Received 26 September 2022, accepted 5 October 2022, date of publication 10 October 2022, date of current version 14 October 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3213081

RESEARCH ARTICLE

Day-Ahead Electricity Price Forecasting Based on Hybrid Regression Model

ALI NAJEM ALKAWAZ[1], (Member, IEEE), ABDALLAH ABDELLATIF[1], (Member, IEEE), JEEVAN KANESAN[1], ANIS SALWA MOHD KHAIRUDDIN[1], AND HASSAN MUWAFAQ GHENI[2]

[1]Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur 50603, Malaysia
[2]Computer Techniques Engineering Department, Al-Mustaqbal University College, Hillah 51001, Iraq

Corresponding author: Jeevan Kanesan (jievan@um.edu.my)

This research work has been supported by Ministry of Higher Education under the Fundamental Research Grant Scheme (FRGS/1/2020/ICT02/UM/02/2).

ABSTRACT Since the deregulation of the power markets, accurate short term Electricity Price Forecasting (EPF) has become crucial in maximizing economic benefits and mitigating power market risks. Due to the challenging characteristics of electricity price, which comprise high volatility, rapid spike, and seasonality, developing robust machine learning prediction tools becomes cumbersome. This work proposes a new hybrid machine learning method for a day-ahead EPF, which involves linear regression Automatic Relevance Determination (ARD) and ensemble bagging Extra Tree Regression (ETR) models. Considering that each model of EPF has its own strengths and weaknesses, combining several models gives more accurate predictions and overcomes the limitations of an individual model. Therefore, the linear ARD model is applied because it can efficiently deal with trend and seasonality variations; on the other hand, the ensemble ETR model is employed to learn from interactions, and thus combining ARD with ETR produces robust forecasting outcomes. The effectiveness of the proposed method was validated using a data set from the Nord Pool electricity market. The proposed model is compared with other models to demonstrate its superiority using performance matrices, such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). Experiment results show that the proposed method achieves lower forecasting errors than other individual and hybrid models. Additionally, a comparative study has been performed against previous works, where forecasting measurement of the proposed method outperforms previous works' accuracy in forecasting electricity price.

INDEX TERMS Electricity price forecasting, electricity market, hybrid regression models, short-term day-ahead prediction, time series analysis.

I. INTRODUCTION

Electricity price is a critical component in the electricity market. The power grid's economical and reliable operation can be ensured beyond accurate Electricity Price Forecasting (EPF), considered a critical part of all participants in the power market competition. In addition, users might control power purchase charges by modifying electricity usage by forecasting electricity prices from the consumer's view. Conversely, producers could formulate an accurate bidding plan using EPF to obtain more significant revenues [1], [2]. Moreover, electricity prices differ from other resources and even commodities because of characteristics such as supply and demand balance, oligopolistic generation, and unexpected decreases in consumption and generation that may cause instability of the electrical grid due to consumption and generation imbalances [3], [4]. These features lead to various significant attributes of energy prices, including daily volatility prices because of the growing use of Renewable Energy Sources (RES), irregular fluctuation, and seasonality variations. However, due to market competition, most of these characteristics could not be accessible to investigators since

The associate editor coordinating the review of this manuscript and approving it for publication was Abdullah Iliyasu.

they are classified and regulated. In order to construct an accurate forecasting price, it is necessary to evaluate several input combinations. By improving the accuracy of power price forecasting, the adverse effects of price instability can be prevented, the system can be stabilized, and economic gains can be realized [2], [5]. Although EPF has a degree of regularity, it also has a range of factors that affect electricity price instabilities, such as historical data price, market design, weather conditions, demand and supply balancing, and participants' bidding strategies. Furthermore, power market decision-making strongly depends on electricity prices, making the forecasting price an essential element for organized and effective electricity market operation. An accurate EPF has many advantages for power consumers and producers to make proper decisions in the market environment; for instance, it can be utilized to optimize electricity storage and reduce energy consumption during peak times. As a result, developing an accurate model to forecast time series electricity prices is quite challenging [6], [7]. During the last few years, many researchers have concentrated on the various machine learning technique's effects on modelling and predicting the price of electricity, particularly in the worldwide market. Typically, two machine learning methods are mainly utilized, the first for energy systems and the latter for EPF. However, the majority of recent techniques employ a variety of deep learning models, such as Deep Neural Networks (DNNs) and Long Short Term Memory (LSTM) [8], [9]. As well as machine learning approaches (e.g., Random Forest (RF), Support Vector Machine (SVM) [10], [11], Naive Bayes, and Extra tree [12], [13]). Additionally, statistical models such as Autoregressive Moving Average (ARMA), Seasonal Autoregressive Integrated Moving Average (SARIMA), and Generalized Autoregressive Conditional Heteroscedasticity (GARCH)) are applied for forecasting electricity price [14], [15], [16].

Various Deep Learning (DL) models have been applied to EPF; for instance, [1] proposed the hourly day-ahead EPF technique LSTM by using a dataset from Pennsylvania, New Jersey, and Maryland (PJM). However, this study is focused on solving network structure and hyperparameters selection of the LSTM model by employing the heterogeneous LSTM model. The sequence Model-Based Optimization (SMBO) algorithm is applied to adjust the hyperparameters of LSTM. Authors in [17] carried out various phases for EPF by utilizing a Deep Neural Network (DNN) with stacked pruning sparse denoising autoencoder to eliminate the noise of the dataset with various supplies. The result of this study indicates improvement in terms of the accuracy of electricity price. However, authors in [18] proposed a hybrid method based on Catboost technique with Bidirectional Long-Short Term Memory (BDLSTM) to forecast electricity prices on a Nord Pool market dataset. The proposed approach shows better performance measurement than other models, such as SVM, LSTM, and Multilayer Perceptron (MLP). Another hybrid technique proposed in [19] as a framework named Adaptive Neuro-Fuzzy Inference System (ANFIS) and ARMA for EPF using Spanish electricity market. Ordered Weighted Average (OWA) is utilized to combine three models to obtain an individual price forecasting model. Another method in [20] has been addressed with a new combination model based on Gated Recurrent Unit (GRU) and LSTM to predict day-ahead EPF on the Turkish data market. Their results highlighted the model's performance which outperformed various neural network structures. Authors in [21] Proposed hybrid models including Convolutional Neural Network (CNN), GRU, and Variational Mode Decomposition (VMD), named SEPNet, using the New York electricity dataset. Comparison results show that CNN and VMD-CNN have better performance over other techniques. Similarly, [14] proposed a DL-based hybrid model for the day-ahead EPF using the PJM data market with three forecasting models combining Deep Belief Network (DBN), LSTM, Recurrent Neural Network (RNN), and CNN to predict EPF. Empirical results show that the proposed method has valuable benefits compared with benchmark techniques. Authors in [22] introduced various neural network structures for forecast electricity prices using the HUPX market dataset compared with the CNN model. The result shows that combining fully connected layers and RNN is best for electricity price prediction. Furthermore, [23] suggested wavelet transforms, and Adam optimized LSTM neural network (WT-Adam-LSTM). This study offered various scenarios to validate the proposed method using datasets from Australia, New South Wales, and France to evaluate the hybrid method's effectiveness. Results indicate that the suggested approach can enhance the accuracy of prediction. In the same context, [16] employed hybrid models SVM, Empirical Wavelet Transform (EWT), and Bi-directional long short-term memory (BiLSTM) to EPF using European Power Exchange Spot (EPEXSPOT). Statistical results show that the proposed hybrid model performs better than other models. Authors in [24] proposed a hybrid approach known as EPNet, which combines CNN with LSTM; this technique is predicted one hour ahead based on the previous 24 hours of the price using the PJM dataset. The performance of their model shows that EPNet outperformed other algorithms in terms of the lowest performances.

