IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 27 May 2024, accepted 18 June 2024, date of publication 26 June 2024, date of current version 3 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3419551

RESEARCH ARTICLE

Electricity Demand Forecasting Using a Novel Time Series Ensemble Technique

HASNAIN IFTIKHAR[1,2], SALVATORE MANCHA GONZALES[2], JUSTYNA ZYWIOŁEK[3], (Associate Member, IEEE), AND JAVIER LINKOLK LÓPEZ-GONZALES[2], (Member, IEEE)

[1]Department of Statistics, Quaid-i-Azam University, Islamabad 45320, Pakistan
[2]Escuela de Posgrado, Universidad Peruana Unión, Lima 15464, Peru
[3]Faculty of Management, Czestochowa University of Technology, 42-200 Czestochowa, Poland

Corresponding authors: Hasnain Iftikhar (hasnain@stat.qau.edu.pk) and Justyna Zywiołek (justyna.zywiolek@pcz.pl)

ABSTRACT Accurate and efficient demand forecasting is essential to grid stability, supply, and management in today's electricity markets. Due to the complex pattern of electric power demand time series, it is challenging to model them directly. Therefore, this research proposes a novel time series ensemble approach to forecast electric power demand in the Peruvian electricity market one month ahead. This approach treats the first preprocessed electricity demand time series for missing values, variance stabilization, normalization, stationarity, and seasonality issues. Secondly, six single time series and three of their proposed ensemble models forecast the clean demand time series. The results indicate that the proposed time ensemble approach is an efficient and precise one-month-ahead forecast for electricity demand in the Peruvian electricity market. Additionally, the final best ensemble forecasting model within the proposed ensemble time series forecasting approach obtained the smallest average accuracy errors, performing statistically significantly better than those mentioned in the best-proposed models in the literature. Lastly, while numerous global studies have been conducted from various perspectives, no analysis has been undertaken using an ensemble learning approach to forecast electric power demand in the Peruvian electricity market.

INDEX TERMS Peruvian electricity market, electricity demand forecasting, time series models, time series ensemble approach, planning and operating power systems.

I. INTRODUCTION

The electricity sector in Peru is regulated by the Supervisory Agency of Investment in Energy and Mining (OSINERGMIN). The agency is responsible for setting electricity prices and ensuring the industry complies with the regulations. On the other hand, the Committee for the Economic Operation of the National Interconnected System (COES) is tasked with coordinating the operation of the National Interconnected System. The COES ensures the electricity supply is reliable and stable by balancing electricity generation and demand. It also determines the optimal dispatch of energy sources and coordinates the expansion of transmission infrastructure. In addition, the COES collects and operates electricity generation, consumption, and system performance data, crucial for system planning, analysis, and decision-making [1].

The National Institute for the Defense of Free Competition and the Protection of Intellectual Property (INDECOPI) is a regulatory body responsible for overseeing competition, consumer rights protection, and the operation of markets, including the electricity sector in Peru. It monitors the market to detect and investigate potential competition law violations and serves as a forum for dispute resolution. In regulated markets, prices are set and overseen by regulatory authorities, such as OSINERGMIN, who formulate a comprehensive tariff structure encompassing electricity generation, transmission, and distribution expenses. The regulated price is strategically crafted to provide consumers with stability and predictability. Any modifications to these prices require regulatory approval. In contrast, free-market prices are influenced by market dynamics, competition, and consumer choices.

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

In a liberalized or free-market environment, consumers can choose their electricity provider from various market players, and prices can vary among providers. Understanding these differences is crucial for avoiding misunderstandings and disputes in Peru's electricity industry [2].

Electricity cannot be stored efficiently, so it must be used as generated. To avoid wasting electricity, it's important not to produce more energy than is needed. The company that generates electricity sends it to the distribution subsidiary, which then distributes it to customers. If too much power is produced but not distributed, it's considered a loss for the company [3], [4], [5], [6]. Accurately predicting customers' energy consumption is crucial to prevent such losses. This helps to minimize errors in production orders and avoid overproduction. However, losses can occur at the distribution subsidiary level and be classified as technical or non-technical. Electrical network and equipment issues cause technical losses, while non-technical losses are due to meter failures, fraudulent behavior, and management failures. The total loss is the sum of technical and non-technical losses [7], [8], [9], [10]. Electricity suppliers conduct annual pilot testing with monthly verification points to maintain a normative loss rate. Accurate monthly predictions of customers' energy use prevent deviations from expected results [11], [12], [13]. Therefore, a new high-accuracy forecasting method has been proposed to improve the electricity suppliers' piloting system.

Electricity demand forecasting has been extensively studied over the last four decades. Researchers have developed various techniques to predict electric power demand, broadly classified into four categories: statistical methods, machine learning models, decomposition-combination techniques, and hybrid approaches [14]. Statistical models, such as autoregessive-based, exponential smoothing, linear, and non-linear regression models, are simple mathematical structures that are easy to implement [15], [16], [17]. For instance, a study [18] conducted in Pakistan used a component-wise forecasting approach to predict electric power consumption one month in advance, dividing the data into deterministic and stochastic components. To forecast the deterministic component, linear (parametric) and nonlinear (nonparametric) regression methods were used, while the stochastic part was modeled using four various time series models. The study found that parametric and nonparametric regression approaches had the highest accuracy with the combined Autorregssive moving average model. In contrast, machine learning algorithms address the most complex nonlinear time series forecasting problems [19], [20], [21]. For example, in a study [22] conducted in Brazil, various time-series forecasting models were applied to industrial electricity consumption, and the multilayer perceptron model was found to have the best forecasting performance.

In the decomposition-combination technique, the original time series data is divided into sub-series to improve performance by creating a more reliable form [23], [24], [25], [26]. For instance, a study [27] analyzing monthly electricity consumption in Pakistan decomposed the original time series into three new subsequent: a long-term trend sub-series, a seasonal sub-series, and a stochastic sub-series. When applied to Pakistan's monthly electric power consumption dataset, which ranges from 1990 to 2020, the proposed framework provided highly accurate and efficient gains, outperforming benchmark approaches and improving the performance of the final aggregate model forecasts. On the other hand, many researchers have also introduced hybrid models by merging the specific features of two or more models to build new models [28], [29], [30], [31]. For example, a study [32] proposed a method for mid-term load forecasting using hybrid statistical models that employ input data representing a load time series' normalized annual seasonal cycle with filtered trend and unified variance. The proposed approach avoids the need to understand the complex time series and has several advantages over an alternative method that does not involve forecasting coding variables. The proposal tested mid-term load forecasting issues for thirty-five European countries and outperformed predecessors Prophet, ETS, and ARIMA by about 13.7%, 17.4%, and 25% in the case of MAPE error. The author claims the proposal can be used for short-term electric power demand forecasting.

This study proposes a new approach for forecasting monthly demand in the Peruvian electricity market. The strategy involves training multiple learners to combine their predictions later to improve accuracy. The goal is to leverage individual learners' diversity and strengths, often generated using different algorithms or subsets of data, to create a more robust and accurate prediction model. Thus, the proposed time series ensemble approach treats the first preprocessed electricity demand time series for missing values, variance stabilization, normalization, stationarity, and seasonality issues. It then uses six standard time series models and three proposed ensemble models to anticipate the clean demand time series, free from any problems. The proposed ensemble models are based on the weighting technique, such as equal weight to single models, in-sample-based weighing (training), and out-of-sample (validation). The proposed novel time series ensemble forecasting approach anticipates monthly power demand using Peruvian electricity market data spanning 2001-January to 2022-December. The primary contributions to this work are as follows:

1) A novel methodology for forecasting month-ahead electric power demand is proposed. It employs a time series ensemble method for increased accuracy and efficiency.
2) The efficiency and accuracy of the proposed ensemble models are compared to various single linear and nonlinear time-series models, which are also used within the proposed forecasting methodology.
3) To evaluate the accuracy and efficiency of the proposed approach, six different accuracy average errors are determined, and a statistical equal forecast test, the Diebold-Marino test, and graphical evaluation using

line plots, bar plots, autocorrelation, and partial correlation plots are performed.

4) The outcomes of the best ensemble model are highly accurate and efficient compared to the best models reported in the literature for month-ahead electric power demand forecasting.
5) This study is the first to propose an electric power demand forecasting strategy for the Peruvian electricity market.
6) The proposed novel time series ensemble forecasting approach can be extended and applied to other energy markets to assess its efficacy. On the other hand, in the future, the proposal of the current research study should be extended to other electricity markets–for instance, the European electricity markets, the United American electricity market, the United Kingdom, the Chinese electricity markets, and so on.

