Energy Demand Forecasting and Optimizing Electric Systems for Developing Countries: A Systematic Review

SAADMAN S. ARNOB[1], ABU ISHA MD. SADOT AREFIN[1,2], AHMED Y. SABER[3], and KHONDAKER A. MAMUN[1,2]

[1]Advanced Intelligent Multidisciplinary Systems (AIMS) Laboratory, United International University), Dhaka 1212, Bangladesh

[2]Department of Computer Science and Engineering, United International University, Dhaka 1212, Bangladesh

[3]Principal Power Engineer (R&D), Operation Technology, Inc. ETAP, CA, USA

Corresponding author: Khondaker A. Mamun (mamun@cse.uiu.ac.bd)

ABSTRACT Currently, developing countries are experiencing a massive shift toward industrialization. Developing countries lack the technical sophistication and infrastructure to encourage low-carbon and sustainable economic growth because of weak public awareness, regulations, and technology. Developing countries must plan the industrialization process for maximum energy efficiency of production, thereby reducing their CO_2 emissions significantly by increasing energy efficiency. This paper attempts to review the current pragmatic methods for forecasting the future load demands from minutes to years ahead in developing countries, following the Preferred Reporting Items for Systematic review and Meta-Analysis Protocols (PRISMA-P). Our primary focus is to provide an optimal forecasting model selection strategy for potential researchers and forecasters. Based on the strengths and weaknesses of the different models, we will discuss the most suitable methods to tailor them to multiple applications and scenarios of load forecasting. The comparison elements are Forecast horizons, Spatio-temporal resolutions, factors affecting the load, different dimensional reduction techniques, model complexity analysis, and the MAPE for error analysis. From the results, We have found ANN hybridized with meta-heuristic techniques to be superior in most of the analysis cases. ANN's ability to handle non-linear data, flexibility, and robustness is why. Consumption data aggregated at the national level can capture trends efficiently. Meteorological and calendar features influence short-term forecasting extensively, whereas economic factors influence long-term load patterns. Finally, we have identified the trends and research gaps from the existing literature, presenting relevant technical recommendations for improvement.

INDEX TERMS Electricity load forecasting, Systematic review, Energy demand forecasting, Artificial neural network, Developing countries, Machine learning, Time series techniques, Long term forecasting, Short term forecasting

I. INTRODUCTION

Load forecasting is how an industrial or utility company forecasts its future consumption. This estimation of future electricity demand influences strategic and operational decisions. The electric utility must accurately predict load to make commitment decisions, reduce spinning reserve capacity, and schedule maintenance plans accordingly [1]. Electric utility companies can avail themselves of load forecasting to minimize cost and operational effects while maximizing revenue.

A massive structural change toward industrialization is currently underway in developing countries. Developing countries lack the technical refinement and infrastructure to encourage low carbon and sustainable economic growth compared to developed countries. Consequently, developing countries that are pro-growth and low-cost are causing an alarming environmental crisis with their low-cost, pro-growth policies [2]. Firstly, developing countries have become increasingly influential in the global growth paradigm over the last three decades. Their economies have transformed from agricultural to energy-driven industries, enjoying higher economic growth over the past three decades [3]. Accordingly, the shares of developing countries in the world GDP, energy consumption, and CO_2 emissions are 19.56, 43.01, and 43.39%, respectively [4]. The findings suggest that developing countries should prioritize reducing CO_2 emissions through energy efficiency. In addition to energy efficiency, improved forecast accuracy contributes to good planning and cost savings.

Many review papers have been conducted on electrical load forecasting models using different approaches and methods and focusing on other criteria for various sectors. What they lack is targeting a specific part of the globe that needs this the most: developing and under-developed countries. The trends in power demand vary between developed and developing economies. In developed countries, demand growth is mostly driven by digitalization and electrification. In developing countries, however, income levels, industrial output, and the services sector have a considerably greater impact on the development of electricity consumption. Therefore, the forecasting techniques for developed countries is not guaranteed to provide a similar outcome for developing countries. The motivation for this particular study is to elaborate on the above-mentioned gap. This article provides a comprehensive overview of the existing electric load forecasting techniques. Table 1 compares the existing review journals and our manuscript.

This paper includes a broader scope of exploration, including load forecasting models. The proposed review discusses the current realistic modeling techniques for forecasting the future load demands from minutes to years ahead in developing countries and the exogenous variables of the forecasting models utilized by numerous researchers and presents relevant technical recommendations for improvement. Based on the strengths and weaknesses of the different models, we will discuss the most suitable methods tailored to multiple applications and scenarios of load forecasting.

The previous systematic reviews primarily provide a generalized solution for load forecasting on a global scale. Our investigation mainly focuses on modeling techniques for developing countries as they need it the most to plan their industrialization correctly. The major contributions of our paper are:

1) Highlights of the top models used for forecasting the energy demand and outlooks of the analysis cases: Data pre-processing, Spatio-temporal resolution, forecasting horizon, model complexity, and error metrics.
2) Statistical analysis of the state-of-the-art models based on various use cases. This type of analysis contributes by providing deeper insights and patterns.
3) We identified the trends observed in this domain for developing countries over the last decade and outlined the potential research gaps and recommendations.
4) Formulation of a unique framework for selecting the optimal strategy for energy demand forecasting based on our analysis and summarized the strength and applications of the models from our analysis.

The outlines for the rest of the paper are as follows - Section II elaborates on the protocols and framework of our systematic review. Section III presents the findings of our systematic review. Afterward, the discussion and recommendations are presented in Section IV and V. Finally, Section VI concludes the article.

II. METHODOLOGY

We followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses Protocol (PRISMA-P), a popular framework widely used across engineering and energy sciences for systematic reviews [6], [11]. The quality of every systematic review depends on the description in the protocol of the investigation's hypothesis, reason for undertaking the investigation, and planned techniques [11]. However, systematic review studies rarely describe the framework of systematic reviews [5]. A detailed, well-described structure for systematic reviews facilitates the understanding and evaluation of the methods adopted. Hence, the PRISMA model was adopted in this study (Figure 1). As shown in Figure 1, the PRISMA presents the flow of information from one stage to another in a systematic review of the literature and gives the total number of the research identified, excluded, and included and the reasons for inclusion and exclusions.

A. STUDY SELECTION STRATEGY

Before identifying the studies to be selected, we need to define the aims, objectives, and research questions in our systematic review. Table 2 defines our study's objectives, scopes, and research questions. The next step is to define the inclusion and exclusion criteria based on which we will select the relevant studies. The inclusion criteria were not too extensive and not too brief. We set the criteria, fulfilling the scope of our study. In Table 3, we defined the selection criteria for commencing our investigation.

TABLE 1. Comparison among existing review articles in the literature and our manuscript.

Reference	Year	Review Type	Target
[5]	2017	Systematic	Identifying scenarios for models suited for their application with a taxonomy for model selection.
[6]	2020	Systematic	Assessing electricity demand forecasting based on methods, time frame, and accuracy metrics.
[7]	2021	Systematic	ANN-based machine learning algorithms and focused on the specific problems tackled by those ANN-based models.
[8]	2021	Systematic	Energy demand modeling regarding: techniques, data, accuracy, energy carriers, sectors, and Spatio-temporal granularity.
[9]	2020	Comprehensive	Single and hybrid type predictive ML models based on short-term load forecasting.
[10]	2021	Narrative	ML approaches address different generation modalities for short-term load forecasting.
Our proposed article	2022	Systematic	Identifying the current pragmatic methods for load demand forecasting from minutes to years ahead in developing countries.

TABLE 2. Formulation of the research question, the objective, and the aim of the study.

Research Statement:	To investigate developing countries' advancement in electricity load forecasting for grid-connected sectors and the various techniques employed for forecasting the future electric load demand for different scenarios.
Objective:	To review the research trends and potential gaps in the current electric load forecasting sector in developing countries and afterward develop a framework from which researchers and forecasters can select the optimal model for forecasting the future load for different scenarios.
Generic Question:	What are the current pragmatic methods for forecasting the future load demands from minutes to years ahead in developing countries, and what are the exogenous variables of the forecasting models?
Research Questions:	RQ1. What are the existing models for forecasting the future load demand in developing countries?
	RQ2. What are the impacts of various exogenous variables in different forecasting horizons?
	RQ3. For high-dimensional data, what are the different dimensionality reduction techniques used in the context of load forecasting?
	RQ4. What was the main focus of the forecasting model (i.e., complexity reduction, adaptability, and higher accuracy)?
	RQ5. What research gaps have not yet been investigated by the research community in developing countries?

B. RESEARCH ARTICLE SELECTION

The potential databases we used for searching the articles are:

- Scopus
- IEEE Xplore
- Wiley online library

TABLE 3. Inclusion and exclusion criteria for article screening.

Inclusion Criteria	Exclusion Criteria
IC1: Articles reviewed by peers in Journals.	EC1 Conference Paper, Review, Conference Review, Reports, Book, Note, Survey, Editorial, and Letter.
IC2: Articles that aim to forecast electric load demand on a large scale from the perspective of a developing country.	EC2: Articles that focus on forecasting electric load demand on Micro-grid and the data validated on a developed country.
IC3: Articles published in the English language.	EC3: Articles published in any language other than English.
IC4: The potential databases for searching records are Scopus, IEEE Xplore, Wiley, ACM, and SpringerLink.	EC4: Articles outside the included database and databases are not available to researchers.
IC5: The publication time frame is from 2012-2022.	EC5: The publication years before 2012.

- ACM library
- SpringerLink

The search keywords embodying our investigations are: "electric* load forecast" , "electric* load predict*", "electric* demand forecast*" and "electric* demand predict*". In a search engine, an asterisk (*) is often used as a wildcard character, which can match any string of characters. For example, the asterisk after the word "electric" would match any string of characters that follows "electric", such as "electricity", "electrical", or "electronic". This allows the search to find more relevant results, as it is not limited to searching for the exact term "electric". Our search for articles is limited to studies published between 2012 and 2022. For our review, we only selected Articles that are published in the English language. We chose the PRISMA flow diagram for this study because it can be used for systematic reviews of various topics and improve the quality and value of the reviews. We screen articles in accordance with the PRISMA framework. When articles are searched,

they are exported as Comma Separated Values (CSV) files with their abstracts. A single file is created by combining all the records from different databases and removing all duplicates. Upon receiving unique articles, we perform the first screening by going over their titles and abstracts. A complete match between the exclusion criteria and the article is automatically dropped at the second screening. Otherwise, the article proceeds to the second screening. The second screening, which considers the full text of articles, includes and excludes them according to the criteria set.

TABLE 4. Generalized search string for the databases.

TITLE-ABS-KEY ('electric load forecast*' OR 'electric* load predicti*' OR 'electric* demand forecast*' OR 'electric* demand predicti*') AND (LIMIT TO (PUBSTAGE , 'final')) AND (LIMIT-TO (SRCTYPE,'j')) AND (LIMIT-TO (SUBJAREA , 'ENGI') OR LIMIT-TO (SUBJAREA , 'COMP') OR LIMIT-TO (SUBJAREA , 'ENER'))

C. DATA EXTRACTION STRATEGY FROM THE INCLUDED STUDIES

For performing the empirical analysis of the studies included, we formulated a data extraction strategy drawing out relevant information. The point of doing this is to reduce the biasness and to generalize the results. Setting pre-defined criteria helps with the extraction process. We aim to systematically evaluate the load forecasting methods for grid-connected sectors in developing countries. So, categorizing the information into segments provides flexibility to the whole process. In Table 5, we selected the features of the articles and grouped them for quantitative analysis and drawing a conclusion from them. We stored the extracted information about the articles in a CSV file for efficient comparison and analysis.

TABLE 5. Data extracted from the selected articles for performing quantitative analysis.

Specifications
Title of the article
Publication year
Forecasting horizon
The Spatio-temporal resolution of the electric load data
Dimensionality reduction techniques
Optimization techniques combined with a forecasting model
Study validation countries
Evaluation metric results
Experiment setups
Methodologies used, the best ones, and the evaluation results for analysis
Factors contributing to forecasting the electric demand and crucial information in the articles

III. RESULTS

In this section, we highlighted the findings of our systematic review. First, we provided an overview of our search results from the different databases. Secondly, we define some analysis cases which will be the basis for our analysis. There, we presented an overview of the analysis cases and finally, we performed quantitative analysis of our cases.