Machine Learning (ML) models have been used to predict electricity prices on different horizons; for instance, Extreme Gradient Boosting (XGBoost) was proposed by [25] on a dataset provided by the Independent Electricity System Operator (IESO) to forecast electricity price. Simulation results indicate that the proposed model predicts better electricity prices than benchmark approaches such as SVM and RF. Likewise, [26] predicted hourly EPF by applying a new learning method based on a generalized Extreme Learning Machine (GELM) using Ontario and Australia data electricity markets. The proposed approach is computationally intensive and provides indeterminate outcomes for substantial datasets. Another hybrid method proposed in [27] by combining the SVM and Kernel Principal Component Analysis (KPCA); however, this approach is concluded with a threshold value. Due to seasonally changing prices in various locations, the proposed model in this study is highly static and location dependent. On the other hand, the authors in [28] introduced a hybrid method to forecast electricity prices applied to historical data from New England. Relevance Vector Machines (RVMs) have been used for individual prediction, and the individual model is combined to formulate linear regression. Results show the hybrid method performs better than the other tested methods. However, the authors in [29] turned their attention to Feature Selection (FS) instead of simplistic training to EPF using PJM regulation zone data, Spain, and the New York electric center. They have used Information Gain (IG) and Mutual Information (MI) as techniques to implement FS however the proposed model is limited to online prediction. In contrast, authors in [30] developed a bi-level ML strategy for EPF, where Linear Regression (LR) is used in the first level to anticipate the wholesale U.S. market's price. In addition, the second level of this study contains a limited optimal model that receives the price signal and then provides feedback on the optimal dispatch solutions to the LR model for EPF of the subsequent level.

Statistical techniques have been proposed to predict electricity prices over various datasets. According to work presented in [31], a functional version of the ARMAHX model based on Hilbert operators is proposed; however, this approach is utilized to assess the Moving Average (MA) terms in practical time series models. The proposed technique is validated using the German and Spanish electricity price market, and the result is compared with other models. Conversely, [32] introduced day-ahead EPF using new integration models ARMAX, Improved Empirical Mode Decomposition (IEMD), ANFIS, and EGARCH in Australian and Spanish data markets. The results reveal that the new integrated model's predicting accuracy exceeds well-known models. Furthermore, [33] focused their study on EPF by suggesting a combination model of ARIMA with another prediction method to enhance residual errors in forecasting the hourly price of the Iberian electricity market. Authors in [34] proposed a hybrid model based on VMD, SARIMA, Simulated Annealing Particle Swarm Optimization (SAPSO), and DBN to predict electricity price using datasets from PJM, Australian, and Spanish markets; however, SAPSO is utilized to optimize DBN and hence can capture irregular variations of electricity price. The proposed model performs better against other models for EPF.

Table 1 summarizes the differences between recent works of forecasting electricity prices in various aspects involving forecasting models, time horizon forecasting, datasets employed, key findings and limitations. Most previous works have employed different ML techniques to forecast electricity prices, such as statistical and conventional models [22], [25]. The time-series electricity price dataset's characteristics include high volatility, rapid spike, and seasonality, making them complicated to forecast using individual techniques such as XGBoost and ANN models. These techniques may provide unsatisfactory results with large forecasting residuals between actual and forecast values. Additionally, they are inefficient at identifying nonlinear time-series behavior and have weak forecasting abilities. Furthermore, several hybrid techniques have been applied to forecast electricity prices in the prior works. However, most researchers focus on combing linear methods with deep learning techniques [18], [20], [24]. Deep learning techniques has complex architecture and consumes large computational resources. In light of this, linear regression with ensemble tree-based models is proposed to achieve better performance in EPF. The present work utilized linear regression models such as Automatic Relevance Determination (ARD) and Ridge combined with ensemble tree base models, including bagging Extra Tree Regression (ETR), Random Forest Regression (RFR) and boosting AdaBoost (ADA). Meanwhile, the real-world Nord Pool electricity market dataset is used to evaluate the proposed method because it is one of the most volatile, seasonal, and rapid spikes electricity markets [35]. Thus, the electricity price utilized in this study can extensively assess the efficiency and applicability of the newly proposed forecasting model. To the authors' knowledge, no other linear technique and tree based ensemble hybrid method in the previous work was used in short-term forecasting electricity prices. The main contributions of this study can be summarized as follows:

1) Development of a new hybrid ML model to forecast day-ahead electricity price using real Nord Pool spot electricity data market to cope with issues of time series data such as volatility and irregular spikes.
2) Integration of linear and tree based ensemble method to resolve complex and nonlinear electricity price fluctuations.
3) Comprehensive comparisons in forecasting accuracy between the proposed method and various traditional ML techniques.

The remaining structure of this paper is organized as follows. Section II describes the general background, including ML techniques and evaluation metrics. The methodology of this study is explained in section III. The forecasting results and dataset exploration are presented in section IV. Section V provides the conclusion of this work.

II. BACKGROUND

This section provides a briefly background, involving the techniques applied in this study and assessment measures for electricity price forecasting.

A. MACHINE LEARNING TECHNIQUES

The primary aim of ML is to create a structure that can learn without explicit programming from the experience. Moreover, ML algorithms are divided into supervised learning and unsupervised learning. In supervised learning, models are trained using a labeled data including the required output. While unsupervised learning does not include output variable which mean the data is not labeled [36].

Moreover, this study evaluates supervised algorithms and classifies them into two phases. The linear regression with

TABLE 1. Highlight the related works of EPF.

References	Single Model	Hybrid Model	Time-Horizon	Dataset Used	Key Findings and Limitations
[25]	XGboost	___	Short term	Ontario	ML technique is employed to increase electricity rates to offload data storage and reduce energy consumption in cloud data centers. However, this study has proposed only a single technique to predict EPF for little samples of time series data. Various hybrid model can be investigated instead of one type.
[29]	FS	___	Short term	PJM, Spanish, New York	The FS technique constructs applicant features' redundancy, relevancy, and interactions. This work has addressed feature selection instead of simplistic training. After testing, this study obtained an MAE of 4.09, but the model can only apply for offline prediction on a vast dataset.
[22]	Dense-LSTM	___	Short term	HUPX	The growing number of market and grid participants based on frequent changes of market conditions, ANN is becoming essential for predictive processes. Dataset used has more features which might impact the prediction performance.
[20]	___	GRU-LSTM	Short term	Turkish	Analysis of RNNs for EPF with emphasize performance of GRUs. However, experimental results validated only for one-day ahead forecasting. Further, the results not constant and can diverge with various seasons, which make the proposed model insufficient.
[27]	___	SVM-KPCA	Short term	ISO, new England	Extracting new features with minimal redundancy which improves SVM results. Due to the use of a large dataset, which involves the cost price of wood, wind, and oil, a significant computation overhead is imposed, contributing to the proposed model's inefficiency.
[24]	___	EPNet	Short term	PJM	The CNN-LSTM (EPNet) model performed better than traditional ML. However, CNN is substantially slower due to an operation such as a max pool. The proposed EPNet model provides huge error rates in real time forecasting with enormous computational complexity.
[17]	___	DNN-SPSDAE	Short term	Australian	The proposed approach could be used to analyses a large amount of input data by separating out the essential elements.
[34]	___	VMD-SAPSO-DBN	Short term	Australian, PJM, Spanish	Although the proposed method's effectiveness can enhance forecasting accuracy, it also has some limitations in terms of computation time and complexity.
[32]	___	ANFIS, ARMAX, EGARCH	Short term	Australian, Spanish	This study demonstrated that the proposed model can drastically improve the day ahead EPF's performance. However, ANFIS, could be highly complex structure and statistical model has lake due to limited capture of nonlinear behavior of electricity changes.
[14]	___	DBN, RNN, LSTM, CNN	Short term	PJM	Deep Learning hybrid framework, anticipated for the following day EPF has considerable bidding potential among market participants and uncertainty risk control. DNNs requires a massive amount of data in order to achieve better results than traditional MLs techniques.
[18]	___	BDLSTM	Short term	Nord Pool	The categorial feature of the BDLSTM proposed model is managed more efficiently and superior to other methods. Despite of effectiveness of the proposed method, it requires enormous time to training and forecasting.
[19]	ANFIS, ANN, ARIMA	___	Short term	Spanish	Combining forecast models to obtain one successful framework with accurate EPF. This work utilized statistical and deep learning models using weekly season data (limited data train).