The remainder of this research study is structured as given: Section II describes the general layout of our proposed novel time series ensemble forecasting approach. Then, in Section III, apply this approach to the Peruvian electricity market demand data. Section IV compares the best-proposed ensemble model to the best models provided in the literature. Section V summarizes our findings, limitations, and recommendations for further study.

II. THE GENERAL LAYOUT OF THE PROPOSED APPROACH

This section explains the proposed time series ensemble technique for one month ahead of electricity demand and demand forecasting in Peru's electricity market. In the proposed time series ensemble technique, the demand time series is first preprocessed by missing value, variance stabilization, normalization, stationary, and seasonality concerns. Second, six different time series models: autoregressive, autoregressive integrated moving averages, exponential smoothing model, nonlinear autoregressive, autoregressive integrated moving averages, neural network autoregressive, and Theta model, and also three proposed ensemble models anticipate the clean demand time series. The details about these steps are in the following subsections.

A. PREPARATION OF RAW DATA

The main goal of this study is to achieve an accurate and efficient one-month-ahead forecast for Peruvian electricity demand. To achieve this objective, we implemented a sequential methodology to model and forecast the electricity demand time series in the Peruvian market. The authors aim to understand the complex characteristics of electricity demand dynamics over time. These characteristics are expected to include missing values, a nonlinear long-run trend, pronounced seasonality, high volatility, nonnormality, and nonstationarity. For instance, see Figure 1(a) for the monthly electricity demand time series from 2001-January to 2022-December. It can show an increasing linear secular trend component in the electric power demand time series. Figure 1(b) shows yearly demand data for the past twenty-two years, which shows a continuously increasing trend in the electric power demand. However, it can be observed that during the 2020 year, there was lower demand, which was the leading cause of the COVID-19 pandemic. Figure 1(c) shows the seasonal box plot for the electricity demand, which shows the variation of demand during the various seasons of the year in Peru. Figure 1(d) illustrates the average monthly demand for the past twenty-two years. It also confirmed a variant of demand due to the different months of the year, which was the leading cause of seasons. Figure 1(e) for the autocorrelation plot of the original electricity demand at sixty lags, and Figure 1(f) for the partial autocorrelation plot of the original electricity demand at sixty lags. These figures clearly illustrate a discernible nonlinear long-run trend and an annual seasonality. Furthermore, nonnormality and nonstationarity are also evident from these visual representations.

The proposed approach addresses these irregularities before modeling and forecasting to achieve high accuracy and efficient forecasts. First, a linear interpolation method replaces a few missing observations with the mean of their three surrounding values. Second, the natural logarithm is applied to stabilize variance and standard deviation. Third, the deterministic characteristics containing a linear long-run trend component and yearly seasonality are removed. To accomplish this, model these deterministic components using the following procedure: Let the time series of the electric power demand time series be donated by $log(X_m)$; m shows the mth month data point. Thus, the dynamics of the log monthly demand times series, $\log(X_m)$, may be described as:

$$log(X_m) = \tau_m + a_m + x_m \tag{1}$$

That is, the $\log(X_m)$ is divided into these components: a long-run linear trend component (τ_m), a yearly seasonality component (a_m), and a residual component (x_m). The (τ_m) component is a function of the series $(1, 2, 3, \ldots, m)$, is estimated by the regression splines method, and dummies capture the annual periodicity: $a_{m,j} = \sum_{j=1}^{12} \zeta_j I_{m,j}$. The variable $I_{m,j}$ is assigned a value of 1 when j refers to the ith month and 0 otherwise. The regression coefficients (ζ_j) associated with these components are determined using the ordinary least square method. Once the estimated deterministic component (long-run trend and annual periodicity) is obtained, the residual or stochastic component can be derived as:

$$x_m = \log(X_m) - (\hat{\tau}_m + \hat{a}_m) \tag{2}$$

Thus, once the electric power demand time series (X_m) is preprocessed (to address the issue of missing values and their imputation, stabilize the variance and standard deviation, and remove the deterministic properties), the next step is to model the remaining residual x_m series; the current work considers six single-time series models and three proposed ensemble

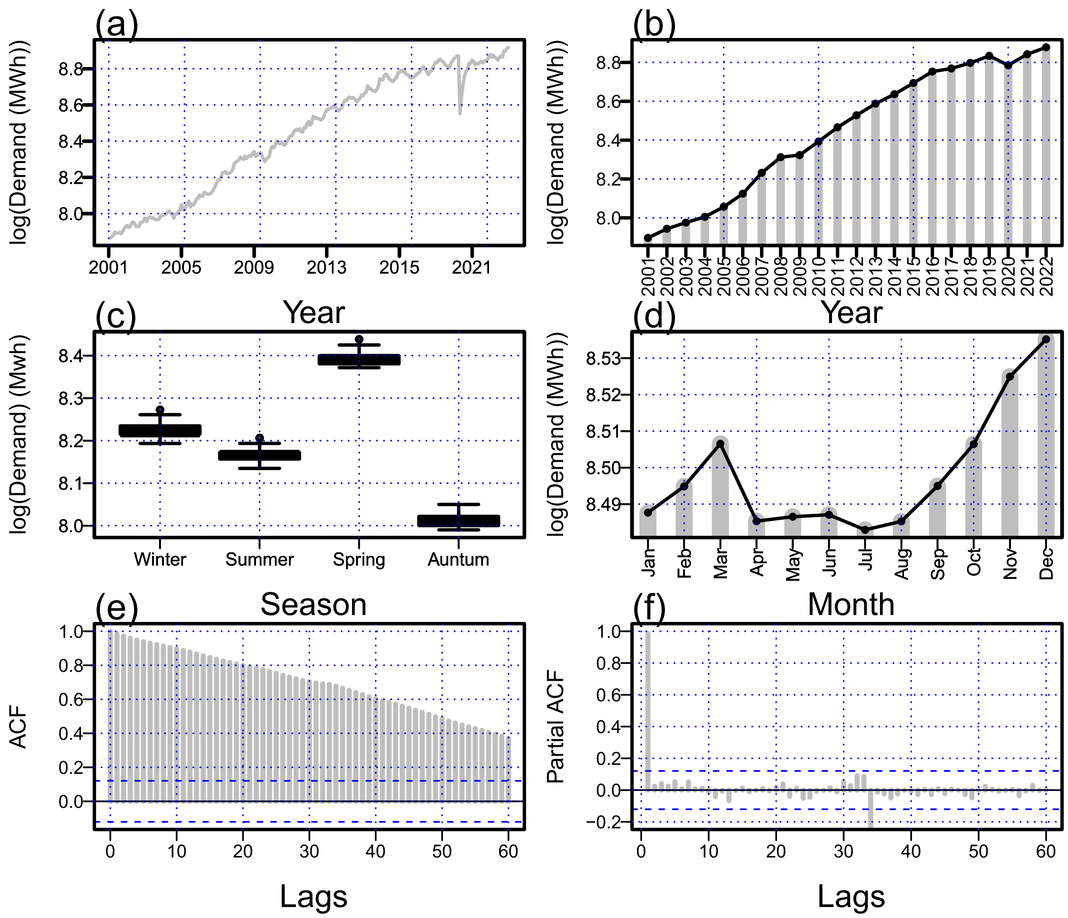

FIGURE 1. Characterization of Peruvian Electric Power Demand (2001-2022): monthly time series plot (a); Yearly electric power demand line plot (b), displays the boxplot of the seasonal electric power demand (c); monthly average demand line plot (d), autocorrelation function plot (e), partial autocorrelation function plot (f).

models. Hence, all forecasting models are described in the coming subsection.

B. FORECASTING MODELS

This section briefly overviews the forecasting models and their proposed ensemble models: the autoregressive model, the simple exponential smoothing model, the autoregressive moving average model, the theta model, the nonparametric autoregressive model, and the neural network autoregressive model.

1) AUTOREGRESSIVE MODEL

A linear and parametric autoregressive (AR) process accounts for x_m's short-term dynamics. It is a linear combination of the preceding time (lag) v observations of x_m, which may be written as follows:

$$x_m = u + \beta_1 x_{m-1} + \beta_2 x_{m-2} + \ldots + \beta_v x_{m-v} + e_m \quad (3)$$

Here, u and β_l ($l = 1, 2, \ldots, v$) are the intercept and slope parameters of the AR process, respectively, while e_m is the white noise process. Based on the visual investigation (autocorrelation and partial autocorrelation plots) and the theoretical analysis (the AIC and BIC measures), the AR model with the first three lags is best suited for modeling x_m.

