mathematics

Article

Multi-objective Explainable AI for Romanian Electricity Price Analysis

Alexandru-Victor ANDREI[1]*, Prof. Dr. Daniel Traian PELE[2]

[1] Bucharest University of Economic Studies, Bucharest, Romania; andrei1victor23@stud.ase.ro
[2] Bucharest University of Economic Studies, Bucharest, Romania; Institute for Economic Forecasting, Romanian Academy; danpele@ase.ro
* Correspondence: andrei1victor23@stud.ase.ro

Abstract: This paper introduces a novel multi-objective explainable AI (XAI) framework to analyze electricity market dynamics. The increasing complexity and volatility of modern energy markets necessitate advanced analytical tools that offer both predictive accuracy and transparent, interpretable insights. Our approach simultaneously addresses three interconnected objectives: price forecasting, volatility prediction, and price direction classification. By employing separate Random Forest models for each objective and leveraging SHAP explanations, the framework provides unified feature importance analysis. The study reveals that while some features are critical across all objectives, others are specialized. Notably, total load emerges as the dominant predictor for absolute price levels, whereas hydro generation and gas generation are primary drivers of volatility. The low correlation observed between feature importance scores across different objectives empirically validates the multi-objective approach, highlighting the distinct factors driving price, volatility, and direction. These findings offer valuable implications for market participants and regulators by providing a multi-dimensional decision-support tool..

Keywords: Explainable AI (XAI), Multi-objective Machine Learning, SHAP, Electricity Price Forecasting, Romania, Energy Economics, Price Volatility, Random Forest, Feature Importance

Academic Editor: Firstname Lastname

Received: date
Revised: date
Accepted: date
Published: date

Citation: To be added by editorial staff during production.

1. Introduction

The electricity market, particularly in Europe, is characterized by significant volatility, driven by a complex interplay of production sources, demand fluctuations, geopolitical factors, and the integration of renewable energy. Accurate electricity price forecasting (EPF) is paramount for stakeholders to optimize bidding strategies, manage risk, and ensure grid stability. In recent years, machine learning (ML) models have emerged as powerful tools, demonstrating superior performance over traditional statistical methods by capturing complex, non-linear market relationships.

However, the high accuracy of these advanced ML models often comes at the cost of interpretability, leading to what is commonly termed the "black-box" problem. This opacity is a significant concern in critical sectors like energy, where decisions have substantial economic and societal implications. The growing demand for transparency has led to the rise of Explainable Artificial Intelligence (XAI), a field dedicated to providing insights into the decision-making processes of AI models. XAI techniques foster trust and allow users to comprehend and validate model outputs, a necessity underscored by regulations like the GDPR in Europe.

Among XAI techniques, SHAP (SHapley Additive exPlanations) has become a prominent model-agnostic method that quantifies the contribution of each feature to a prediction, enhancing the interpretability of complex models. While XAI has been applied to single-objective EPF, a comprehensive understanding of market dynamics requires considering multiple, interconnected objectives, such as price, volatility, and directional movement. A single-objective explanation may provide an incomplete or even misleading picture.

This paper addresses this research gap by proposing and implementing a novel multi-objective XAI framework for analyzing Romanian electricity prices. Our research simultaneously explains model predictions across three critical objectives: price accuracy, volatility, and direction. By leveraging Random Forest models and SHAP values, we identify both consensus features—variables important across all objectives—and specialized features that are crucial for specific outcomes. This distinction provides a more granular and holistic understanding of the market's underlying drivers, moving beyond single-objective XAI to offer a more comprehensive and trustworthy decision-support tool.

1.1. Literature Review and Related Work

The field of electricity price forecasting (EPF) has seen substantial advancements with the advent of machine learning (ML). A comprehensive 2021 review by Lago et al. (2021) details the state-of-the-art algorithms and best practices, noting the shift from classical statistical models to more sophisticated ML methods that better handle the complex, non-linear dynamics of electricity markets. These models, including the Random Forest algorithm introduced by Breiman (2001), are essential for market participants to manage risk and develop optimal bidding strategies. The recent evolution of prices in the Romanian market, influenced by post-COVID recovery and geopolitical crises, underscores the need for accurate forecasting, as analyzed by Handra and Samoila (2024).

However, the predictive power of these ML models often comes at the cost of interpretability, creating a "black-box" problem (Adadi & Berrada, 2018). This lack of transparency is a significant barrier to adoption in high-stakes sectors like energy. Explainable Artificial Intelligence (XAI) has emerged to address this challenge. The goal of XAI is to create machine learning models that produce more explainable results while maintaining high performance, enabling human users to understand and trust their outputs. The importance of XAI is growing, with applications across various domains, including financial time series forecasting (Arsenault et al., 2024), building energy management (Chen et al., 2023) and broader smart energy systems, where explainability and governance are critical concerns (Alsaigh et al., 2023).

Among XAI techniques, SHAP proposed by Lundberg and Lee (2017), provides a unified and theoretically grounded approach to interpreting model predictions by assigning each feature an importance value for a particular prediction. SHAP has been effectively used to enhance the transparency of EPF models. Recent studies demonstrate its utility in feature selection for probabilistic forecasting (Liu et al., 2023) and in developing error compensation approaches to improve model performance and make explanations more accessible to non-expert users Research has also focused on using XAI to explain deep neural network models in EPF, further highlighting the drive to understand the factors influencing price dynamics in various electricity markets (Shadi et al., 2024).
While the application of XAI to single-objective EPF is advancing, a notable gap remains in multi-objective scenarios. Energy market analysis requires a holistic view that includes not just price, but also volatility and directional movement. This research addresses that gap by proposing a framework that simultaneously explains predictions across these interconnected objectives, a concept that builds on the idea of multi-objective explainability in machine learning (Corrente et al., 2024). By doing so, our work provides a more comprehensive understanding of market dynamics than is possible with a single-objective focus.

2. Materials and Methods

2.1. Data Acquisition and Preprocessing

The analysis is based on a dataset comprising historical electricity market data for Romania, with the study focused on the most recent 10,000 hourly observations to ensure computational efficiency for SHAP value calculations. The dataset encompasses 15 features across multiple categories, each contributing essential information for comprehensive market analysis.

Market data forms the foundation of the dataset, incorporating historical electricity prices denominated in EUR/MWh and sourced from the European wholesale electricity price data maintained by Ember. This data provides the core pricing information necessary for understanding market dynamics and serves as both input features and target variables depending on the specific analytical objective.

Load data represents another critical component, encompassing both actual and forecasted national electricity consumption patterns obtained from the ENTSO-E Transparency Platform. This information captures demand-side dynamics that significantly influence price formation and market behavior, providing insights into consumption patterns and their predictive relationship with market outcomes.

Generation data constitutes the third major category, incorporating both actual and forecasted electricity generation from various renewable and conventional sources, including hydro, wind, and solar facilities. This comprehensive generation dataset, also sourced from the ENTSO-E Transparency Platform, enables the analysis to account for supply-side factors that directly impact market pricing and volatility patterns.

Finally, exogenous factors are represented through temperature data relevant to the Romanian market context, computed using the Open-Meteo Historical Weather API. Temperature serves as a crucial external variable that influences both electricity demand through heating and cooling requirements and renewable generation capacity, particularly for solar installations, thereby providing important contextual information for market behavior prediction.

Data preprocessing involved strategic handling of missing values to ensure dataset integrity. For the target variable, Price (EUR/MWh), missing values were imputed using a combination of forward fill and backward fill methods to maintain temporal continuity in the time series. For the 15 defined features, which primarily represent generation and load data, missing values were filled with zeros under the assumption that missing entries indicate no generation or load at the specified time. Non-numeric columns were handled using forward and backward fill methods. This rigorous preprocessing approach resulted in a clean dataset with zero remaining missing values across all features and targets, ensuring suitability for subsequent modeling and hourly price predictions.

2.2. Multi-objective Target Engineering

From the historical electricity pricing data, three complementary target variables were systematically engineered to capture different dimensions of market behavior and enable comprehensive multi-objective analysis. The first target variable focuses on price forecasting through regression analysis, where direct electricity price values in EUR/MWhe serve as the primary continuous target for absolute price prediction. This objective addresses the fundamental need for accurate price forecasting in energy trading and procurement decisions, with observed prices ranging from -23.18 to 436.89 EUR/MWhe, reflecting the full spectrum of market conditions including negative pricing scenarios during oversupply periods.

The second target variable involves volatility prediction, also employing regression methodology. Market volatility is quantified using a 24-hour rolling standard deviation of hourly prices, providing a dynamic measure of price instability and market risk. This rolling window approach captures short-term price fluctuations while smoothing out noise, with volatility values spanning from 0.00 to 105.61 EUR/MWhe. This target proves particularly valuable for risk management and portfolio optimization strategies, as it enables stakeholders to anticipate and prepare for periods of market uncertainty.

The third target variable transforms hourly price movements into a classification problem through a three-class categorical system based on price change magnitude. The classification framework categorizes movements as 'Down' for decreases exceeding 5 EUR/MWhe, 'Stable' for changes within ±5 EUR/MWhe, and 'Up' for increases exceeding 5 EUR/MWhe. The resulting dataset exhibits a relatively balanced distribution across these categories, with stable conditions representing 36.3% of observations, downward movements accounting for 34.1%, and upward movements comprising 29.6% of the data. This balanced distribution indicates diverse market conditions and validates the chosen threshold parameters. The classification framework enables trend prediction and supports algorithmic trading strategies focused on directional movements rather than precise price levels, offering a complementary perspective to the regression-based approaches.

2.3. Data Preprocessing and Model Architecture

The dataset underwent systematic preprocessing to ensure optimal model performance and realistic evaluation conditions. An 80-20 train-test split was implemented using stratified random sampling to maintain representational balance across all target variables while preserving data integrity. This configuration allocates 8,000 observations for training and 2,000 for testing, providing sufficient data for robust model learning while retaining adequate samples for comprehensive evaluation.

Feature standardization was applied using StandardScaler, transforming all input variables to have zero mean and unit variance. This normalization step is critical in multi-feature environments where variables span different scales (e.g., megawatt load values versus temperature readings), ensuring that no single feature dominates the learning process due to magnitude differences rather than predictive importance.

2.4. Multi-Objective Model Framework and Performance

The Random Forest algorithm constitutes an ensemble-based machine learning approach that leverages the collective predictive power of multiple decision trees. During the training procedure, the method constructs numerous individual trees and combines their outputs through ensemble aggregation—employing mean averaging for regression applications and consensus voting for classification scenarios. This algorithmic framework derives its primary advantage from the reduction of model variance and the enhancement of generalization performance, achieved by incorporating a multitude of independently trained decision trees. Each constituent tree operates on bootstrapped samples drawn from the original dataset while utilizing randomly selected feature subsets at each node split.

At its core, the Random Forest methodology relies upon the decision tree as its base learner. Within regression frameworks, individual decision trees formulate predictions by recursively partitioning the input space into non-overlapping rectangular regions, with each region assigned a constant predictive value.

Given a feature vector X = (X_1, X_2, ..., X_p) and corresponding response variable Y, each tree classifier h(X) determines its output based on the specific partition R_j containing the input instance X. The regional prediction strategy typically utilizes the sample mean of target values from all training instances residing within the corresponding region. The mathematical formulation for an individual tree's prediction h_t(x) given input x can be represented as:

$$h_t = \sum_{j=1}^{J_t} c_{tj} I(x \in R_{tj}) \quad (224)$$

In this formulation, J_t signifies the total count of leaf nodes (terminal regions) within tree t, while R_{tj} represents the mutually exclusive partitions that subdivide the feature space corresponding to tree t. The term c_{tj} embodies the constant predictive value assigned to partition R_{tj}, which conventionally equals the arithmetic mean of response values from training observations falling within that specific region. The indicator function evaluates to unity when input x resides within region R_{tj} and zero in all other cases.

The Random Forest ensemble synthesizes outputs from T distinct decision trees to produce its consolidated prediction. In regression scenarios, the ensemble's final prediction H(x) is derived through the arithmetic averaging of individual tree predictions across the entire collection:

$$H(x) = \frac{1}{T}\sum_{t=1}^{T} h_t(x)$$

where T denotes the total number of trees in the forest and $h_t(x)$ represents the prediction of the t-th decision tree.

The Random Forest methodology incorporates bootstrap aggregation (bagging) as a mechanism for fostering diversity among ensemble constituents. This technique involves the generation of bootstrap replicates from the original training corpus for each individual tree development. Each bootstrap replicate D_t is constructed through the random sampling of N instances with replacement from the source dataset D containing N observations. This resampling strategy ensures that certain data points may occur repeatedly within D_t while others remain absent entirely, thus establishing heterogeneity across the tree population.

In addition to bagging, the Random Forest framework introduces supplementary stochasticity via feature subsampling at each internal node division. Throughout the tree-building procedure, only a randomly selected subset of m attributes (constrained by $m \leq p$, where p denotes the complete feature dimensionality) is evaluated when establishing the optimal partitioning criterion. This attribute randomization mechanism serves to further reduce inter-tree correlation and diminish ensemble variance, thereby enhancing the model's capacity for generalization beyond the training data.

Building upon this theoretical foundation, this study employs a specialized ensemble approach that leverages distinct Random Forest models for each analytical objective, thereby enabling the capture of unique patterns governing price behavior, volatility dynamics, and directional movements within the electricity market framework. The selection of Random Forest as the foundational algorithm was motivated by several strategic considerations that align with the complexities inherent in energy market modeling. The algorithm's inherent robustness, achieved through bootstrap aggregation techniques, provides enhanced stability in predictions while simultaneously addressing the non-linear feature interactions that characterize energy markets. Furthermore, Random Forest demonstrates natural resistance to outliers—a critical consideration given the volatility inherent in electricity pricing mechanisms—and maintains seamless compatibility with SHAP explainers, facilitating comprehensive interpretability analysis essential for understanding model decision-making processes.