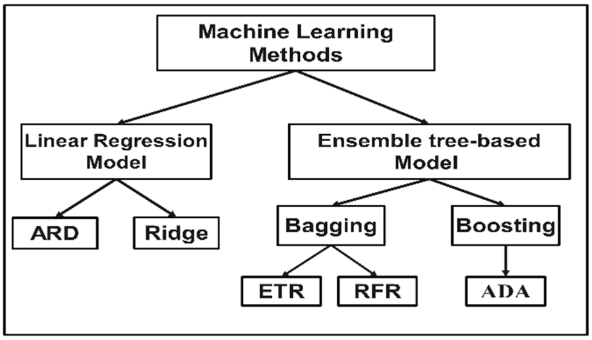

FIGURE 1. Machine learning methods used in this study.

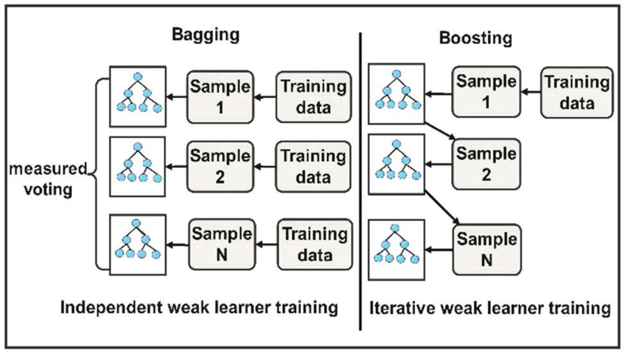

FIGURE 2. Workflows of boosting and bagging technique.

two models, Ridge, ARD, and ensemble tree-based models including bagging ETR and RFR models and boosting model ADA as indicated in Fig. 1. The following subsections describe each of them briefly.

1) LINEAR REGRESSION MODEL

Linear regression model is a supervised ML approach in which the model determines the linear connection between dependent and independent variables. However, LR can be divided into simple linear with a single independent variable and linear with several independent variables, and the model must identify a linear relationship. The LR analysis model is formulated as follows:

$$Y = \beta_0 + \beta_1 X_1 \ldots\ldots\ldots + \beta_i X_i + \varepsilon \tag{1}$$

where, X_i represents the independent variable and Y denotes the dependent variable. Whereas β_i estimated slop coefficient and β_0 is the intercept and i denotes samples of data for multiple LR. The random error component is defined as ε. The unobservable error component accounts for the failure of data to lie on the straight line and represents the difference between the true and observed realization of Y. Moreover, Ridge and ARD models have been broadly used as LR techniques that utilize the least square to suit a LR model employing a method that regularizes the measurement estimator to zero. Ridge is constructed to suit several regression models when the independent variables demonstrate multicollinearity. Multicollinearity refers to the scenario in which the X variables are associated with one another, which commonly results in erroneous estimations of regression model coefficients when ordinary least squares are used. The ARD is a new learning technique that can learn the relevance of the components and then delete unnecessary components to eliminate overfitting [37], [38].

2) ENSEMBLE TREE-BASED MODEL

The ensemble technique is a popular paradigm of the ML approach that combines a collection of learners instead of employing individual learners to forecast unidentified targets. Each learner's output values are combined using a voting process to forecast the final class label. The fundamental objective of ensemble learning is to create a more accurate classifier composed of several learners. Meanwhile, various techniques (bagging and boosting) have been established and implemented in experimental data and compared. Bagging generates several bootstraps from the training dataset, and a unique prediction pattern is developed for each bootstrap as shown in Fig. 2. Therefore, bagging can enhance the stability of models by reducing variance and avoiding the issue of overfitting. Random Forest (RF) has recently gained popularity as a bagging approach. In addition, during training, RF generates many decision trees using regression techniques and calculates an average prediction as an output. Similar to RF, Extra Tree (ET) combines various decision trees during training and generates a mean forecast as an output. The critical difference is that each decision tree uses the whole training subset and the splitting of the decision tree is random [39]. Boosting is another technique which several classifiers are built from primary samples, and weak classifiers are combined to build robust classifiers. Furthermore, ADA has widely used as a boosting ML technique that combines numerous powerless learners into a particular classifier through weighted linear combination. ADA uses a learning system progressively to reweight samples of the original training data.

Consequently, the risk of selecting misclassified samples for the training set rises, and a greater proportion of instances are correctly classified [40], [41]. However, boosting is a continual process of constructing classifiers enhanced by the weights of weak classifiers from previous rounds, which contributes to reducing dataset volatility and variability.

B. PERFORMANCE INDICES OF MODEL EVALUATION

Various criteria are utilized to evaluate the proposed regression model's performance, as discussed in [42]. Moreover, three statistical indices are employed to forecast the performance of the electricity price, comprising Mean Square Error (MSE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). The calculation formula of these indices is given as follows:

$$MAE = \frac{1}{N} \int_{i=1}^{N} |X_i - y_i| \tag{2}$$

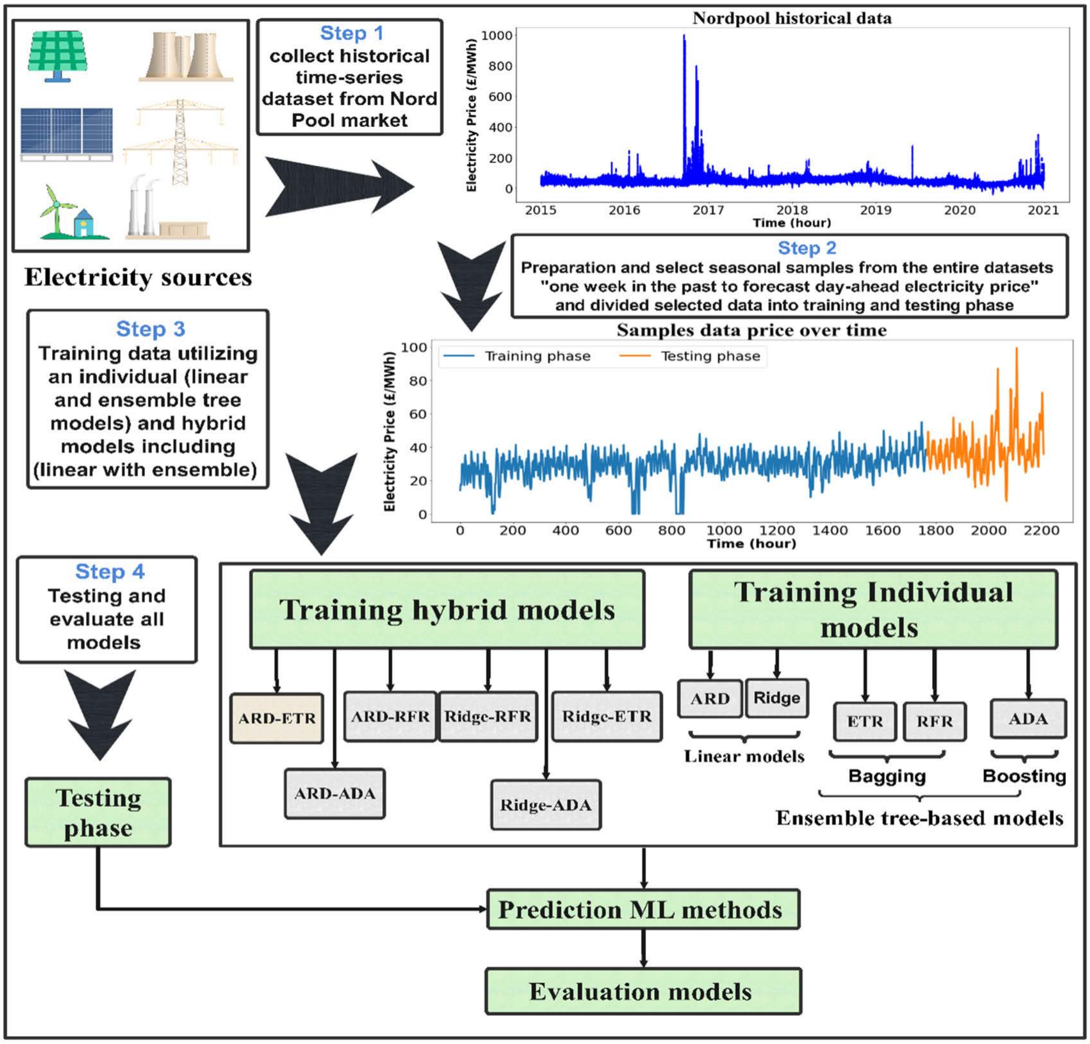

FIGURE 3. The overall flowchart of forecast day-ahead electricity price applied in this study.

$$MSE = \frac{1}{N}\int_{i=1}^{N}(X_i - y_i)^2 \tag{3}$$

$$RMSE = \sqrt{\frac{1}{N}\int_{i=1}^{N}(X_i - y_i)^2} \tag{4}$$

where X_i denotes the actual value, y_i denotes the predicted electricity price value at time i and N represents the number of testing samples. In general, lesser MSE, MAE, and RMSE scores demonstrate accurate prediction, which arises when the predicted value, $y_i$is close to the actual value X_i.

III. METHODOLOGY

This section describes the implementation of the proposed forecasting methods to predict day-ahead electricity price by adopting individual and hybrid models as shown in Fig. 3. Firstly, time-series data is collected every hour from the Nord Pool spot market, the European power exchange transmission system [43]. Secondly, data selection is performed for the training and testing process. In the training process, several ML models including single and hybrid are applied for comparison purposes. In order to evaluate the performance of each ML model, a set of the test dataset is applied. Then obtain final predication model. The analysis of each forecasting model is briefly discussed in section IV.

A. TIME SERIES DATA DESCRIPTION

A time series is a collection or sequence of observable data organized chronologically and in equally spaced periods, such as daily or hourly. However, one of the essential uses of time series analysis is the creation of suitable models to predict future events based on known past events for observed time series and the subsequent use of these models to perform accurate time series forecasting. Furthermore, time series data can be visualized and analyzed to find the most effective component, for example a trend that describes the observation of downwards or upwards patterns over an extended

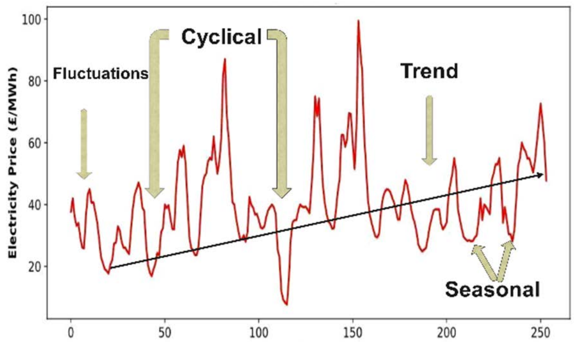

FIGURE 4. The time-series data components.

period. Seasonality variations occur regularly; for instance, electricity consumption is high throughout the day and low at night. The cyclical component is considered over the long-term prospect. Additionally, irregular effects impact of any random events [44]. These time-series data components are exhibited in Fig. 4. The historical time series dataset used in this study has only measurement price information and it has approximately 2200 instances collected every one hour. The description of the dataset including time length and data preparation is discussed in the next section.

B. THE PROPOSED HYBRID FORECASTING METHOD

ARD is a linear model which highly comparable to Bayesian Ridge Regression. However, the ARD regression model leads to a sparser coefficient which poses a different prior; it also drops the spherical Gaussian distribution for centered elliptic Gaussian distribution. This means each coefficient can be drowned from Gaussian distribution, centered on zero and a precision. Conversely, Bayesian Ridge Regression has its standard deviation. ARD allows the selection of relevant features, which should prevent the overfitting of models. Using all the benefits of the ARD, it is theoretically possible to enhance the accuracy of short-term forecasting of electricity price time series [45].

The ensemble bagging ETR is an ML technique that extends the RF algorithm and is less prone to overfit a dataset. ETR employs a similar framework as RF and uses a random selection of features for training each base estimator. However, it randomly chooses the optimal feature and value for splitting the node. ETR trains each regression tree using the complete training dataset. In contrast, RF trains the model with a bootstrap replica [39]. Moreover, ETR enhances the model's functionality, reduces errors, and forecasts spikes by learning from interactions to produce an accurate prediction result. The organization structure of ETR model is shown in Fig. 5. Moreover, ML linear model ARD and tree-based bagging model ETR are combined to forecast day-ahead electricity price ARD-ETR. The reason behind selecting this type of ML model is that ARD is developed to capture the general trends and seasonality. The ETR improves the model's performance by reducing errors and predicting spikes

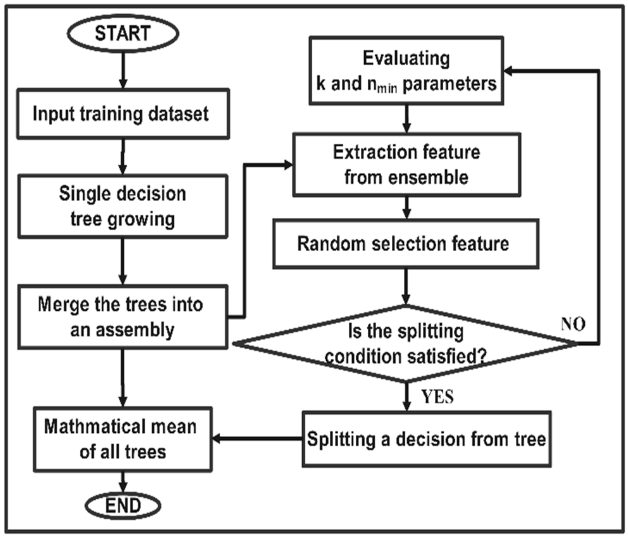

FIGURE 5. The flow chart of ETR algorithm.

by learning from interactions to produce accurate forecasting. Hence, it is justified that combining linear regression with ensemble tree model can provide a more accurate forecast and overcome the shortcomings of a particular model.

During the training phase, the parameters of model i were optimized to minimize a loss function L_i expressed as:

$$L_i = (X_i - y_i)^2 \tag{5}$$

where X_i denotes the actual outcome from model i and y_i is the predicted one. ML models are not ideal. For instance, the ARD prediction deviates from the actual value $ARD_{predict} = X_i + \varepsilon_1$. A second model ETR was trained to forecast the residuals ε_1 by reducing loss to decrease the deviation ε_1. The loss is defined as follows:

$$L_2 = (ETR_{predict} - \varepsilon_1)^2 \tag{6}$$

where $\varepsilon_1 = ARD_{predict} - X_1$. The final EPF of prediction of the hybrid model is expresses as:

$$\begin{aligned} EPF &= ARD_{predict} - ETR_{predict} = (X_1 + \varepsilon_1) - (\varepsilon_1 + \varepsilon_2) \\ &= X_1 + \varepsilon_2 \end{aligned} \tag{7}$$

The experimental results, in agreement with ensemble learning, showed that the hybrid error $\varepsilon_2$is less than ε_1. Hence, time-series difficulties could be eliminated using the robustness of these models to forecast electricity price. The flowchart in Fig. 6 illustrates the proposed hybrid method. Detailed steps are given as follows:

1) Samples of historical time series datasets with an hourly time step were collected from the Nord Pool spot electricity market price.
2) Past week values were selected to forecast day-ahead electricity price by dividing the data samples into 80% as training to establish the forecasting model, while the remining 20% as testing for the evaluation model.

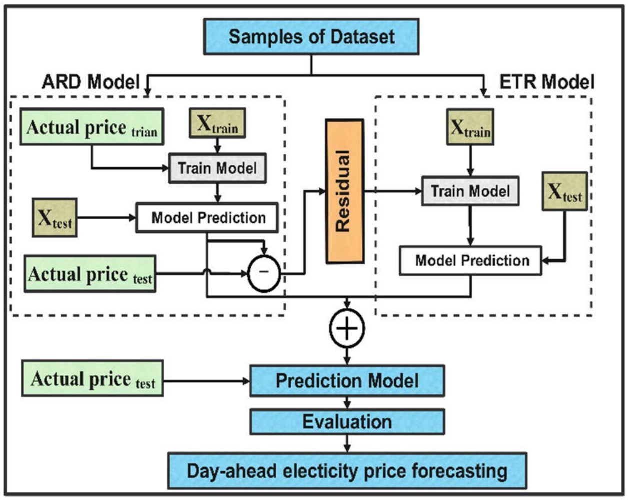

FIGURE 6. Flow chart of the proposed hybrid method.