2) THE EXPONENTIAL SMOOTHING MODEL

The Exponential Smoothing Model (ESM) is a group of forecasting models that apply exponentially decreasing weights to previous observations. It is a time-series forecasting model that uses a weighted average of past observations to predict the future value of a variable. The ES model assumes that a variable's future value depends on its past values, with greater emphasis placed on recent values than on older ones. The ESM model can be expressed as follows:

$$x_{m+1} = \alpha \cdot x_m + (1 - \alpha) \cdot x_{m-1} \quad (4)$$

In the given equation, x_{m+1}, x_m, and x_{m-1} are the actual values of the filters demand time series at times m+1, m, and m−1. At the same time, α is the smoothing parameter determining the weight assigned to the most recent observation.

3) AUTOREGRESSIVE MOVING AVERAGE MODEL

The autoregressive moving average (ARMA) model is a strong strategy that considers the target variable's previous values and integrates pertinent information using moving average terms. The ARMA model describes the behavior of the current study variable, x_m, using the preceding r terms and the delayed residual values. By considering both the AR and the MA process, the ARMA model provides a comprehensive framework for describing the dynamics of the variable in question. The model may be expressed as follows:

$$x_m = u + \sum_{l=1}^{v} \beta_l x_{m-l} + \sum_{o=1}^{r} \eta_o e_{m-o} + e_m \tag{5}$$

Here, u denotes the intercept, β_l ($l = 1, 2, \ldots, v$) and η_o ($o = 1, 2, \ldots, r$) are the parameters of the AR and the MA components, respectively, and e_m is the white noise process, having zero mean ($\mu = 0$) and variance σ_e^2. Based on the visual investigation (autocorrelation and partial autocorrelation plots) and the theoretical analysis (the AIC and BIC measures), the ARMA(3,2) model is the best model for the eclectic power demand (x_m) forecasting.

4) THE THETA MODEL

The Theta Model is a forecasting method that predicts future values based on the average change in the time series data. It involves calculating the average change between consecutive time points and extrapolating it into the future. The equation for the Theta Model is given by:

$$x_{m+1} = \frac{1}{n} \left(x_m + x_{m-1} + \ldots + x_{m-n+1}\right) \tag{6}$$

In the last equation, x_{m+1}, x_m, x_{m-1}, and x_{m-m+1} are the actual values of the filtered electricity demand time series at times m+1, m, m−1, and m-n+1. Here, n denotes the number of past values used in the average.

5) THE NONPARAMETRIC AUTOREGRESSIVE MODEL

The nonparametric autoregressive model (NPAR) presents an alternative to the conventional parametric AR model, departing from the latter's reliance on specific mathematical equations to elucidate the relationship between past and future values. In contrast, NPAR models employ flexible and adaptive techniques, such as kernel regression or spline functions, to capture dynamic patterns in the data without explicit parameter estimation. These models are distinguished by their flexibility, absence of predefined parameters, emphasis on local relationships, and reliance on data-driven structures to address intricate and nonlinear dependencies within time series data. In this model, the relationship between the variable x_m and its previous values is not restricted to a specific parametric form, thereby allowing for nonlinear associations.

$$x_m = u_1(x_{m-1}) + u_2(x_{m-2}) + \ldots + u_n(x_{m-n}) + \varepsilon_m \tag{7}$$

The relationship is represented as a series of smoothing functions, denoted as u_j ($j = 1, 2, \ldots, n$), which describe the association between x_m and its previous values. This study uses cubic regression splines to represent the smoothing functions, and the model employs the first three lags for Nonparametric Additive Regression modeling.

6) THE NEURAL NETWORK AUTOREGRESSIVE MODEL

The Neural Network Autoregressive (NNA) model is a machine learning approach that uses past observations to predict future values in a time series. This is done by analyzing a mathematical function that considers the previous values, denoted by $x_{m-1}, x_{m-2}, \ldots, x_{m-n}$, where n is the time delay parameter. During training, the backpropagation and steepest descent approaches minimize the difference between predicted and actual values. When forecasting, the autoregression order is determined, which indicates the number of previous values needed to predict the current value of the time series. The NNA is then trained using a dataset that reflects the autoregression order, and the number of input nodes is determined based on this order. These input nodes represent past lagged observations in univariate time series forecasting. The NNA's output provides predicted values. However, choosing the number of hidden nodes often requires trial and error and lacks a theoretical basis. Careful consideration is necessary to prevent overfitting when selecting the number of iterations. In this study, an NNA design of (4, 2) is utilized, expressed as $x_m = f(x_{m-1})$, where $x_m = (x_{m-1}, x_{m-2}, x_{m-3}, x_{m-4})$ represents past values of the cleaned monthly demand time series (x_m), and f denotes a neural network with four hidden nodes in a single layer.

7) 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. Within the scope of this study, ensemble techniques are initially employed to compute weights for the results derived from individual models. Consequently, the proposed ensemble encompasses three distinct weighting strategies: a) equitable distribution of weight among all single models, denoted as Ensemble (A); b) weight assignment based on training average accuracy errors (1), designated as Ensemble (B); and c) weight assignment based on validation mean accuracy measures, denoted as Ensemble (C). The model allocates greater weight to the ensemble model for training and validation datasets with lower mean accuracy errors, while models exhibiting higher mean accuracy errors contribute comparatively less weight to the ensemble. Notably, the model weights assume small positive values and their accumulation equates to one, signifying the percentage of reliance or anticipated performance on each model.

Thus, after estimating the long-run trend component and annual periodicity using the spline regression method and dummy variables discussed above, the next step is forecasting the remaining part (x_m) using six single and three proposed ensemble models as discussed above. Thus, this work can

TABLE 1. Mean evaluation errors.

S.No	Error	Formula
1	MAE	$\frac{1}{M}\sum_{m=1}^{M}\lvert x_m - \hat{x}_m\rvert$
2	MASE	$\frac{1}{M}\left[\frac{\lvert x_m-\hat{x}_m\rvert}{\frac{1}{M-1}\sum_{m2}^{M}\lvert x_m - x_{M-1}\rvert}\right]$
3	MAPE	$\frac{1}{M}\sum_{m=1}^{M}\left\lvert\frac{x_m-\hat{x}_m}{x_m}\right\rvert$
4	RMSE	$\left[\left[\sum_{m=1}^{M}\right]\frac{(x_m-\hat{x}_m)^2}{M}\right]^{0.5}$
5	RMSLE	$\left[\frac{1}{M}\sum_{m=1}^{M}[\log(x_m+1)-\log(\hat{x}_m+1)]^2\right]^{0.5}$
6	RRSE	$\left[\frac{\sum_{m=1}^{M}(x_m-\hat{x}_m)^2}{\sum_{m=1}^{M}(x_m-\overline{x}_m)^2}\right]^{0.5}$

obtain the monthly electric power demand for the next month's forecast as follows:

$$\hat{X}_{m+1} = \exp(\hat{\tau}_m + \hat{a}_m + \hat{x}_m) \tag{8}$$

C. EVALUATION CRITERIA

This study examines two evaluation criteria for the proposed time series ensemble forecasting technique: accuracy average errors and an equal forecast accuracy test. Primarily, Table 1 presents the accuracy average errors, outlining the formulas for computing each metric. The metrics encompass the mean absolute error (MAE), the mean absolute percent error (MAPE), the mean scaled absolute error (MASE), the root mean squared error (RMSE), the root relative squared error (RRSE), and the root mean log squared error (RMSLE).

In the given table, x_m denotes observed values, while $\hat{x}_m$ represents forecasted electricity demand for the mth observation ($m = 1, 2, \ldots, 48 = M$). The MAE calculates the average absolute differences between the expected and actual values. It is not as sensitive to outliers. However, MAPE calculates the average of the absolute percentage differences between the model's predictions and the actual values. Therefore, this metric expresses the average error as a percentage of the actual value. MAPE penalizes negative errors more heavily (when the predicted value exceeds the actual). This is because the percentage error cannot exceed 100% for very low predictions, while there is no upper limit for higher predictions. Consequently, MAPE tends to favor models that underestimate rather than overestimate. Meanwhile, the MASE is a metric for forecast accuracy. It is the mean absolute error of the forecast values divided by the mean absolute error of the in-sample one-step naive forecast.