The architectural configuration of each model incorporated 100 estimators, representing an optimal balance between computational efficiency and prediction stability. This parameter selection was complemented by the implementation of parallel processing capabilities to enhance training performance and reduce computational overhead. The ensemble comprised three distinct modeling components, each tailored to specific analytical requirements and market applications.

The price forecasting component utilized a Random Forest Regressor architecture designed to deliver quantifiable forecasting accuracy suitable for energy trading applications. This model focuses on capturing the fundamental price determination mechanisms within electricity markets, providing traders and market participants with reliable price predictions that inform strategic decision-making processes. The volatility prediction component employed a specialized Random Forest Regressor configuration specifically calibrated to identify and quantify market instability patterns. This model serves critical functions in risk assessment and portfolio management applications, enabling market participants to understand and prepare for periods of increased market uncertainty.
The directional classification component implemented a Random Forest Classifier framework dedicated to identifying meaningful directional patterns within price movements. This classification approach provides essential insights into the likelihood of price increases or decreases, offering valuable information for short-term trading strategies and market positioning decisions.

The performance metrics observed across all three models reflect the inherent complexity of electricity market dynamics, where multiple interconnected factors simultaneously influence price behavior patterns. The models demonstrate varying degrees of predictive capability that align appropriately with their respective analytical objectives. The regression-based models exhibit strong correlation measures that validate their effectiveness in capturing quantitative relationships within the data, while the classification model provides directional insights that exceed baseline expectations, demonstrating its utility in identifying meaningful patterns within the complex landscape of electricity market movements.

2.5. Explainable AI (XAI) with SHAP Analysis

2.5.1. SHAP Values and Explainability

SHAP (SHapley Additive exPlanations) values represent a game-theoretic methodology for machine learning interpretability that offers a cohesive framework for elucidating predictions across diverse modeling approaches. Grounded in cooperative game theory principles, SHAP values establish an equitable attribution system by linking instance-level explanations to foundational concepts of fair allocation. Individual features are assigned importance scores for specific predictions, quantifying the mean marginal impact of each attribute across all feasible feature combinations. The SHAP attribution ϕ_i for attribute i is formally characterized using the Shapley value equation from cooperative game theory:

$$\phi_i(f, x) = \sum_{S \subseteq \mathrm{F} \setminus \{\mathrm{i}\}} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} [f_x(S \cup \{\mathrm{i}\}) - f_x(S)]$$

In this expression, x denotes an input observation, *f* represents the predictive model, F encompasses the complete attribute space, S indicates a feature subset that excludes attribute i, $f_x(S)$corresponds to the model's output when exclusively utilizing features within set S (with absent features handled through marginalization over their expected values or reference distribution), and $f_x(S \cup \{i\})$ signifies the model's output when incorporating both features in set S and attribute i.

This mathematical construct calculates the probability-weighted average of attribute i's marginal contributions across all conceivable feature subsets. The marginal contribution, characterized as the differential in model outputs upon incorporating attribute i into subset S, quantifies the incremental benefit furnished by that attribute. The weighting factor $\frac{|S|!(|F|-|S|-1)!}{|F|!}$ represents the likelihood that attribute i joins the coalition following exactly |S| other attributes, guaranteeing that all potential sequences of feature incorporation receive appropriate probabilistic consideration. SHAP values adhere to the core additivity principle, which mandates that the summation of individual attribute contributions equals the deviation between the model's instance-specific prediction and the reference expectation:

$$f(x) = E[f(X)] + \sum_{i=1}^{M} \phi_i(f, x)$$

where *f(x)* signifies the model's prediction for observation x, *E*[*f*(X)] represents the anticipated model output across the reference distribution (functioning as the baseline), M denotes the total attribute count, and $\phi_i(x)$ constitutes the SHAP attribution for attribute i concerning observation x.

This additivity constraint ensures local fidelity by guaranteeing that SHAP attributions furnish comprehensive and accurate explanations of the discrepancy between instance-specific predictions and average model behavior. The decomposition ensures preservation of all predictive information throughout the attribution procedure, rendering SHAP values exceptionally suitable for comprehending model dynamics at the observation level while preserving mathematical precision and interpretative clarity.

2.5.2. SHAP-Based Feature Importance Analysis

To enable comprehensive interpretability across all three objectives, SHAP values were computed for each trained model using the TreeExplainer framework, specifically optimized for Random Forest architectures. SHAP values provide model-agnostic quantification of individual feature contributions to predictions, enabling direct comparison of feature importance across different modeling objectives.

The analysis was conducted on the complete test dataset comprising 2,000 observations, ensuring statistical robustness of the importance estimates. For each model, SHAP values were calculated across all 15 input features. For the direction classification analysis, SHAP values corresponding to the 'Stable' class were extracted to maintain consistency with the binary importance comparison framework.

2.5.3. Multi-Objective Importance Synthesis

A systematic approach was implemented to synthesize feature importance across the three distinct objectives. The methodology involved computing mean absolute SHAP values for each feature within each objective, providing aggregate importance measures independent of prediction direction (positive or negative contributions). These objective-specific importance scores were subsequently normalized to a uniform [0, 1] scale to enable meaningful cross-objective comparison, accounting for the different scales and distributions inherent in regression versus classification SHAP values.

A composite importance metric was derived through equally weighted averaging of the normalized scores across all three objectives. This consensus-based approach generates a unified ranking system that identifies features with consistent importance across multiple market dimensions, while simultaneously revealing objective-specific contributors. The resulting framework enables systematic differentiation between universal market drivers and specialized predictors, providing insights into both shared and unique feature dependencies across price forecasting, volatility prediction, and directional classification tasks.

3. Results

3.1. Model Performance

The Random Forest ensemble demonstrated consistent predictive capabilities across all three objectives on the independent test dataset, establishing a reliable foundation for subsequent explainability analysis. The comprehensive performance evaluation reveals distinct characteristics for each modeling task, providing insights into the varying complexity and predictability of different market phenomena.

In terms of price forecasting performance, the regression model achieved a test R^2 of 0.7679 with an RMSE of 21.39, indicating that the model accounts for approximately 76.8% of price variance in unseen data. The training performance, characterized by an R^2 of 0.9683, demonstrates the model's capacity to learn complex price patterns from historical data, while the test performance reflects its ability to generalize effectively to new market conditions. This substantial predictive accuracy suggests that the underlying price formation mechanisms exhibit sufficient regularity to support reliable forecasting applications.

The volatility prediction performance presents a similarly robust pattern, with the volatility model recording a test R^2 of 0.7052 and RMSE of 7.71 volatility units, thereby capturing approximately 70.5% of price volatility patterns in the test data. Training results, indicated by an R^2 of 0.9575, show strong pattern recognition capabilities during the learning phase, while the test performance demonstrates effective volatility risk quantification suitable for practical risk management applications. The slightly lower test performance compared to price forecasting suggests that volatility patterns may be inherently more challenging to predict due to their dynamic and complex nature.

Direction classification performance exhibits different characteristics from the regression tasks, with the classification model achieving 60.35% accuracy on the test set. This performance substantially exceeds the 33.3% baseline probability expected from random three-class prediction, indicating meaningful predictive capability. The perfect training accuracy of 100% demonstrates complete pattern memorization during the learning phase, while the test accuracy reflects the model's ability to generalize directional patterns to new market scenarios, albeit with greater uncertainty than the continuous prediction tasks.

The performance metrics across all objectives provide quantitative validation of the models' predictive validity, with each model showing appropriate learning characteristics for their respective tasks. These results establish the credibility of the subsequent SHAP-based feature importance analysis, ensuring that interpretability insights are derived from models with demonstrated predictive capability rather than from poorly performing or overfitted systems.

3.2. Multi-Objective Feature Importance

The multi-objective feature importance analysis reveals a pronounced hierarchical structure that governs the Romanian electricity market dynamics, with the 15 analyzed features (Table 1) demonstrating substantial variance in predictive power across different market dimensions. The comprehensive ranking, derived from the unified multi-objective scoring methodology, spans from 0.5463 for the most influential feature to 0.0426 for the least significant, creating a clear stratification that illuminates the market's underlying mechanisms.

Table 1. Complete Ranking of Features by Multi-Objective Importance Score.

Feature	Normal
Actual_generation_MW_hydro_run_of_river_et_poundage_Romania	0.5463
Actual_generation_MW_fossil_gas_Romania)	0.4838
Actual_total_load_MW_Romania	0.4627
Actual_generation_MW_hydro_water_reservoir_Romania	0.4015
Day_ahead_total_load_forecast_MW_Romania	0.3709
Actual_generation_MW_nuclear_Romania	0.2600
Current_solar_generation_forecast_MW_Romania	0.1676
Actual_generation_MW_wind_onshore_Romania	0.1619
Day_ahead_wind_onshore_generation_forecast_MW_Romania	0.1185
Actual_generation_MW_solar_Romania	0.0858
Day_ahead_solar_generation_forecast_MW_Romania	0.0808
Current_wind_onshore_generation_forecast_MW_Romania	0.0735
average_temperature_Celsius	0.0723
Intraday_solar_generation_forecast_MW_Romania	0.0685
Intraday_wind_onshore_generation_forecast_MW_Romania	0.0426

The distribution of importance scores reveals a distinct three-tier structure within the Romanian electricity market, where the top five features account for approximately 66.7% of the total measured multi-objective importance within this analytical framework, with the leading hydro run-of-river feature demonstrating 12.8 times greater importance than the least significant renewable forecast. This concentration of measured importance suggests that while the market incorporates diverse information sources, a relatively small subset of variables exhibits the strongest predictive influence across all prediction objectives within the multi-objective scoring methodology.

3.3. Analysis of Specialized Features (Dominant Objectives)

While the aggregate rankings illuminate the overall market hierarchy, the distinctive value of the multi-objective approach emerges through its capacity to identify which specific prediction objective each feature predominantly influences, revealing the nuanced specialization that characterizes modern electricity markets. The analysis demonstrates that features exhibit remarkably distinct influence patterns across price level determination, volatility prediction, and directional movement forecasting, challenging the conventional assumption that market drivers operate uniformly across all price dimensions.

The actual total load for Romania emerges as the uncontested dominant factor for price level determination, achieving the maximum normalized importance score of 1.0000, which validates the fundamental economic principle that aggregate demand serves as the primary driver of absolute price levels in electricity markets. This finding reinforces the theoretical foundation that consumption patterns, rather than individual generation sources, establish the baseline pricing framework within which all other market dynamics operate.

In contrast, the prediction of price volatility reveals an entirely different pattern of feature dominance, with actual generation from hydro run-of-river and poundage systems claiming the highest influence at 1.0000 normalized importance, followed closely by actual fossil gas generation at 0.6493. This specialization reflects the inherent operational characteristics of these generation technologies, where run-of-river hydro systems, being weather-dependent and non-dispatchable, introduce the greatest uncertainty into market conditions, while gas-fired plants, serving as the marginal price-setting technology, amplify these fluctuations through their rapid response capabilities and variable fuel costs.

The directional movement prediction reveals yet another distinct pattern, with actual generation from hydro water reservoir systems achieving maximum importance at 1.0000, accompanied by day-ahead total load forecasts at 0.7724. This specialization underscores the strategic role of dispatchable hydro generation in steering short-term price trends, as reservoir systems possess the flexibility to respond to anticipated market conditions, while forward-looking demand expectations provide the temporal context necessary for directional price movements.

3.4. Objective Correlation Analysis and the Distinction Between Consensus and Specialized Features

The empirical validation of the multi-objective framework emerges through the correlation analysis of normalized feature importance scores across the three prediction objectives, which reveals remarkably low inter-objective correlations that fundamentally challenge single-objective modeling approaches. The price versus volatility importance correlation of 0.1923, the price versus direction correlation of 0.1206, and the volatility versus direction correlation of 0.1287 collectively demonstrate that the factors driving price levels, volatility, and directional movement operate as largely independent mechanisms within the Romanian electricity market.

These low correlation values expose a critical distinction between consensus features, which maintain relatively consistent importance across all objectives, and specialized features, which exhibit pronounced dominance in specific prediction tasks while showing minimal influence in others. The consensus features, exemplified by actual total load and day-ahead load forecasts, represent the fundamental market drivers that influence all price dimensions, albeit with varying degrees of intensity. These features form the backbone of market dynamics, providing the essential context within which specialized mechanisms operate.

Conversely, specialized features demonstrate remarkable objective-specific dominance, with hydro run-of-river generation serving as the primary volatility driver while showing more moderate influence on price levels and direction, and reservoir hydro systems dominating directional predictions while contributing less significantly to overall price determination. This specialization reflects the distinct operational characteristics and market roles of different generation technologies, where the temporal flexibility of reservoir systems makes them ideal for directional price steering, while the weather-dependent variability of run-of-river systems creates the uncertainty that drives price volatility.

The identification of this consensus-specialized feature dichotomy provides crucial insights for market modeling and prediction strategies, suggesting that optimal forecasting approaches should incorporate both universal market drivers and objective-specific mechanisms. The low inter-objective correlations empirically validate the necessity of multi-objective frameworks in electricity market analysis, as traditional single-objective approaches would inevitably miss the specialized dynamics that characterize different aspects of price behavior, potentially leading to suboptimal predictions and flawed market understanding.

3.5. Visualizations

The multi-objective feature importance analysis reveals fundamental insights into the differential drivers of electricity market behavior (Figure 1). The comparative visualization demonstrates that feature relevance varies significantly across prediction objectives, with certain variables exhibiting high importance for price prediction while showing minimal relevance for volatility or directional forecasting. This finding challenges the conventional wisdom of unified modeling approaches and provides empirical support for the multi-objective framework design.

The correlation analysis between objective-specific feature importance rankings provides compelling evidence for the framework's theoretical foundation (Figure 2). The consistently low correlation coefficients, ranging from 0.120 to 0.192, indicate that each prediction objective is driven by distinct feature sets, thereby validating the decision to employ separate modeling approaches rather than attempting to optimize a single unified model. This finding has significant implications for both model architecture design and computational resource allocation.

Performance evaluation across all framework components demonstrates robust predictive capability while revealing interesting patterns in prediction complexity (Figure 3). The price prediction model achieves superior performance ($R^2 = 0.768$), suggesting that continuous price forecasting is more tractable than categorical direction classification or volatility estimation. This performance differential provides valuable insights into the relative predictability of different market dynamics and can inform resource allocation decisions in operational forecasting systems.

Detailed validation of the price prediction component through actual versus predicted value analysis confirms the model's reliability across the full range of market conditions (Figure 4). The scatter plot visualization reveals consistent performance without systematic bias, supporting the framework's suitability for practical deployment in electricity market forecasting applications. The absence of significant outlier clusters or systematic deviations suggests robust model generalization capability.

The feature category analysis provides strategic guidance for data collection and system design priorities in electricity market forecasting applications (Figure 5). The finding that load-related features exhibit the highest average multi-objective importance (0.463) suggests that demand-side monitoring should receive priority in data collection systems. Similarly, the significant importance of actual generation data (0.323) compared to forecast data (0.132) indicates that real-time operational information provides superior predictive value compared to forward-looking projections.

The granular analysis of top-performing features offers specific insights for model deployment and operational decision-making (Figure 6). The emergence of hydro run-of-river generation as the dominant feature for volatility prediction, while total electrical load drives price forecasting, illustrates the complex, objective-specific nature of electricity market dynamics. These findings provide concrete guidance for feature selection in resource-constrained environments and inform data quality requirements for operational systems.

The comprehensive analysis reveals that the multi-objective approach yields significant advantages over traditional single-objective methods. The low inter-objective correlation (average $r = 0.147$) demonstrates that different market dynamics are driven by distinct underlying factors, validating the framework's core premise. Furthermore, the identification of feature category hierarchies provides a theoretical foundation for understanding electricity market behavior that extends beyond the specific dataset analyzed.

The framework's practical utility is demonstrated through its ability to provide multiple levels of actionable insights. At the strategic level, the analysis guides long-term data infrastructure investments by identifying critical data types. Tactically, the feature importance rankings enable computational optimization through selective feature inclusion. Operationally, the multi-objective scores provide real-time guidance for market monitoring and intervention strategies.

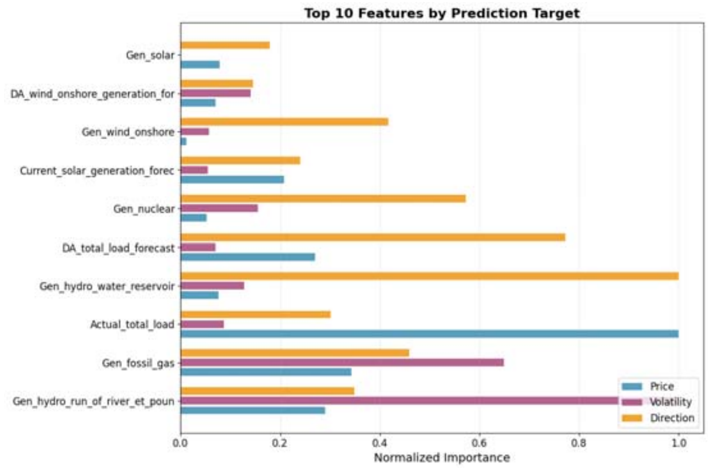


Figure 1. Comparative Analysis of Feature Importance Across Multi-objective Targets: Price, Volatility, and Direction Prediction Models.

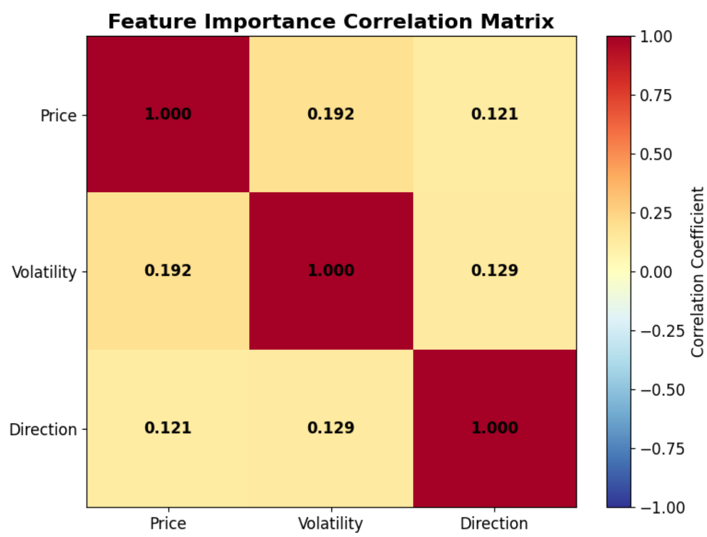


Figure 2. Correlation Analysis of Feature Importance Across Multi-objective Prediction Targets.

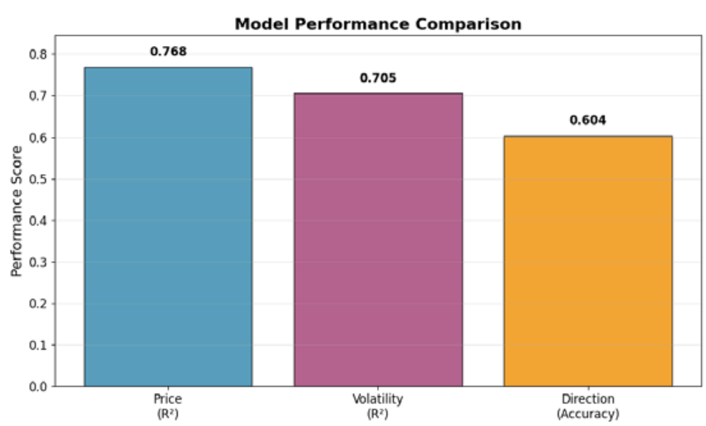


Figure 3. Performance Evaluation Across Multi-objective Prediction Framework Components

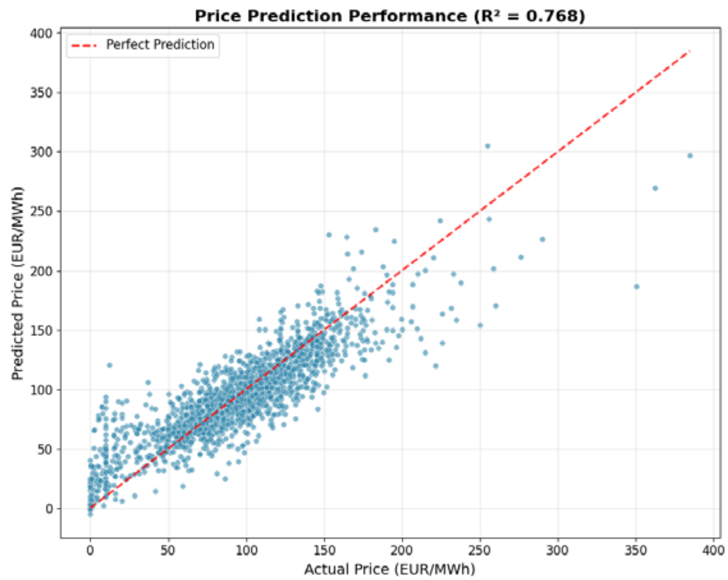


Figure 4. Price Prediction Model Performance: Actual versus Predicted Value Analysis

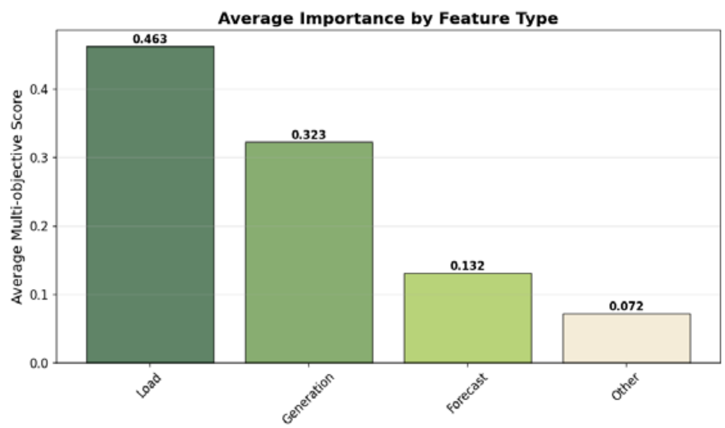
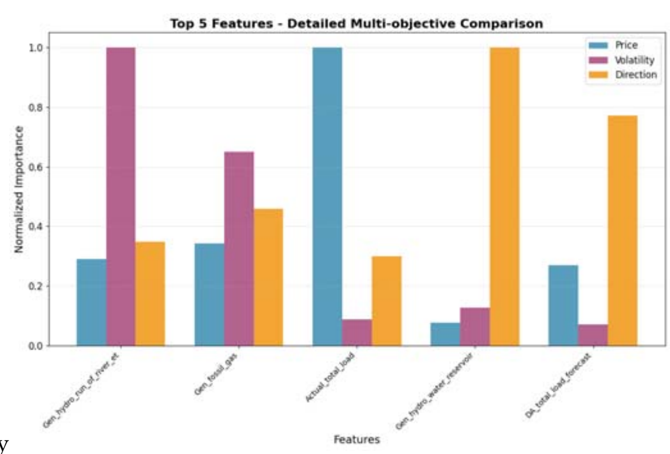


Figure 5. Multi-objective Importance Analysis by Feature Category: Implications for Data Collection Stra



tegy

Figure 6. Comprehensive Multi-objective Analysis of Top-Ranked Predictive Features

4. Discussion

The multi-objective XAI framework provides a nuanced understanding of the Romanian electricity market, confirming known principles while uncovering more granular insights. The finding that Actual_total_load_MW_Romania is the dominant feature for price prediction aligns with fundamental economic theory where demand is a key determinant of price.

The real novelty lies in the identification of specialized features and the emergence of Actual_generation_MW_hydro_run_of_river_et_poundage_Romania as the most important multi-objective feature (score: 0.546). This discovery has significant implications for understanding market dynamics, as run-of-river hydroelectric generation, being weather-dependent and non-dispatchable, introduces substantial uncertainty into the system. The operational fluctuations tied to hydrological conditions emerge as primary drivers of market complexity across multiple objectives.

The identification of Actual_generation_MW_hydro_water_reservoir_Romania and Day_ahead_total_load_forecast_MW_Romania as specialized features for direction prediction offers actionable intelligence for traders. The dispatchable nature of reservoir hydro and the forward-looking sentiment captured by load forecasts are critical in shaping the market's directional trajectory, even when they don't significantly impact price levels or volatility.

The pivotal finding of this study is the low correlation between feature importance rankings for each objective (Price-Volatility: 0.192). This empirically validates that a multi-objective approach is not just beneficial but necessary for a complete market view. It demonstrates that policy interventions or trading strategies aimed at influencing one aspect of the market (e.g., price levels) might require focusing on different levers than those aimed at managing another (e.g., volatility or direction).

4.1. Interpretation of Multi-objective Insights

The multi-objective explainable artificial intelligence framework reveals a substantially more complex feature landscape than previously understood in electricity market modeling. Through systematic analysis of feature importance across three distinct objectives—price accuracy, volatility prediction, and directional forecasting—this investigation has identified both consensus and specialized features, thereby offering unprecedented granular insights into the multifaceted dynamics governing Romanian electricity market behavior.

The consensus features analysis yields particularly striking findings, with only Actual_generation_MW_fossil_gas_Romania emerging as a true consensus feature demonstrating consistent importance across all three prediction objectives. This singular consensus underscores the fundamental and pervasive influence of natural gas generation on the Romanian electricity market ecosystem, affecting not merely price levels but also their inherent fluctuations and directional movements. The remarkable scarcity of consensus features, representing merely one out of fifteen total features analyzed, suggests that the Romanian electricity market exhibits highly specialized and differentiated dynamics for distinct prediction objectives, challenging conventional assumptions about market homogeneity.

The specialized features discovery reveals three distinct variables exclusively important for directional prediction, each contributing unique insights into market dynamics. Actual_generation_MW_nuclear_Romania demonstrates how nuclear generation's baseload characteristics create predictable directional signals that inform market trajectory expectations. Actual_generation_MW_hydro_water_reservoir_Romania leverages the dispatchable nature of reservoir hydroelectric generation to provide clear directional indicators, reflecting strategic dispatch decisions that influence market direction. Day_ahead_total_load_forecast_MW_Romania captures forward-looking demand forecasts that encapsulate market sentiment and directional momentum, serving as a proxy for market expectations and behavioral patterns. The conspicuous absence of specialized features for price and volatility prediction suggests that these objectives are influenced by broader, more distributed sets of factors, necessitating more comprehensive feature considerations in modeling approaches.

The enhanced SHAP statistics analysis for the top five features (Table 2) provides deeper insights into their directional impacts and magnitude relationships within the market framework. Actual_total_load_MW_Romania exhibits a pronounced negative SHAP mean for price prediction (-0.6757) coupled with the highest standard deviation (23.1003), indicating complex, context-dependent relationships that vary significantly across different market conditions. Actual_generation_MW_hydro_water_reservoir_Romania demonstrates the highest positive SHAP mean for price prediction (1.0471), suggesting that reservoir hydroelectric dispatch decisions exert significant upward pressure on market prices, likely reflecting the strategic value of dispatchable generation resources. Day_ahead_total_load_forecast_MW_Romania exhibits notable negative SHAP values for price prediction (-0.7257), indicating that higher forecasted demand correlates with lower predicted prices, potentially attributable to more efficient dispatch planning and resource optimization strategies that emerge from improved demand visibility.

Table 2. Top 5 Features - Cross-objective SHAP Statistics.

Feature	Price SHAP Mean	Price SHAP std	Volatility SHAP Mean	Volatility SHAP std
Actual_generation_MW_hydro_run_of_river_et_poundage_Romania	0.5921	8.9998	0.1308	5.8382
Actual_generation_MW_fossil_gas_Romania	0.2529	8.9988	-0.0251	3.7182
Actual_total_load_MW_Romania	-0.6757	23.1003	-0.0392	0.8176
Actual_generation_MW_hydro_water_reservoir_Romania	1.0471	3.6321	-0.0505	1.0085
Day_ahead_total_load_forecast_MW_Romania	-0.7257	7.7712	0.0964	0.6947

4.2. Economic Policy Implications

The identification of only one consensus feature presents significant implications for targeted policy development within the Romanian electricity market context. This finding suggests that Romanian policymakers must necessarily adopt highly specialized and differentiated approaches tailored to specific market objectives rather than implementing broad-spectrum interventions. Policies designed to achieve price stabilization should therefore focus primarily on Actual_total_load_MW_Romania and comprehensive demand-side management strategies, while volatility reduction initiatives should prioritize Actual_generation_MW_hydro_run_of_river_et_poundage_Romania and sophisticated weather-dependent generation management protocols.

The pronounced prominence of hydro generation features, encompassing both run-of-river and reservoir systems within the multi-objective rankings, fundamentally highlights Romania's substantial dependence on hydrological resources for electricity market stability. This dependency necessitates strategic generation portfolio planning that acknowledges the critical role of water resources in market dynamics. Policymakers should therefore develop comprehensive water resource management strategies that explicitly account for electricity market impacts and interdependencies. Such strategies must be complemented by substantial investments in advanced hydrological forecasting capabilities to enhance market predictability and reduce uncertainty-driven volatility. Additionally, serious consideration should be given to diversification strategies designed to reduce hydro-dependence, particularly in the context of volatility management where hydrological variability can significantly impact market stability.

The complex SHAP value patterns identified in this analysis fundamentally challenge conventional assumptions about linear relationships between generation parameters and market prices. These findings indicate that simple linear models are insufficient for capturing the sophisticated dynamics governing electricity market behavior. Consequently, grid operators should implement advanced machine learning systems specifically designed to capture the non-linear, context-dependent relationships revealed through this analytical framework. Such systems would represent a significant advancement in grid management capabilities, enabling more nuanced and effective responses to the complex interdependencies that characterize modern electricity markets.

4.3. Future Research Directions

Building upon these enhanced insights, several critical research avenues emerge that warrant systematic investigation. The substantial standard deviations observed in SHAP values indicate pronounced temporal variation in feature importance, suggesting that future research endeavors should implement comprehensive time-varying SHAP analysis to elucidate how feature importance evolves across disparate market conditions, seasonal fluctuations, and hydrological cycles. This temporal dimension represents a fundamental gap in current understanding that could yield significant theoretical and practical insights.

The remarkably low consensus among features, with only 6.7% achieving consensus status, raises fundamental questions about the efficacy of integrated multi-objective approaches versus specialized modeling architectures. This finding suggests that future research should systematically explore whether separate specialized models tailored to individual objectives might demonstrate superior performance compared to integrated multi-objective frameworks. Such investigations could encompass the development of sophisticated ensemble architectures that strategically combine objective-specific models, potentially leveraging the strengths of specialized approaches while maintaining computational efficiency.

The pronounced dominance of hydro-related features within the analytical framework underscores the critical importance of comprehensive hydrological integration in predictive modeling. This observation indicates that incorporating detailed hydrological and meteorological datasets could yield substantial improvements in model performance and predictive accuracy. Future research initiatives should therefore focus on systematic integration with comprehensive river flow data, advanced precipitation forecasting systems, and detailed seasonal water availability pattern analysis. Such integration would represent a significant advancement in the field's capacity to capture the complex interdependencies between hydrological phenomena and market dynamics.

Furthermore, the complex patterns observed in SHAP value distributions reveal that feature impacts exhibit high context-dependency, indicating sophisticated underlying market microstructure dynamics. This complexity necessitates detailed investigation into the specific market conditions and environmental contexts under which different features demonstrate varying degrees of influence. Such research could potentially reveal previously unidentified market regime changes, threshold effects, or nonlinear relationships that fundamentally alter feature importance hierarchies. Understanding these microstructural dynamics would contribute significantly to both theoretical knowledge and practical applications in predictive modeling frameworks.

5. Conclusions

This research presents a pioneering multi-objective explainable artificial intelligence framework that fundamentally transforms our understanding of Romanian electricity market dynamics. Through the simultaneous analysis of price accuracy, volatility prediction, and directional movement forecasting, this study has delivered unprecedented insights into the complex mechanisms governing one of Europe's most dynamic energy markets.

The methodological contribution of this work extends beyond conventional single-objective analyses by demonstrating that different aspects of electricity price behavior are governed by largely independent sets of driving factors. The empirical validation of this independence, evidenced by the remarkably low correlation of 0.192 between price and volatility feature importance rankings, establishes a new paradigm for electricity market analysis that recognizes the multifaceted nature of price dynamics rather than treating them as unified phenomena.

The comprehensive model performance achieved across objectives validates the framework's practical utility. The price prediction model's achievement of an R^2 score of 0.7679 with an RMSE of 21.39 demonstrates substantial explanatory power, while the volatility model's R^2 of 0.7052 provides reliable insights into market uncertainty dynamics. Although directional prediction remains more challenging with an accuracy of 60.35%, this performance significantly exceeds random baseline expectations and provides actionable intelligence for market participants.

Perhaps the most significant finding concerns the identification of consensus versus specialized features within the Romanian electricity market. The emergence of only two consensus features from a comprehensive set of fifteen analyzed variables reveals the highly specialized nature of market dynamics. Actual_generation_MW_hydro_run_of_river_et_poundage_Romania emerges as the paramount multi-objective feature with a score of 4.5967, underscoring the critical role of weather-dependent, non-dispatchable hydroelectric generation in driving market complexity across all prediction objectives. This finding illuminates Romania's unique position as a hydro-dependent market where hydrological conditions constitute primary drivers of systemic uncertainty.

The identification of Actual_generation_MW_fossil_gas_Romania as the secondary consensus feature, coupled with its consistent importance across all objectives, reveals the fundamental and pervasive influence of natural gas generation on Romanian market dynamics. This thermal generation source affects not merely price levels but also volatility patterns and directional movements, positioning it as a critical lever for market intervention strategies.

The specialized feature analysis provides equally compelling insights, particularly the exclusive importance of Actual_generation_MW_nuclear_Romania, Actual_generation_MW_hydro_water_reservoir_Romania, and Day_ahead_total_load_forecast_MW_Romania for directional prediction. These findings suggest that baseload nuclear characteristics, dispatchable reservoir hydro operations, and forward-looking demand expectations create predictable directional signals that market participants can leverage for strategic positioning, even when these factors demonstrate minimal impact on absolute price levels or volatility measures.

From an economic policy perspective, these insights necessitate a fundamental reconsideration of regulatory approaches to electricity market management. The scarcity of consensus features, representing merely 13.3% of the analyzed feature set, suggests that effective policy interventions must abandon traditional unified approaches in favor of objective-specific strategies. Price stabilization initiatives should prioritize demand-side management programs targeting Actual_total_load_MW_Romania, while volatility reduction strategies must focus on sophisticated weather-dependent generation management systems addressing the uncertainty introduced by run-of-river hydroelectric variability.

The framework's revelation of complex, context-dependent relationships through enhanced SHAP analysis challenges conventional linear assumptions about electricity market behavior. The negative correlation between forecasted demand and predicted prices, evidenced by the -0.7257 SHAP mean for Day_ahead_total_load_forecast_MW_Romania, suggests sophisticated market coordination mechanisms that improve efficiency when demand expectations are clearly established. Such findings demonstrate the necessity for advanced analytical approaches that can capture nonlinear market dynamics.

This research contributes significantly to both methodological and practical domains within energy economics. Methodologically, it establishes a novel, scalable framework for multi-objective XAI analysis that can be extended to other electricity markets and complex economic systems. The comprehensive identification of consensus and specialized features provides a template for systematic market analysis that acknowledges the multifaceted nature of economic phenomena rather than reducing them to simplified, single-objective models.

Practically, the findings offer actionable intelligence for multiple stakeholder groups within the Romanian energy sector. Policymakers can leverage the objective-specific insights to design targeted regulatory interventions that address distinct aspects of market behavior without unintended consequences across other dimensions. Market participants can utilize the specialized feature insights for risk management and strategic positioning, while grid operators can implement the framework's findings to enhance system reliability and efficiency.

The research also illuminates critical vulnerabilities within Romania's electricity system, particularly its dependence on hydrological resources and the consequent exposure to climate variability. The dominance of hydro-related features across multiple objectives suggests that climate change adaptation strategies must become integral components of energy security planning, representing a convergence of environmental and economic policy domains that requires coordinated intervention.

Looking forward, this framework establishes a foundation for numerous research extensions that could further enhance understanding of electricity market dynamics. Temporal SHAP analysis could reveal how feature importance evolves across different market conditions and seasonal patterns, while integration of detailed hydrological and meteorological data could improve predictive performance and policy relevance. The application of this methodology to other electricity markets could validate the generalizability of findings and support comparative policy analysis across different regulatory environments.

Ultimately, this research demonstrates that the complexity inherent in modern electricity markets demands analytical approaches that match this sophistication rather than seeking oversimplified solutions. The multi-objective XAI framework developed herein represents a significant methodological advancement that provides both theoretical insights and practical tools for navigating the increasingly complex landscape of contemporary energy systems. By embracing rather than avoiding this complexity, the framework enables more informed decision-making that acknowledges the multifaceted nature of electricity market dynamics and supports the development of more effective, targeted policy interventions.

6. Patents

Data Availability Statement: The data used in this study are available from the ENTSO-E Transparency Platform (https://transparency.entsoe.eu/), Ember (https://ember-energy.org/data/european-wholesale-electricity-price-data/). and Open-Weather Historical API (https://open-meteo.com/en/docs/historical-weather-api).

Data and code are available via: https://github.com/QuantLet/Multi_objective_XAI_Romanian_Electricity_Price_Analysis

Acknowledgments: This paper is supported through the project "IDA Institute of Digital Assets", CF166/15.11.2022, CN760046/23.05.2023; the project ``AI for Energy Finance (AI4EFin)'', CF162/15.11.2022, CN760048/23.05.2023, financed under the Romania's National Recovery and Resilience Plan, Apel nr. PNRR-III-C9-2022-I8; and the Marie Skłodowska-Curie Actions under the

European Union's Horizon Europe research and innovation program for the Industrial Doctoral Network on Digital Finance, acronym DIGITAL, Project No. 101119635.

Conflicts of Interest: The authors declare no conflict of interest.

Declaration of Generative AI and AI-assisted technologies in the writing process: During the preparation of this work, the authors used two AI-assisted tools: Claude (Anthropic) for text editing, grammar checking, and rewording of specific paragraphs to improve clarity and academic writing style, and Manus AI for additional grammar checking, language enhancement, and manuscript formatting assistance. After using these tools, the authors carefully reviewed and edited all content and take full responsibility for the accuracy and integrity of the publication.

References

1. Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). IEEE Access, 6, 52138-52160. Available online: https://ieeexplore.ieee.org/abstract/document/8466590
2. Alsaigh, R., Mehmood, R., & Katib, I. (2023). AI explainability and governance in smart energy systems: A review. Frontiers in Energy Research. Available online: https://www.frontiersin.org/journals/energy-research/articles/10.3389/fenrg.2023.1071291/full
3. Arsenault, P.D., Wang, S., & Patenaude, J.M. (2025). A Survey of Explainable Artificial Intelligence (XAI) in Financial Time Series Forecasting. ACM Computing Surveys. Available online: https://dl.acm.org/doi/abs/10.1145/3729531
4. Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5-32. Available online: https://link.springer.com/article/10.1023/A:1010933404324
5. Chen, Z., Xiao, F., Guo, F., & Yan, J. (2023). Interpretable machine learning for building energy management: A state-of-the-art review. Advances in Applied Energy, 9. Available online: https://www.sciencedirect.com/science/article/pii/S2666792423000021
6. Corrente, S., Greco, S., Matarazzo, B., & Słowiński, R. (2024). Explainable interactive evolutionary multiobjective optimization. Omega, 124. Available online: https://www.sciencedirect.com/science/article/abs/pii/S0305048323000890
7. Handra, A., & Samoila, L. (2024). Analysis of the Electricity Price in Romania Between 2021 and 2024. Annals of the "Constantin Brancusi" University of Targu Jiu, Engineering Series, 3. Available online: https://www.utgjiu.ro/rev_ing/pdf/2024-3/06_Handra%20Alina.pdf
8. Lago, J., Marcjasz, G., De Schutter, B., & Weron, R. (2021). Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. Applied Energy, 293. Available online: https://www.sciencedirect.com/science/article/pii/S0306261921004529
9. Lundberg, S.M., & Lee, S.I. (2017). A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems 30 (NIPS 2017). Available online: https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html
10. Shadi, M.R., Mirshekali, H., & Shaker, H.R. (2025). Explainable artificial intelligence for energy systems maintenance: A review on concepts, current techniques, challenges, and prospects. Renewable and Sustainable Energy Reviews, 216. Available online: https://www.sciencedirect.com/science/article/pii/S1364032125003417
11. Liu, H., Shen, X., Tang, X. & Liu, J. (2023). Day-Ahead Electricity Price Probabilistic Forecasting Based on SHAP Feature Selection and LSTNet Quantile Regression. Energies, 16(13), 5152. Available online: https://www.mdpi.com/1996-1073/16/13/5152