3) Final EPF is predicted form the trained ARD-ETR model.
4) Comparison and evaluation of the proposed method and other individual and hybrid methods utilized in this work.

IV. RESULTS AND DISCUSSION

This section describes the data exploration and the experimental results. Moreover, this section benchmarks the proposed model with several state-of-the-art models to justify the robustness of the proposed work. Comparative analysis is concisely explained at the end of this section.

A. DATA EXPLORATION AND EXPERIMENTAL SETUP

The proposed forecasting model was applied in this study utilizing a real electricity price dataset collected from the Nord Pool spot electricity market [43]. Each day on the Nord Pool spot market, there are 24 hourly observations separated by one hour. Furthermore, historical time series data covers a period from 2015-2021, which combine into a single CSV file. Fig. 7 shows the data as a time series to provide an overview of the entire dataset.

Moreover, the essential statistics characteristics of the entire time series data show that the mean, maximum, minimum, and standard deviation values are 43.6502, 999, −38.8, and 18.95, respectively. As demonstrated in Fig. 7, the entire dataset has irregular changes, followed by trend and seasonality variations. It can be seen that the price fluctuated significantly and suffered from price spikes. Moreover, the maximum price approximately reaches above 990 £/MWh. Conversely, the dataset has some negative values that generally occur when the supply offered exceeds the demand; these actions usually happen in the middle of the day, when generators (i.e., wind, large-scale solar, and coal-fired generators) compete for energy dispatch. Even if the whole dataset has only price measurement, Fig. 7. depicts specific high price

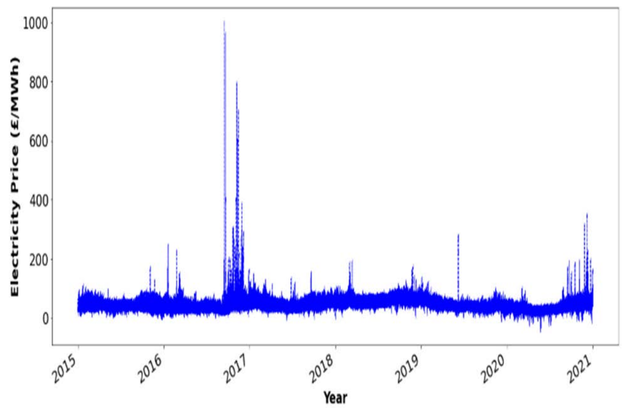

FIGURE 7. Historical datasets as a time series.

points in various periods. In order to entirely complete the performance test in this study, samples from different seasons over the entire dataset have been selected and used in price forecasting. Moreover, as mentioned in Section III, a time-series dataset is a chronological sequence of observations recorded at regular intervals. Besides, supervised learning includes (x) as input patterns and (y) as output patterns, hence, the algorithm can learn how to forecast the (y) from the (x). Thus, time-series dataset requires reframing into supervised learning by shifting the data into the past to predict the value in the future. In this study, weeks' values in the past are selected to predict one day- ahead.

The entire selected dataset is split into two segments; 80% training phase is applied to develop the prediction model, while the remining 20% testing phase is utilized for evaluation, as illustrated in Fig. 8. However, the time series data has irregular fluctuation over some period of time. This might occur due to the combination of surging energy demand, fuel supply disruptions, and global shortages of oil, gas, and coal that have affect global energy prices [25].

Sliding validation for time-series forecasts to fine-tune the model has been carried out during the training phase, where the algorithms are continuously trained K times [46]. In this instance, a K value of 5 is selected as indicated in Fig.8. Since sequential samples are correlated in time series, a standard train/test split that assumes the sample's independence does not make sense. Instead, sliding window validation allows testing of predictive performance at various correlated time steps. This mimics a possible utilized model in which the model is retrained as more data is collected, followed by a prediction of the electricity price ahead. All experiments in this study are executed five times and performed using python 3.8.5 environment on 1.8 GHz PC i7-8565U, 8GB memory and, NVIDIA GeForce MX150.

Furthermore, effectively forecasting day-ahead electricity prices has many uses for power producers and consumers to make proper decisions in a market-oriented environment; firstly, it can be used to optimize electricity storage. Secondly, it enables demand-side flexibility to reduce consumption in

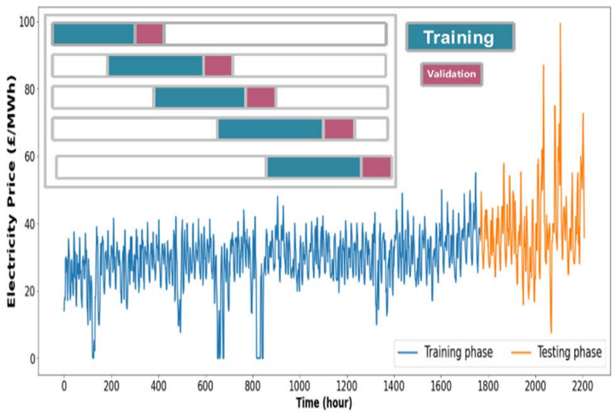

FIGURE 8. Training and testing phase as a time series dataset.

on-peak times. And last but not least, it can facilitate maximizing economic benefits and reducing risks to the power market [35].

B. THE FORECASTING RESULTS

To assess the performance of the proposed hybrid ML method, a comparison of all ML forecasting methods, including individual and hybrid models, is conducted. Table 2 and Table 3 elaborate on the statistical forecasting results for individual and hybrid regression models. As shown in Table 2, the ensemble bagging ETR regression model exhibited the highest testing outcomes in terms of MAE, RMSE, and MSE of 2.99, 4.36, and 19.03, respectively, followed by another bagging model RFR with an MAE value of 3.43, and RMSE, MSE of 4.94, 24.37. Then linear model such as the ridge model indicates that MAE performs better measurement than RMSE, and MSE with 4.06, 6.24, and 38.95, respectively, while ARD implies that MAE, RMSE, and MSE are marginally inferior to the ridge model with 4.1, 6.26, and 39.24 respectively. However, in terms of RMSE, MAE, and MSE rankings, ADA scored the worst compared to other ML models. The experiment results indicate that the LR model performs better than the ensemble boosting method in forecasting measurement. The reason is that the LR performs remarkably well for linearly separable time-series datasets and manages overfitting effectively by using sliding validation and dimensionality reduction techniques. On the other hand, the tree-based bagging models achieved better results due to ease of implementation, handles overfitting, and reduces the variance within the learning algorithm, and increase model's accuracy.

Moreover, Table 3 elaborates the performance accuracy measurement of the proposed ARD-ETR compared to ARD-RFR, ARD-ADA, Ridge-ETR, Ridge-RFR, and Ridge-ADA methods of day-ahead EPF in terms of MSE, MAE, and RMSE with minimum value of 16.7, 2.03, and 3.09 respectively. An individual linear model with an ensemble bagging method has the best integration model to forecast electricity prices. The reason is that the linear model can manage

TABLE 2. Statistical forecasting measurement of individual model.

Model	MSE (£/MWh)	RMSE (£/MWh)	MAE (£/MWh)
RFR	24.37	4.94	3.43
ARD	39.24	6.26	4.1
Ridge	38.95	6.24	4.06
ETR	19.03	4.36	2.99
ADA	45.67	6.76	5.16

TABLE 3. Statistical forecasting measurement of hybrid model.

Model	MSE (£/MWh)	RMSE (£/MWh)	MAE (£/MWh)
ARD-RFR	24.53	4.95	3.33
Ridge-RFR	24.11	4.91	3.30
Ridge-ETR	17.24	4.15	2.83
ARD-ADA	38.32	6.19	4.85
Ridge-ADA	37.71	6.14	4.81
the proposed ARD-ETR	**16.7**	**3.09**	**2.03**

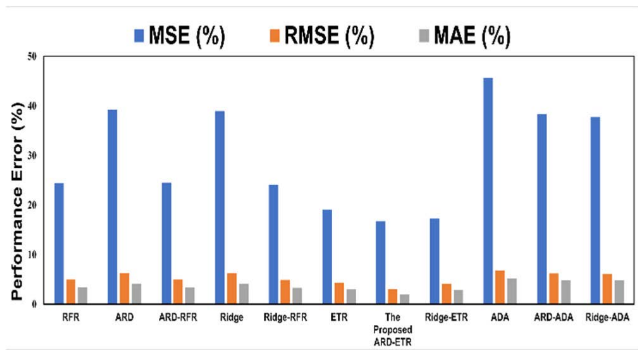

FIGURE 9. Forecasting results of the proposed hybrid model with other methods.

time-series issues such as trend and seasonality cycles. Besides, it has extraordinarily linear separable datasets with remarkably managed overfitting. In contrast, the ensemble bagging model tackles the irregular fluctuation of price over time and can effectively learn from interactions with low variance. The forecasting accuracy of ML methods compared to the proposed hybrid model is visualized and presented in Fig. 9.

It can be noticed that the proposed hybrid method ARD-ETR is much higher than ML models and other hybrid approaches with minor testing MSE, RMSE, and MAE. As a result, the proposed hybrid method has been demonstrated to be a successful strategy for predicting day-ahead electricity price forecasting. In order to explain the distinction between forecasting curves generated by various hybrid methods, 24 forecasted samples of different days of the forecasted data are plotted to verify the proposed hybrid model. Fig. 10

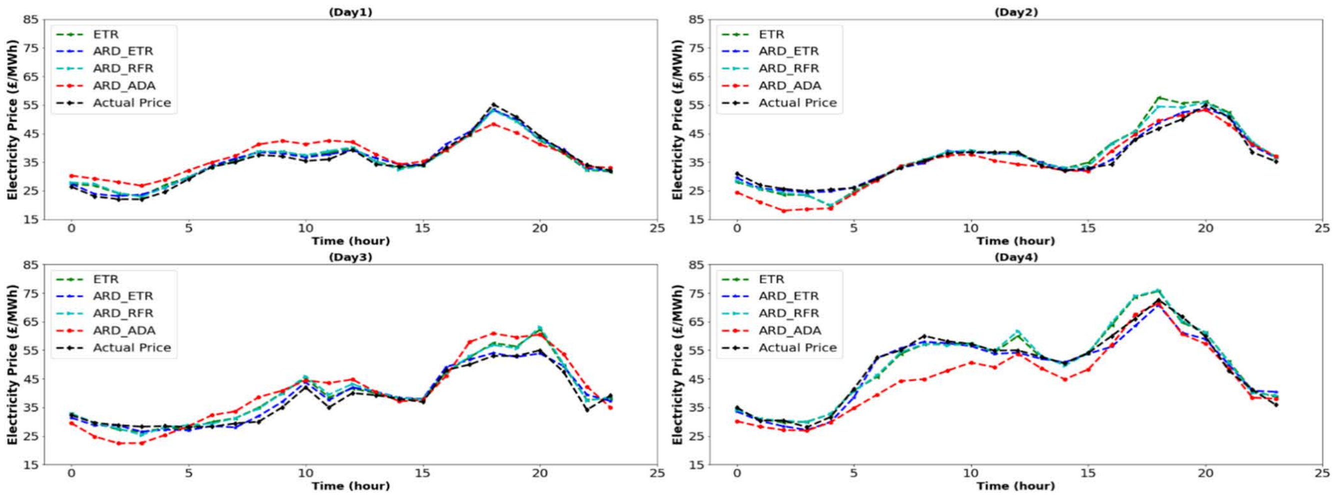

FIGURE 10. Forecasting results of proposed hybrid model using different methods (different test data: day (1), day (2), day (3), day (4)).

depicts the forecasting results of individual ETR and other hybrid models. It can be noticed that ARD-ADA performance is weaker in EPF; it is powerless to grasp the actual electricity price, resulting in high predicted measurement. ETR, has more efficient forecasting results than the hybrid ARD-RFR model; it can grasp the actual data price with fewer errors. However, the forecasted results of the proposed hybrid ARD-ETR method are incredibly close to the actual electricity price curve in contrast to other ML approaches. The following index is presented to evaluate the improvement of the proposed method [34]:

$$P_{index} = \frac{N_p - N_o}{N_o} \tag{8}$$

where P_{index} index denotes the assessment index, N_o represents the error value of other models and N_p is the residual error of the proposed model.

Table 4 displays the evaluation outcomes of various ML approaches where the hybrid model proposed has greatly improved.

For instance, by comparing with RFR, ARD, ARD-RFR, Ridge, Ridge-RFR, ETR, Ridge-ETR, ADA, ARD-ADA, and Ridge-ADA, we find that the RMSE of the proposed method has been reduced by 37.44%, 50.63%, 37.57%, 50.48%, 37.06%, 29.12%, 25.54%, 54.28%, and 50.08%, the MAE of the proposed model has been reduced by 40.81%, 50.48%, 39.03%, 50%, 38.48%, 32.1%, 28.26%, 60.65%, 58.14%, and 57.79%, while the MSE of the proposed model has been diminished by 31.47%, 57.44%, 31.92%, 57.12%, 30.73%, 12.24%, 3.13%, 63.43%, 56.41%, and 55.71%. This result shows that the proposed hybrid method can perfectly capture time-series data price difficulties. Hence, the proposed method can produce more accurate EPF.

C. COMPARATIVE ANALYSIS

This study proposes the ARD-ETR method for day-ahead EPF compared with other techniques. Moreover, hourly time series data is obtained from Nord Pool electricity market. Empirical results indicate that the proposed hybrid strategy has outperformed other methods in terms of performance with the lowest MAE, RMSE, and MSE values compared to other approaches. Fig. 11 demonstrates the reduction in testing MAE and RMSE by four hybrids algorithms relative to the created ensemble based ETR model. In addition, the proposed hybrid approach ARD-ETR has provided a more significant decrease in testing MAE with (32.1) and RMSE with (29.12) compared to other hybrid methods. The comparison forecasting results of the proposed hybrid ARD-ETR technique with the results of prior works are listed in Table 5. The proposed ARD-ETR model attained the lowest MAE and RMSE values with 2.03 (£/MWh) and 3.09 (£/MWh) respectively. Two testing metrics, RMSE and MAE values of the XGboost model in [25], are utilized on Ontario dataset between November and December with 9.25 (£/MWh) and 3.74 (£/MWh), respectively. In contrast, the EPNet method

TABLE 4. Performance evaluation results of different ML models (%).

Model	P_{RMSE}	P_{MAE}	P_{MSE}
RFR	37.44	40.81	31.47
ARD	50.63	50.48	57.44
ARD-RFR	37.57	39.03	31.92
Ridge	50.48	50	57.12
Ridge-RFR	37.06	38.48	30.73
ETR	29.12	32.1	12.24
Ridge-ETR	25.54	28.26	3.13
ADA	54.28	60.65	63.43
ARD-ADA	50.08	58.14	56.41
Ridge-ADA	49.67	57.79	55.71

TABLE 5. Comparison of the forecasting results of the proposed hybrid method to previous works.

Forecasting Technique	Reference	Year	Testing Metrics (£/MWh)	Dataset Used
XGboost	[25]	2020	RMSE: 9.25 MAE: 3.74	Ontario
Dense-LSTM	[22]	2022	RMSE: --- MAE: 14.438	HUPX market
BDLSTM	[18]	2022	RMSE: 34.99 MAE: 22.186	Nord Pool
FS	[29]	2017	RMSE: 18.9 MAE: 4.09	PJM, Spanish, New York
GRU-LSTM	[20]	2018	RMSE: 11.99 MAE: 5.71	Turkish
SVM-KPCA	[27]	2019	RMSE: 10.21 MAE: 18.97	ISO, new England
EPNet	[24]	2018	RMSE: 14.2 MAE: 8.84	PJM
VMD-DBN	[34]	2020	RMSE: 3.28 MAE: 2.07	Australia, PJM, Spanish
SSA-DELM	[47]	2022	RMSE: 4.7 MAE: 3.8	Nordic market
SDR-MASES-SPSDAE	[17]	2021	RMSE: 4.12 MAE: 11.76	Australia
ARD-ETR	**Present study**	--	**RMSE: 3.09** **MAE: 2.03**	**Nord Pool market**

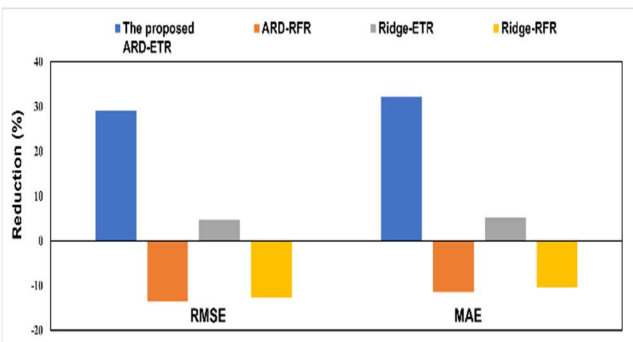

FIGURE 11. Forecasting results of the proposed hybrid model with other methods.

in [24] attained testing MAE value of 8.84 (£/MWh) and RMSE value of 14.2 (£/MWh). In addition, the obtained results in [18] using BDLSTM performs poorly results compared to the proposed ARD-ETR method using the Nord Pool data market with RMSE of 34.99 (£/MWh) and MAE of 22.186 (£/MWh). The proposed model is slightly superior to the testing results in [34]. Hence, the proposed hybrid method has demonstrated better results with lower testing error values for EPF compared to other approaches in other works as shown in Table 5.

V. CONCLUSION AND FUTURE WORK

In recent decades, various forecasting electricity price methods have been developed. Due to the difficult characteristics of electricity price, which include high volatility, rapid spike, and seasonality within various periods of samples, which can affect the prediction of short-term electricity prices, developing robust machine learning forecasting tools becomes cumbersome. A new hybrid ML technique is proposed in this paper to forecast day-ahead electricity prices based on linear regression and ensemble tree bagging method ARD-ETR. The electricity price dataset is collected as a time series from the Nord Pool spot market, which is utilized to validate the efficiency of the proposed method. Moreover, the proposed method can deal with the time series characteristics difficulties by utilizing linear regression model ARD with ensemble tree-based bagging model ETR. The historical dataset is converted into a supervised learning method by taking one week's value in the past and divided into the training phase to establish the forecasting model and the test phase for the evaluation model. Empirical results demonstrate that the proposed ARD-ETR hybrid technique has achieved the best performance in terms of MAE, RMSE, and MSE compared to the individual and other hybrid approaches used in this paper, where the results reveal that the ARD-ETR method attained the lowest MAE, RMSE, MSE values (£/MWh) with 2.03, 3.09 and 16.7 respectively. Further, the hybrid ARD-ETR method achieved the lowest MAE scores and RMSE compared to other methods utilized in the prior works for the day-ahead EPF. Thus, a newly proposed hybrid method shows better improvement evaluated with the benchmark's technique and a high reduction in testing MAE value (£/MWh) with 32.1 and RMSE value (£/MWh) with 29.12 against other hybrid models. Moreover, integrated linear ARD with bagging ensemble tree ETR model exhibited more robustness and practical to be applied in the day-ahead EPF. Further study with different regressors based on DL models such as stack autoencoders (SAEs) and LSTM can be employed to better forecasting results.

REFERENCES

[1] S. Zhou, L. Zhou, M. Mao, H.-M. Tai, and Y. Wan, "An optimized heterogeneous structure LSTM network for electricity price forecasting," *IEEE Access*, vol. 7, pp. 108161–108173, 2019.
[2] Z. Yang, L. Ce, and L. Lian, "Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods," *Appl. Energy*, vol. 190, pp. 291–305, Mar. 2017.
[3] U. Ugurlu, O. Tas, and U. Gunduz, "Performance of electricity price forecasting models: Evidence from Turkey," *Emerg. Markets Finance Trade*, vol. 54, no. 8, pp. 1720–1739, May 2018.
[4] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Appl. Energy*, vol. 221, pp. 386–405, Jul. 2018.
[5] A. Hayfavi and I. Talasli, "Stochastic multifactor modeling of spot electricity prices," *J. Comput. Appl. Math.*, vol. 259, pp. 434–442, Mar. 2014.
[6] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *Int. J. Forecasting*, vol. 30, no. 4, pp. 1030–1081, Dec. 2014.
[7] J. Nowotarski and R. Weron, "Recent advances in electricity price forecasting: A review of probabilistic forecasting," *Renew. Sustain. Energ Rev.*, vol. 81, no. 1, pp. 1548–1568, Jan. 2018.
[8] P. H. Kuo and C. J. Huang, "A high precision artificial neural networks model for short-term energy load forecasting," *Energies*, vol. 11, no. 1, p. 213, Jan. 2018.
[9] F. Ascione, N. Bianco, C. De Stasio, G. M. Mauro, and G. P. Vanoli, "Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: A novel approach," *Energy*, vol. 118, pp. 999–1017, Jan. 2017.
[10] H. Jiang, Y. Zhang, E. Muljadi, J. J. Zhang, and D. W. Gao, "A short-term and high-resolution distribution system load forecasting approach using support vector regression with hybrid parameters optimization," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3341–3350, Jul. 2018.
[11] J.-P. Liu and C.-L. Li, "The short-term power load forecasting based on sperm whale algorithm and wavelet least square support vector machine with DWT-IR for feature selection," *Sustainability*, vol. 9, no. 7, p. 1188, Jul. 2017.
[12] C. C. Yang, C. S. Soh, and V. V. Yap, "A non-intrusive appliance load monitoring for efficient energy consumption based on naive Bayes classifier," *Sustain. Comput., Informat. Syst.*, vol. 14, pp. 34–42, Jun. 2017.
[13] S. Ottesen, A. Tomasgard, and S.-E. Fleten, "Prosumer bidding and scheduling in electricity markets," *Energy*, vol. 94, pp. 828–843, Jan. 2016.
[14] R. Zhang, G. Li, and Z. Ma, "A deep learning based hybrid framework for day-ahead electricity price forecasting," *IEEE Access*, vol. 8, pp. 143423–143436, 2020.
[15] D. Bissing, M. T. Klein, R. A. Chinnathambi, D. F. Selvaraj, and P. Ranganathan, "A hybrid regression model for day-ahead energy price forecasting," *IEEE Access*, vol. 7, pp. 36833–36842, 2019.
[16] H. Cheng, X. Ding, W. Zhou, and R. Ding, "A hybrid electricity price forecasting model with Bayesian optimization for German energy exchange," *Int. J. Electr. Power Energy Syst.*, vol. 110, pp. 653–666, Sep. 2019.
[17] L. Sun, L. Li, B. Liu, and S. Saeedi, "A novel Day-ahead electricity price forecasting using multi-modal combined integration via stacked pruning sparse denoising auto encoder," *Energy Rep.*, vol. 7, pp. 2201–2213, Nov. 2021.
[18] F. Zhang, H. Fleyeh, and C. Bales, "A hybrid model based on bidirectional long short-term memory neural network and catboost for short-term electricity spot price forecasting," *J. Oper. Res. Soc.*, vol. 73, no. 2, pp. 301–325, Oct. 2022.
[19] A. Darudi, M. Bashari, and M. H. Javidi, "Electricity price forecasting using a new data fusion algorithm," *IET Gener., Transmiss. Distrib.*, vol. 9, no. 12, pp. 1382–1390, Sep. 2015.
[20] U. Ugurlu, I. Oksuz, and O. Tas, "Electricity price forecasting using recurrent neural networks," *Energies*, vol. 11, no. 5, p. 1255, May 2018.
[21] C. Huang, Y. Shen, Y. Chen, and H. Chen, "A novel hybrid deep neural network model for short-term electricity price forecasting," *Int. J. Energy Res.*, vol. 45, no. 2, pp. 2511–2532, Sep. 2021.
[22] M. Pavićević and T. Popović, "Forecasting day-ahead electricity metrics with artificial neural networks," *Sensors*, vol. 22, no. 3, p. 1051, Jan. 2022.
[23] Z. Chang, Y. Zhang, and W. Chen, "Electricity price prediction based on hybrid model of Adam optimized LSTM neural network and wavelet transform," *Energy*, vol. 187, Nov. 2019, Art. no. 115804.
[24] P.-H. Kuo and C.-J. Huang, "An electricity price forecasting model by hybrid structured deep neural networks," *Sustainability*, vol. 10, no. 4, p. 1280, Apr. 2018.
[25] S. Albahli, M. Shiraz, and N. Ayub, "Electricity price forecasting for cloud computing using an enhanced machine learning model," *IEEE Access*, vol. 8, pp. 200971–200981, 2020.
[26] M. Rafiei, T. Niknam, and M. H. Khooban, "Probabilistic forecasting of hourly electricity price by generalization of ELM for usage in improved wavelet neural network," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 71–79, Feb. 2017.
[27] K. Wang, C. Xu, Y. Zhang, S. Guo, and A. Y. Zomaya, "Robust big data analytics for electricity price forecasting in the smart grid," *IEEE Trans. Big Data*, vol. 5, no. 1, pp. 34–45, Mar. 2019.
[28] M. Alamaniotis, D. Bargiotas, N. G. Bourbakis, and L. H. Tsoukalas, "Genetic optimal regression of relevance vector machines for electricity pricing signal forecasting in smart grids," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2997–3005, Nov. 2015.
[29] O. Abedinia, N. Amjady, and H. Zareipour, "A new feature selection technique for load and price forecast of electrical power systems," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 62–74, Jan. 2017.
[30] Q. Zhang, K. Dehghanpour, Z. Wang, and Q. Huang, "A learning-based power management method for networked microgrids under incomplete information," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1193–1204, Mar. 2020.
[31] J. P. Gonzalez, A. M. S. Roque, and E. A. Perez, "Forecasting functional time series with a new Hilbertian ARMAX model: Application to electricity price forecasting," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 545–556, Jan. 2018.
[32] J.-L. Zhang, Y.-J. Zhang, D.-Z. Li, Z.-F. Tan, and J.-F. Ji, "Forecasting day-ahead electricity prices using a new integrated model," *Int. J. Electr. Power Energy Syst.*, vol. 105, pp. 541–548, Feb. 2019.
[33] R. A. Chinnathambi, A. Mukherjee, M. Campion, H. Salehfar, T. Hansen, J. Lin, and P. Ranganathan, "A multi-stage price forecasting model for day-ahead electricity markets," *Forecasting*, vol. 1, no. 1, pp. 26–46, Jul. 2018.
[34] J. Zhang, Z. Tan, and Y. Wei, "An adaptive hybrid model for short term electricity price forecasting," *Appl. Energy*, vol. 258, Jan. 2020, Art. no. 114087.
[35] A. N. Alkawaz, J. Kanesan, A. S. M. Khairuddin, C. O. Chow, and M. Singh, "Intelligent charging control of power aggregator for electric vehicles using optimal control," *Adv. Electr. Comput. Eng.*, vol. 21, no. 4, pp. 21–30, Nov. 2021.
[36] E. Esenogho, K. Djouani, and A. M. Kurien, "Integrating artificial intelligence Internet of Things and 5G for next-generation smartgrid: A survey of trends challenges and prospect," *IEEE Access*, vol. 10, pp. 4794–4831, 2022.
[37] V. Renkens and H. Van Hamme, "Automatic relevance determination for nonnegative dictionary learning in the Gamma-Poisson model," *Signal Process.*, vol. 132, pp. 121–133, Mar. 2017.
[38] B. Carrera and K. Kim, "Comparison analysis of machine learning techniques for photovoltaic prediction using weather sensor data," *Sensors*, vol. 20, no. 11, p. 3129, Jun. 2020.
[39] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Mach. Learn.*, vol. 63, no. 1, pp. 3–42, 2006.
[40] G.-W. Cha, H.-J. Moon, and Y.-C. Kim, "Comparison of random forest and gradient boosting machine models for predicting demolition waste based on small datasets and categorical variables," *Int. J. Environ. Res. Public Health*, vol. 18, no. 16, p. 8530, Aug. 2021.

[41] M. H. D. M. Ribeiro and L. Dos Santos Coelho, ''Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series,'' *Appl. Soft Comput.*, vol. 86, Jan. 2020, Art. no. 105837.

[42] A. Botchkarev, ''A new typology design of performance metrics to measure errors in machine learning regression algorithms,'' *Int. J. Info. Knowl. Manag.*, vol. 14, p. 45, Jan. 2019.

[43] *Nord Pool Historical Market Data Set*. Nord Pool. Accessed: Mar. 2022. [Online]. Available: https://www.nordpoolgroup.com/historical-market-data/

[44] M. Mudelsee, ''Trend analysis of climate time series: A review of methods,'' *Earth-Sci. Rev.*, vol. 190, pp. 310–322, Mar. 2019.

[45] D. Wipf and S. Nagarajan, ''A new view of automatic relevance determination,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 20, 2007, pp. 1–8.

[46] Y. Bengio and Y. Grandvalet, ''No unbiased estimator of the variance of K-fold cross-validation,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 16, 2003, pp. 1–8.

[47] Y. Zhang, P. Tao, X. Wu, C. Yang, G. Han, H. Zhou, and Y. Hu, ''Hourly electricity price prediction for electricity market with high proportion of wind and solar power,'' *Energies*, vol. 15, no. 4, p. 1345, Feb. 2022.

ALI NAJEM ALKAWAZ (Member, IEEE) received the Bachelor of Engineering (B.Eng.) and Master of Engineering Science (M.Eng.Sci.) degrees in control systems and mechatronics engineering from the University of Technology, Baghdad, Iraq, in 2014 and 2018, respectively. He is currently pursuing the Ph.D. degree in electrical engineering with the Universiti Malaya, Malaysia. His research interests include machine learning, control systems (robotics and automation), electric vehicles, green technology, and sustainable energy.

ABDALLAH ABDELLATIF (Member, IEEE) received the M.Sc. degree in control systems from the Universiti Tun Hussein Onn Malaysia, in 2019. He is currently pursuing the Ph.D. degree in control systems with the Department of Electrical Engineering, Universiti Malaya. His research interests include data mining, mainly working machine learning and deep learning.

JEEVAN KANESAN received the degree in electrical engineering from the UTM, in 1999, and the M.Sc. and Ph.D. degrees from the USM, Penang, in 2002 and 2006, respectively. He joined the Department of Electrical Engineering, Universiti Malaya, in 2008. He worked as an Engineer with Carsem Semiconductor, Ipoh. He worked as a Research and Development Engineer at Intel, Penang, for 2.5 years. He has published more than 70 peer-reviewed journals. His research interests include optimization and machine learning.

ANIS SALWA MOHD KHAIRUDDIN received the Bachelor of Electrical Engineering degree from the Universiti Tenaga Nasional, Malaysia, in 2008, the Master of Computer Engineering degree from the Royal Melbourne Institute of Technology (RMIT), Australia, in 2010, and the Ph.D. degree in electrical engineering from the Universiti Teknologi Malaysia, in 2014. She is currently working as a Senior Lecturer at the Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Malaysia. Her research interests include expert system (machine learning, deep learning), agriculture robotics and automation (classification, control system, intelligent system) and image and pattern recognition. More details can be found at: https://umexpert.um.edu.my/anissalwa.

HASSAN MUWAFAQ GHENI received the Bachelor (B.Sc.) of Electrical and Electronic Engineering from the Department of Electrical Engineering, Babylon University, Hilla, Iraq, in June 2016, and the master's degree from the Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Malaysia, in February 2018. He is currently a Lecturer at the Department of Computer Techniques Engineering, Al-Mustaqbal University College. His research interests include optical communication, the IoT, wireless sensor networks, communications, V2V systems, and artificial intelligence.