On the other hand, the RMSE is the square root of the mean square error. Mathematically, it measures the standard deviation of the error. Due to its ease of comprehension, RMSE is widely used to compare the performances of different models. However, the RRSE is relative to what it would have been if a simple predictor had been used. More specifically, this simple predictor is just the average of the actual values. Thus, the relative squared error takes the total squared error and normalizes it by dividing it by the total squared error of the simple predictor. Taking the square root of the relative squared error reduces the error to the same dimensions as the predicted quantity. Similarly, the RMSLE is calculated by applying a log to the actual and the predicted values and then taking their differences. RMSLE is robust to outliers where the small and the large errors are treated evenly. It penalizes the mode more if the predicted value is less than the actual value, while the model is less penalized if the expected value is more than the actual value. It does not penalize high errors due to the log. Hence, the model has a more significant penalty for underestimation than overestimation. This can be helpful in situations where we are not bothered by overestimation, but underestimation is unacceptable. Consequently, diminishing values for MAE, MASE, MAPE, RMSE, RRSE, and RMSLE generally signify heightened predictive accuracy of the model.

Second, a statistically equal forecast test, the Diebold-Marino (DM) test [33], is performed to evaluate the forecasting ensemble time series proposed approach. In the literature, It is used to evaluate time series forecasting models, determining whether the forecast errors from one model are statistically different from another model's forecast errors [34], [35], [36]. To perform the DM test, the forecast errors of each model are calculated using a loss function. Then, a statistical value is computed by comparing the errors of each model. The test statistic is based on the difference between the mean squared errors of the two models. Suppose the test statistic is above a certain threshold and the p-value is below a significance level ($\alpha = 0.05$). In that case, the forecasts from one model are significantly better than the other model. For instance, calculate the forecast errors for both models. Forecast errors ($e_d = x_m - \hat{x}_m$) are the differences between the observed values (x_m) and the forecasted values ($\hat{x}_m$). Compute the mean difference ($\bar{w}$) of the forecast errors: $\bar{w} = \frac{1}{M}\sum_{m=1}^{M}(e_{1d} - e_{2d})$. Where: e_{1d} and e_{2d} are the forecast errors from Model 1 and Model 2 at time d, respectively, and D is the number of observations. Next, calculate the variance of the differences, such as $\sigma_d^2 = \frac{1}{M}\sum_{m=1}^{M}(e_{1d} - e_{2d} - \bar{w})^2$. Thus, the DM test statistic, DM $= \frac{\bar{w}}{\sqrt{\sigma_d^2}}$. Finally, the Null and alternative hypothesized generally state as H_0: There is no difference in forecast accuracy between the two models (H_0: $\bar{w} = 0$) Vs. H_A: The two models differ in forecast accuracy (H_A: $\bar{w} \neq 0$). Hence, the null hypothesis implies that there is not a statistically significant difference in forecast accuracy between the models. In contrast, the alternative hypothesis suggests a significant difference in forecast accuracy between the two models.

To complete this section, the main steps, including the developed time series ensemble forecasting approach, are listed below in bullet form. In addition, Figure 2 provides a visual representation of the procedural flow.

- In the first step, the electric power demand time series (X_m) is preprocessed (to address the issue of missing values and their imputation, stabilize the variance and standard deviation, and remove the deterministic properties), discussed in detail in section II-A.
- In the second step, this work divides the stochastic (short-run dynamic) electric power time (x_n) into three parts: training (in-sample), validation (evaluation), and testing (out-of-sample) datasets. Let $x_n; n = 1, 2, \ldots, N$ (264) is a stochastic (short-run dynamic) electric power time series. The training (in-sample) dataset is $y_t; t = 1, 2, \ldots, T(180)$, the validation (evaluation) dataset is $w_v; v = 1, 2, \ldots, V(36)$, and the testing (out-of-sample) dataset is $z_m; m = 1, 2, \ldots, M(48)$ where N (N = T+V+M) is the total data points.
- The train data is modeled in the third step using single models, such as the AR, ARMA, ESM, NPAR, Theta, and NNA models.
- In the fourth step, calculate the one-month-ahead electricity demand forecast using the expanding window technique. The forecast values, $\hat{x}^j_{N-(T+V+M)}$ for $j = 1, 2, \ldots, 6$, are obtained by the models listed in step 3.
- In the fifth step, the output of a primary ensemble method is mathematically described by Equation 9.

$$\hat{x}^j_{N-(T+V+M)} = \sum_{j=1}^{6} \omega_i \hat{x}^j_{N-(T+V+M)} \tag{9}$$

  Where ω_i, are obtained by three weighting strategies: a) equal weight to all single models and denoted by (Ensemble (A)); b) weight assigned based on training mean accuracy measures (see Table 1) and denoted by (Ensemble (B)); c) weight assigned based on validation mean accuracy measures and denoted by (Ensemble (C)). The lower accuracy mean errors model assigns more weight to the ensemble model in training and validation data sets. In contrast, the model with the model with the highest accuracy has fewer errors than the ensemble model. However, the model weights are small positive values, and their sum equals one, indicating the percentage of trust or expected performance from each model. Thus, obtain the one-month-ahead forecast values using Equation (9) for the Ensemble (A), Ensemble (B), and Ensemble (C) models.
- In the sixth step, evaluate the model based on average accuracy errors and an equal forecast statistical test (see details in II-C).

III. CASE STUDY ANALYSIS

This study analyzed the monthly electric power demand in the Peruvian electric power market from January 2000 to December 2022. The data set comprises 264 data points, with twelve for each twenty-two years. To create a reliable data model for forecasting, the dataset was divided into three parts: a training part for model estimation, a validation part for model validation and hold-out-sample forecast, and a testing part for out-of-sample forecast. The training used data from 2001-January to 2014-December, consisting of 180 observations. The period from 2015-January to 2018-December, with 48 observations, was used for model validation. Finally, the period from 2019-January to 2022-December, with 48 observations, was used for model testing. To analyze the electric power demand time series database, this work computes the descriptive statistics (smallest, 25%, 50%, arithmetic mean, mode, variance, standard deviation, 75%, and largest values) listed in Table 2. The second column in this table contains information about the original electric power demand without treatment. The third column contains the natural log, the original electric power demand. It is seen that after taking the natural log, the variance and standard deviation are stabilized. On the other hand, normality is also achieved, as confirmed by the mean, median, and mode, which have approximately the same values. In addition, after capturing the deterministic part (the linear long-run trend and the yearly seasonality components), the remaining series (x_m) have no evidence of seasonality and nonstationarity issues. Therefore, once the electric power price time series addresses all the essential treatments (missing values, variance or standard deviation stabilization, normality, seasonality, and stationary issues), proceed further to model and forecast the filtered electric power demand time series. This work used six single time series models and three of their proposed ensemble models to forecast the filtered electric power demand time series. Therefore, the ensemble technique of the proposed time series compares the nine overall models. The performance of six single time series models and the proposed three ensemble models within these nine models will be checked. Compute various average errors, including the MAE, the MASE, the MAPE, the RMSE, the RRSE, and the RMSLE, to evaluate the performance of the proposed time-series ensemble forecasting approach. The results of the average errors for nine models can be found in Table 3.

TABLE 2. Descriptive statistics.