A. SEARCH RESULT

A total of 1841 articles are identified throughout all the databases using the generalized search string. Out of all the databases, Scopus yielded the highest article count at 808. All the other databases like SpringerLink, IEEE Xplore, Wiley and ACM digital library returned a similar number of articles ranging from 200-300. Afterward, we isolated 116 duplicate articles from the databases. However, most of the articles that ACM search engine returned were completely unrelated with 5 articles passing the initial screening. The search result that Scopus had were pretty much on point as majority of them passed the initial screening. A total of 841 articles passed the first screening with Scopus encompassing majority of the articles. Next, following the eligibility criteria in Table 3, we performed a full-text analysis of the initially screened articles. The final number of articles included in our systematic review was 111. Figure 1 illustrates the PRISMA flow diagram of our article selection. The article selection stages for all the databases are provided in Table 6.

TABLE 6. Statistics on the selection of articles for the review in different stages of screening from the databases.

Database	No. of articles after database search	No. of articles after 1st screening	No. of articles after 2nd screening
Scopus	808	395	73
IEEE Xplore	291	149	14
Wiley	298	153	11
SpringerLink	260	139	13
ACM Digital	204	5	1
Total	1861	841	111

B. OVERVIEW OF THE ANALYSIS CASES

For the unbiased and generalized assessment of the forecasting methods, it is necessary to draw out some analysis cases. These cases will draw out critical patterns on when and under what circumstances we should use the models. The analysis cases we will use for quantitatively synthesizing the forecasting models are drawn in Table 7. The properties of the models in the 111 included studies are summarized and analyzed using the analysis cases defined in Table 7. This will provide us with a better insight into the model selection process. After the analysis, the articles are sorted using structured tables for easier accessibility.

TABLE 7. Cases for quantitative analysis of the forecasting models.

Analysis Case	Description	Possible Values
Model inputs	Predictor variables affecting the load consumption curve	Historic load, calendar information, weather, economy, demographics, environment, prices
Data pre-processing techniques	Techniques used for pre-processing the raw data	Outlier removal, feature selection, and extraction techniques
Forecasting horizon	Timespan forecasted by the models	Very short-term, short-term, mid-term, and long-term
Forecasting resolution	Time intervals used by the models for prediction	sub-hourly, hourly, daily, weekly, monthly, and yearly
Spatial resolution	The smallest entity of energy consumption	National, regional and city level
Forecasting techniques	Generalized techniques used for forecasting	ANN, RNN, Stochastic time series, Machine learning models, and Fuzzy logic and clustering models
Study validation countries	Countries where the experiment was validated	Developing countries
Model complexity	Computational time and the tools used for the forecasting setup	Python, TensorFlow, MATLAB, and temporal values
Prediction accuracy	Evaluation metrics for assessing model's performance	MAPE values
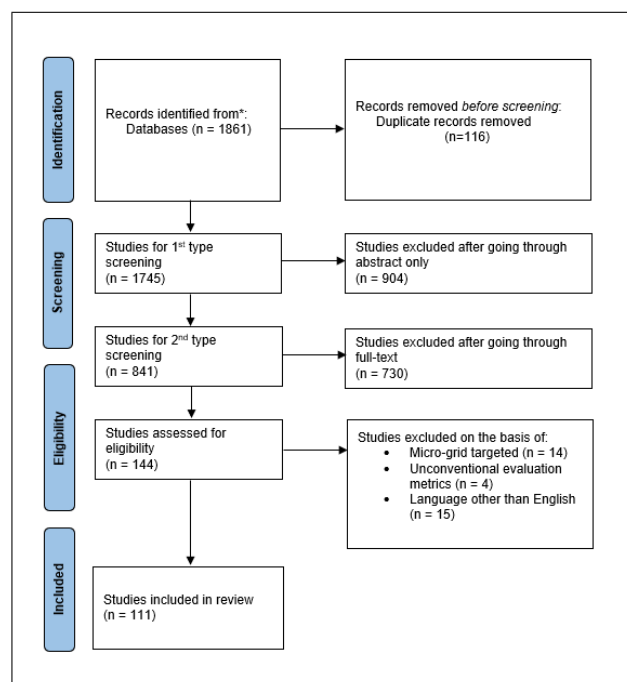
FIGURE 1. PRISMA workflow for the systematic search of the literature.

1) Forecasting horizon

Time frame is one of the most critical characteristics of load forecasting [12]. We categorized electricity load forecasting into four main types based on the forecast horizon. They are:

1) Very Short Term Load Forecasting (VSTLF):
   VSTLF method predicts electricity load in the next few minutes to hours [13]. Real-time safety analysis and monitoring of power systems are its main applications. It is difficult to accurately pre-

dict very short-term load factors because electrical loads are heterogeneous and random [14]. Energy Management Systems can perform economic dispatch by monitoring load frequency when dealing with load forecasts with short time horizons.

2) Short Term Load Forecasting (STLF):
   STLF is a method of forecasting loads over the next one-day period to one week in advance [15]. Its main uses are economic dispatch, fuel purchase scheduling, fuel allocations, and real-time control. STLF is essential for power systems as it helps make critical decisions like unit commitment and load switching [16].

3) Mid Term Load Forecasting (MTLF):
   MTLF is concerned with predicting electricity load in the coming weeks or months [17]. Its purpose is to save operating costs by planning maintenance scheduling and coordinating dispatching loads [18]. For the upcoming months, the price fixing and meeting demands, MTLF is essential. This will improve supply reliability.

4) Long Term Load Forecasting (LTLF):
   LTLF method is used to plan power systems over the long run [19]. The forecasting time range is three years up to decades. A better understanding of how the system works and how economic and technological changes can affect the electricity market is essential for long-term forecasting.

Figure 2 shows the applications of load forecasting over different time horizons.

From Figure 3, we can infer that most forecasting methods are mainly short and long-term focused. Predictions for the very short-term and mid-term are uncommon. It is because the industry needs short-term primarily and long-term forecasts. Short-term predictions are often used to respond to electricity demand

FIGURE 2. Electric load forecasting applications based on forecast horizon methods.

rapidly, while long-term predictions are used to develop strategies and plans.
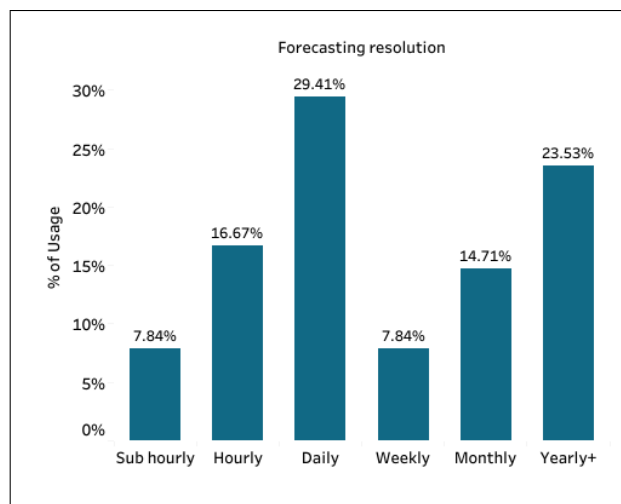
2) Forecasting resolution

Forecasting resolution is the time intervals considered for sequence-to-sequence prediction [8]. Single-step ahead forecasting predicts the following load values from historical load values. Multi steps ahead forecasting involves forecasting multiple sequence values in a time series. Power utility planners and demand controllers rely heavily on multi-step forecasting to ensure necessary generation for the next few hours or even days [20].

Figure 4 illustrates that day-ahead predictions are most commonly accessed. Since production depends on intra-day electricity market negotiations, this is especially relevant. Next comes annual consumption prediction and hour-ahead predictions. Hour-ahead predictions are significant for optimal economic scheduling of generating capacity and generator unit commitment. On the other hand, year-ahead forecasts are crucial for energy policymakers. System planning and staff recruitment also dramatically relies on the Year ahead forecasting.
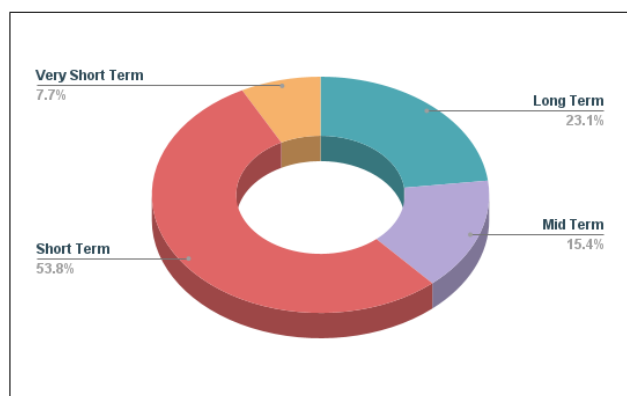
FIGURE 3. Distribution of forecasting methods' usage based on the forecast horizons.

FIGURE 4. Distribution of forecasting methods' usage based on the forecast resolutions.
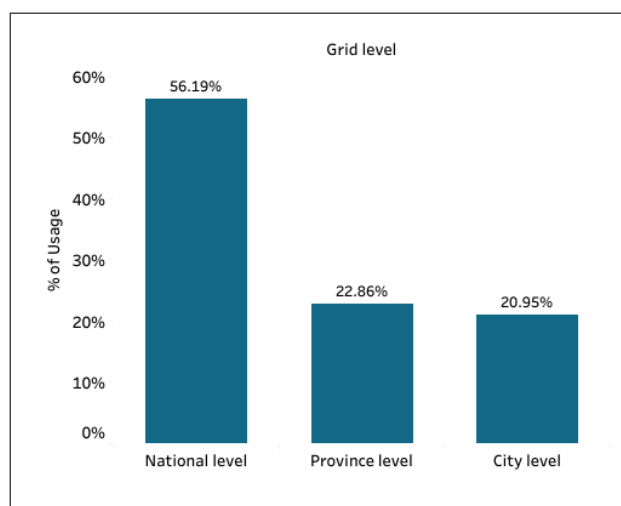
FIGURE 5. Distribution of research based on the spatial granularity.

3) Spatial resolution

It is becoming increasingly difficult to predict load profiles due to external factors such as weather, emergencies, stochastic human activities, etc. Their correlation might be explained by the distribution of load in space scales and the historical load curve. To predict future electric behavior on a larger scale, we must consider both time and space regions [21]. Throughout the articles included, the spatial resolution is determined by the smallest entity consuming the most electrical energy. In our study, we had three levels for investigating the detail of models. They are national level, province level, and city level. Figure 5 shows the distribution of entity-level consumption data used for forecasting.

The majority of the analyzed literature focuses on consumption at the national level/country level. District and city-level consumption come next with a similar implementation rate. Our study concentrates on large-

scale forecasting, so we excluded studies with consumption at residential, commercial, and appliance levels. The majority of forecasting focuses on consumption at the national level.

4) Predictor/Input variables

Electricity and power forecasting are not always linear, especially in developing countries where energy demand and consumption are highly erratic due to a shortage of resources [22]. Multiple factors come into play when energy forecasting is concerned. External factor-independent models cannot explore the effect external factors have on the prediction and thus produce prediction errors. Therefore the majority of the time, factor-dependent models outperform univariate models. A forecasting model inhibits different explanatory variables for each application according to its forecast horizon.

People productively use power when they have to work or study on weekdays; however, due to the internet growth of gadget use, power consumption does not decrease significantly since people keep in touch with each other through social media on weekends. On national holidays, some people seem to go out for a holiday, so there is a decrease in power consumption.

The exogenous variables contributing to better prediction accuracy for VSTLF are temperature and historical load [13]. Calendar features like an hour of the day, workday, weekend and holiday, day of the week, Week number in a year, and Month number in a year are sometimes incorporated for VSTLF [1]. In Table 8, the classification of the contributing factors is illustrated.

In the case of STLF, the factors primarily responsible for influencing consumers' electricity consumption patterns are mostly load, calendar, and meteorological variables. Weather parameters have a high correlation with electricity consumption. In a study by Mir et al. [12], air conditioning generally consumes more electricity during hot weather due to increased temperatures. Conversely, in winter, electricity consumption tends to fall as temperatures drop. Nevertheless, the increase in load demand in summers is more significant than in winters, resulting in higher demand for utilities. In the case of calendar features, they mainly incorporate the effects of human activities on electricity consumption. In addition, weekends tend to be less demanding than weekdays. In another study by Han et al. [23], due to the long-range dependence on the electric load and the occurrence of certain factors once a year, such as festivals, holidays, and seasons, the load data from the same events can be used to predict the future load for events like the Spring Festival and Mid-Autumn Festival in China.

