Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Ensemble learning with anomaly detection for accurate day-ahead electricity price forecasting

Faheem Jan
Bacha Khan University
Musaad S. Aldhabani
University of Tabuk
Izatmand Haleemzai
izatmandhaleemzi@gmail.com

Polytechnical University of Kabul
Ahmed M. Zidan
King Khalid University
Mehwish Tahir
Shaheed Benazir Bhutto Women University

Article

Keywords: Electricity Price Forecasting, Anomaly Detection, Ensemble Learning, Machine Learning, STL Decomposition.vector autoregressive, autoregressive moving average, autoregressive neural network, k-nearest Neighbors

Posted Date: April 21st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9130354/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Ensemble learning with anomaly detection for accurate day-ahead electricity price forecasting

FAHEEM JAN*[1], MUSAAD S. ALDHABANI[2], IZATMAND HALEEMZAI†[3], AHMED M. ZIDAN[4], and MEHWISH TAHIR[5]

[1]Department of Mathematics and Statistics, Bacha Khan University, Charsadda, Pakistan; fahimjan@bkuc.edu.pk

[2]Department of Mathematics, Faculty of Science, University of Tabuk, P.O.Box741, Tabuk 71491, Saudi Arabia; maldhaban@ut.edu.sa

[3]Department of Higher Mathematics, Kabul Polytechnic University, Kabul Afghanistan; izatmandhaleemzi@gmail.com

[4]Department of Mathematics, College of Science, King Khalid University, P.O. Box: 9004, 61413 Abha, Saudi Arabia; ahmoahmed@kku.edu.sa

[5]Department of Mathematics, Shaheed Benazir Bhutoo Women University Peshawar, KP, Pakistan; mehwishtahir@sbbwu.edu.pk

Abstract

Timely and accurate electricity price forecasting is essential for the efficient operation of competitive electricity markets. However, predicting day-ahead electricity prices remains challenging due to market volatility, anomalous price observations, and complex nonlinear relationships in electricity price dynamics. This study proposes a novel ensemble forecasting framework for predicting day-ahead electricity prices in the German electricity market. The framework integrates classical time-series models, machine-learning approaches, and hybrid models as base learners, and combines their forecasts using three weighting strategies based on validation performance, model

*Corresponding author: fahimjan@bkuc.edu.pk

†Corresponding author: izatmandhaleemzi@gmail.com

diversity, and error minimization. To enhance forecasting robustness, six years of hourly electricity market data were preprocessed to detect and normalize anomalous observations. Experimental results demonstrate that the proposed ensemble framework consistently outperforms individual classical time-series models, standalone machine-learning models, and hybrid approaches. Among the benchmark models, hybrid models show better predictive performance than classical time-series methods but remain inferior to the proposed ensemble strategies. The findings highlight the importance of anomaly treatment and intelligent model combination for improving electricity price forecasting accuracy. The proposed framework provides valuable insights and practical forecasting tools for energy traders, grid operators, and policy-makers involved in electricity market decision-making.

Keywords: Electricity Price Forecasting, Anomaly Detection, Ensemble Learning, Machine Learning, STL Decomposition.vector autoregressive, autoregressive moving average, autoregressive neural network, k-nearest Neighbors

1 Introduction

To have an efficient and secure deregulated energy market, it is important that electricity prices should be forecasted accurately. An accurate assessment of what the firms must pay for power and the direction and magnitude of its volatility plays a role in many types of decision making, including trading, risk management, and grid stability [26]. Forecasting electricity prices is a complicated task because electricity is a unique commodity having its specific features: it cannot be stored, and demand for electricity must match supply instantaneously; this leads to a time series with dramatic volatility, sharp price spikes, complex seasonal patterns (daily, weekly, yearly), and the potential for excessive shocks accounting for weather, fuel prices, renewable generation, etc, [30; 14]. While many statistical model types exist to model electricity price data (e.g., ARIMA), the nonlinearities, compounded with the high frequency of data have encouraged researchers in the field of electricity markets to turn to machine learning model types [13; 16]. These models (e.g., neural networks and ensemble methods) have a greater ability to identify these nonlinearities, especially when large datasets are created. This research program aims to develop a method of forecasting that carefully considers and tackles the unique challenges facing electricity markets when determining price forecasting, through careful data preprocessing and careful model selection and implementation also contributes to the research efforts to make reliable forecasts with some accuracy of what will take place in the fluctuating German electricity market. Research along with forecasting model applications have garnered substantial attention from various research domains during the previous decades. The development of multiple forecasting models has emerged during recent years to estimate electricity prices. A substantial number of investigations along with numerous experiments have emerged from different perspectives to create advanced electrical power system management and electricity market forecasting models for electricity prices.The research in [25] examined data preprocessing algorithms for forecasting while [29] focused on feature selection methods. A review of the related literature reveals that almost all models can be regarded as hybrid forecasting models. The hybrid electricity price forecasting (EPF) model introduced by [28] combines linear and nonlinear prediction functions to analyze three real-world electricity markets. The new hybrid model developed by [5] achieved accurate EPF predictions after testing with four case studies from the European Power Exchange Spot. The adaptive hybrid model developed by [29] proved its effectiveness through testing with three actual electricity

market datasets. Studies using hybrid forecasting methods that incorporate the learn from strong to offset weaknesses concept progress by utilizing the unique advantages of multiple methods because of the identified deficiencies in single forecasting models and the current trends in EPF research. New and innovative perspectives in EPF research emerge through these approaches.

Continuing on this base, it becomes evident that reliable EPF is challenged by anomalies, i.e. extreme price spikes or drops of the market, which are induced by the disruption of the market, supply shocks, or changes of the policies. When not mitigated, these problems can lead to a distortion of model estimation and highly exaggerated forecast errors, as well as misrepresented estimation uncertainty [21]. Although classical time series models such as ARIMA and seasonal decomposition can be useful to explain linear trends and regular patterns of seasons, they are usually not successful in explaining such unusual events and non-linear dynamics of the modern electricity market. Due to this, they are largely not useful as benchmarks particularly under highly volatile conditions where the unmodeled anomalies could greatly reduce the forecast accuracy and operational reliability. Thus, the methods of identifying and dealing with anomalies have become essential to enhancing the accuracy of classical and novel forecasting strategies. This explains why the complex hybrid frameworks that have been mentioned above are necessary.

The last years were characterized by significant advancement in applying machine learning (ML) methods to electricity price forecasting (EPF). The methods have superseded classical autoregressive and naive standards. Artificial neural networks (ANN), k-nearest neighbors (KNN) and autoregressive neural networks (ARNN) have been found to be more effective in capturing complex price patterns [19]. It is becoming popular as ensemble learning methods that are combinations of multiple individual models. They leverage mutual strengths and contribute to reducing the variance of predictions when forecasting is used at the same time with others of its kind [1; 3]. These are ensemble and ML-based forecasting methods that are logical extensions of previous forecasting methods.

According to the review by Weren, the approaches to electricity EPF have been studied throughout the years [26]. These modern techniques might be more computationally intensive, as demonstrated in the study by [14] but the power of computers is increasing, so this cost should no longer be a concern. Deep learning (DL) as a component of EPF is also fairly recent, and the research started gaining momentum in 2016 or later on [17]. This study is based on the previous studies on hybrid and ensemble forecasting whereby emphasis has been put on long term trends, seasonal trends, and random sub-series to form very accurate forecasts.

Hybrid models which take after other approaches have also been created [7; 9]. A single study [18] developed a mid-term load forecasting model which employed input of normalized seasonal cycles and trends. This model was superior to others such as Prophet, ETS and ARIMA by 13.7%, 17.4% and 25% percent mean absolute percentage error (MAPE). This method has the prospects of both mid-term and short-term electric power demand prediction.

Graph neural networks are applied in models such as Price FM, which considers spatial dependence between European markets resulting in the creation of more robust probabilistic predictions [27]. Another form of regime-conditional neural processes (R-NP) is a regime detection used with multi-criteria optimization. This can deal with the structural breaks and regime shifts that are important in maintaining some accuracy during turbulent market times [6] Fueled by these developments and the practical issues in predicting the price of electricity the present paper examines classical time series models, machine learning algorithms, and an emerging ensemble framework. This model employs

various weighting models on the ground of training and validation performance using hourly data 2018-2023. This is aimed at bridging the gap between theory and practice observing the multifaceted nature of the electricity market today with a combination of both proven and emerging forecasting techniques.

In addition, STDF has been implemented in smart grids in cities of Australia. This shows that ANNs are powerful in terms of generalization of this purpose[**?**]. Nevertheless, it has a number of disadvantages. These are the risk of overfitting, slow convergence rate and easily getting stuck in a local optimum. There is an increase in renewable energy sources, and that complicates the process of predicting smart grid loads. To have good forecasting results, more advanced technologies are required. The recent developments of these techniques emphasize valuable input by researchers like, but not limited to, [23; 20; 22; 8; 12; 10].

Their innovative work is pushing the field forward, making it important to stay updated on these significant developments. The main objects of the current study is:

- To analyze short-term electricity price dynamics within the Munich, German day-ahead market, using hourly data for the entire period from January 2018 to December 2023.
- To evaluate the forecasting performances of conventional time series models including autoregressive and naive methods.
- To assess machine learning predictive abilities such as k-nearest neighbors (KNN), artificial neural networks (ANNs), and autoregressive neural networks (ARNNs).
- The proposed an ensemble forecasting framework where outputs of several individual models would be integrated.
- Design and benchmark three weighting strategies for ensembles:
  1. Equal weights (Ensemble A)
  2. Weights based on inverse training errors (Ensemble B)
  3. Weights based on validation accuracy (Ensemble C)
- Rank the forecasting methods considering empirical performance and identify the most effective approach for short-term EPF.
- Provide insights on the advantages of the validation-based weighting schemes in improving forecasting accuracy in the context of competitive electricity markets.

A variety of forecasting techniques are employed to achieve these objectives. Traditional statistical models such as regression analysis, time series models like AR, and naïve methods are widely used for their trend and pattern recognition robustness. On the machine learning front, methods including k-nearest neighbors (KNN), artificial neural networks (ANN), and autoregressive neural networks (ARNN) are applied due to their ability to model complex nonlinear relationships within large datasets. Moreover, The ensemble proposed here incorporates three weighing schemes: Ensemble A (equal weights); Ensemble B (weights based on inverse training errors); and Ensemble C (weights based on validation accuracy), approaches that blend statistical and machine learning techniques are increasingly popular, offering enhanced accuracy by leveraging the strengths of both methodologies.

The rest of the study follows in Section 2, and that section discusses preparing the data. Section 3 discusses the models suggested in this paper. Section 4 discusses the methods of

treatment of anomalies,Section 5 discusses the alternative methods. In Section 6 discusses measures of quality and implementation of the methodology and discusses the results. It includes also discussions of the electricity price data, and measures of quality. In Section 7 we present the results and discussion. Lastly the paper concludes with some remarks (Section 8) about the concluding remarks and ideas for future research.

2 PREPARATION OF RAW DATA

This study focuses on achieving a good, accurate, and efficient one-month-ahead forecast for Germany electricity price. In Line with this goal a sequential approach was conducted to model and forecast the electricity price time series in Germany. A sequential approach was developed as the goal was to better understand the unique complexities of forecasting electricity prices considering the tendency towards complex structure with missing values, a nonlinear long-run trend, strong seasonality, high volatility, non-normality, and non-stationarity which are highly problematic for static forecasting. The main goal of this study is to achieve an accurate and efficient one-year-ahead forecast for Munich, Germany electricity price. The authors conducted a step-by-step process to analyze and predict Munich, German electricity prices using time series data for this purpose. The authors investigate the intricate patterns which govern electricity price movement throughout different time periods. The observed characteristics will include anomalies in the data points when using the moving filter on price.

This is the new strategy that will address the price series anomalies. These abnormalities will be identified and substituted with normal values. This will guarantee a more accurate set of data. Once this meticulous manipulation is done, then the series of prices will be anomaly-free. This will allow for more accurate analysis and decision-making. The second stage is to take the natural logarithm of the series which would reduce variance and heteroscedasticity associated with the time series. The third stage is an explicit treatment of the deterministic components that we will model and detrend which in this case is a linear long run trend and yearly seasonality in the data. We will denote the log of the hourly price series as $Log(Y)_j$; j shows the j^{th} hour data point. Thus, the dynamics of the log hourlyprice times series, $Log(Y)_j$, may be described as:

$$log(Y_j) = exp(\xi_j + \rho_j + y_j) \tag{2.1}$$

The $log(Y_j)$ represents the sum of three components which include the long-run linear trend component (ξ_j), the yearly seasonality component (ρ_j), and the residual component (y_j). The (y_j) component depends on the series values $(1, 2, 3, \cdots, j)$ and is estimated through regression splines and dummies which capture the annual periodicity $\rho_j, k = \sum_{k=1}^{24} \phi_k \Psi_{j,k}$. The variable $\Psi_{j,k}$ holds the value of 1 for the k^{th} hour and 0 for all other hours. The regression coefficients (ϕ_k) which correspond to these components are computed through the ordinary least square method. The residual or stochastic component can be computed after obtaining the estimated deterministic component (long-run trend and annual periodicity) as follows:

$$y_j = log(Y_j) - (\hat{\xi}_j + \hat{\rho}_j) \tag{2.2}$$

The preprocessing of electric power price time series (Y_j) involves first handling missing values through imputation and then stabilizing variance and standard deviation before removing deterministic properties before modeling the residual y_j series. This study examines five individual time series models together with three ensemble models for the

modeling of the residual y_j series after completing the preprocessing of electric power prices. The upcoming subsection presents details regarding all forecasting models that have been included in this study.

Table 1: Descriptive Statistics of Germany electricity price (MW)

Statistic	Value
Minimum	0.10
1st Quartile	42.83
Median	53.57
Mean	56.60
3rd Quartile	67.83
Maximum	324.20
Standard Deviation	19.65

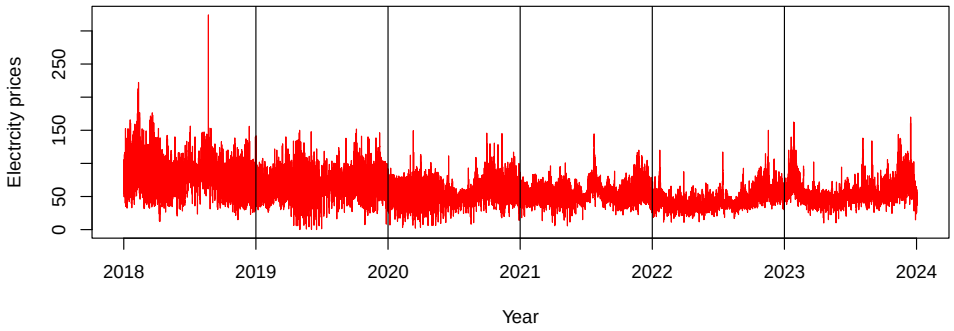


Figure 1: electricity prices: The hourly filtered electricity prices is shown

3 The proposed ensemble models

At its core, an ensemble technique integrates outcomes from various models, each meticulously calibrated before unity. This approach capitalizes on the inherent strengths of individual models while compensating for their inherent limitations. The initial step of this investigation uses ensemble methods to create weights which integrate outcomes from individual models. The proposed ensemble system contains three separate weighting approaches which consist of Ensemble (A) that provides equal weights to all single models and Ensemble (B) which uses training average accuracy errors (1) and Ensemble (C) which uses validation mean accuracy measures. The model assigns higher weights to the ensemble model when training and validation datasets have reduced mean accuracy errors while models with higher mean accuracy errors receive lower weights for their ensemble contribution. The model weights exist as small positive values and their sum equals one thus representing the percentage reliance or expected performance on each model. After estimating the long-run trend component together with annual periodicity using spline regression and dummy variables the following step involves forecasting y_j with five single and three proposed ensemble models. The following equation shows the method to calculate hour electric power price for upcoming hours forecast in this research.

$$\hat{Y}_j + 1 = exp(\hat{\xi}_j + \hat{\rho}_j + \hat{y}_j) \tag{3.1}$$

4 Alternative methods

In this section we discussed some of the popular classical time series models and machine learning models. The results of the suggested and aternative models are compared.

4.0.1 AutoRegressive (AR) model

An AR model is core of TS analysis, that regresses the current value of a variable on its previous values. The univariate AR model predicts a variable by its own observations from the past. The AR model of order P, AR(p), describes the current value as a mixture of j preceding values and a random error term. Here is the mathematical representation:

$$y_j = v + \sum_{l=1}^{p} \delta_l y_{j-l} + \varepsilon_j \tag{4.1}$$

Where y_j is the current value, v is a constant term, the parameters of the model are δ_l, and ε_j is the error term. The model is therefore able to capture temporal dependencies in the data, rendering it relevant in several applications including, but not limited to economics, engineering, and environmental science. The current values are identified and given as per the formulations in equations mentioned above. Current value based on previous data with stimulus by the number of days. In most of the studies AR(7) is used for example [11; 21], So in this paper we also use AR(7).

4.0.2 The naïve Benchmark

This section describes a naïve predictive model classified in similar-day techniques and, according to the research, is the best naïve approach in accuracy [11]. The methodology can be described as follows:

1. Identification of Reference Day: That is to forecast the Monday. Now for the preceding day (Sunday), denote all metrics as y^*.
2. Compiling Similar Days: Collect this data from other Sundays apart from y^* to form a pool of comparable days.
3. Calculated MAE:Calculate the MAE for y^* and all other Sundays in the validation set, then create a vector of MAE values.
4. Optimal Match: Identify the Sunday that corresponds to the minimum MAE, then data from next Monday will show what will happen on the target Monday. This procedure will go for all the days of a week. This would mean that the day before the intended day behaves like all the other days of such a type and thus improves one's forecast performance.

4.1 k-Nearest Neighbors

k-Nearest Neighbors (KNN) is a well-known non-parametric technique in machine learning, effective for solving both classification and regression tasks. The core concept of KNN is based on the idea of learning by analogy. When estimating the target value for an unseen instance, the model looks for the k nearest neighbors within the training set. This process involves computing the distance between the new instance and all other points in the set. The Euclidean distance is a frequently used measure that calculates the straight-line

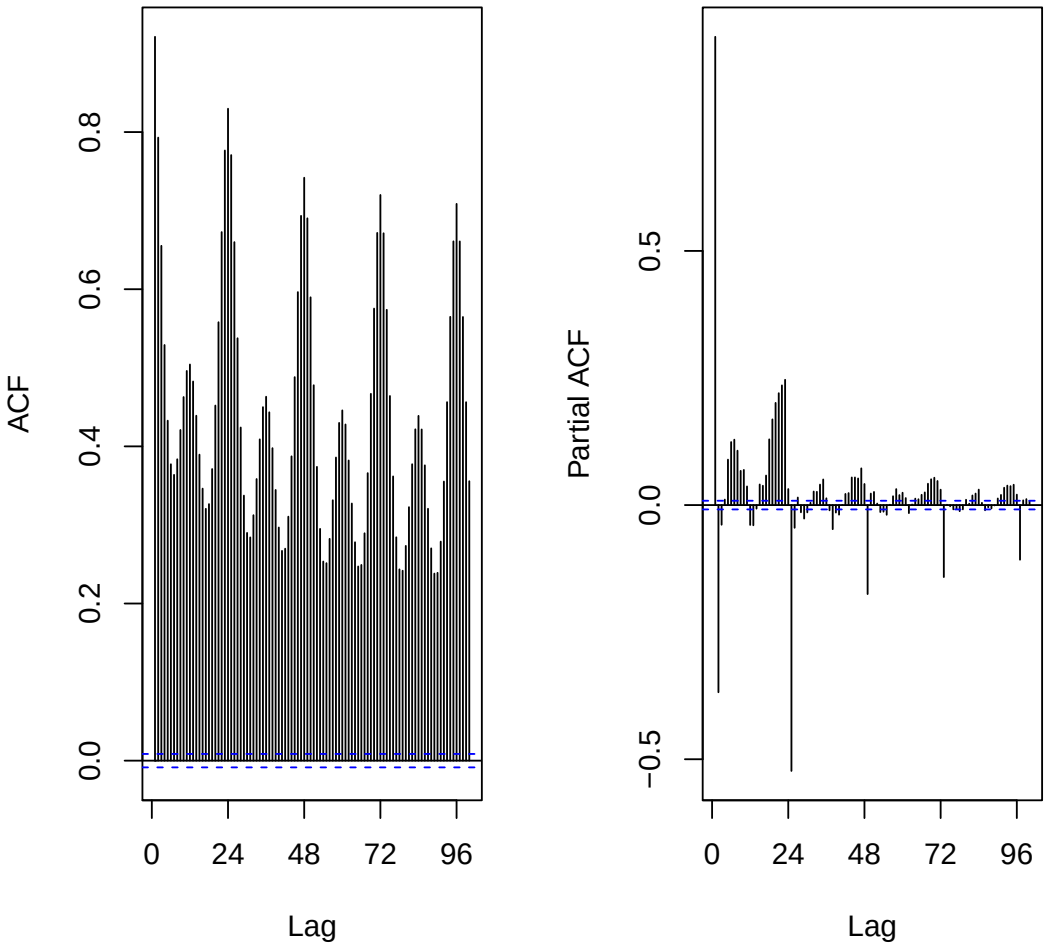


Figure 2: ACF and PACF Plots. Visualizing the strong daily seasonality (peaks at 24-hour intervals) and short-term autocorrelation in the hourly electricity price series.

distance between two points in a multi-dimensional space. For instance, when comparing two time series $y = (y_1, y_2, \ldots, y_j)$ and $T_l = (T_{l1}, T_{l2}, \ldots, T_{lj})$, the distance between the two time series can be calculated as:

$$d(y, T_l) = \sqrt{\sum_{k=1}^{j} (y_l - T_{lk})^2}$$

where, y_l and T_{lk} represent the demand values at the l-th interval for the time series y and T_l, respectively.

After identifying the nearest neighbors, the model uses the mean of the target values of the nearest neighbors to estimate the target value of the new instance. A cross-validation (CV) technique was used to identify the optimum values of the lag and the k value in this research. This is necessary as it assists in finding the values that minimise model error, but not lead to overfitting or underfitting. Through experimentation with variations of lag and k, the experiment found that a lag of 3 and k value of 7 resulted in the most accurate predictions on the provided dataset and gave the best tradeoff between the complexity of

the model and its performance in prediction.

4.2 Artificial Neural Networks

Artificial Neural Networks (ANNs) are a type of machine learning method inspired by how the human brain processes information. They have several layers of connected nodes that look like biological neurons. These networks are designed to replicate the brain's learning and decision-making processes. Information flows into the system through the input layer, moves through one or more intermediate layers, and ultimately reaches the output layer. Each node in the network receives signals from other nodes, processes these inputs, and sends a resulting output. The strength of the connections between nodes, referred to as weights, determines the degree of influence one node has over another. A widely used type of ANN is the feedforward neural network, where data flows in one direction, from the input layer to the output layer, without any recurrent feedback loops.

Model	**MAPE**	**MAE**	**RMSE**
AR	4.63	6.45	7.22
naïve	8.72	9.45	10.94
KNN	4.88	7.23	8.25
ANN	4.68	6.63	7.35
ARNN	4.43	6.17	7.03
Ensemble(A)	3.58	5.63	5.25
Ensemble(B)	3.37	5.43	5.15
Ensemble(C)	3.28	5.35	5.07

Table 2: Out-of-sample forecasting performance comparison between AR, naïve,KNN,ANN, ARNN,Ensemble(A),Ensemble(B), and Ensemble(C) usingMAPE, MAE, and RMSE

4.2.1 Autoregressive Neural Networks

An Autoregressive Neural Network (ARNN) is a deep learning model that combines the principles of AR models with neural networks to capture complex temporal dependencies in time series data. The fundamental idea behind ARNN is the same as the AR models, wherein historical values are invested in as an input to predict upcoming values. Nonlinear relationships however can be captured by ARNN unlike the traditional AR models, and therefore ARNN is suitable with the complex time series data. In the ARNN(j, l) architecture in particular, p lagged inputs and k hidden layer neurons are used. This architecture is being trained to do one step ahead prediction using the latest past note

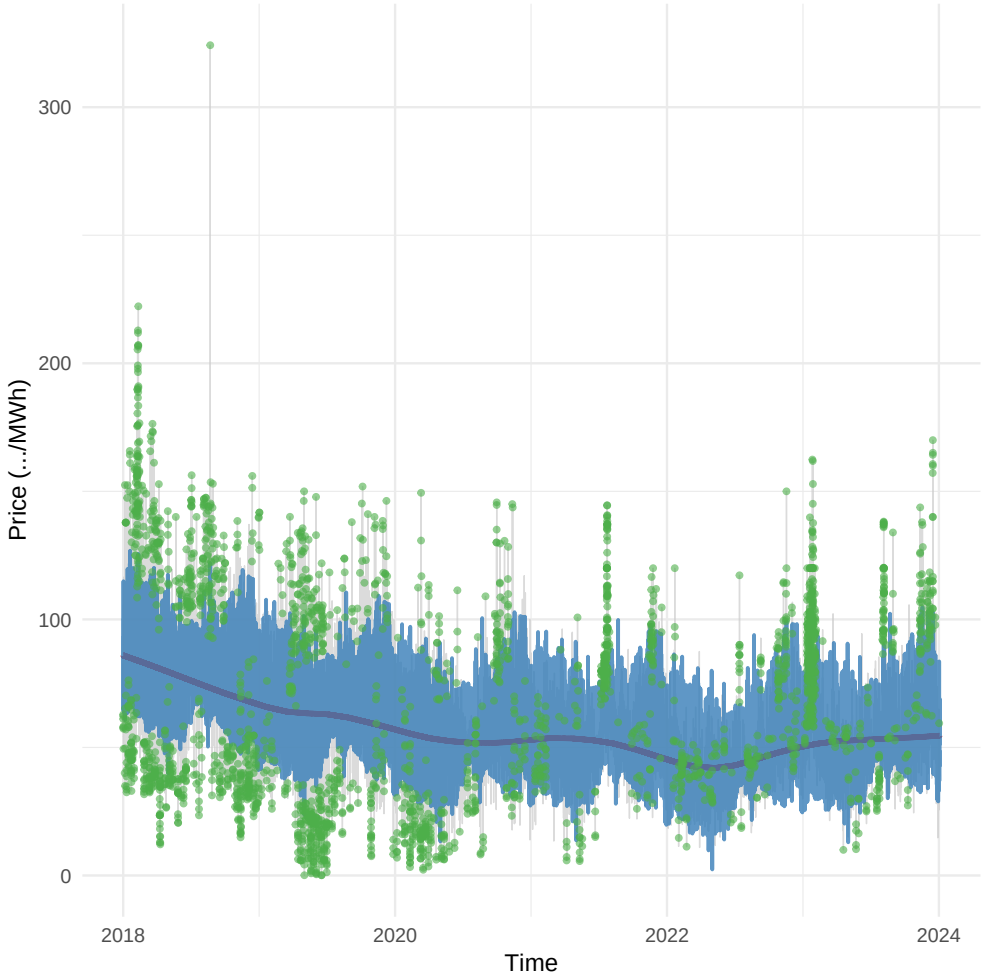


Figure 3: The original time series appears in light gray. The extracted components are highlighted in color: a strong upward trend in red, consistent seasonal patterns in blue, and significant outlier events in green. This visualization clearly shows the long-term direction, predictable cycles, and unexpected volatility in the market.

as input. The ARNN model has certain similarities to the ARIMA models but operates without having any stationarity requirements. It is also more flexible and can be used with various kinds of time series data. The equation of the model is as follows:

$$y_{jt} = g\left(\sum_{i=1}^{N} w_{ji} y_{i(t-1)}\right) + \kappa x_{j(t-1)}^{T} \epsilon_t$$

In which the notation y_{jt} denotes the response of the node j and time t and w_{ji} denotes the strength of the connection between node j and node i. This expression denotes the external influences, which is represented by the product of the preceding year terms: the outside impacts denoted by the term $\kappa x_{j(t-1)}^{T}$, the nonlinear activation term $g(\cdot)$, and the shock expression t by the term ϵ_t. ARNN model is optimized by estimations that include profile least squares estimation and local linear approximation. Such techniques are useful in modeling the unclear linking function and also giving good estimates of the parameters. The model is repeatedly trained using history data to both make one-step-ahead and multi-step predictions. The ARNN(3,7) model is applied in this research. This implies that three observations were taken in the past and the hidden layer had seven neurons as an input and this was determined by trial and error.

Table 3: Daily forecasting errors of all models

Method	Error	Days of the Week						
		Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
AR	MAPE	4.55	3.98	3.15	4.09	3.77	4.88	3.88
naïve		5.77	6.57	7.11	5.60	7.91	8.87	9.31
KNN		4.83	3.68	4.95	4.26	4.26	5.31	5.43
ANN		4.61	3.42	4.03	3.59	3.89	4.59	5.06
ARNN		4.29	3.76	3.02	3.94	3.62	3.78	3.47
Ensemble (A)		3.00	3.06	3.17	2.94	2.88	3.26	3.17
Ensemble (B)		3.09	3.04	2.87	2.84	2.77	3.13	3.04
Ensemble (C)		3.05	3.01	2.82	2.77	2.72	3.06	2.92
AR	MAE	5.99	6.52	7.09	6.20	7.82	6.01	5.01
naïve		7.27	8.96	9.76	7.20	9.92	9.40	8.67
KNN		6.11	7.11	7.29	8.16	8.08	7.99	7.40
ANN		5.98	6.85	6.02	5.70	7.16	5.33	4.68
ARNN		5.68	6.32	6.92	7.61	5.83	4.89	4.93
Ensemble (A)		4.40	5.28	5.16	5.20	4.25	4.56	4.27
Ensemble (B)		4.30	5.16	5.07	5.13	5.14	4.42	4.10
Ensemble (C)		4.27	5.10	5.03	5.04	5.03	4.29	4.01
AR	RMSE	7.09	6.81	6.88	8.23	7.17	6.92	8.00
naïve		11.63	10.66	10.76	9.22	10.09	11.33	11.47
KNN		7.51	7.71	7.30	9.46	8.39	7.06	8.63
ANN		6.29	6.41	6.26	7.99	7.03	6.54	7.84
ARNN		6.88	6.70	6.45	8.14	6.92	6.35	7.79
Ensemble (A)		5.02	5.76	5.27	6.24	5.42	5.66	6.47
Ensemble (B)		4.93	5.70	5.20	6.14	5.36	5.57	6.36
Ensemble (C)		4.91	5.69	5.18	6.06	5.28	5.54	6.27

Table 4: Monthly forecasting errors of all models

Month	Errors	AR	naïve	KNN	ANN	ARNN	Ensemble (A)	Ensemble (B)	Ense
	MAPE	6.96	7.23	7.10	6.93	6.77	6.28	6.17	
January	MAE	7.56	7.94	7.69	7.51	7.38	7.14	7.03	
	RMSE	9.22	9.84	9.36	9.30	9.04	8.66	8.56	
	MAPE	6.38	7.65	7.18	6.85	6.21	6.06	5.83	
February	MAE	6.78	7.42	7.29	7.06	6.34	6.03	5.85	
	RMSE	8.59	9.40	8.89	8.75	8.22	8.08	7.86	
	MAPE	5.98	6.93	6.23	6.15	5.62	5.11	4.92	
March	MAE	6.81	7.92	7.14	7.06	6.59	5.99	5.67	
	RMSE	7.23	8.14	7.45	7.32	7.10	6.88	6.74	
	MAPE	4.99	6.56	5.85	5.72	4.63	4.15	3.98	
April	MAE	5.58	6.68	5.94	5.64	5.27	4.78	4.52	
	RMSE	6.34	7.84	7.19	7.07	6.16	5.23	5.15	
	MAPE	5.79	7.32	6.29	6.11	5.71	5.45	5.28	
May	MAE	6.12	7.55	7.16	7.03	6.02	5.56	5.34	
	RMSE	6.49	7.44	6.88	6.72	6.23	5.34	5.08	
	MAPE	5.98	6.69	5.34	5.17	5.46	4.45	4.34	
June	MAE	6.34	7.53	6.52	6.46	6.21	5.94	5.73	
	RMSE	7.37	8.80	7.29	7.10	7.13	6.54	6.39	
	MAPE	5.23	6.02	5.77	5.66	5.12	4.45	4.16	
July	MAE	6.47	7.83	6.63	6.51	6.25	5.67	5.88	
	RMSE	7.06	7.94	7.35	7.24	6.90	5.94	5.73	
	MAPE	3.45	5.69	4.84	4.73	3.17	2.45	2.26	
August	MAE	5.88	7.57	6.44	6.14	5.77	4.01	3.78	
	RMSE	6.09	7.51	6.85	6.66	5.79	4.44	3.99	
	MAPE	6.39	7.91	7.20	7.10	6.18	5.34	5.19	
September	MAE	6.56	8.94	6.99	6.81	6.34	5.14	5.03	
	RMSE	6.21	7.63	6.75	6.64	6.03	5.27	5.12	
	MAPE	5.67	7.34	5.90	6.78	5.40	4.44	4.19	
October	MAE	6.16	8.75	6.82	6.73	5.82	4.23	4.07	
	RMSE	6.89	8.54	7.16	7.03	6.33	4.63	4.34	
	MAPE	5.11	7.35	5.84	5.71	4.75	3.88	3.56	
November	MAE	6.28	8.52	6.71	6.55	5.25	5.04	4.84	
	RMSE	5.39	6.66	6.22	6.14	5.08	3.88	3.74	
	MAPE	5.23	7.89	5.63	5.58	4.57	3.29	3.09	
December	MAE	6.22	9.24	6.57	6.39	5.86	4.55	4.45	
	RMSE	5.84	9.38	6.48	6.27	5.41	4.34	4.23	

5 Anomalies detectation method

The literature often calls anomalies by different names in various fields. These include outliers, novelty, flaws, unusual observations, extreme values, change points, intrusions, misuses, exceptions, aberrations, surprises, peculiarities, odd values, and contaminants [4; 24]. Anomalies can greatly impact forecasting and estimation. Time series models are particularly challenging to estimate and forecast because extreme values frequently occur

in time series data. A crucial step in the data cleaning process for time series analysis is identifying anomalies and replacing them with normal values. This often improves parameter estimation and the performance of forecasting models. Anomalies in data related to electricity consumption or prices can be caused by various events, such as sudden spikes in demand, power plant outages, severe weather, and more. The box plot of 60 min electricity prices is plotted in Figure 4. One can see that there are anomalies present in the prices series. Many strategies are generally used to identify and replace these anomalies [21]. The moving filter on price method proposed by [2] is used to treat them in this study.

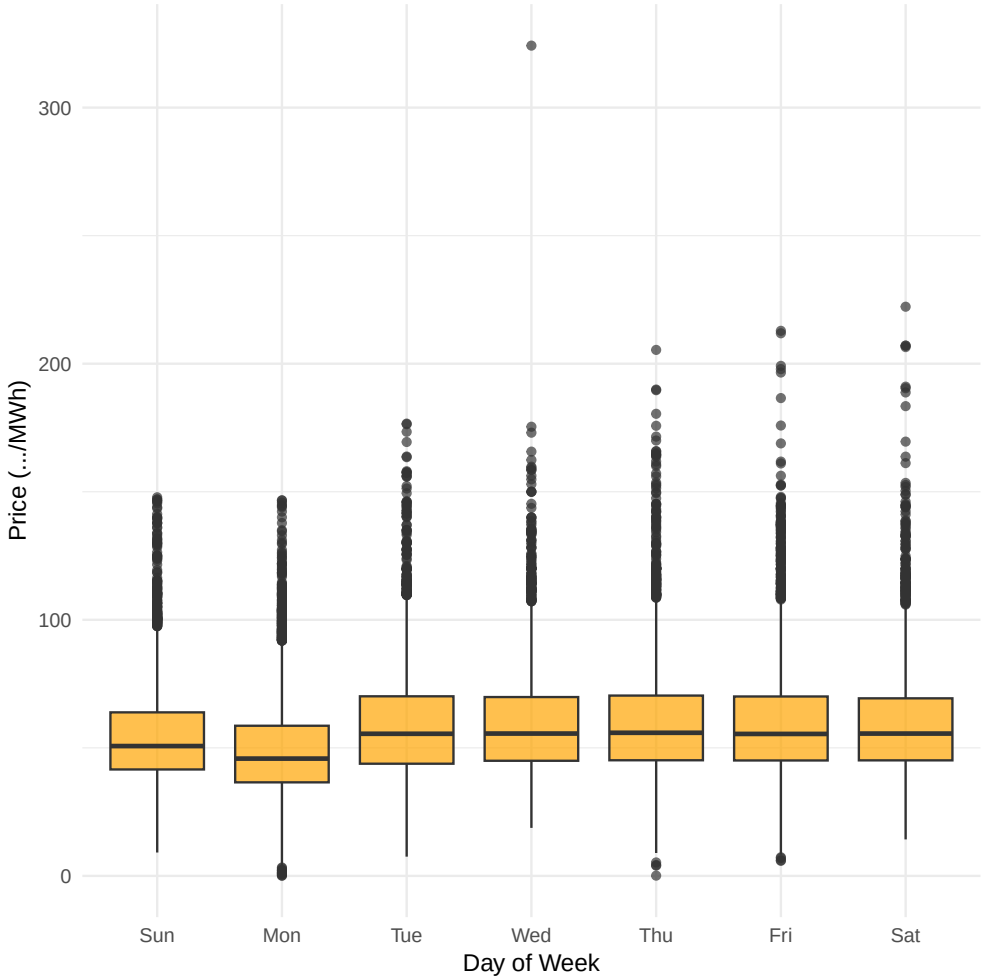


Figure 4: Eleectricity price distribution by day of week

5.1 Moving filter on price

The Standard Deviation Filter on price (SFP) is a technique that builds on the concept of the Moving Filter on price (MFP). While the MFP operates using a rolling window of a constant width, the SFP utilizes the entire time series.

If we denote the window width as s and the total number of observations in the series as v, the original time series is divided into $w = v/s$ segments during the first step. The initial segment of the time series is then processed using the MFP. Anomalies are identified as price values whose absolute difference from the estimated mean $(\widehat{\mu})$ exceeds a defined multiple of the estimated standard deviation (SD). Using the MFP filter, the subset of spikes y_j^o found is obtained as

$$y_j^o = \{y_j : |y_j - \widehat{\mu}| \geqslant 1.64 \cdot SD(y)\}$$

After the initial analysis, the window moves forward by a set number of observations, and anomalies are detected by applying the filtering process to the next window. This process continues until all w components have been examined. Once the anomalies are identified, the median value of the corresponding window is used to construct a price series that is free from anomalies.

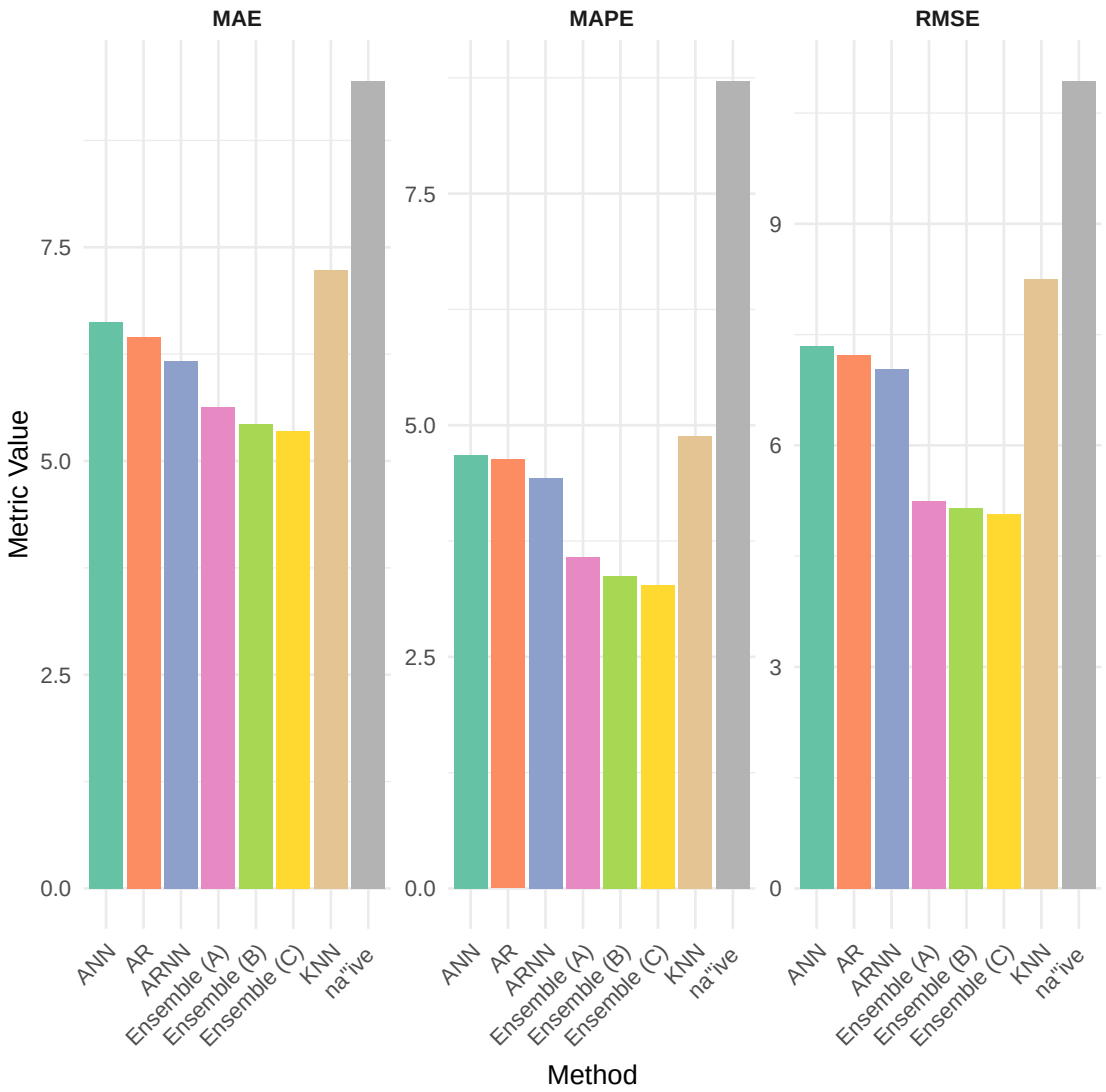


Figure 5: Performance Comparison of Methods

6 Quality Measures

In this section various quality indicators are computed to compare the models and the electricity data are presented in this section. We turn the projected curves into 60-min electricity price data to make the comparison easier. The mean absolute percentage error (MAPE), mean absolute error (MAE), and root mean square error (RMSE) are all used to evaluate forecast accuracy. The eletricity price data are also presented in this section.

6.1 Mean Absolute Percentage Error

To calculate the MAPE, divide the average absolute error for each period by the observed values and multiply by 100. The MAPE is useful for evaluating accuracy when a prediction variable is large [15]. The MAPE gives a relative measure of forecast error. It shows the difference between the forecasted and observed values. Mathematically, the MAPE can be expressed as:

$$\text{MAPE} = \text{mean}(|\frac{y_{j,m} - y_{j,m}}{y_{j,m}}|) \times 100$$

where $y_{j,m}$ is the observed data and $y_{j,m}$ is the predicted one-day-ahead electricity load for $j^{th}(j = 1, 2, \ldots, 365)$ day and m^{th} $(m = 1, 2, \ldots, 24)$ price period.

6.2 Mean Absolute Error

To make sure that negative numbers do not cancel out positive numbers, we calculate the Mean Absolute Error (MAE) by averaging the absolute differences between the forecasted and actual values for the same time period as follows:

$$\text{MAE} = \text{mean}(|y_{j,m} - y_{j,m}|)$$

6.3 Root Mean Square Error

The scale of the dependent variable affects the Root Mean Square Error (RMSE) and causes it to increase. This method compares forecasts from different models applied to the same dataset. To calculate the RMSE, take the square root of the average of the squared differences between the actual and forecasted values. Mathematically, it is defined as follows:

$$\text{RMSE} = \sqrt{\text{mean}(y_{j,m} - \hat{y_{j,m}})^2}$$

6.4 Out-of-sample electricity price forecasting

The day-ahead electricity market in Munich, Germany, is one of the largest and most active in Europe. It shows significant volatility because of the high use of renewable energy sources, price changes, and regulatory shifts. The data was taken from the $(https://data.open-power-system-data.org)$. The main goal of this study is to model and forecast 60-minute electricity price data from the Munich market, covering January 2018 to December 2023. To ensure effective modeling and forecasting, we divided the data into two separate groups: one for model estimation and the other for out-of-sample forecasts.

- Validation data: From 1^{st}, January 2018 to 31^{st}, December 2022 (43,824 observations,covering j=1826 days)
- Testing data: From 1^{st}, January 2023 till 1^{st}, December 2023 (8,760 observations, covering j=365 days.

Using the moving or rolling window method for our one-day-ahead out-of-sample forecasts gives us a strong and flexible way to understand electricity prices. This study aims to shed light on past market behavior and to lay the groundwork for future forecasting and decision-making in the energy sector.

The models are estimated using the expanding window approach, and the annual one-day out-of-sample forecast is derived. The graph 1, the discrete 60-min electricity pricefrom 1^{st} January 2018 to 31^{st} December 2023. 4 indicates that electricity prices fluctuate across different days of the week, with noticeable variations between weekdays and weekends. Prices tend to peak on certain days, suggesting higher demand or market activity, while relatively lower values are observed on others, highlighting potential off-peak periods. The graph 2 denotes the acf, pacf used for the order section of AR,ARNN and KNN models. The graph 5 shows that, the Ensemble methods (particularly Ensemble B and C) consistently demonstrate the lowest error values across all three metrics (MAE, MAPE, RMSE), indicating they are the most accurate forecasting models for this dataset. The naïve and KNN methods appear to be the least accurate, consistently showing the highest error values. The performance of the ANN, AR, and ARNN models falls in the middle range of the results. The graph 3 The STL decomposition reveals a clear, steady upward trend in electricity prices from 2018 to 2024. It also shows a strong, consistent seasonal pattern that repeats every year. Finally, the remainder component indicates significant short-term volatility, pointing to unpredictable events causing sharp price spikes and dips outside of the normal trend and seasonality.

7 Result and Discussion

This section presents a comprehensive assessment of the out-of-sample forecasting performance of the described models. Initially, the price series outliers are detected by applying a moving filter to the price series. Following this, the filtered price series is decomposed into two components: deterministic and stochastic. The deterministic component captures seasonal patterns such as annual cycles, weekly fluctuations, bank holiday effects, and long-term trends, and is modeled using a generalized additive model (GAM). For the stochastic component, both the proposed and competitive models, such as AR, naïve, KNN,ANN ,ARNN, Ensemble(A),Ensemble(B) and Ensemble(C), are utilized to generate one-day-ahead out-of-sample forecasts. Based on the table 3, the Ensemble models (A, B, and C) consistently achieve the lowest error values on the estimation for each weekday per evaluation metric (MAPE, MAE, RMSE). This indicates that they have the greater robustness and accuracy. Among them, Ensemble (C) performs extremely well, and produces the least errors on most of the days, proving its effectiveness in modeling daily pattern phenomenon. Correspondingly, the naïve method is the worst one at all times, while ARNN proves to be the best individual model, although still low as compared to ensemble methods. As found from the monthly forecast results using the table 4, the Ensemble models like A, B, and C have shown a consistent trend across all months and metrics MAPE, MAE, RMSE to yield the least values of error. Among all these, the most promising has been Ensemble (C). With the exception of some months, the nave model has been consistently the worst performing among all others, while the ARNN model usually beats the individual models but still lags behind the ensemble methods. This shows the strength and accuracy of ensemble methods in modeling the capture of the monthly seasonal and variation profile of electricity prices in general.

8 Conclusion

This study provided a solid framework for EPF by simultaneously addressing the dual challenges of data quality and model selection. First, it determined that when anomalous price data are estimated in a model, forecasts are unreliable; thus, it is impossible to make

utility decisions based on forecasts we cannot trust. As a solution, the study employed a moving filter method to find outliers, then replaced the outlier estimates with their normal values. This was a critical data cleaning process that happened prior to model estimation and is a process that the study felt improved the integrity of the input data significantly. The ensemble models proposed performed best when modelled on the cleaned dataset, but especially Ensemble (C) model. Overall, the Ensemble (C) model outperformed all other benchmark models including the classical (AR, naïve) and machine learning (KNN, ANN, ARNN) models across all metrics and prediction horizons evaluated.

The consistent outperformance of the ensemble approach lends support to the conclusion that when each of the individual forecasts has an error in them, successfully combining forecasts attenuates the error of each of the forecasts with the combined forecast thus providing a more reliable and robust forecast. This study emphasizes that successful forecasting, in volatile markets such as electricity, cannot rely solely on modelling sophistication, but is inheretly reliant on data cleaning and prefiltering. Future work must also include external variables such as renewable generation and load data, and develop dynamic ensemble forecasting methods that capture the effect of shocks to the electricity market or sudden changes in prices so that they better improve forecasts for energy trading and grid operations.

Future research can incorporate external variables, like renewable energy generation or weather data, to further improve the predictive ability of themodels. The framework outlined in this paper can also be applied and tested to forecasting in other volatile energy markets to assess its applicability. Lastly, there can be further dynamic weighting methods for the ensemble models, that could include real-time learning to improve response to abrupt changes in the market.

Conflict of Interest

The authors declare that they have no conflict of interest regarding the publication of this paper.

Data availability statement

The data are available on the request of the reviewer from https://data.open-power-system-data.org.

Funding

This research did not receive funding

References

[1] Amor, S. B., Möbius, T., and Müsgens, F. (2024). Bridging an energy system model with an ensemble deep-learning approach for electricity price forecasting. *arXiv preprint arXiv:2411.04880*.

[2] Borovkova, S. and Permana, F. J. (2006). Modelling electricity prices by the potential jump-diffusion. In *Stochastic Finance*, pages 29–36. Springer.

[3] Cerjan, M., Krželj, I., Vidak, M., and Delimar, M. (2013). A literature review with statistical analysis of electricity price forecasting methods. *Eurocon 2013*, pages 56–63.

[4] Chandola, V., Banerjee, A., and Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3):1–18.

[5] Cheng, H., Ding, X., Zhou, W., and Ding, R. (2019). A hybrid electricity price forecasting model with bayesian optimization for german energy exchange. *International Journal of Electrical Power & Energy Systems*, 110:653–666.

[6] Das, A. and Schlüter, S. (2025). Regime-aware conditional neural processes with multi-criteria decision support for operational electricity price forecasting. *arXiv preprint arXiv:2508.00040*.

[7] Fan, G.-F., Wei, X., Li, Y.-T., and Hong, W.-C. (2020). Forecasting electricity consumption using a novel hybrid model. *Sustainable Cities and Society*, 61:10–23.

[8] Gajowniczek, K. and Zbkowski, T. (2024). Electricity forecasting on the individual household level enhanced based on activity patterns. *PLoS ONE*, 19(4):17–40.

[9] Hajirahimi, Z. and Khashei, M. (2023). Hybridization of hybrid structures for time series forecasting: A review. *Artificial Intelligence Review*, 56(2):1201–1221.

[10] Jan, F., Iftikhar, H., Tahir, M., and Khan, M. (2025). Forecasting day-ahead electric power prices with functional data analysis. *Frontiers in Energy Research*, 13:77–89.

[11] Jan, F., Shah, I., and Ali, S. (2022). Short-term electricity prices forecasting using functional time series analysis. *Energies*, 15(9):23–34.

[12] Jiang, B. et al. (2025). Model based transfer learning for efficient short-term load forecasting model training. *IEEE Transactions on Power Systems*.

[13] Kou, P., Liang, D., Gao, L., and Lou, J. (2015). Probabilistic electricity price forecasting with variational heteroscedastic gaussian process and active learning. *Energy Conversion and Management*, 89:298–308.

[14] Lago, J., Marcjasz, G., De Schutter, B., and 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*, 293:69–83.

[15] McKenzie, J. (2011). Mean absolute percentage error and bias in economic forecasting. *Economics Letters*, 113(3):259–262.

[16] Nowotarski, J. and Weron, R. (2018). Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Renewable and Sustainable Energy Reviews*, 81:48–56.

[17] Panapakidis, I. P. and Dagoumas, A. S. (2016). Day-ahead electricity price forecasting via the application of artificial neural network based models. *Applied Energy*, 172:132–151.

[18] Pełka, P. (2023). Analysis and forecasting of monthly electricity demand time series using pattern-based statistical methods. *Energies*, 16(2):8–17.

[19] Raaj, S., Gupta, V., Singh, V., and Shukla, D. P. (2024). A novel framework for peak flow estimation in the himalayan river basin by integrating swat model with machine learning based approach. *Earth Science Informatics*, 17(1):211–226.

[20] Santos, D. R. et al. (2023). Explainable artificial intelligence for definition of inputs in neural networks and k-nearest neighbors forecasting of electricity consumption. *IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*.

[21] Shah, I., Jan, F., and Ali, S. (2022). Functional data approach for short-term electricity demand forecasting. *Mathematical problems in engineering*, 2022(1):67–79.

[22] Singh, R. et al. (2024). A hybrid k-means and knn approach for enhanced short-term load forecasting. *Energy Reports*, 10:1123–1135.

[23] Syed, M. et al. (2024). Application of artificial neural network for peak load forecasting in a 150 kv power system. *Journal of Electrical Engineering and Technology*, 19(1):123–135.

[24] Toth, E. and Chawla, S. (2018). Group deviation detection methods: A survey. *ACM Computing Surveys (CSUR)*, 51(4):1–11.

[25] Wang, D., Luo, H., Grunder, O., Lin, Y., and Guo, H. (2017). Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and bp neural network optimized by firefly algorithm. *Applied Energy*, 190:390–407.

[26] Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International journal of forecasting*, 30(4):1030–1041.

[27] Yu, R., Gu, C., Stiasny, J., Wen, Q., Dilov, W. S., Qi, L., and Cremer, J. L. (2025). Pricefm: Foundation model for probabilistic electricity price forecasting. *arXiv preprint arXiv:2508.04875*.

[28] Zhang, C., Zhou, J., Li, C., Fu, W., and Peng, T. (2017). A compound structure of elm based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting. *Energy Conversion and Management*, 143:360–376.

[29] Zhang, J., Tan, Z., and Wei, Y. (2020). An adaptive hybrid model for short term electricity price forecasting. *Applied Energy*, 258:40–87.

[30] Ziel, F. and Weron, R. (2018). Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks. *Energy Economics*, 70:396–420.