Statistic	Original Series (X_n)	log (X_n)
Minimum	2606.78	7.87
25%	3348.12	8.12
50%	4900.20	8.50
Mean	4906.28	8.45
Mode	2606.78	7.87
75%	6405.83	8.77
Variance	2307069.60	0.11
Standard Deviation	1518.90	0.33
Maximum	7467.45	8.92

Table 3 shows that the Ensemble (C) model produced the best forecasting results compared to all nine models within the proposed time series ensemble forecasting approach. For instance, the average accuracy errors for the Ensemble

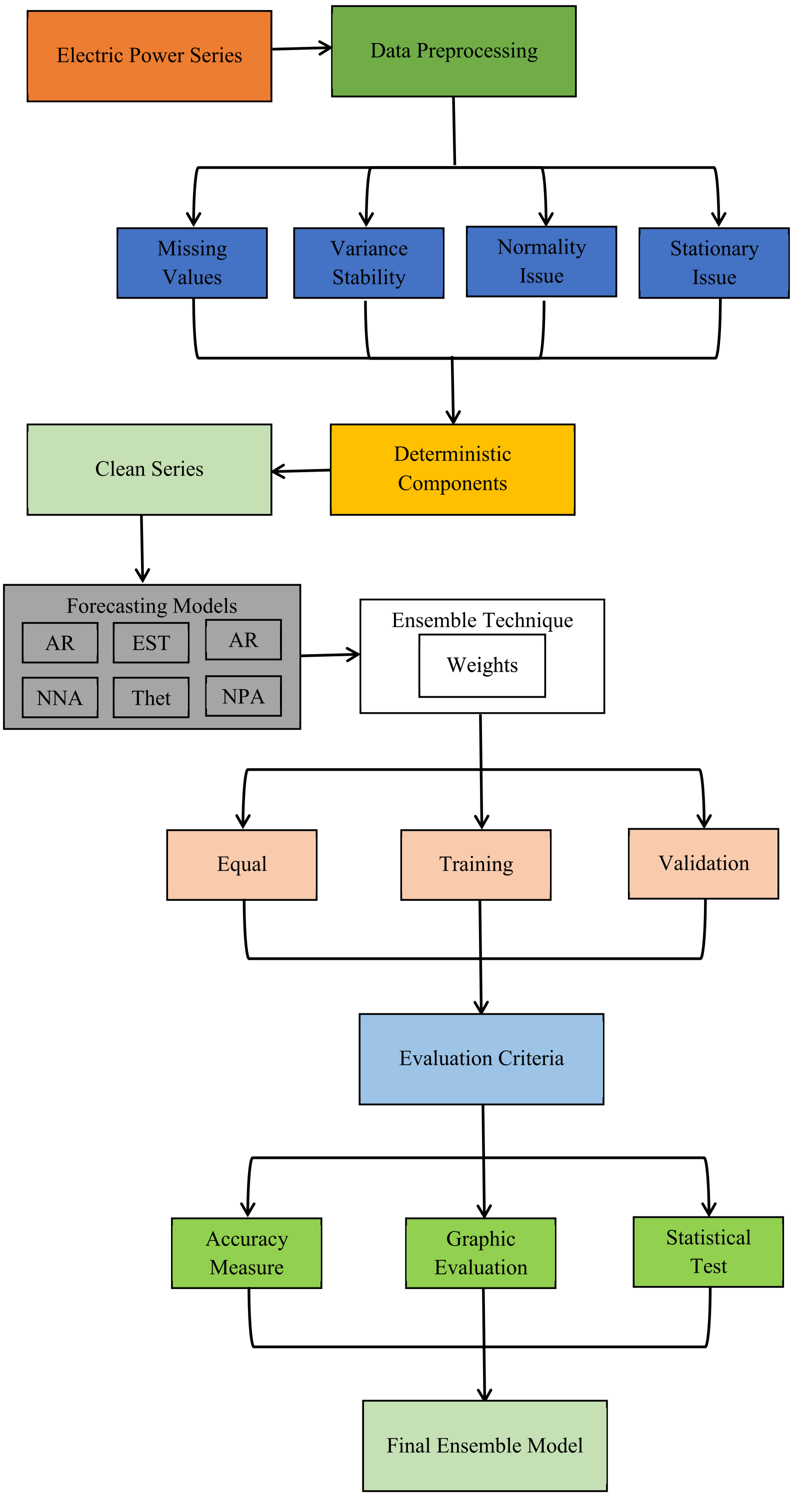

FIGURE 2. Peruvian electricity market: The proposed time series ensemble approach layout.

(C) model are the following: MAPE = 0.01304, MAE = 79.18273, MASE = 0.63671, RMSE = 253.32520, RRSE = 0.69545, and RMLSE = 0.04130. However, the Ensemble (B) and ARMA models produced the second and third-best forecasting results such as (the Ensemble (B) model: MAPE = 0.01466, MAE = 91.26190, MASE = 0.73383, RMSE = 259.38770, RRSE = 0.71209, and RMLSE = 0.04225) and (the ARMA model: MAPE = 0.01647, MAE = 101.45942,

TABLE 3. One month ahead electric power demand forecasting accuracy mean errors.

Model	MAPE	MAE	MASE	RMSE	RRSE	RMLSE
AR	0.01796	110.82196	0.89112	281.06050	0.77159	0.04562
ARMA	0.01647	101.45942	0.81583	272.91580	0.74923	0.04419
ESM	0.01778	109.99298	0.88445	279.88020	0.76835	0.04536
NNA	0.03381	215.01260	1.72891	401.01730	1.10091	0.06433
NPAR	0.02737	167.28372	1.34512	457.41930	1.25575	0.09853
Theta	0.01786	110.36110	0.88741	282.16820	0.77463	0.04572
Ensemble (A)	0.01920	118.19376	0.95039	289.35830	0.79437	0.04752
Ensemble (B)	0.01466	91.26190	0.73383	259.38770	0.71209	0.04225
Ensemble (C)	0.01304	79.18273	0.63671	253.32520	0.69545	0.04130

TABLE 4. Diebold-mariano test: comparing one model to others.

Models	AR	ARMA	ESM	NNA	NPAR	Theta	Ensemble (A)	Ensemble (B)	Ensemble (C)
AR	0.000	0.861	0.117	0.918	0.936	0.859	0.929	0.066	0.040
ARMA	0.139	0.000	0.126	0.919	0.936	0.828	0.933	0.067	0.045
ESM	0.883	0.874	0.000	0.918	0.936	0.871	0.930	0.066	0.044
NNA	0.082	0.081	0.082	0.000	0.937	0.081	0.090	0.074	0.041
NPAR	0.064	0.064	0.064	0.063	0.000	0.064	0.064	0.065	0.044
Theta	0.141	0.172	0.129	0.919	0.936	0.000	0.933	0.067	0.042
Ensemble (A)	0.071	0.067	0.070	0.910	0.936	0.067	0.000	0.067	0.055
Ensemble (B)	0.934	0.933	0.934	0.926	0.935	0.933	0.933	0.000	0.071
Ensemble (C)	0.936	0.935	0.936	0.929	0.936	0.935	0.935	0.939	0.000

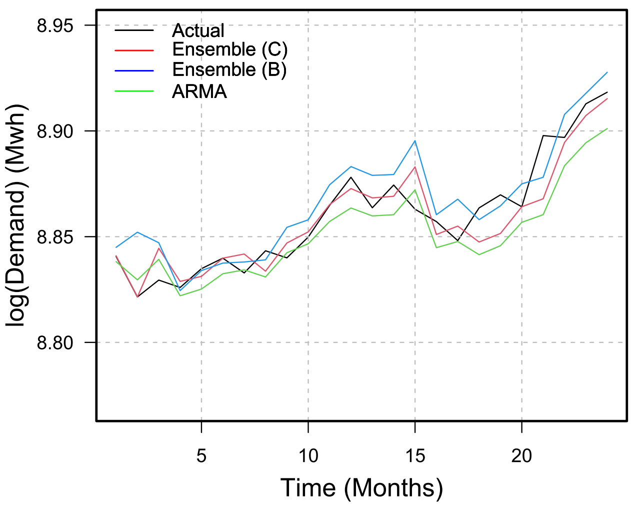

FIGURE 3. Comparison of original and forecasted peruvian electricity demand: Ensemble (C), Ensemble (B), and ARMA Models (24 Months).

MASE = 0.81583, RMSE = 272.91580, RRSE = 0.74923, and RMLSE = 0.04419). In contrast to this comparison, when comparing only the single time series models, the ARMA model shows the best forecasting results amongst all single time series models. However, when comparing only the proposed ensemble models, the Ensemble (C) model produced the lowest average errors. Therefore, the Ensemble (C) model shows highly accurate and efficient monthly electric power demand forecasting for the Peruvian Electricity Market within the proposed forecasting ensemble technique. After calculating the performance metrics (MAE, MASE, MAPE, RMSE, RRSE, and RMLSE), we used the Diebold-Mariano (DM) test to statistically assess the superiority of models within the proposed ensemble technique (see Table 4 for p-values). Our analysis indicates a 5% significance level–the performance of nine forecasting models, including six base models and three proposed ensemble models. Statistical analysis (the DM test) revealed that the Ensemble (C) model achieved statistically superior performance across all models. Notably, the Ensemble (C) model also showed strong results, outperforming seven other models. These findings confirm Ensemble (C)'s accuracy as the most reliable model for one-month-ahead electric power demand forecasting within the scope of this study. These results support the conclusion that the Ensemble (C) model offers the most accurate and reliable one-month-ahead electricity demand forecasts among the models considered.

In addition to the above, this comparative analysis also performed a graphical analysis to validate the current work proposed Ensemble (C) model's superiority further. Figure 3 visually compares the actual and forecasted electric power demand for the top three models: the Ensemble (C), the Ensemble (B), and the ARMA. The Ensemble (C) model's forecasts closely track the actual demand, demonstrating its exceptional accuracy. Additionally, the authors examined the correlogram plots (autocorrelation and partial autocorrelation) of the residuals for these three models (Figure 4). The

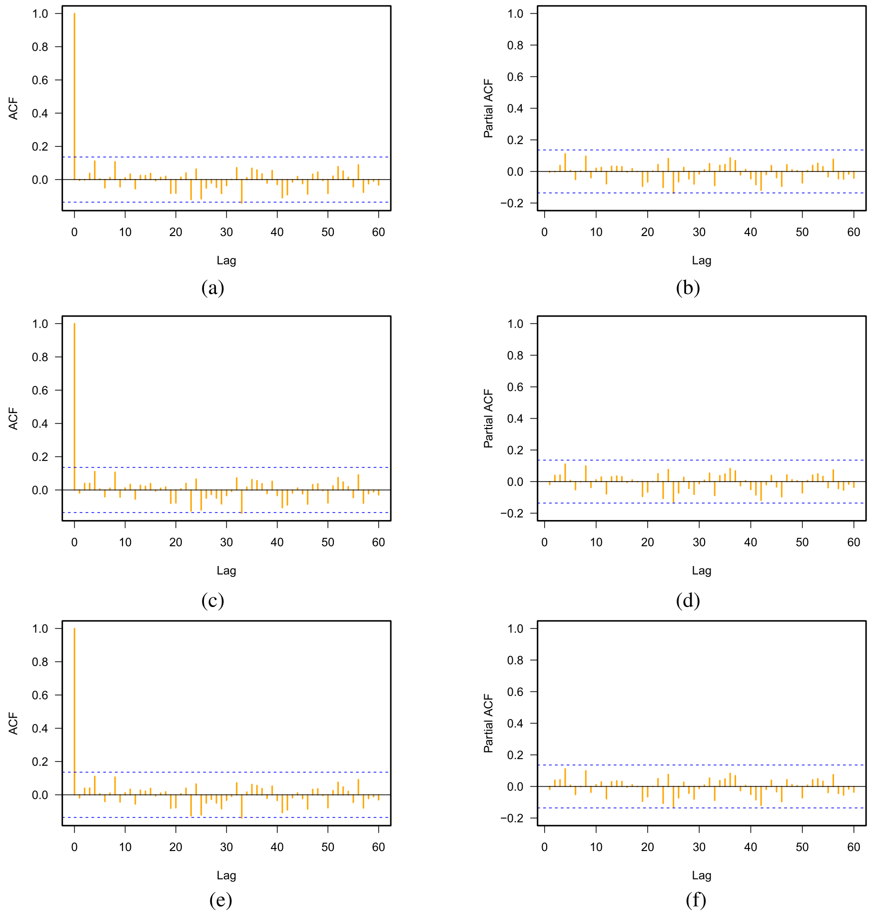

FIGURE 4. Autocorrelation function and partial autocorrelation plots for the three best models among all nine considered models: Ensemble (C) model (a-b), Ensemble (B) model (c-d), and ARMA model (e-f).

TABLE 5. Peruvian electricity market: One-month-ahead average accuracy metrics of the best proposed (Ensemble (C)) model versus those reported as the best models in the literature.

Model	MAPE	MAE	MASE	RMSE	RRSE	RMLSE
Ensemble (C)	0.01304	79.18273	0.63671	253.32520	0.69545	0.04130
The NNAR model [17]	0.02204	98.82730	0.86671	275.31220	0.92995	0.07364
The proposed method 1 [37]	0.01854	93.95030	0.83794	272.45220	0.89695	0.07040
the DR-SFGM model [29]	0.01754	89.85030	0.74794	268.45220	0.87795	0.06740
The MLP model [22]	0.02104	96.97230	0.85101	273.46720	0.91475	0.07149
The NP-ARMA [18]	0.01704	88.54030	0.73651	266.23620	0.86555	0.06620
The GSM model [38]	0.01804	91.93830	0.83630	271.30620	0.89445	0.06980

absence of significant autocorrelation in the residuals of all models indicates that they have been sufficiently whitened, signaling satisfactory model performance.

In conclusion, the combination of accuracy metrics (MAE, MASE, MAPE, RMSE, RRSE, and RMLSE), statistical testing (DM test), and graphical analysis provides compelling evidence for the superiority of our proposed ensemble forecasting approach in one-month-ahead Peruvian electric power demand forecasting. Specifically, the Ensemble (C) model consistently generates the most precise forecasts compared to this study's other single and the proposed ensemble models.

IV. DISCUSSION

Once the best model (Ensemble (C)) was this work obtained among all nine considered nine models by considering three evaluation criteria: 1. accuracy average errors (MAPE, MAE, MASE, RMSE, RRSE, RMLSE); 2. an equal forecast accuracy test (the DM test); and 3. the graphical assessment (line plot, dot plot, and correlogram plots). Next, this section compares our study's best model, Ensemble (C), with the best models proposed in the literature. The current work found that this work is the super best model to be highly comparable with the considered methods. Table 5 numerically and Figure 5 graphically empirically compare

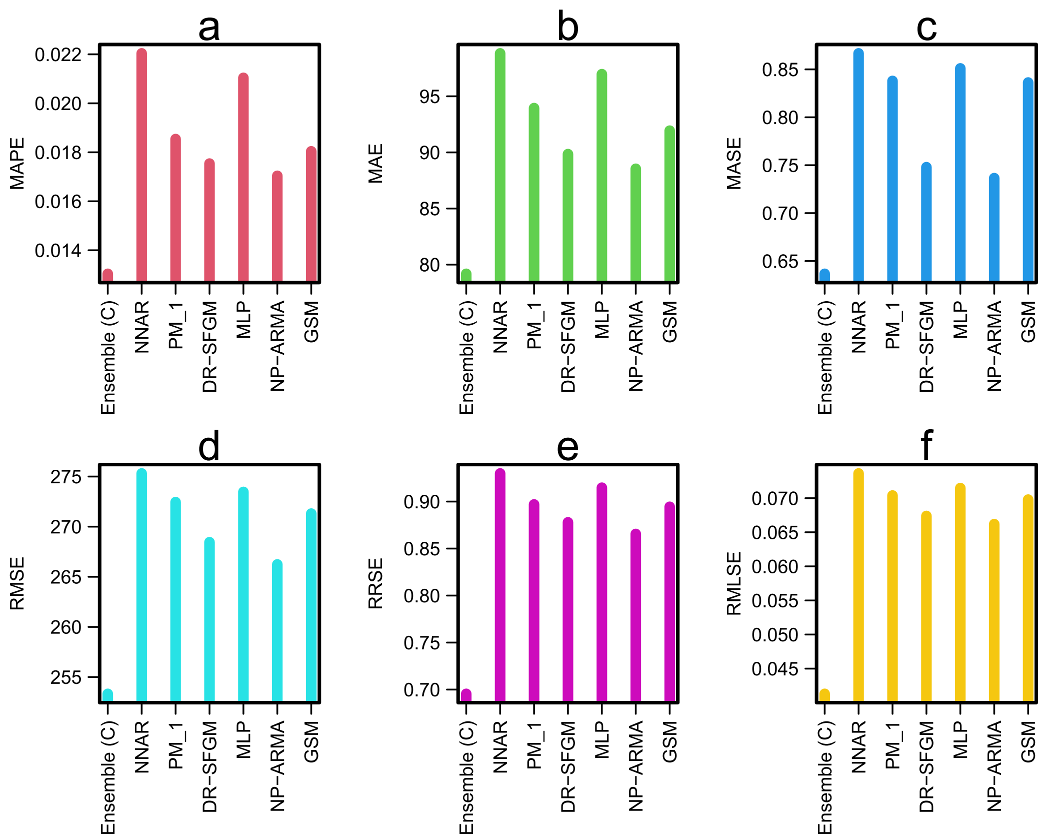

FIGURE 5. Peruvian electricity market: one-month-ahead average accuracy metrics (a-MAPE, b-MAE, c-MASE, d-RMSE, e-RRSE, and f-RMSLE) of the best proposed (Ensemble (C)) model versus those reported as the best models in the literature.

our model with the other models proposed by researchers. This work applied the best model proposed by [17], the neural network artificial autoregressive model (NNAR) model, to our dataset and calculated their accuracy average errors. The accuracy average error values reported by [17] for their proposed best model were: MAPE = 0.02204, MAE = 98.82730, MASE = 0.86671, RMSE = 275.31220, RRSE = 0.92995, RMSLE = 0.07364. These values were higher than the accuracy average error values of our best model, Ensemble (C), which are: MAPE = 0.01304, MAE = 79.18273, MASE = 0.63671, RMSE = 253.32520, RRSE = 0.69545, and RMLSE = 0.04130. Another study proposed a final best model (method 1) [37], and the authors computed their prediction average errors, which were: MAPE = 0.01854, MAE = 93.95030, MASE = 0.83794, RMSE = 272.45220, RRSE = 0.89695, RMLSE = 0.07040. These values were also higher than our Ensemble (C) model's forecasting average errors. Similarly, in reference [29], the best-proposed model (DR-SFGM) was applied to our dataset, and the performance metrics were obtained, such as 0.01754, 89.85030, 0.74794, 268.45220, 0.87795, and 0.06740 for the MAPE, MAE, MASE, RMSE, RRSE, RMLSE, respectively; which were comparatively higher than our best (Ensemble (C)) model. Furthermore, In work [22], the proposed best model (MLP) used our data, and the average accuracy measures they calculated were also comparatively much higher than our Ensemble (C) model. Another study proposed the best model [18], the nonlinear regression autoregressive moving average (NP-ARMA) model, and we applied it to our dataset. The accuracy metrics obtained were also relatively higher than our best (Ensemble (C)) model. In reference [38], the best proposed (GSM) was applied to our dataset, and the accuracy metrics computed were worse than those obtained with our best model (Ensemble (C)). Hence, this study's best ensemble (Ensemble (C)) model outperformed all other studies' best models in forecasting one-month-ahead electric power demand for the Peruvian Electricity Market Pool market, as shown numerically (Table 5 and graphically (Figure 5). Furthermore, highly accurate and efficient monthly electric power demand forecasting provides numerous advantages, including practical short–and medium-term strategic forecasting for lower operational and maintenance costs, heightened stock and demand management, increased system reliability, and future reserves. Monthly demand forecasting assists in decreasing risks and making well-versed economic conclusions that affect return margins, revenue, supply allocation, expansion budding, inventory accountancy, functioning expenses, people, and total disbursement.

V. CONCLUSION

Understanding electricity demand is crucial for informed decisions regarding infrastructure investment, pricing strategies, and system reliability. Analyzing historical data is essential for forecasting studies, which provide insights into trends, seasonality, and peak demand periods. This study aims to uncover demand evolution, providing a foundation for accurate forecasting and aiding private sector entities, regulatory bodies, and stakeholders in making informed decisions. By understanding demand evolution, authorities and private entities can implement measures to maintain a stable electricity supply and ensure system reliability. Therefore, this study proposes a novel time series ensemble technique specifically tailored for electricity demand forecasting in the Peruvian market. The proposed time series ensemble forecasting technique uses the first preprocessed electricity demand time series for missing values, variance stabilization, normalization, stationarity, and seasonality issues. Secondly, six single time series and three of their proposed ensemble models forecast the clean demand time series. Based on the results obtained, it can be inferred that the proposed time ensemble approach was an efficient and accurate one-month-ahead forecast for electricity demand in the Peruvian electricity market. In addition, the final best ensemble forecasting model (Ensemble (C)) produces the lowest average accuracy error, performing statistically significantly better than those mentioned in the best-proposed forecasting model's literature. Finally, while numerous global studies have been conducted from various perspectives, no analysis has been undertaken using an ensemble learning approach to forecast electricity demand for the Peruvian electricity market.

However, this work only focuses on the Peruvian electricity market; in the future, the proposal of the current research study should be extended to other electricity markets–for instance, the European electricity markets, the United American electricity market, the United Kingdom, the Chinese electricity markets, and so on. In addition, the current work proposal relies on only single linear and nonlinear time-series models; it might use other models like machine learning and deep learning in future projects within the current proposal.

REFERENCES

[1] M. Juanpera, B. Domenech, L. Ferrer-Martí, A. Garzón, and R. Pastor, "Renewable-based electrification for remote locations. Does short-term success endure over time? A case study in Peru," *Renew. Sustain. Energy Rev.*, vol. 146, Aug. 2021, Art. no. 111177.

[2] E. V. Guevara, "Competition and wholesale electricity market: The monitoring task assigned to the Peruvian electricity coordinator (COES)," *Revista IUS ET Veritas*, vol. 29, no. 61, pp. 94–112, 2020.

[3] I. Shah, H. Iftikhar, S. Ali, and D. Wang, "Short-term electricity demand forecasting using components estimation technique," *Energies*, vol. 12, no. 13, p. 2532, Jul. 2019.

[4] S. Li, X. Zhao, W. Liang, M. T. Hossain, and Z. Zhang, "A fast and accurate calculation method of line breaking power flow based on Taylor expansion," *Frontiers Energy Res.*, vol. 10, Jul. 2022, Art. no. 943946.

[5] H. Iftikhar, N. Khan, M. A. Raza, G. Abbas, M. Khan, M. Aoudia, E. Touti, and A. Emara, "Electricity theft detection in smart grid using machine learning," *Frontiers Energy Res.*, vol. 12, Mar. 2024, Art. no. 1383090.

[6] Z. Liu, J. Feng, and L. Uden, "Technology opportunity analysis using hierarchical semantic networks and dual link prediction," *Technovation*, vol. 128, Dec. 2023, Art. no. 102872.

[7] Y. Lei, C. Yanrong, T. Hai, G. Ren, and W. Wenhuan, "DGNet: An adaptive lightweight defect detection model for new energy vehicle battery current collector," *IEEE Sensors J.*, vol. 23, no. 23, pp. 29815–29830, Dec. 2023.

[8] Y. Duan, Y. Zhao, and J. Hu, "An initialization-free distributed algorithm for dynamic economic dispatch problems in microgrid: Modeling, optimization and analysis," *Sustain. Energy, Grids Netw.*, vol. 34, Jun. 2023, Art. no. 101004.

[9] M. Shirkhani, J. Tavoosi, S. Danyali, A. K. Sarvenoee, A. Abdali, A. Mohammadzadeh, and C. Zhang, "A review on microgrid decentralized energy/voltage control structures and methods," *Energy Rep.*, vol. 10, pp. 368–380, Nov. 2023.

[10] X. Zhang, L. Gong, X. Zhao, R. Li, L. Yang, and B. Wang, "Voltage and frequency stabilization control strategy of virtual synchronous generator based on small signal model," *Energy Rep.*, vol. 9, pp. 583–590, Apr. 2023.

[11] H. Iftikhar, "Modeling and forecasting complex time series: A case of electricity demand," M.S. thesis, Dept. Statist., Quaid-i-Azam Univ., Islamabad, Pakistan, Islamabad, Pakistan, 2018.

[12] X. Lin, Y. Wen, R. Yu, J. Yu, and H. Wen, "Improved weak grids synchronization unit for passivity enhancement of grid-connected inverter," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 10, no. 6, pp. 7084–7097, Dec. 2022.

[13] P. Li, J. Hu, L. Qiu, Y. Zhao, and B. K. Ghosh, "A distributed economic dispatch strategy for power–water networks," *IEEE Trans. Control Netw. Syst.*, vol. 9, no. 1, pp. 356–366, Mar. 2022.

[14] I. Shah, H. Iftikhar, and S. Ali, "Modeling and forecasting electricity demand and prices: A comparison of alternative approaches," *J. Math.*, vol. 2022, pp. 1–14, Jul. 2022.

[15] M. Elsaraiti, G. Ali, H. Musbah, A. Merabet, and T. Little, "Time series analysis of electricity consumption forecasting using ARIMA model," in *Proc. IEEE Green Technol. Conf.*, Apr. 2021, pp. 259–262.

[16] O. O. Omogoroye, O. O. Olaniyi, O. O. Adebiyi, T. O. Oladoyinbo, and F. G. Olaniyi, "Electricity consumption (KW) forecast for a building of interest based on a time series nonlinear regression model," *Asian J. Econ., Bus. Accounting*, vol. 23, no. 21, pp. 197–207, Oct. 2023.

[17] S. Krstev, J. Forcan, and D. Krneta, "An overview of forecasting methods for monthly electricity consumption," *Tehnički Vjesnik*, vol. 30, no. 3, pp. 993–1001, 2023.

[18] I. Shah, H. Iftikhar, and S. Ali, "Modeling and forecasting medium-term electricity consumption using component estimation technique," *Forecasting*, vol. 2, no. 2, pp. 163–179, May 2020.

[19] A.-D. Pham, N.-T. Ngo, T. T. Ha Truong, N.-T. Huynh, and N.-S. Truong, "Predicting energy consumption in multiple buildings using machine learning for improving energy efficiency and sustainability," *J. Cleaner Prod.*, vol. 260, Jul. 2020, Art. no. 121082.

[20] M. Khalil, A. S. McGough, Z. Pourmirza, M. Pazhoohesh, and S. Walker, "Machine learning, deep learning and statistical analysis for forecasting building energy consumption—A systematic review," *Eng. Appl. Artif. Intell.*, vol. 115, Oct. 2022, Art. no. 105287.

[21] A. González-Briones, G. Hernández, J. M. Corchado, S. Omatu, and M. S. Mohamad, "Machine learning models for electricity consumption forecasting: A review," in *Proc. 2nd Int. Conf. Comput. Appl. Inf. Secur. (ICCAIS)*, May 2019, pp. 1–6.

[22] F. L. C. da Silva, K. da Costa, P. C. Rodrigues, R. Salas, and J. L. López-Gonzales, "Statistical and artificial neural networks models for electricity consumption forecasting in the Brazilian industrial sector," *Energies*, vol. 15, no. 2, p. 588, Jan. 2022.

[23] H. Iftikhar, J. E. Turpo-Chaparro, P. Canas Rodrigues, and J. L. López-Gonzales, "Day-ahead electricity demand forecasting using a novel decomposition combination method," *Energies*, vol. 16, no. 18, p. 6675, Sep. 2023.

[24] N. Carbo-Bustinza, H. Iftikhar, M. Belmonte, R. J. Cabello-Torres, A. R. H. De La Cruz, and J. L. López-Gonzales, "Short-term forecasting of ozone concentration in metropolitan lima using hybrid combinations of time series models," *Appl. Sci.*, vol. 13, no. 18, p. 10514, Sep. 2023.

[25] H. Iftikhar, A. Zafar, J. E. Turpo-Chaparro, P. Canas Rodrigues, and J. L. López-Gonzales, "Forecasting day-ahead Brent crude oil prices using hybrid combinations of time series models," *Mathematics*, vol. 11, no. 16, p. 3548, Aug. 2023.

[26] W. Ren, N. Jin, and L. OuYang, "Phase space graph convolutional network for chaotic time series learning," *IEEE Trans. Ind. Informat.*, vol. 20, no. 5, pp. 7576–7584, May 2024.

[27] H. Iftikhar, J. E. Turpo-Chaparro, P. Canas Rodrigues, and J. L. López-Gonzales, "Forecasting day-ahead electricity prices for the Italian electricity market using a new decomposition—Combination technique," *Energies*, vol. 16, no. 18, p. 6669, Sep. 2023.
[28] G.-F. Fan, X. Wei, Y.-T. Li, and W.-C. Hong, "Forecasting electricity consumption using a novel hybrid model," *Sustain. Cities Soc.*, vol. 61, Oct. 2020, Art. no. 102320.
[29] S. Ding, Z. Tao, R. Li, and X. Qin, "A novel seasonal adaptive grey model with the data-restacking technique for monthly renewable energy consumption forecasting," *Expert Syst. Appl.*, vol. 208, Dec. 2022, Art. no. 118115.
[30] Z. Hajirahimi and M. Khashei, "Hybridization of hybrid structures for time series forecasting: A review," *Artif. Intell. Rev.*, vol. 56, no. 2, pp. 1201–1261, Feb. 2023.
[31] R. Wang and R. Zhang, "Techno-economic analysis and optimization of hybrid energy systems based on hydrogen storage for sustainable energy utilization by a biological-inspired optimization algorithm," *J. Energy Storage*, vol. 66, Aug. 2023, Art. no. 107469.
[32] P. Pełka, "Analysis and forecasting of monthly electricity demand time series using pattern-based statistical methods," *Energies*, vol. 16, no. 2, p. 827, Jan. 2023.
[33] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Statist.*, vol. 13, no. 3, p. 253, Jul. 1995.
[34] H. Iftikhar, M. Khan, M. S. Khan, and M. Khan, "Short-term forecasting of monkeypox cases using a novel filtering and combining technique," *Diagnostics*, vol. 13, no. 11, p. 1923, May 2023.
[35] X. Lin, Y. Liu, J. Yu, R. Yu, J. Zhang, and H. Wen, "Stability analysis of three-phase grid-connected inverter under the weak grids with asymmetrical grid impedance by LTP theory in time domain," *Int. J. Electr. Power Energy Syst.*, vol. 142, Nov. 2022, Art. no. 108244.
[36] H. Iftikhar, M. Khan, J. E. Turpo-Chaparro, P. C. Rodrigues, and J. L. López-Gonzales, "Forecasting stock prices using a novel filtering-combination technique: Application to the Pakistan stock exchange," *AIMS Math.*, vol. 9, no. 2, pp. 3264–3288, 2024.
[37] M. Meng, D. Niu, and W. Sun, "Forecasting monthly electric energy consumption using feature extraction," *Energies*, vol. 4, no. 10, pp. 1495–1507, Sep. 2011.
[38] W. Zhou, H. Tao, J. Chang, H. Jiang, and L. Chen, "Forecasting Chinese electricity consumption based on grey seasonal model with new information priority," *Sustainability*, vol. 15, no. 4, p. 3521, Feb. 2023.

HASNAIN IFTIKHAR received the B.Sc. degree in computer science and the M.Sc. degree in statistics from the University of Peshawar, Pakistan, and the M.Phil. degree in statistics from Quaid-i-Azam University, Islamabad, Pakistan. Currently, he is employed in a variety of roles at various institutions, including as a Lecturer with the Faculty of Mathematics, City University of Information and Science Technology, Peshawar, Pakistan; a Subject Specialist (Statistics) with the Education Department, Peshawar; and a Young Advisor with the Universidad Peruana Unión, Peru. He has reviewed many international journals, including IEEE Access, Elsevier, Springer, Sage, the Royal Society, and Universal Publisher. His research interests include time series forecasting, regression analysis, energy economics, biostatistics, and applied statistics.

SALVATORE MANCHA GONZALES received the B.Eng. degree in metallurgical engineering from the Universidad Nacional Mayor de San Marcos, Peru, and the master's degree in management operations and supply chain from the Universidad Privada del Norte, Peru. He was the Technical Director's Assistant of CAPESA. He was a Production Supervisor with the Steel Grinding Ball Production Unit, MEPSA. He has been employed as an Engineer with the Metallurgical Laboratory, SEMAN FAP. He has recently completed a second specialty study in applied statistics at the Universidad Peruana Union. During his time at MEPSA, his research interests included modeling the effects of heat treatment process parameters on the mechanical properties of steel grinding balls, continuous cooling of liquid steel in a refractory ladle, heat transfer in the holding furnace for liquid steel, data analysis of electricity demand, electricity prices, and other industrial datasets. At SEMFAP, he is involved in testing the mechanical properties of aircraft components and heat treatments. He is also an International Member of the ASTM and a Follower of Committee E11 on Quality and Stat.

JUSTYNA ZYWIOŁEK (Associate Member, IEEE) received the master's degree in metallurgy, in 2010. In 2014, she defended her doctoral dissertation with honors in management science, specializing in security science. Since January 2015, she has been an Assistant Professor with the Faculty of Management, Technical University of Czestochowa. She is currently a respected Researcher in Poland and abroad for her experience in management issues, such as knowledge and information management and their security. She has published over 160 articles in well-reputed international journals and actively participated in presenting research results at various international conferences. She is also a member of Intellectual Capital Association (ICAA) and EU-OSHA and in organizations operating in Poland promoting information and knowledge security.

JAVIER LINKOLK LÓPEZ-GONZALES (Member, IEEE) received the B.S. degree in statistical and informatics engineering from the Universidad Peruana Unión (UPeU), Peru, the M.Sc. degree in metrology from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Brazil, and the Ph.D. degree in statistics from the Universidad de Valparaiso (UV), Chile. He is currently a Full Professor with the Universidad Peruana Unión. In addition, he is qualified as a RENACYT Researcher. His main research interests include pattern recognition in machine learning, air pollution with deep learning techniques, and time series with singular spectrum analysis.