For MTLF and LTLF, economic variables are incorporated. As the economy develops rapidly, electricity demand is expected to increase. For example, a larger population and growing industry will affect the trends in industrial electricity consumption. The electricity price

similarly affects the consumption pattern in the long run. Consumers can be significantly reduced when the price is higher than the capacity of enterprises [24]. As a second factor, price affects the distribution of intensive electricity consumption in the industrial sector.

5) Data pre-processing techniques

Real-world data is always noisy, inconsistent, and incomplete. If the raw data is not pre-processed and fits into forecasting models, it is likely to give higher prediction errors. It is impossible to produce accurate and good results with dirty data. In essence, dirty data is transformed into clean data. Data pre-processing mainly deals with this noise, outliers, and missing values. To improve the performance of the forecasting model, data pre-processing consists mainly of 4 steps: data cleaning, data aggregation, dimensionality reduction, and data transformation [25]. Especially for fitting those data into machine learning algorithms, data pre-processing is essential. Data pre-processing is divided into four stages:

- Data Cleaning
- Data Aggregation
- Dimensionality Reduction
- Data Transformation

a: Techniques to deal with outliers, noise, and missing values

Data cleaning mainly deals with missing data, outliers, and noisy data. Jawad et al. [26] interpolated the missing data rather than discarding them, which helps in obtaining real-time data patterns. They used the method of 'Means by nearby points'. Other ways are filling the values with the mean or the last observed value. In some cases, the duplicate rows are dropped. A famous method for removing outliers is using the interquartile range, where IQR = Q_3 - Q_1, representing the interquartile range. In an IQR, outliers are data that lie outside the upper and lower edge lines [27]. A study used the extreme value constraint method (EVCM) to deal with outliers [28]. Zhang et al. [21] proposed a k-medoids clustering approach to remove the effect of outliers and noise. A local regression filtering technique can also be used to remove outliers [29]. However, Zheng et al. [30] replaced the outliers with an average value of adjacent points.

b: Techniques to transform the data

Data normalization prevents errors and weights from being dispersed too widely during training. In most cases, the data must be scaled within the [0-1] range or [-1 1]. In a study by Houimli et al. [31], Mapminmax normalization method was adopted to normalize data. Cai et al. [32] normalized the samples using the z-score normalization method. In most cases, min-max normalization is done. Normalization can reduce the error

TABLE 8. **Input variable classification to be incorporated into forecasting models**

Input Variables	Possible values
Load variables	Current load, Historic load, Lagged load, and previous loads
Meteorological variables	Temperature, wind speed, pressure, humidity, heat index, precipitation, dew point temp, skin temp, soil temp, and surface thermal radiation downwards.
Calendar variables	Hour of the day, workday, weekend and holiday, Day of the week, Week number in a year, Month number in a year, seasons, post-holiday, and bridging holiday.
Economic variables	GDP, Total retail sales of consumer goods, Total merchandise exports and imports, Gross industrial output value, Population, Import, Industrial rate, Second industry proportion, and Fixed assets investment.
Energy prices	Real-time electricity price and generation cost.

percentage by 5.99% to 7.85% [25]. Data normalization also affects the runtime of the models, which is critical for short-term load forecasting. In a study, the influence of various data normalization methods (z-score, min-max, max, decimal, sigmoid, and softmax) on the run time of the models was studied [33]. According to that study, decimal methods had the shortest run time of 32 seconds, but the softmax normalization method gave the lowest RMSE. Copula notion is another method adopted [34], where data normalization is achieved by using Copula, a pre-processing technique that enhances peak load forecasting precision. Calendar features are mostly categorical variables and not continuous in nature. When working with a regression model, calendar features need to be recoded before they can be used in the model. This is because most regression algorithms can only handle numerical data. Zeng et al. used six binary features to encode the calendar index, which stands for Monday–Saturday [35].

c: Techniques to reduce dimensionality
Dimensionality reduction is a must for improving model performance and reducing computational time. In machine learning, the concept of the 'Curse of Dimensionality' is genuine. Several algorithms have trouble training effective models when the number of features relative to observations is huge. The "Curse of Dimensionality" refers to this phenomenon. Features are crucial inputs to the forecasting models for getting better predictions. Dimensionality reduction has mainly two sub-categories [36]:

1) Feature Selection techniques
2) Feature Extraction techniques

d: Feature selection techniques
Through feature selection, model complexity can be reduced, computational efficiency can be enhanced, and generalization error reduced by selecting a subset of features from the original set. Three different methodological categories can be used to classify the feature selection methods: 1) Filtering, 2) Embedded, and 3)
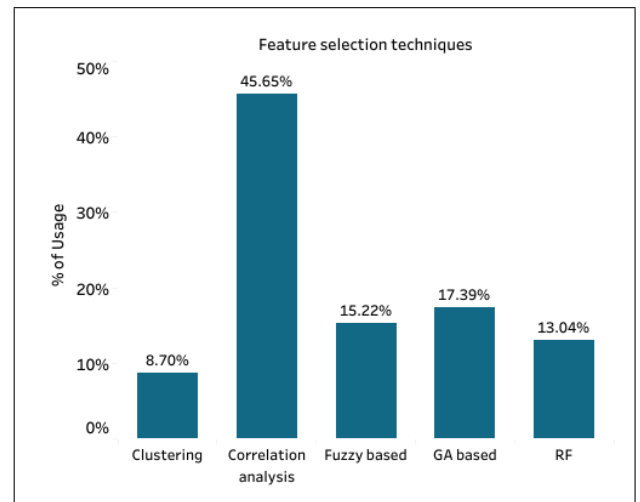
FIGURE 6. **Feature selection techniques used in the forecasting model.**

Wrapper. The primary distinctions between these approaches are as follows: 1) the degree to which the process of feature selection is handled independently or whether it is incorporated into the learning algorithm as a whole; 2) the evaluation metrics; 3) the level of computational complexity; and 4) the ability to identify redundant features and interactions between features.

According to the bar chart in Figure 6, Correlation analysis is mainly prescribed with a percentage of 45.65%. Other techniques are RF-based, GA-based, fuzzy-based, and clustering techniques.

Wrapper methods use prediction accuracy (out-of-sample) as a quality criterion for evaluating the appropriateness of a feature subset by conducting an exhaustive search on a large pool of candidate features [37]. This contrasts the filter method, which uses only in-sample data. Meta-heuristics algorithms are particularly used for their optimum searching ability. Genetic algorithm (GA) is one of the most commonly used feature selection techniques. In the article by Shiekhan et al. compared GA-ACO with other feature selection methods (ACO, PCA) and found that GA-

ACO outperforms them [38]. Sarkar et al. compared two feature selection techniques: Pearson correlation coefficient and Recursive Feature Elimination (RFE). The findings there suggest RFE to be more accurate as it aims to find best performing feature subset and ranks the features based on the order of elimination [39]. In another study, HCSA is adopted for the feature selection process. Essentially, the best inputs-output relationship is modeled using the HCSA method, and the outputs of all models are combined to forecast the future [40]. ANFIS is another powerful fuzzy-based feature selection technique adopted [41], [42]. A genetic algorithm (GA) is one of the most commonly used feature selection techniques.

Filter methods might be viewed as a more straightforward and expedient alternative to wrappers. To evaluate the appropriateness of each feature, they simply analyze its statistical relationship based on evaluation criteria such as correlation analysis (CA) or mutual information (MI) as proxies for the model's performance metric. Pearson correlation coefficient is one of the most used methods to remove redundant features by analyzing the feature correlation matrix. Javed et al. explored various Exploratory Data Analysis (EDA) methods used to select features [43]. Auto-Correlation Analysis is one of the techniques they used for de-trending the time series. In order to demonstrate the significance of previous data, [44] applied the partial auto-correlation function (PACF) to the relationship between historical and current load. MI is ideal for the 'information content' evaluation of features in complex classification tasks since it assesses the arbitrary dependence between random variables. However, a significant redundancy issue has arisen because MI does not properly account for the interdependence between candidate variables [45]. Gao et al. proposed partial mutual information (PMI) based filter to overcome the limitation of MI [45]. It must be noted, however, that the statistical regression algorithms are insufficient to capture the space and time variations and non-linear patterns of electrical loads in the case of short-term load forecasting [43]. Clustering divides input data points into different groups based on some similarity measures. Using cluster analysis, data points belonging to the same group should be highly similar, and points belonging to other groups should not be highly similar. Bedi et al. applied K-means clustering for the electricity data to find similar consumption patterns between groups of months [25].

Embedded methods combine filter methods with wrapper methods. It is implemented by algorithms that have their own method of feature selection built-in. LASSO, tree-based, and permutation importance (PI)-based feature selection are among the most well-known examples of these techniques. Different feature subsets are challenging to predict with SVR and ANN, which makes feature selection difficult [27]. The Random Forest
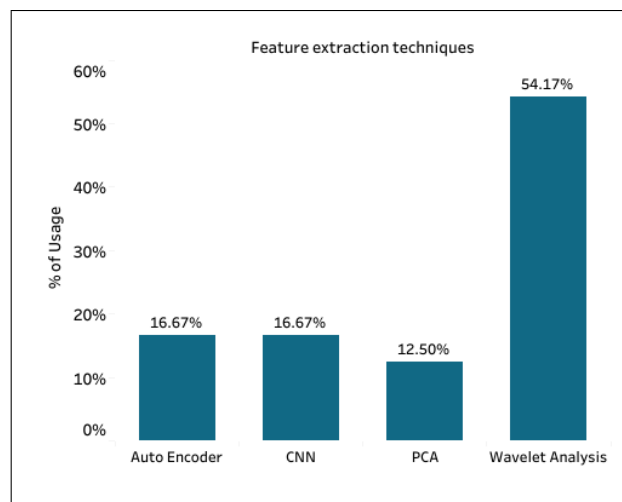

FIGURE 7. Feature extraction technique used in the forecasting model.

(RF) method of feature selection for high-dimensional feature sets reduces the number of iterations, increasing the search strategy's efficiency [27], [46]. When the original feature set is utilized for training a RF model, the permutation importance (PI) of each feature for the prediction can be determined throughout the training process. On this premise, the sequential backward search (SBS) approach can be applied to choose the ideal features. Thus, when RF is utilized for load forecasting, complicated feature selection methods are not required [27]. During the model creation process, the authors utilized a feature selection technique in order to create an accurate and reasonably straightforward model. Elastic nets provide the capacity to pick aggregated or pooled variables, a feature that lasso regression lacks [12].

A concise summary of the methods is presented in Table 9. Table presents the methods and assesses the pros and cons of the methods based on various factors affecting the performance of the feature selection methods.

e: Feature extraction techniques

While feature selection keeps the more relevant variables from the source data, new features are generated using the feature extraction technique by linearly combining the existing elements. It is all about capturing the same information with fewer features. From Figure 7, we observed that five different techniques were mainly used for extracting useful features from the data source in the last decade. The decomposition technique based on wavelet analysis comes at the top by quite a large margin. Others, including Convolutional Neural Networks (CNN), Principal Component Analysis (PCA), Auto-Encoder, etc., have lesser usage.

Decomposition techniques help estimate the time series separately. There are various decomposition tech-

TABLE 9. Assessment of different feature selection methods based on various factors. The number of stars represents the levels of factors considered.

| Methods	Factors				Remarks
	Computation cost	Risk of over fitting	Generalization	Runtime	
Filter	*	*	*	***	1. No interaction with classification model. 2. Ignores feature dependencies.
Wrapper	***	***	**	*	1. Not feasible for a large number of features. 2. Includes feature dependencies.
Embedded	**	**	***	**	1. Interacts with the classification model. 2. Gives better error generalization with more data.

niques in existence. They are discrete wavelet transform (DWT) [47], wavelet packet transform (WPT) [48], empirical mode decomposition (EMD) [13], [25], ensemble empirical mode decomposition (EEMD) [40], [49], complementary ensemble empirical mode decomposition (CEEMD) [50], variational mode decomposition (VMD) [50], and dynamic mode decomposition (DMD) [28]. Wavelet Transform (WT) provides a powerful means of managing non-stationary load behavior by converting a time series into a set of constitutive components. Saleh et al. [47] used DWT to decompose the ARIMA residuals. Its main advantage is to capture the non-linear factors. Another advantage is its ability to retain both frequency and spatial information in the temporal resolution giving DWT the edge over Fourier analysis [51]. The application of dynamic mode decomposition (DMD) is limited to a certain extent by the sample dimension [28] as it is based on error correction. Again, EMD limits mode aliasing, leading to false distribution of time-frequency. EEMD has been proposed as a solution to EMD's pattern mixing problem. The decomposition rate of EEMD is slow, and the method cannot eliminate all white noise. However, a comparison between EMD and wavelet analysis shows that EMD is more effective than wavelet decomposition in the VSTLF, as the MAPE increases by 9.62 and 5.68%, respectively [13]. A comparison between all the decomposition techniques for multi-step ahead forecasting shows that, for one-step, three-steps, and five-step forecasts, VMD ensemble models give the best performance [50]. Nevertheless, for specific use cases, the result can be different. For example, wavelet packet decomposition (WPD) gives better results than DWT [48]. Though fundamentally the same, the detail and approximation coefficients are decomposed in WPD. This helps extract relevant high-frequency components, which is crucial for peak demand forecasting.

PCA and sensitivity analysis is done to overcome the limitations of statistical regression analysis. Cheng et al. [13] verified that using PCA on the model boosted the MAPE performance by 6.9%. A variant of autoencoder that got the limelight was a variational autoencoder (VAE). VAE adds a few restrictions to the encoding process, resulting in roughly a standard normal dis-

tribution of the generated latent vector. Yang et al. [50] compared VAE and A.E. and observed that in the region of Nanjing, MAPE decreased by 61.5%, 23.8%, and 32.1% in one-, three-, and five-step-ahead forecasts using VAE. CNN cascades features of different levels, effectively extracting relationships between continuous and discontinuous data and fusing all the levels for a compelling feature vector, which can be used to make predictions [52].

6) Evaluation metrics

To evaluate a model's performance, forecasting accuracy is crucial. Various error functions can determine the performance of forecasting and prediction algorithms [43]. To assess the model's performance, the following evaluation metrics are used: Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Square Error (MSE), and R-squared value. Figure 8 shows the distribution of the metrics used for evaluating the performance of models. There is no doubt that MAPE is the most widely used metric [28]. Figure 8 also demonstrates that the articles in our investigation have MAPE as the highest implemented evaluation metric.
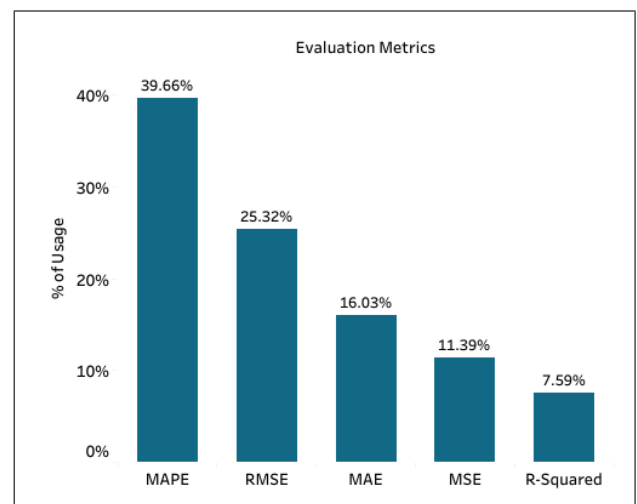
FIGURE 8. Implementation of various evaluation metrics.
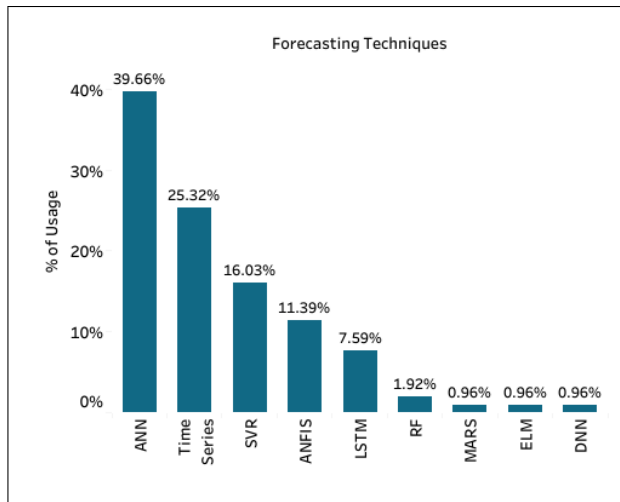
FIGURE 9. Distribution of different forecasting models included in our study.

7) Energy demand forecasting models

Of the 111 studies included, nine unique forecasting models have been highlighted. Table 10 illustrates the categorization of the five unique models.

According to Figure 9, the distribution of the models is illustrated. From Figure 9, we observed that ANN had been implemented as the most used forecasting model, roughly (37.5%) out of all the papers. The time series techniques occupy the second position (18.27%). They are followed closely by SVR models (16.35%) and the LSTM techniques (13.46%), with ANFIS coming at last (9.62%). The other models are in insignificant proportions. Indeed we can conclude that ANN and stochastic time series models are famous among researchers as they prefer their usability and applications in the load forecasting domain. This is why they are considered top models in the field. Despite their low prevalence, SVR, LSTM, and ANFIS have an established framework on which several studies have built.

a: Artificial Neural Network (ANN models)

- Single models

Traditional modeling techniques have difficulty modeling complex relationships and non-linear relationships. Due to its ability to solve the problem, ANN is more prevalent among others. Multilayer perceptrons (MLP) and backpropagation (BP) are used to design the ANN. ANN is composed of three layers: input, hidden, and output layers. There has always been a challenge in creating an ANN model for peak load forecasting because choosing the input variables, constructing the model, and obtaining training data are important considerations [53]. A successful ANN relies heavily on determining the appropriate network structure. Primary consideration should be given to the number of hidden neurons and the selection of hidden and output layer activation functions

[54]. Another study adopted a new method of selecting the hidden layer neurons of feed-forward neural networks [55]. Azadeh et al. [16] proposed a seasonal ANN model for STLF that can easily handle non-linear and complex relationships, use exogenous variables as inputs, and handle corrupted data more effectively than regression models. One limitation of ANN is the slower convergence to the actual value [22]. Using a backpropagation algorithm, an auto-encoder aims to make its output values as close as possible to its input values. This technique has some limitations as well. To tackle the overfitting problem proposed by the SDAE technique, a Denoising Auto-Encoder (DAE) is adopted, which corrupts the input data to prevent the over-fitting problem. Still, the gradient vanishing problem remains. Liu et al. [56] proposed a stacked DAE model which uses a greedy algorithm to pre-train and fine-tune to fix the gradient vanishing problem. BPNN is well known for its strong generalization capability and quick convergence toward actual data. However, for multistage and weather fluctuations, the issue of low accuracy still exists. Wang et al. [57] removed this limitation by integrating a maximum deviation similarity criterion (MDSC) of time segment to BPNN. Another comparison between SARIMA and ANN shows that, even though the error trend appears to be the same, ANN with calendar data and previous loads outperforms SARIMA. Regarding predicting with a high degree of accuracy, calendar data and load feature sets work best on ANN [58]. The multivariate forecasting success is again observed in a study [59] comparing ANN with LSTM in LTLF. The usage and applicability of ANN in STLF are significant as the research community heavily relies on this method. This concludes that for unit commitment, fuel allocation, economic scheduling of generation, maintenance scheduling from demand side management, etc., ANN is pretty much reliable.

- Hybrid models

Computing efficiency, complexity, and high error percentage are often disadvantages of single methods for load forecasting. Hybrid load forecasting methods and models have been developed over the years to achieve greater accuracy with a lower error rate by overcoming the limitations of single models. The purpose of hybrid models is to make forecasting more accurate, efficient, and effective by combining two or more different methods.

In MLP-based ANN architecture, one of the significant disadvantages is that it requires large MLP structures to train the actual data [9]. Various meta-heuristic methods are integrated into ANN for tuning hyperparameters. In 2021, Golilarz et al. [60] implemented PSO-ANN for STLF. Particle Swarm Optimization (PSO) also has reduced runtime and usage, i.e., faster convergence than other meta-heuristic techniques. Ardadkani et al. [19] proposed an improved PSO-ANN

TABLE 10. Classification of the forecasting models.

Forecasting models	Included techniques
ANN	ANN, MLP, DNN, DBN, and other hybrid models of ANN
RNN	LSTM, GRU, BiGRU, BiLSTM, and other hybrid models of RNN
Stochastic time series	Time series techniques like A.R., MA, ARMA, ARIMA, SARIMA, grey modeling, etc.
Machine learning models	Different kernel-based models of SVR, LSSVM, and other hybrid variants of SVR
Fuzzy techniques	ANFIS, Fuzzy c-means, KmFuzz, and other hybrid variants of ANFIS

for long-term electrical energy consumption by introducing a mutation operator to overcome the limitations. Another case study in Jordan [61] used a grey wolf optimizer (GWO) to optimize the neural network. GWO can control the optimal convergence rate and find the optimal hidden node numbers. fNN-GWO outperformed other meta-heuristic algorithms (fNN-GA, fNN-PSO, fNN-ACO, fNN-ES, and fNN-WOA). A modified GWO-ANN was proposed by Rizk et al. [62], overcoming the limitations of GWO-ANN. The main advantage is lower convergence time and error. The model's run time is lowest at 39.5s compared to other hybrid algorithms. For VSTLF applications, the ANN-LM hybrid approach was proposed to determine the optimal structure for neural networks [31]. The MAPE values ranged between 1.1%-3.4%, which confirms the performance of this algorithm as evaluated by evaluation indicators. For forecasting short-term power demand or load from developing countries such as Ghana, where power stability is a significant concern, the study [63] proposes the DWT-PSO-RBFNN method to remove the irregularities within a time series through a decomposition technique like DWT and optimizes the parameters for the prediction stage. It is capable of handling volatile and complex non-linear time series data. A unique approach, ANN-GA-WT, was proposed for day-ahead peak load forecasting for Iran's national grid, [64], where the signal is decomposed through WT. Then GA was used for hyperparameter tuning the neural network. This resulted in a training time of 95s instead of 50 mins and a lower error. Another feature selection-based hybrid approach was proposed to deal with complex non-linear time series data [38]. MLP is used for hourly load prediction here, and GA-ACO is used for feature selection. Several factors are considered in this study that influence load, including the weather and calendar features. Another experiment validated in Thailand [65] proposed ANN-GA-PSO where PSO-GA was implemented in response to the limitations of BP, where it can get trapped in local minima. The introduction of GA helps PSOs escape local minima, even if they get stuck in them. ANN applications in LTLF are not as rich as STLF. The author et al. [24] propose a hybrid PSO–GA-RBF for long-term forecasting. By self-generating a unique RBF model, it eliminates the drawbacks of traditional learning methods, such as

identifying problems and selecting parameters for RBFs. Wu et al. [66] proposed an online training algorithm that uses a rolling mechanism to save computational costs. There is no need to select the inputs for the network when choosing the single multiplicative neuron structure (SMN). The forecasting model's parameters can also be updated when a new set of observed data arrives (being online) so that it can accommodate, over time, changes in the model's parameters. The only limitation is that they work on small datasets.

b: Recurrent Neural Network (RNN models)
- Single models

Recently, the application of RNN has been exponentially rising due to its effortless handling of long-term dependence, which is crucial for time series forecasting. A recurrent neural network (RNN) is a form of neural network suitable for sequential input data such as time series. RNN can retain the memory of what it has processed through temporal feedback loops in contrast to Feed-forward neural networks (FFNNs). However, RNNs have limitations that are followed up and fixed in their extensions, like LSTM, GRU, BiLSTM, BiGRU, etc. In 1997, Hochreiter and Schmidhuber proposed long short-term memory (LSTM), which many research groups modified. LSTM is a variant of RNN which acts as a solution for vanishing and exploding gradients. GRU is modeled similarly to deal with the same issue [67]. Another variant is BiLSTM which utilizes both forward and backward propagation in contrast to conventional LSTM models. When forecasting seven days in Turkey, the algorithm has a R^2 of 0.73, down from 0.94 for one day ahead [67]. However, the study also showed the limitation of LSTM in LTLF and multivariate forecasting [68]. Also, Large-scale power data, which requires many neurons, cannot be predicted using LSTM's step-by-step prediction method [68]. NARX (non-linear autoregressive neural network with exogenous inputs) is another variant of RNN used for long-term forecasting of load [69]. Compared to other techniques under different study periods and forecasting horizons, the NARX technique proved superior to other econometric approaches in all studied cases. Another comparison between LSTM and ARMA, ARIMA, and SARIMA shows that LSTM has the slightest deviation in prediction, which is only 3.2%,

opting for its applicability in generation and distribution planning [70]. Phyo et al. [29] observed in their experiment that LSTM outperformed the Deep belief network (DBN) with a yearly average MAPE of 3.79% only through smoothing out raw data.

- Hybrid models

RNN models are a better choice for time series data but still have some limitations, which a hybrid model fixes. RNN is a viable option for univariate modeling using only its time lags. Validated in Bangladesh, Rafi et al. [71] adopted a hybrid model of CNN-LSTM for STLF, which utilizes the advantages of both CNN and LSTM. A study suggested a hybrid CNN-BiLSTM-AM model overcome the multivariate feature usage limitations of LSTM [72]. BiLSTM filters invalid data using the CNN layer into the max-pooling layer to choose the top features from the captured data features through feature maps with activation functions. AM (Attention mechanism) optimization limits computation resources to vital parts. Another similar model CNN-BiGRU was adopted by Xuan et al., [46], which gives a higher prediction as GRU has lower parameters affecting the prediction than LSTM. To deal with the limitations of multivariate load forecasting issues in LSTM, Javed et al. [73] proposed a hybrid SRDCC-BiLSTM with improved generalization capability for multi-step and multivariate STLF. Compared to CNN-LSTM, the proposed approach exhibits 35% greater accuracy and can capture local trends in electrical load patterns. To model the periodicity and volatility of electricity consumption data, Zheng et al. presented a Time-Frequency Variational Autoencoder (TFVAE) that simultaneously extracts frequency features and inherent temporal features outperforming other hybrid auto encoder LSTM methods [30]. For electricity demand time series forecasting, EMD-LSTM uses EMD to capture trends for non-stationary signals and deal with the non-linear features (weather, location, timing, etc.) of electricity time series data [25]. Another similar but improved approach was introduced, namely, VAE-VMD-LSTM, where in addition to reducing the data dimensions, VAE contributes to extracting adequate information from components by incorporating it into the decomposed sequence by VMD, outperforming EMD-LSTM [50]. For MTLF, since RNNs do not scale linearly, in addition to the vanishing gradient problem that still exists for LSTMs, they are ineffective when processing large time-span load series. Han et al. [23] proposed a two-hybrid approach between TD-CNN and C-LSTM. The temporal correlation is extracted directly using CNN in this case. It overcomes the limitations of LSTM mentioned above while retaining most of the original data by rebuilding a new short sequence from the actual load series. This hybrid approach can forecast from 1 week to 1 year ahead with a computation time of 51

minutes, whereas conventional LSTM requires 216 minutes. For LTLF applications and multivariate analysis, it seems like LSTM is still struggling [59]. However, LSTM is a viable option for univariate modeling using only its time lags.

c: Machine learning models

According to our studies, most machine learning models are kernel-based techniques like SVR, LS-SVM, boosting, bootstrap, ensemble approaches, etc. In 1995, Vpakian (1998) introduced a support vector machine (SVM) algorithm for classification problems which can also be applied for regression using loss functions [74]. This usability for regression problems gave rise to support vector regression (SVR) models. SVR can easily handle complex and non-linear data. SVR can easily avoid trapping in local minima, a common problem in neural networks [74]. Compared with neural networks, SVR is more reliable in achieving the global optimum when working with systems with a limited amount of historical data [75]. In a study, [74] performed on the data from the Iranian national grid, a performance comparison between ANN and SVR shows that for hourly demand forecasting, SVR outperforms ANN on different days with different patterns. Zeng et al. succeeded in the integration of calendar and weather features into a hybrid. The main problem is that the SVR model improves peak load forecasting accuracy [35]. SVR model prediction accuracy is affected by parameter settings. In a study by Zhang et al. [76], a hybrid SVR-GA model was proposed adding weather and electricity price data to their inputs. Tran et al. [33] analyzed the effects of data normalization methods on grid search-based optimization of SVR. A softmax normalization was selected in this study based on the results showing that it can significantly improve the accuracy of the grid search algorithm for SVR models during training and testing. Kavousi-Fard et al. proposed a modified FA-SVR which outperformed ARMA, ANN, SVR-GA, SVR-HBMO, SVR-PSO, and SVR-FA by increasing the diversity in the population of fireflies [77]. The decomposition technique can be adapted further to improve the optimization of the meta-heuristic tools. Such an adaption was proposed by Niu et al. [40], where they used EEMD to decompose the signal that outperformed ANN, ELM, SVM, SVM–EMD, and SVM–EEMD. The experiment was validated on four grids proving its robustness and reliability. Boosting models generated higher MAE and MAPE values than bagging models and were thus unsuitable for real-time scenarios. LS-SVM is a hybrid SVM model that uses the least square as a loss function to optimize SVM. For long-term forecasting, Compared to the MLR and ANN models, the LSSVM model achieved better results by 1.70 and 0.88% [22]. LSSVM models incorporated optimization techniques like FOA, HSA, CSA, and MFO [78], [79].

Jawad et al. proposed a unique solution to select the least cost electric load forecasting model using correlated weather variables [26]. MLR, SVR with different kernels, kNN, Random Forest, and AdaBoost were the models implemented for choosing the least cost forecasting models by minimizing RMSE, MAE, and MAPE. SVRr-based electric load forecasting is proven to have the best performance and can save PKR 0.313 million per day when used for the Lahore region in Pakistan. In one of the studies by Golilarz et al. [60], a hybrid WSVM is proposed where wavelet decomposition is performed to improve the data and a sparrow search algorithm is used for optimization. For capturing the long-term trend of the data, the model must be good at non-linear curve fitting. Wang et al. proposed a hybrid DE-SVR model to overcome the curve overfitting of BPNN [80]. The proposed model captures the trend of the annual load curve pattern, quickly proving its reliability in LTLF applications.

d: Fuzzy logic and clustering models

In fuzzy logic systems, fuzzy numbers and sets can be defined in semantic variables to address imprecision in input and output variables [9]. In many studies, data clustering is done through fuzzy techniques. Liu et al. [81] proposed an improved version of Fuzzy C-means designed to solve the uniform effect of FCM and give a significant performance boost through robust load profiling. For VSTLF applications, Pati et al. [14] developed an incomplete fuzzy decision system an hour ahead of load forecasting using rough sets to extract relevant domain knowledge from a load-lowering computation time. To deal with uncertainties in variables like calendar and weather, a type-2 fuzzy approach is proposed [82]. Fuzzifying the dataset generates fuzzy sets, with the total number of fuzzy sets equaling the number of optimal partitions. Forecasting and de-fuzzing are then accomplished using the weighted average method. Jain et al. [83] proposed KmFuzz, which removes the limitation of finding the optimal number of partitions for clustering with an accuracy of 98%. Adaptive neuro-fuzzy inference system (ANFIS) has the upper hand on ANN in handling complex load patterns. ANFIS is easy to design and is more robust compared to conventional ANN. A case study performed in India showed that ANFIS provided a MAPE of 0.55589%, proving more accurate than other state-of-the-art models [84]. ANFIS is generally more suitable for large datasets. For LTLF applications, to tackle the non-linear behavior of input variables, a hybrid ANFIS-GEP approach is developed where GEP and ANFIS effectively combine the ability to select model structures and features in the proposed approach [41]. ANFIS-GA is another hybrid technique to deal with the dependence of clustering on the radius of the cluster by minimizing the SRC value [85]. Panda et al. adopted a hybrid EMD-PSO-ANFIS, which utilized feature selection to impact forecasting accuracy positively [86].

e: Stochastic time series models

Several time-series analysis techniques are around to analyze and forecast time series, including Auto Regressive Integrated Moving Average (ARIMA), Auto Regressive Moving Average (ARMA), and Seasonal Auto Regressive Integrated Moving Average (SARIMA). Compared to other approaches, the structure is simpler, faster, and faster to develop. This approach does not consider external factors that strongly influence consumption. As load profiles no longer follow a homogeneous evolution, ARIMA models no longer work well beyond a threshold forecast horizon. The order of ARIMA models is selected using the Akaike information criterion (AIC) [87]. ARIMA model is mainly used to forecast the linear part of the load data due to the linear correlation configuration. By using probabilistic forecasts, the planner can evaluate whether the system will have the ability to generate more electricity in the future than it does today. It can be calculated when the system can exceed its current maximum capacity in such a case [88]. A probabilistic (Bayesian structural time series) approach is more appropriate for places with unpredictable load trends [88]. If results seem counterintuitive, BSTS modeling enables one to investigate the origin of the problem by determining the contribution of each component. As a solution to the seasonality issue for time series data, SARIMA has been widespread. Since the transmission peak load has a periodic distinction, SARIMA can be used for mid to long-term forecasting. A hybrid of SARIMA-GPR showed acceptable potential for Cambodia's mid-term load forecasting [89]. The main drawback of SARIMA is that it cannot differentiate between working days and holidays. However, there seems to be an advantage to SARIMA in recovering quickly from holiday effects [58].

Grey modeling is another stochastic technique used for predicting the future load pattern. They are mostly adopted for LTLF applications. The doctrinal methods are GM (1,1) and GM (1,2). A significant advantage of grey modeling is that it describes a system's behavior and reveals the continual process of change in it. Gray forecasting theory requires only tiny amounts of data. Modeling gray differentiation and time-varying differences with GM (1,1) allows the definition of a gray system of "one order, one variable [90]. The main disadvantage of this method is that, when a forecasting horizon is large, it uses only predicted values for predicting the next period. Therefore, Hamzacebi et al. [91] proposed an optimized GM (1,1) model and validated the robustness of the model in Turkey. For complex patterns, Lee et al. integrated dynamic programming into forecasting [92].

8) Model complexity

In a forecasting system, the computational time is the total number of executions it takes to run a model [16]. A computer's processing power, data length, and forecasting method are the main factors that affect execution time. When choosing the best forecasting methods based on similar performance, the computation time of the model could be the deciding factor, especially for lower-middle-income countries where resources are somewhat scarce. In contrast to artificial neural networks, hybridization achieves better generalization results since ANNs suffer from noise, bias, variance, and inefficiency [65]. ANN, on the other hand, can achieve improved results in less time if these problems are resolved. LSTM is a variant of a deep neural network with a higher number of layers. This increases the complexity of the models and affects the computation time. In one experiment [59], the total computation time of the LSTM model was 2685.22s. In another study [71], the total computation time was 25 minutes, roughly 1500s. The inefficacy of LSTM in MTLF and LTLF applications was once again observed in a study [23], where it took roughly 50-80 minutes to forecast month and year-ahead values. Therefore it is not recommended to use LSTM without sufficient processing power. Time series techniques are computationally inexpensive due to their simplicity. As a result of a study conducted [93], long-term forecasts took 1.35 seconds and daily predictions for four years took 153.35 seconds. SVR models are computationally expensive due to higher convergence time and the selection of optimal parameters. However, when hybridized by optimization algorithms, the limitations can be removed. Integrating DE and FOA with SVR reduced the time to 78s and 17s [79], [80].

C. QUANTITATIVE ANALYSIS OF THE CASES

This section illustrates the statistical analysis of the cases in Table 7 with bar and box plots.

1) Research time frame analysis

In Figure 10, we observe that load forecasting research has boomed since 2017. Recent advancements in smart grid technology can be one of the reasons for that. A robust global economy and higher heating and cooling demands in some parts drove a two-fold increase in global energy consumption in 2018. With renewable energy incorporated into the grid, carbon emissions are reduced. Solar and wind energy open limitless opportunities for clean energy as the world introduces renewable energy sources. In 2020, every world sector took a hit due to the pandemic (COVID-19), especially the GDP sector. As GDP is an essential factor affecting the energy sector, researchers all around started to contribute to it so that the world (especially the developing countries) can quickly get back on its feet.
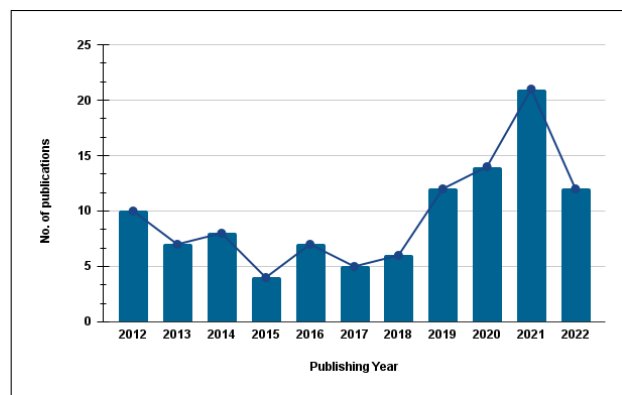

FIGURE 10. Research trends of electric demand forecasting in developing countries since 2012.
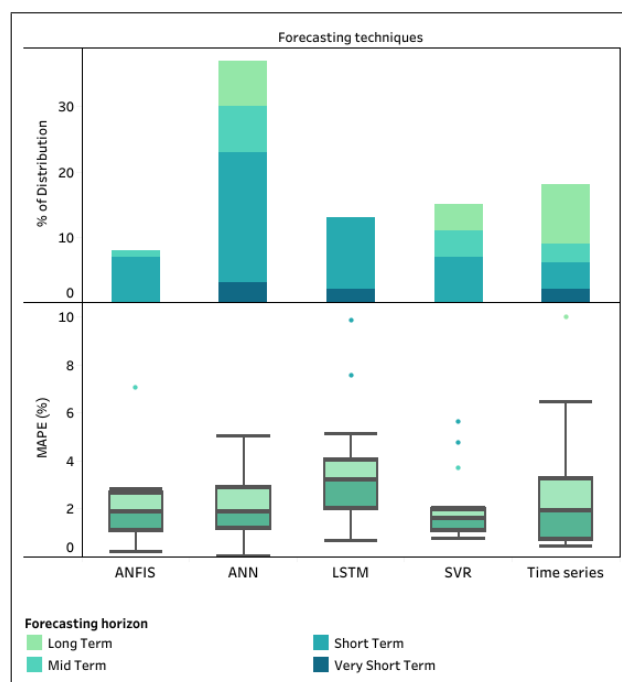

FIGURE 11. Forecasting model usage based on the forecast horizons.

2) Forecasting technique analysis

Starting with this section, a statistical comparison has been conducted based on the number of appearances the forecasting models had over the last decade. Previously, we considered the top 5 models based on many appearances for comparison cases. For further analysis, we created a box plot of MAPE values by categories of techniques validating our previous result. All five models show MAPE below 4%. From the bar chart in Figure 11, we can see that all the techniques have STLF applications. ANN models show an overall greater implementation rate over all horizons. LSTM and ANFIS provide us with a robust model for STLF applications only. SVR shows implementation over all horizons at a low MAPE. ANN, Time series techniques, and SVR models are reliable for mid-term load forecasting. For

long-term forecasting, ANN and Time series techniques provide us with more accurate models. Finally, ANN and LSTM models are somewhat reliable for very short-time forecasting.

3) Forecasting tools analysis

Forecasting tools become an essential factor when it comes to reproducing the experiment. For utility companies, implementing an online forecast system automatically updates the training dataset with real-time electricity demand observations each day so that the execution time can be increased. Table 11 illustrates the tools used in the review papers. While most articles do not specify the tools used, MATLAB is the most convenient tool. Limited articles adopted TensorFlow for modeling and training the deep learning models, specifically LSTM.

**TABLE 11.** Tools used for performing the forecasting experiment.

Tools	Number of appearances
MATLAB	36
TensorFlow	13
Not specified	62

4) Input factors analysis

In this section, the impact of different inputs combination is presented. From Figure 12, it is pretty evident that the historical load data will always be there in every combination. For LTLF applications, economic factors affect forecasting accuracy significantly in ANN and ANFIS models. Otherwise, the load variables alone successfully model techniques like ANN, SVR, and time series. For MTLF applications, meteorological and calendar features are seamlessly incorporated into all the models other than ANFIS. ANFIS can utilize the economic variable in this horizon to some extent. For STLF applications, it shows a strong correlation with meteorological features and calendar features for both ANN and SVR models. ANN model especially shows its flexibility by introducing most of the input combinations. LSTM models can only forecast using the historical load due to their long-term dependency. Exogenous variables are rarely used in time series analysis. For VSTLF applications, all the exogenous variables other than economic variables affect the forecasting accuracy of the models. Most models are designed to introduce exogenous variables for better performance, but they are not necessary for the model to function correctly.

5) Spatio temporal analysis

Besides representing the level of detail in the forecasting models, Spatio-temporal properties are essential criteria for model selection since the entity level of consumption
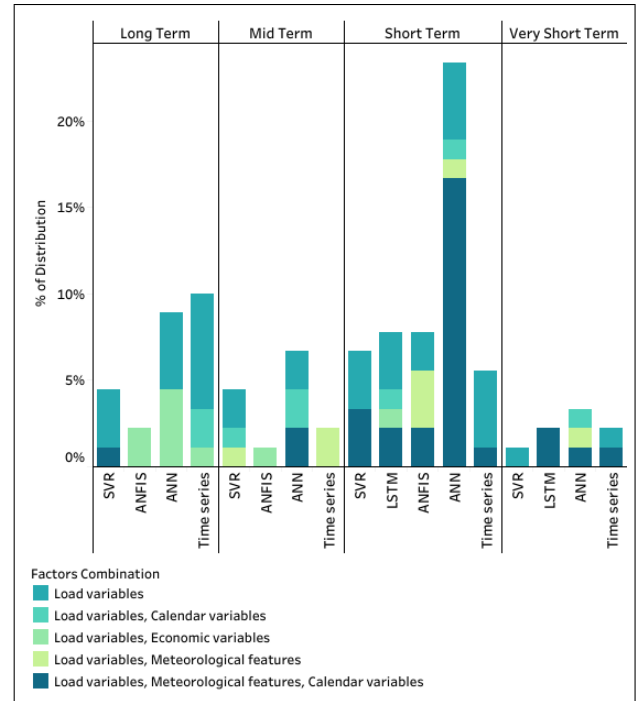
FIGURE 12. Exogenous variable usage based on the forecasting horizons.

data significantly affects the resolution and horizon of a model's outputs. From Figure 13, the influence of spatial detail on the forecasting model and horizon is visualized.

Based on an analysis of the spatial resolution in Figure 13, it can be concluded that there is a general tendency to model energy consumption on a national scale for all forecast horizons. Except for short-term forecasting, which emphasizes every level of consumption, relatively few approaches focus on regional and city-level consumption. There is a strong tendency for ANN techniques to be used at any level, but mainly at the national level. It demonstrates how flexible this technique is. For LTLF applications, ANN models have been scaled to the national level in half of the cases through extensive socio-economic data. However, time series methods are widely used nationally, reflecting their ability to capture trends at large geographic scales. Concise time horizons are the only timescales for which LSTM models are implemented.

Figure 14 dictates that forecasting on a daily and annual basis is more prevalent in the analysis of forecasting time resolution. Having daily steps allows for the inclusion of effects of human activities impacting consumption, such as daily routines, while still managing the amount of data. This could be because daily steps offer a reasonable compromise between the level of detail, availability, and data quantity for most projects. For the case of yearly steps, the socio-economic factors tracked yearly can be incorporated for accurate long-term analysis. Sub-hourly to week-ahead forecasting is
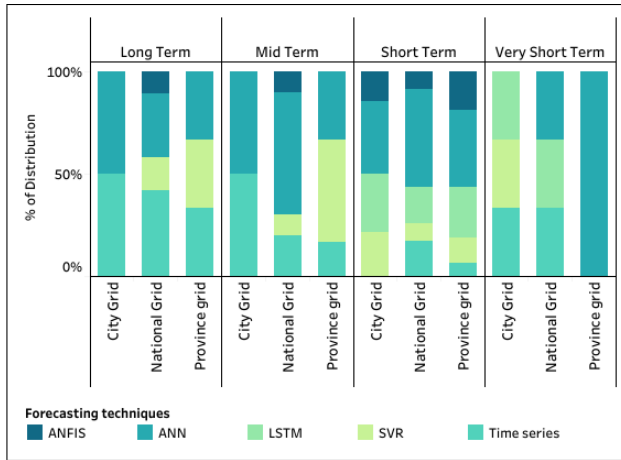
IEEE *Access*

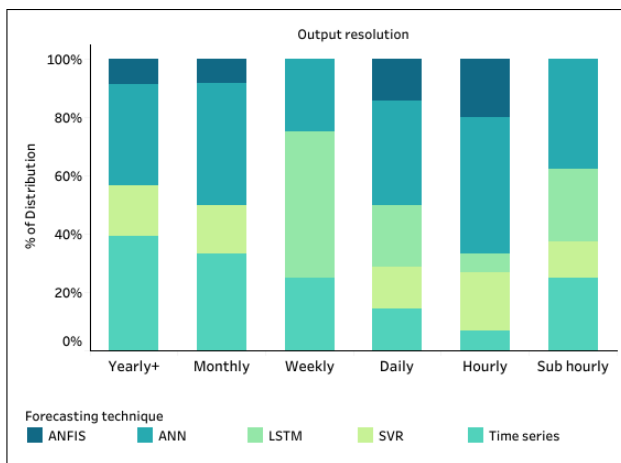**FIGURE 13. Forecasting technique usage based on the spatial granularity.**

FIGURE 14. Forecasting technique usage based on temporal resolutions.

effective using RNN techniques like LSTM. Additionally, Figure 14 shows that time series techniques are more commonly used as time steps increase, whereas AI techniques usage remains the same more or less. This study again shows how flexible ANN approaches are, regardless of the time resolution.

6) Forecasting accuracy analysis

For our particular investigation, we chose MAPE as our performance metric. From Figure 15, we can observe that consumption data from the national level produces the lowest error consistently over all horizons. Data aggregated at the city or area level gives a bit higher MAPE. Based on their smoothness and tendency to display seasonal or trend-related patterns, aggregated loads at the level of countries are more accessible to predict than lower-level aggregated loads. Human behavior and its influence on the load consumption pattern depend on varying degrees of randomness, which naturally produces lower precision for disaggregated loads.

The boxplot in Figure 16 shows that for increasing
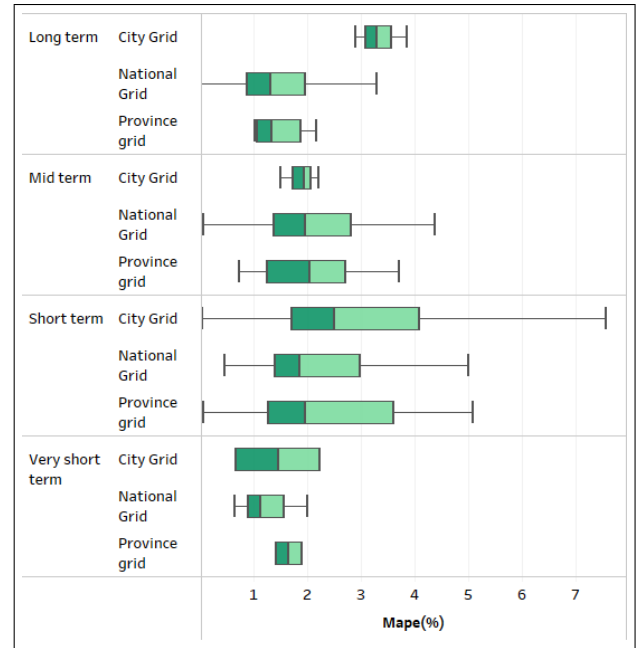
FIGURE 15. Error % based on the consumption levels.
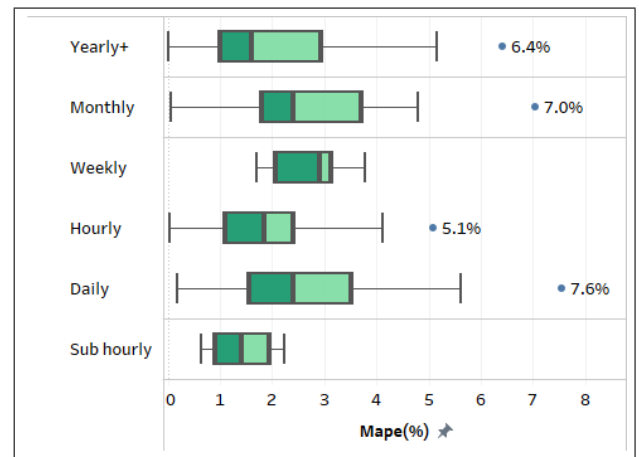
FIGURE 16. Error % based on the forecasting time resolution levels.

time steps, the forecasting model performance is more or less the same. However, the error is lower for yearly time steps due to smoothness and its tendency to show seasonal patterns. For daily or lower time steps, the introduction of various exogenous variables affecting the load curve is incorporated, producing an overall lower error in these temporal resolutions [72], [94].

7) Study validation country analysis

From Figure 17, we can see an unequal distribution of research performed in different countries. China is leading the research of energy sector forecasting among developing countries. In contrast to other middle countries, China has already invested heavily in renewable energy and sustainability. Because of this, China boasts
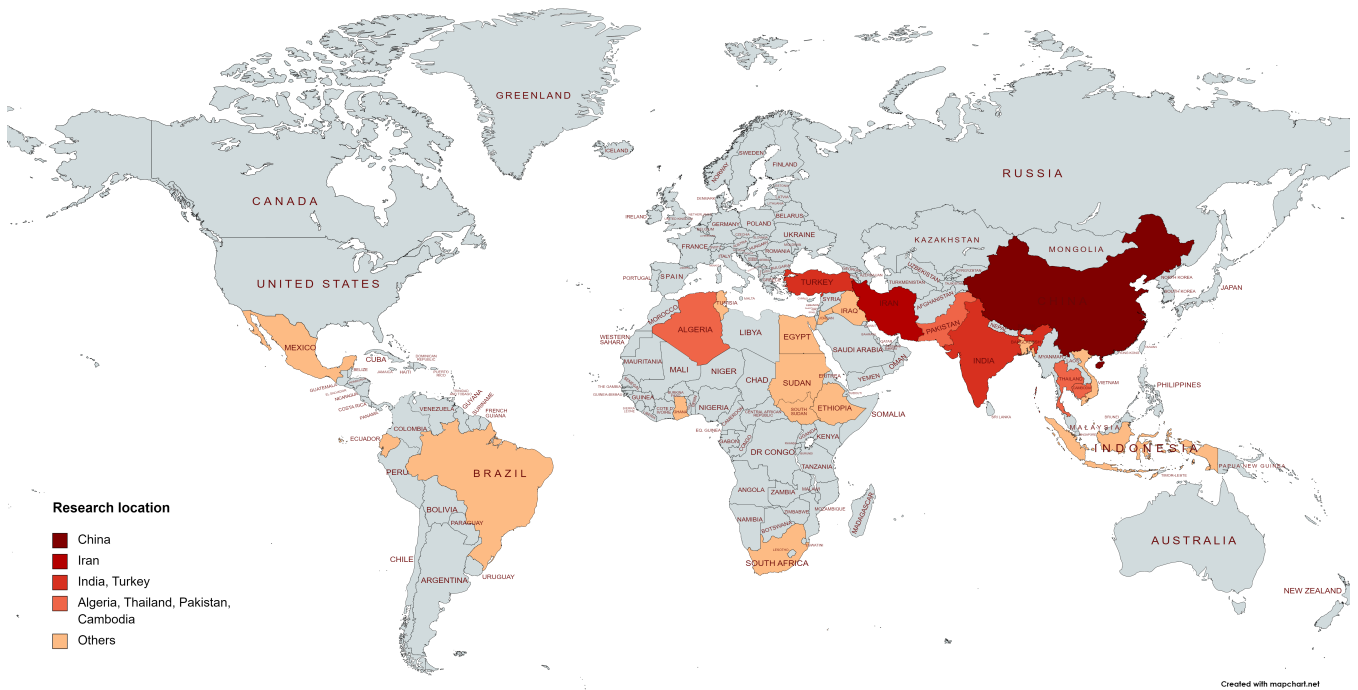
FIGURE 17. Geographic colored mapping of the study conducted on developing countries around the world. The red color's intensity represents the research distribution in that particular country. The light orange color represents the country in the singularity zone regarding research on energy forecasting.

exceptional wind power resources and is the world's largest producer of renewable energy electricity. While countries like China, Iran, Turkey, India, etc. have rapid progress in this sector, several countries are in the singularity region.

IV. DISCUSSION

Large-scale forecasting is not as easy as it seems due to numerous factors influencing the model's complexity, robustness, and accuracy. The analysis done in this study revolves around different cases. Our investigation thoroughly analyzes all the different energy demand forecasting models under various scenarios. For ease of analysis, we chose the models most prevalent among the research community for the analysis cases. The models which are most frequently adopted is a proof of their flexibility and robustness. An essential aspect of decision-making is the accuracy of prediction, which allows stakeholders to understand the performance of the methods and making choice of the model more plausible. We chose to compare the model's performance using MAPE, which does not necessarily dictate which model is the best, as there is no universal metric for evaluating a model's performance. Table 13 outlines the strength and applications of the forecasting models summarized in our analysis. Table 12 and 14 provide a comprehensive forecasting model selection framework for various scenarios in UMICs and LMICs based on our

analysis.

As stated before, the whole point of this systematic review is that we can answer the generic question we formulated. For example, let us generate a scenario based on which we can answer that question. Sample scenario: Suppose we want to forecast the hourly electricity demand on a national level for a developing country like Bangladesh. From Table 12, we can evaluate the scenario step by step:

1) Bangladesh is a lower middle-income country.
2) Beforehand, it is recommended to deal with outliers or missing data and perform data normalization always.
3) Now, referring to Table 12, we can see both univariate and multivariate options available for an hour ahead forecasting on a national level.
4) The ANN and SVR models perform well with correlation analysis for dimensionality reduction.
5) Now, two approaches are available: ANN and SVR. Both are computationally expensive. We know that computation time is a crucial aspect of hourly or daily forecasting. So any of the two models can be utilized with the high processing unit. They can always refer to the corresponding article for more details on the model structure.

So the solution to the scenario is utilizing the SVR or ANN model with historical load consumption data as the input features and using correlation analysis as a

TABLE 12. A framework for forecasting strategy selection for lower-middle-income countries.

Output resolution	Spatial resolution	Exogenous factors	Techniques used	Dimensionality reduction techniques	Reference
Sub hourly	National level	Meteorological and Calendar features	ANN	Correlation analysis	[31]
	Regional level	Calendar features	ANN	Not necessary	[95]
Hourly	National level	Historical load data	SVR	Correlation analysis	[74]
		Meteorological and Calendar features	ANN	Correlation analysis	[16]
Daily	City level	Meteorological and Calendar features	ANFIS	Fuzzy rule-based selection	[82]
		Meteorological features	ANFIS	GA	[14]
		Meteorological, Calendar and Energy Price features	SVR	Correlation analysis	[26]
	National level	Historical load data	SVR	Wavelet Transform	[47]
		Meteorological and Calendar features	ANN	GA	[64]
		Meteorological features	ANFIS	Fuzzy rule-based selection	[48]
	Regional level	Meteorological and Calendar features	ANFIS	Fuzzy rule-based selection	[84]
Weekly	National level	Historical load data	LSTM	Not necessary	[96]
		Meteorological and Calendar features	ANN	Elastic net	[12]
Monthly	National level	Meteorological and Calendar features	ANN	Correlation analysis	[43]
		Meteorological features	Time series techniques	Not necessary	[69]
		Meteorological, Socioeconomic and Calendar features	Time series techniques	Not necessary	[89]
Yearly+	National level	Historical load data	Time series techniques	Correlation analysis	[97]
		Socioeconomic features	ANN	Not necessary	[98]
	Regional level	Historical load data	ANN	Not necessary	[53]

feature reduction technique. For the model detail, they can use the references from Table 12.

KEY FINDINGS

- STLF and LTLF applications are on the rise due to their versatility. While short-term forecasting is necessary for energy purchase and operations scheduling, long term focuses on financial and distribution planning in the long run. Similarly, Day-ahead and annual forecasting are on the rise. Day-ahead is most relevant in the case of optimal generator unit commitment, whereas annual forecasting

helps in the system and economic planning.
- When it comes to LTLF applications, economic factors play a prominent role in forecasting accuracy. Forecasting accuracy for mid to short-term applications is affected to some degree by meteorological and calendar features. Historical load data alone can produce good results in time series techniques.
- Across all forecast horizons, energy consumption is typically modeled at the national level. A disaggregated load has a lower degree of precision due to human behavior and data granularity. For long-term forecasting, data aggregated at the city

TABLE 13. Forecasting model's strength and applications. The number of stars represents the strength levels. (* = good, ** = better, *** = best).

Techniques	Strengths and applications							References
	Complexity	Accuracy	Robustness	VSTLF	STLF	MTLF	LTLF	
LSTM	***	**	***	***	***	X	X	[23], [29], [34], [46], [50], [52], [59], [71]–[73], [96]
SVR	**	***	**	*	***	***	***	[17], [22], [26], [28], [33], [35], [40], [51], [74], [76]–[80], [99], [100]
ANN	**	***	***	**	***	***	***	[12], [13], [16], [19], [24], [25], [31], [38], [42], [43], [45], [53]–[58], [60]–[67], [70], [81], [95], [98], [101]–[112]
Time series	*	***	*	**	***	**	***	[47], [69], [87]–[94], [97], [113]–[120]
ANFIS	**	***	**	*	***	X	**	[14], [15], [41], [48], [49], [82]–[85]

level gives a higher error. Daily steps allow human activities that influence consumption, such as daily routines, to be considered. Since seasonal patterns tend to show up for yearly time steps, the error is lower because of the smoothness and seasonality.

- Data normalization is essential for increasing accuracy and reducing computation time. It is recommended to use data normalization beforehand fitting into the model. Dimensional reduction plays a significant role in improving the accuracy and computation time of the forecasting models. Most techniques adopt correlation analysis for feature reduction. Clustering-based grouping helps in detecting load curve trends.

- For multi-step forecasting, VAE and VMD-based feature extraction techniques significantly reduce the MAPE. Peak load forecasting is performed more efficiently using WPD than DWT.

- Hybrid models have been observed to overcome the limitations of single models, though they sometimes introduce complex structures. Time series techniques are computationally inexpensive compared to other models due to having a non-complex structure.

- Among all the optimization techniques, PSO is the most conventional and provides minimal MAPE. In addition to having faster computation time and ease of implementation, PSO is superior because of its ability to converge more quickly than other optimization techniques.

- Broadly speaking, ANN is the most conventional technique due to its flexibility and ability to handle non-linear data. ANN can be incorporated with optimization techniques to reduce its complexity and computational cost limitation. Time series methods are most commonly applied at a national level because they capture trends at large geographic scales.

- Surprisingly, the simulation modeling technique, a commonly used method for load forecasting, has shown scarcity in developing countries. The most probable reason for this is the computational requirements of simulation models. This can also be a barrier in a developing country with limited access to advanced computing equipment. These models often require significant computational resources to run, which may not be available in such a setting.

TABLE 14. A framework for forecasting strategy selection for upper-middle-income countries.

Output resolution	Spatial resolution	Exogenous factors	Dimensionality reduction techniques	Techniques used	Reference
Sub hourly	City level	Meteorological and Calendar features	CNN	LSTM	[121]
	National level	Historical load data	Phase-space reconstruction	Time series techniques	[94]
	Regional level	Meteorological features	PCA	ANN	[13]
Hourly	City level	Meteorological and Calendar features	Wavelet Transform	SVR	[51]
		Meteorological features	RF	ANN	[106]
	National level	Historical load data	Not necessary	Time series techniques	[93]
		Meteorological features	Not necessary	ANFIS	[15]
	Regional level	Meteorological features and Energy Price	RF selection and CNN based extraction	LSTM	[46]
		Meteorological and Calendar features	RF	ANN	[81]
Daily	National level	Historical load data	Not necessary	LSTM	[29]
		Meteorological and Calendar features	Correlation analysis	ANN	[32]
	Regional level	Historical load data	Decomposition techniques	ANFIS	[86]
		Calendar features	Auto Encoder	LSTM	[50]
		Meteorological and Calendar features	Clustering	ANN	[57]
Weekly	City level	Socioeconomic features	Not necessary	LSTM	[23]
Monthly	City level	Meteorological and Calendar features	PMI	ANN	[45]
		Meteorological features	Clustering	Time series techniques	[118]
	National level	Historical load data	Not necessary	ANN	[101]
		Calendar features	GA	ANN	[65]
	Regional level	Calendar features	Not necessary	SVR	[99]
Yearly+	City level	Socioeconomic features	GA	ANN	[24]
	National level	Historical load data	Not necessary	ANN	[66]
			Not necessary	Time series techniques	[92]
		Calendar features	Not necessary	Time series techniques	[116]
		Meteorological and Calendar features	Not necessary	SVR	[100]
		Socioeconomic features	Correlation analysis	SVR	[22]
	Regional level	Historical load data	Not necessary	SVR	[80]

V. RECOMMENDATIONS

Based on the findings of the previous sections, this section discusses quite a few areas that could be explored in future research.

PERFORMANCE METRIC

The lack of a universal evaluation metric is something that should be explored. Every evaluation metric has a different focus. For example, R^2 determines how well a model fits the dependent variables and considers the overfitting problem. RMSE, on the other hand, determines the goodness of fit. So each metric has its unique traits. So it is recommended to use a specific set of metrics to compare the models correctly.

MODEL ADAPTABILITY

So far, minimal literature exists on adaptive models, which will automatically adjust with the change of model parameters and input data. Adaptation to load behavior should be possible quickly in the model. It is vital to ensure that the model is suitable for various situations. It is also crucial that the model be reliable, as it should be free of unreasonable forecasts caused by exceptional circumstances. It is advised to focus on the effortless adaptability of the model to an energy management system.

SPATIAL DISTRIBUTION OF RESEARCH

The majority of the research is conducted in China, Iran, Turkey, and India. Most of the other developing countries are in the singularity region. The forecasting models work differently for different countries as the load consumption pattern, climate, and economy are inconsistent. It is highly suggested that future researchers from all over the world explore this research gap in their respective countries.

MODEL REPRODUCIBILITY

Reproducibility is another major limitation of the current studies. Most of the literature does not specify the experiment setups, including the computational time, tools used, source code, and the computer specifications in which the experiment is validated. This drawback makes the reproducibility of the models harder. Researchers should consider making their source code and experiment setup details publicly accessible.

VI. CONCLUSION

The planning cycle of an electric utility starts with forecasting energy demand and consumption. Developing countries face complex challenges in forecasting demand and energy as the grid is experiencing a change of shape as technology turns a demand-driven energy system into a generation-driven system. Several factors contribute to the problem, including a high load growth rate, new technologies at the consumer level and differences in consumption levels and modes between cities. Because of limited safety margins and increased risk levels, forecasting energy demand has become more critical than ever for the future grid.

Our aim is to identify the energy demand forecasting methods relevant to the various use cases from a developing country's perspective, identify the potential research gaps, and offer recommendations to future researchers and forecasters. To this end, we performed a systematic review of energy demand forecasting articles over the last decade following the PRISMA protocol. The prevailing models identified so far in our investigation were: SVM, ANN, LSTM, ANFIS, and time series techniques. To comprehensively assess the models, we concocted various analysis cases designed to generate deeper insights. ANN has the most overall appearances for different use cases due to its flexibility and ability to handle non-linear data. For a better outlook on the performance of the models in the real-time systems, we considered the model complexity alongside the evaluation metrics. A framework was fabricated based on the findings to determine the optimal energy forecasting method for different scenarios with references. For a better understanding, we showcased the strength and applications of the forecasting models. We expect this article to be of significance to future forecasters and researchers in the domain of energy demand forecasting in developing countries.

For future works, we intend to explore the existing adaptive load forecasting models and how well they perform with changing conditions in smart grid environments. As developing countries are transitioning towards smart grid adaptations, the research can have a lasting impact and can be used to inform and guide future work in the field.

APPENDIX A

In this section, the works of literature are categorized based on the analysis cases, making it easier for the researchers to access the reviewed studies directly. Table 15 grouped the literature based on factors affecting the load for various models. Finally, in Table 16 and 17 the articles are sorted based on Spatio-temporal granularity.

TABLE 15. Classification of the literature according to the model and all possible combinations of input variables.

Models	Factors Contributing	References
ANFIS	Load variables	[83]
	Load and Economic variables	[41], [49], [85]
	Load and Meteorological features	[14], [15], [48]
	Load, Meteorological and Calendar variables	[82], [84]
ANN	Load variables	[42], [53]–[55], [60], [62], [63], [66], [101], [102]
	Load and Calendar variables	[65], [95], [103], [104]
	Load and Economic variables	[19], [24], [98], [105]
	Load and Meteorological features	[13], [106]
	Load, Meteorological and Calendar variables	[12], [16], [31], [38], [43], [45], [56], [57], [61], [64], [81], [107]–[112]
	Load, Meteorological, Calendar and Economic variables	[58]
	Load, Meteorological, Energy Price and Calendar variables	[25], [67], [70]
LSTM	Load variables	[29], [71], [96]
	Load and Calendar variables	[50]
	Load and Economic variables	[23]
	Load, Meteorological and Calendar variables	[34], [59], [72], [73]
	Load, Meteorological, Energy Price and Calendar variables	[52]
	Load, Energy Price and meteorological features	[46]
SVR	Load variables	[28], [33], [40], [74], [77]–[80]
	Load and Calendar variables	[99]
	Load and Economic variables	[22]
	Load and Meteorological features	[17]
	Load, Meteorological and Calendar variables	[35], [51], [76], [100]
	Load, Meteorological, Energy Price and Calendar variables	[26]
Time series techniques	Load variables	[47], [87], [90]–[94], [97], [113]–[115]
	Load and Calendar variables	[88], [116]
	Load and Economic variables	[117]
	Load and Meteorological features	[69], [118]
	Load, Meteorological and Calendar variables	[119], [120]
	Load, Meteorological, Calendar and Economic variables	[89]

TABLE 16. Classification of the literature based on forecasting resolution.

Forecast resolution	References
Sub hourly	[13], [31], [33], [59], [94], [95], [119], [121]
Hourly	[1], [15], [16], [21], [27], [39], [44], [46], [51], [60], [74], [76], [81], [83], [93], [103], [106], [112]
Daily	[14], [25], [26], [28], [29], [32], [35], [38], [42], [47], [48], [50], [52], [56], [57], [63], [64], [71], [73], [77], [82], [84], [86], [87], [107], [108], [110], [111], [114], [120]
Weekly	[12], [23], [34], [58], [67], [70], [96], [102]
Monthly	[17], [20], [43], [45], [49], [54], [65], [69], [78], [89], [99], [101], [104], [118], [122]
Yearly+	[19], [22], [24], [41], [53], [55], [62], [66], [79], [80], [85], [88], [90]–[92], [97], [98], [100], [105], [113], [115]–[117]

TABLE 17. Classification of the literature based on spatial resolution.

Spatial resolution	References
City level	[14], [23]–[27], [33], [35], [44], [45], [51], [52], [56], [73], [82], [102], [106], [108], [112], [118], [119], [121]
National level	[1], [12], [15]–[17], [19]–[22], [29], [31], [32], [41]–[43], [47]–[49], [54], [55], [58]–[60], [62]–[67], [69], [71], [74], [77], [79], [85], [87]–[94], [96]–[98], [100], [101], [103]–[105], [110], [113]–[117], [122]
Regional level	[13], [28], [34], [38], [39], [46], [50], [53], [57], [70], [76], [78], [80], [81], [83], [84], [86], [95], [99], [107], [111], [120]

REFERENCES

[1] L. Alhmoud, R. Abu Khurma, A. M. Al-Zoubi, and I. Al-jarah, "A real-time electrical load forecasting in jordan using an enhanced evolutionary feedforward neural network," Sensors, vol. 21, no. 18, 2021.

[2] H. Yanikkaya, "Trade openness and economic growth: A cross-country empirical investigation," Journal of Development Economics, vol. 72, no. 1, pp. 57–89, 2003.

[3] K. Sohag, M. Al Mamun, G. Uddin, and A. Ahmed, "Sectoral output, energy use, and co2 emission in middle-income countries," Environmental Science and Pollution Research, vol. 24, no. 10, pp. 9754–9764, 2017.

[4] Y. Cho, D. N. Margolis, and D. A. Robalino, "Labor markets in low and middle income countries : trends and implications for social protection and labor policies," The World Bank, Social Protection Discussion Papers and Notes 67613, 2012. [Online]. Available: https://ideas.repec.org/p/wbk/hdnspu/67613.html

[5] C. Kuster, Y. Rezgui, and M. Mourshed, "Electrical load forecasting models: A critical systematic review," Sustainable Cities and Society, vol. 35, pp. 257–270, 2017.

[6] I. K. Nti, M. Teimeh, O. Nyarko-Boateng, and A. F. Adekoya, "Electricity load forecasting: a systematic review," Journal of Electrical Systems and Information Technology, vol. 7, no. 1, p. 13, 2020.

[7] A. Román-Portabales, M. López-Nores, and J. J. Pazos-Arias, "Systematic review of electricity demand forecast using ann-based machine learning algorithms," Sensors, vol. 21, no. 13, 2021.

[8] P. A. Verwiebe, S. Seim, S. Burges, L. Schulz, and J. Müller-Kirchenbauer, "Modeling energy demand - a systematic literature review," Energies, vol. 14, no. 23, 2021.

[9] A. A. Mamun, M. Sohel, N. Mohammad, M. S. Haque Sunny, D. R. Dipta, and E. Hossain, "A comprehensive review of the load forecasting techniques using single and hybrid predictive models," IEEE Access, vol. 8, pp. 134 911–134 939, 2020.

[10] A. Azeem, I. Ismail, S. M. Jameel, and V. R. Harindran, "Electrical load forecasting models for different generation modalities: A review," IEEE Access, vol. 9, pp. 142 239–142 263, 2021.

[11] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting, and D. Moher, "The prisma 2020 statement: an updated guideline for reporting systematic reviews," BMJ, vol. 372, 2021. [Online]. Available: https://www.bmj.com/content/372/bmj.n71

[12] A. A. Mir, Z. A. Khan, A. Altmimi, M. Badar, K. Ullah, M. Imran, and S. A. A. Kazmi, "Systematic development of short-term load forecasting models for the electric power utilities: The case of pakistan," IEEE Access, vol. 9, pp. 140 281–140 297, 2021.

[13] L. Cheng, Y. Bao, L. Tang, and H. Di, "Very-short-term load forecasting based on empirical mode decomposition and deep neural network," IEEJ Transactions on Electrical and Electronic Engineering, vol. 15, no. 2, pp. 252–258, Feb. 2020.

[14] U. Pati, P. Ray, and A. R. Singh, "An intelligent approach towards very short-term load forecasting," International Journal of Emerging Electric Power Systems, vol. 23, no. 1, pp. 59–72, 2022.

[15] H. H. Çevik and M. Çunkaş, "Short-term load forecasting using fuzzy logic and anfis," Neural Computing and Applications, vol. 26, no. 6, pp. 1355–1367, Aug. 2015.

[16] A. Azadeh, S. F. Ghaderi, M. Sheikhalishahi, and B. P. Nokhandan, "Optimization of short load forecasting in electricity market of iran using artificial neural networks," Optimization and Engineering, vol. 15, no. 2, pp. 485–508, Jun. 2014.

[17] H. Alirezaei, A. Salami, and M. Mohammadinodoushan, "A study of hybrid data selection method for a wavelet svr mid-term load forecasting model," Neural Computing and Applications, vol. 31, no. 7, pp. 2131–2141, 2019.

[18] C. Hamzaçebi, H. A. Es, and R. Çakmak, "Forecasting of turkey's monthly electricity demand by seasonal artificial neural network," Neural Computing and Applications, vol. 31, no. 7, pp. 2217–2231, Jul. 2019.

[19] F. Ardakani and M. Ardehali, "Long-term electrical energy consumption forecasting for developing and developed economies based on different optimized models and historical data types," Energy, vol. 65, pp. 452–461, 2014.

[20] P. NOP and Z. QIN, "Transmission system load forecasting using neural network approach, case study: Electricité du cambodge (edc)," in 2020 IEEE 3rd Student Conference on Electrical Machines and Systems (SCEMS), 2020, pp. 329–333.

[21] W. Zhang, G. Mu, G. Yan, and J. An, "A power load forecast approach based on spatial-temporal clustering of load data," Concurrency and Computation: Practice and Experience, vol. 30, no. 23, p. e4386, Dec. 2018.

[22] F. Kaytez, M. C. Taplamacioglu, E. Cam, and F. Hardalac, "Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines," International Journal of Electrical Power and Energy Systems, vol. 67, pp. 431–438, 2015.

[23] L. Han, Y. Peng, Y. Li, B. Yong, Q. Zhou, and L. Shu, "Enhanced deep networks for short-term and medium-term load forecasting," IEEE Access, vol. 7, pp. 4045–4055, 2019.

[24] S. Yu, K. Wang, and Y.-M. Wei, "A hybrid self-adaptive particle swarm optimization–genetic algorithm–radial basis function model for annual electricity demand prediction," Energy Conversion and Management, vol. 91, pp. 176–185, 2015.

[25] J. Bedi and D. Toshniwal, "Empirical mode decomposition based deep learning for electricity demand forecasting," IEEE Access, vol. 6, pp. 49 144–49 156, 2018.

[26] M. Jawad, M. S. A. Nadeem, S.-O. Shim, I. R. Khan, A. Shaheen, N. Habib, L. Hussain, and W. Aziz, "Machine learning based cost effective electricity load forecasting model using correlated meteorological parameters," IEEE Access, vol. 8, pp. 146 847–146 864, 2020.

[27] N. Huang, G. Lu, and D. Xu, "A permutation importance-based feature selection method for short-term electricity load forecasting using random forest," Energies, vol. 9, no. 10, 2016.

[28] X. Kong, C. Li, C. Wang, Y. