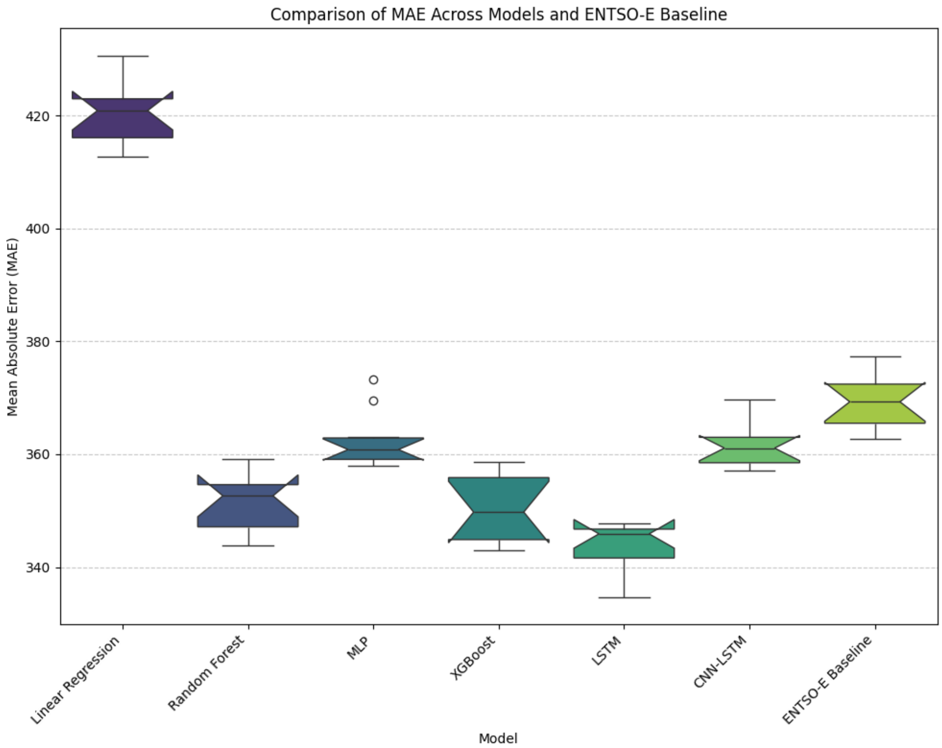

Figure 5.1: MAE bloxplot of models.

present the summary plots of SHAP and LIME values respectively, ranking the features by their average impact on the model output and providing both global and local explanations.

SHAP:

The SHAP analysis reveals that solar position variables are the most influential predictors. In particular, apparent zenith angle and zenith angle at recent time steps (t23 and t22) dominate the feature importance ranking. This is consistent with the physical relationship between solar angles and the availability of irradiance, thus confirming that the model heavily relies on astronomical factors to estimate energy output. High zenith values (when the sun is low in the sky) generally drive negative SHAP values, reducing predicted irradiance, while lower zenith values (sun closer to overhead) contribute positively. This alignment with physical intuition increases stakeholder trust, as the model's decision process reflects well-understood solar mechanics rather than opaque data correlations.

Following solar angles, radiation-related variables such as shortwave radiation, diffuse radiation, and direct normal irradiance also show significant contributions, while temporal features like hour of the day are moderately important. The prominence of these features demonstrates that the model captures both deterministic astronomical cycles and stochastic atmospheric effects. For system operators, this provides confidence that the forecasts are grounded in measurable, interpretable drivers of PV output, which facilitates adoption and operational decision-making. Regulators and policy makers can also benefit from this transparency, as it ensures that the forecasting process can be audited and verified against known physical principles.

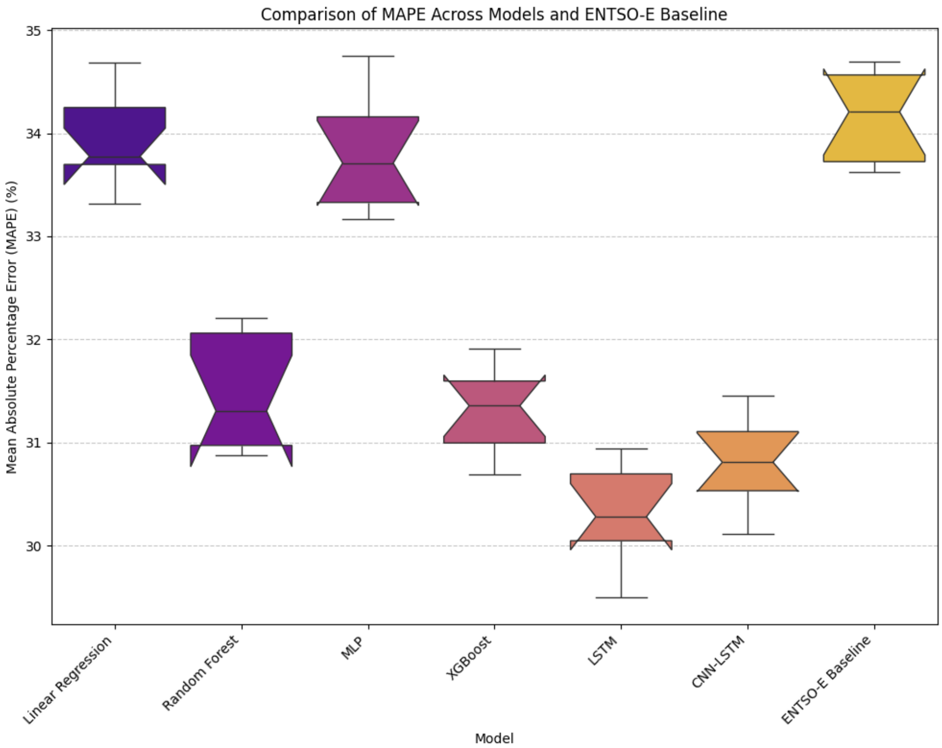

Figure 5.2: MAPE bloxplot of models.

LIME:

To complement the global interpretability provided by SHAP, LIME was applied to a single test prediction at 16:00 on 8 June 2024. The local explanation shows that zenith angle and apparent zenith angle again had the strongest positive contribution to the forecasted irradiance, reinforcing the insights from SHAP. Shortwave radiation at t22 and t23 also contributed positively, while direct normal irradiance (t22) acted as a negative adjustment in this instance, slightly lowering the predicted output. The consistency between SHAP and LIME further strengthens transparency, as both global and local explanations point to the same dominant features.

From a stakeholder perspective, LIME adds value by clarifying individual predictions. For plant operators, being able to trace how specific atmospheric or geometric conditions influenced a given forecast increases accountability and enables human oversight of automated decisions. For consumers and prosumers, this interpretability can help build confidence that fluctuations in energy supply are the result of real-world conditions rather than hidden biases in the model. For regulators, LIME provides a practical auditing tool for case-by-case verification of forecast reliability.

The XAI results demonstrate that the model's reasoning is accurate and interpretable in ways that support human-centered AI principles. Trust is reinforced by the alignment of dominant features with established physical relationships. Transparency is achieved through SHAP's global feature ranking and LIME's instance-level explanations, both of which make the "black box" model auditable. Accountability and human oversight are supported by enabling stakeholders understand and question forecasts when necessary.

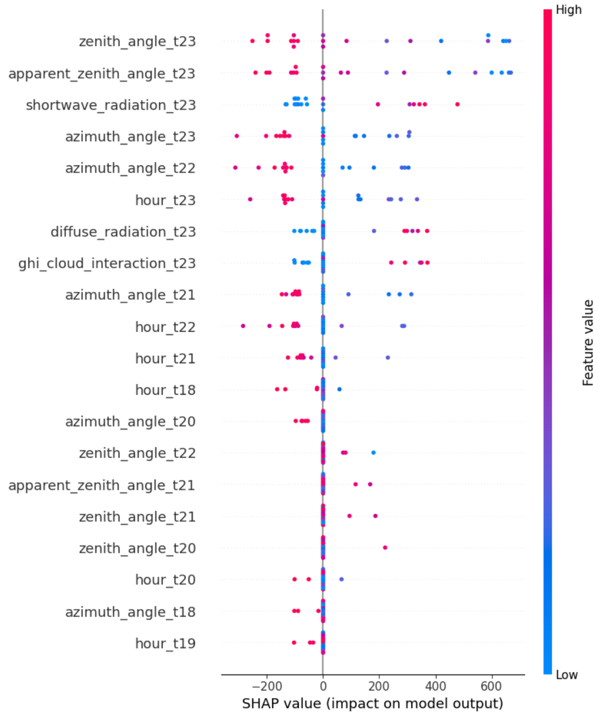

Figure 5.3: SHAP analysis.

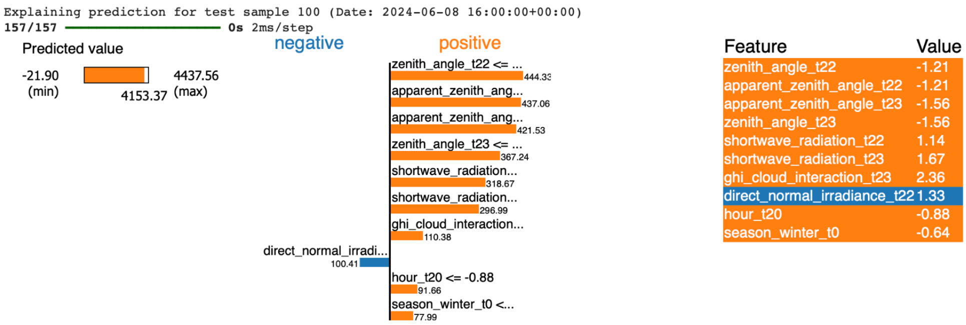

Figure 5.4: LIME analysis.

5.3 Economic impact

Accurate electricity forecasting is not only essential for operational efficiency but also critical for minimizing financial penalties imposed for deviations from scheduled energy delivery. In European energy markets, these penalties can reach substantial amounts when forecasts deviate from actual consumption or production. This section evaluates the financial implications of forecasting errors in day-ahead photovoltaic (PV) energy forecasting in France, focusing on the economic performance of the Long Short-Term Memory (LSTM) model compared to the ENTSO-E baseline.

The implementation of the LSTM model for forecasting demonstrates a significant economic advantage compared to the ENTSO-E forecast. Over the evaluation period (2024–2025), the ENTSO-E forecast incurred a net penalty fee of €217,694,531.05, reflecting the financial cost of its forecast inaccuracies. In contrast, the LSTM model reduced the net penalty to €81,071,268.94. This represents a difference of €136,623,262.11, corresponding to a 62.76% reduction in penalties compared to the conventional forecast. Fig. 5.5 shows a chart comparing the performances of the ENTSO-E and LSTM model.

The reduction in penalties underscores the direct economic impact of adopting the LSTM model.

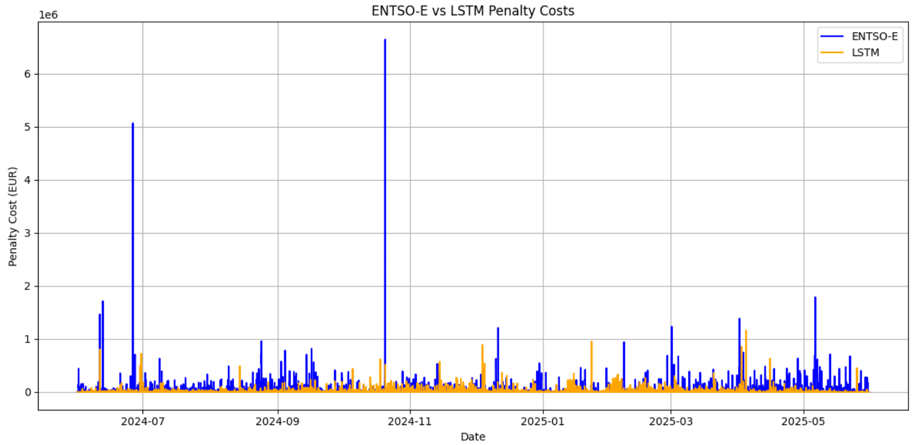

Figure 5.5: Chart showing ENTSO-E imbalance costs compared to LSTM model.

5.4 Discussion

This study investigated the application of advanced ML and DL models to improve day-ahead PV energy production forecasting, with the broader aim of enhancing economic efficiency, transparency, and sustainability in energy markets. The research was guided by four primary questions: how ML and DL models can be effectively designed for superior forecasting accuracy, which methodologies best quantify forecast errors, the extent to which these capabilities improve economic sustainability, and how Explainable AI (XAI) can promote transparency and support human decision-making.

Six models were implemented and compared, including linear regression, random forest (RF), multi-layer perceptron (MLP), XGBoost, LSTM, and CNN-LSTM. Tested across the 2024–2025 period, each model performance was compared to the ENTSO-E baseline using error metrics metrics. The LSTM model consistently outperformed the ENTSO-E baseline, and other ML models. With a lower forecasting error, the LSTM model minimizes financial penalties to be paid by suppliers by 62.76%.

These results align with findings from previous studies that have demonstrated the effectiveness of LSTM models in capturing temporal dependencies in PV forecasting [47][48][49]. However, unlike many earlier works which focus solely on technical accuracy [49], this research explicitly links forecasting improvements to financial outcomes, further strengthening the hypothesis proposed by Weron [5] that accurately forecasting energy production for energy markets, and simultaneously predicting the potential errors in these forecasts, can significantly enhance the sustainability of PV systems .

The use of XAI distinguishes this study from prior work. While most forecasting studies rely on black-box deep learning models [48][49][50][51], this thesis employed SHAP and LIME to provide explanatory insights. This approach advances the argument made by Arrieta et al. [8], who emphasize that transparency and trustworthiness are crucial for the adoption of AI in high-stakes domains such as energy markets. In doing so, this work addresses transparency and accountability requirements increasingly emphasized in the context of the EU AI Act (Annex XII) [56].

Despite these contributions, several limitations should be acknowledged. First, the study utilized historical data from France, potentially limiting generalizability to other regions with different climatic or market conditions. Prior studies [11][12], have shown that local irradiance variability can strongly influence model performance, suggesting the need for cross-country validation. Secondly, the implementation industry-specific data could capture site-specific variability and revenue. Incorporating such data could improve both the technical accuracy and the economic realism of the analysis. Finally, computational demands are higher for deep learning models compared to conventional statistical methods. While this echoes concerns raised by Wu et al. [57] regarding scalability and fairness in AI applications, it also highlights an important ethical dimension: resource-intensive models may not be equally accessible to all stakeholders

5.5 self-appraisal

I am highly satisfied with the project's completion, particularly the LSTM's economic impact and the human-centred approach (XAI implementation). The project required practical and technical approaches: this ranged from having to study EU energy markets, studying how weather conditions affect pv production, obtaining datasets from energy transparency platforms (ENTSO-E) and Open Meteo, performing rigorous data preprocessing, feature engineering, model training, designing, hypertuning, and evaluation. If given a similar project again, I would make sure to incorporate company-specific data and using a more expanded dataset with more variables. I would also explore hybrid modeling approaches that combine physical and statistical methods.

In conclusion, I am proud of the work completed. The results demonstrate both technical and practical economic relevance, which satisfies my research objectives. While there is always room for refinement, the project has equipped me with valuable skills in human-centred AI and industry level insight.

6 Conclusion and Future work

This study advanced day-ahead photovoltaic (PV) energy forecasting in France using the ENTSO-E and Open Meteo datasets, addressing four research questions on model design, error quantification, economic sustainability, and transparency. Six models were built: linear regression model, random forest model, MLP, XGBoost, LSTM & CNN-LSTM. The findings provide clear evidence that LSTM model outperform the ENTSO-E baseline and the other models. The economic analysis demonstrated a transformative impact, with the LSTM model reducing net penalty fees by €136,623,262.11 (62.76%) compared to the baseline's €217,694,531.05.

Another key contribution was the integration of XAI methods. This is a critical factor for the adoption of AI in the energy sector, where trust and interpretability are just as important as accuracy. By making the black box nature of LSTM model more transparent, stakeholders can better understand the drivers of forecasts and make informed operational decisions. This human-centered focus ensures that advanced AI tools are not only technically effective but also ethically usable.

Future work should incorporate industry-specific data and explore the adoption of probabilistic forecasting which allows uncertainty quantification and provide operators with confidence intervals rather than point forecasts. This could improve risk management in energy markets. Future works can also focus on forecasting PV generation for the intra-day and balancing market. Additionally, the application of federated learning techniques would allow the sharing of forecasting models across operators and countries without exposing sensitive data, aligning with EU data governance standards. Finally, extending the economic analysis to include lifecycle impacts of PV deployment, integration into ancillary services markets, and the role of storage optimization would provide a more holistic perspective on the economic sustainability of renewable energy systems.

Bibliography

[1] European Commission. 2050 long-term strategy. https://climate.ec.europa.eu/eu-action/climate-strategies-targets/2050-long-term-strategy_en, 2020. Accessed: 2025-08-24.

[2] United Nations. Sustainable development goals. https://sdgs.un.org/goals, 2015. Accessed: 2025-08-24.

[3] Tao Hong, Pierre Pinson, and Shu Fan. Global energy forecasting competition 2012. *International Journal of Forecasting*, 30(2):357–363, 2014.

[4] Cyril Voyant, Gilles Notton, Soteris Kalogirou, Marie-Laure Nivet, Christophe Paoli, Fabrice Motte, and Alexis Fouilloy. Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, 105:569–582, 2017.

[5] Rafał Weron. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030–1081, 2014.

[6] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de Pison, and F. Antonanzas-Torres. Review of photovoltaic power forecasting. *Solar Energy*, 136:78–111, 2016.

[7] Juan M. Morales, Antonio J. Conejo, Henrik Madsen, Pierre Pinson, and Marco Zugno. *Integrating Renewables in Electricity Markets*. Springer, 2014 edition, 2014.

[8] Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, Sergio Gil-Lopez, Daniel Molina, Richard Benjamins, Raja Chatila, and Francisco Herrera. Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information Fusion*, 58:82–115, 2020.

[9] European Network of Transmission System Operators for Electricity (ENTSO-E). Entso-e transparency platform. https://transparency.entsoe.eu, 2015. Accessed: 2025-08-24.

[10] Patrick Zippenfenig. Open-meteo.com weather api. https://open-meteo.com, 2023. Accessed: 2025-08-24.

[11] Sobrina Sobri, Sam Koohi-Kamali, and Nasrudin Abd Rahim. Solar photovoltaic generation forecasting methods: A review. *Energy Conversion and Management*, 156:459–497, 01 2018.

[12] Dazhi Yang, Jan Kleissl, Chris Gueymard, Hugo Pedro, and Carlos Coimbra. History and trends in solar irradiance and pv power forecasting: A preliminary assessment and review using text mining. *Solar Energy*, 168, 02 2018.

[13] U.S. Department of Energy. Solar photovoltaic technology basics. https://www.energy.gov/eere/solar/solar-photovoltaic-technology-basics, 2025. Accessed: 2025-08-24.

[14] Enel X. What is a photovoltaic system and how does it work? https://corporate.enelx.com/en/question-and-answers/how-does-a-photovoltaic-system-work, 2025. Accessed: 2025-08-24.

[15] Arka360. What are the basic components of the photovoltaic system? https://arka360.com/ros/basic-components-of-photovoltaic-system, 2025. Accessed: 2025-08-24.

[16] Energy Education. Energy education main page. https://energyeducation.ca/encyclopedia/Main_Page, 2025. Accessed: 2025-08-24.

[17] International Energy Agency (IEA). Solar photovoltaic (pv) — iea energy system renewables. https://www.iea.org/energy-system/renewables/solar-pv, 2025. Accessed: 2025-08-24.

[18] IEA Photovoltaic Power Systems Programme (IEA-PVPS). Snapshot of global pv markets 2025. https://iea-pvps.org/wp-content/uploads/2025/04/Snapshot-of-Global-PV-Markets_2025.pdf, April 2025. 13th edition providing preliminary 2024 market data.

[19] MarketsandMarkets. Building integrated photovoltaic market: Global forecast to 2029. https://www.marketsandmarkets.com/Market-Reports/bipv-market-509.html, January 2025. Report Code EP 1720. Updated January 23, 2025. Accessed: 2025-08-24.

[20] Atie Asadikia, Abbas Rajabifard, and Mohsen Kalantari. A systems perspective on national prioritisation of sustainable development goals: Insights from australia. *Geography and Sustainability*, 4(3):255–267, 2023.

[21] European Commission. Electricity market design. https://energy.ec.europa.eu/topics/markets-and-consumers/electricity-market-design_en, 2025. Accessed: 2025-08-24.

[22] Jacint Balaguer. Cross-border integration in the european electricity market. evidence from the pricing behavior of norwegian and swiss exporters. *Energy Policy*, 39(9):4703–4712, 2011.

[23] Giulia Ragosa. The political economy of electricity market coupling: Comparing experiences from europe and the united states. *The Electricity Journal*, 37(7):107430, 2024.

[24] European Union. Wholesale electricity and gas markets – eu oversight rules. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=legissum:180201_1, n.d. Accessed: 2025-08-24.

[25] European Parliament. Internal energy market. https://www.europarl.europa.eu/factsheets/en/sheet/45/internal-energy-market, 2025. Accessed: 2025-08-24.

[26] Yashar Ghiassi-Farrokhfal, Rodrigo Belo, Mohammad Reza Hesamzadeh, and Derek Bunn. Optimal electricity imbalance pricing for the emerging penetration of renewable and low-cost generators. *Manufacturing & Service Operations Management*, 25(5):1966–1983, 2023.

[27] Jens Baetens, Joannes Laveyne, Greet Van Eetvelde, and Lieven Vandevelde. Imbalance pricing methodology in belgium: Implications for industrial consumers. In *2020 17th International Conference on the European Energy Market (EEM)*, pages 1–6, 2020.

[28] R. Van der Veen et al. Balance responsibility and imbalance settlement in northern europe. Technical report, SINTEF, 2009. Accessed: 2025-08-24.

[29] V. Laitsos, G. Vontzos, P. Paraschoudis, E. Tsampasis, D. Bargiotas, and L. H. Tsoukalas. The state of the art electricity load and price forecasting for the modern wholesale electricity market. *Energies*, 17(22):5797, 2024.

[30] Louise Dalton and Stella Mavrommati. Eu reform of electricity market design – part one: The commission's proposed reform of the remit framework. https://www.grip.globalrelay.com/eu-reform-of-electricity-market-design-the-commissions-proposed-reform-of-the-remit-framework-part-one/, May 2023. Accessed: 2025-08-24.

[31] John Duffie and William Beckman. *Solar Engineering of Thermal Processes*. 01 2006.

[32] Jay A. Kratochvil, William E. Boyson, and David L. King. Photovoltaic array performance model. 2004.

[33] M K Fuentes. A simplified thermal model for flat-plate photovoltaic arrays. Technical report, Sandia National Labs., Albuquerque, NM (USA), 05 1987.

[34] Tina Jakaša, Ivan Andročec, and Petar Sprčić. Electricity price forecasting — arima model approach. In *2011 8th International Conference on the European Energy Market (EEM)*, pages 222–225, 2011.

[35] Ângela Paula Ferreira, Jenice Gonçalves Ramos, and Paula Odete Fernandes. A linear regression pattern for electricity price forecasting in the iberian electricity market. *Revista Facultad de Ingeniería, Universidad de Antioquia*, (93):117–127, 2019.

[36] Jwaone Gaboitaolelwe, Adamu Murtala Zungeru, Abid Yahya, Caspar K. Lebekwe, Dasari Naga Vinod, and Ayodeji Olalekan Salau. Machine learning based solar photovoltaic power forecasting: A review and comparison. *IEEE Access*, 11:40820–40845, 2023.

[37] Alexandra I. Khalyasmaa, Stanislav A. Eroshenko, Valeriy A. Tashchilin, Hariprakash Ramachandran, Teja Piepur Chakravarthi, and Denis N. Butusov. Industry experience of developing day-ahead photovoltaic plant forecasting system based on machine learning. *Remote Sensing*, 12(20):3420, 2020.

[38] K. Bhatia, R. Mittal, J. Varanasi, and M. M. Tripathi. An ensemble approach for electricity price forecasting in markets with renewable energy resources. *Utilities Policy*, 70:101185, 2021.

[39] Ali Shiri, Mohammad Afshar, Ashkan Rahimi-Kian, and Behrouz Maham. Electricity price forecasting using support vector machines by considering oil and natural gas price impacts. In *2015 IEEE International Conference on Smart Energy Grid Engineering (SEGE)*, pages 1–5, 2015.

[40] Nishant Saxena, Rahul Kumar, Yarrapragada KSS Rao, Dilbag Singh Mondloe, Nishikant Kishor Dhapekar, Abhishek Sharma, and Anil Singh Yadav. Hybrid knn-svm machine learning approach for solar power forecasting. *Environmental Challenges*, 14:100838, 2024.

[41] D. Singhal and K. S. Swarup. Electricity price forecasting using artificial neural networks. *International Journal of Electrical Power & Energy Systems*, 33(3):550–555, 2011.

[42] Martina Radicioni, Valentina Lucaferri, Francesco De Lia, Antonino Laudani, Roberto Lo Presti, Gabriele Maria Lozito, Francesco Riganti Fulginei, Riccardo Schioppo, and Mario Tucci. Power forecasting of a photovoltaic plant located in enea casaccia research center. *Energies*, 14(3):707, 2021.

[43] Francesco Nicoletti and Piero Bevilacqua. Hourly photovoltaic production prediction using numerical weather data and neural networks for solar energy decision support. *Energies*, 17(2):466, 2024.

[44] Luca Massidda, Fabio Bettio, and Marino Marrocu. Probabilistic day-ahead prediction of pv generation: A comparative analysis of forecasting methodologies and of the factors influencing accuracy. *Solar Energy*, 271:112422, 2024.

[45] Sravankumar Jogunuri, Josh F. T., Albert Alexander Stonier, Geno Peter, Jayakumar Jayaraj, and Vivekananda Ganji. Random forest machine learning algorithm based seasonal multi-step ahead short-term solar photovoltaic power output forecasting. *IET Renewable Power Generation*, 2024.

[46] Wassila Tercha, Sid Ahmed Tadjer, Fathia Chekired, and Laurent Canale. Machine learning-based forecasting of temperature and solar irradiance for photovoltaic systems. *Energies*, 17(5):1124, 2024.

[47] Fouzi Harrou, Farid Kadri, and Ying Sun. Forecasting of photovoltaic solar power production using lstm approach. In *Advanced Statistical Modeling, Forecasting, and Fault Detection in Renewable Energy Systems*, chapter 3, pages 51–72. Elsevier, 2020.

[48] Maria Konstantinou, Stefani Peratikou, and Alexandros G. Charalambides. Solar photovoltaic forecasting of power output using lstm networks. *Atmosphere*, 12(1):124, 2021.

[49] Arif Ozbek, Alper Yildirim, and Mehmet Bilgili. Deep learning approach for one-hour ahead forecasting of energy production in a solar-pv plant. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 44(4):10465–10480, 2022.

[50] Ali Agga, Ahmed Abbou, Moussa Labbadi, and Yassine El Houm. Short-term self consumption pv plant power production forecasts based on hybrid cnn-lstm, convlstm models. *Renewable Energy*, 177:101–112, 2021.

[51] Wencheng Liu and Zhizhong Mao. Short-term photovoltaic power forecasting with feature extraction and attention mechanisms. *Renewable Energy*, 226:120437, 2024.

[52] Diaa Salman, Cem Direkoglu, Mehmet Kusaf, and Murat Fahrioglu. Hybrid deep learning models for time series forecasting of solar power. *Neural Computing and Applications*, pages 1–18, 2024.

[53] R. Machlev, L. Heistrene, M. Perl, K. Y. Levy, J. Belikov, S. Mannor, and Y. Levron. Explainable artificial intelligence (xai) techniques for energy and power systems: Review, challenges and opportunities. *Energy and AI*, 9:100169, 2022.

[54] Ali Öter and Betül Ersöz. Artificial intelligence assisted solar energy forecasting by explainability approaches with lime and shap. *El-Cezeri Journal of Science and Engineering*, 12(2):205–212, 2025.

[55] Chukwuebuka Joseph Ejiyi, Dongsheng Cai, Dara Thomas, Sandra Obiora, Emmanuel Osei-Mensah, Caroline Acen, Francis O. Eze, Francis Sam, Qingxian Zhang, and Olusola O. Bamisile. Comprehensive review of artificial intelligence applications in renewable energy systems: Current implementations and emerging trends. *Journal of Big Data*, 12:169, 2025.

[56] European Union. Annex xii: Transparency information referred to in article 53(1), point (b)–technical documentation for providers of general-purpose ai models to downstream providers. https://artificialintelligenceact.eu/annex/12/, 2024. Accessed: 2025-09-03.

[57] Carole-Jean Wu, Ramya Raghavendra, Udit Gupta, Bilge Acun, Newsha Ardalani, Kiwan Maeng, Gloria Chang, Fiona Aga Behram, James Huang, Charles Bai, Michael Gschwind, Anurag Gupta, Myle Ott, Anastasia Melnikov, Salvatore Candido, David Brooks, Geeta Chauhan, Benjamin Lee, Hsien-Hsin S. Lee, Bugra Akyildiz, Maximilian Balandat, Joe Spisak, Ravi Jain, Mike Rabbat, and Kim M. Hazelwood. Sustainable AI: environmental implications, challenges and opportunities. *CoRR*, abs/2111.00364, 2021.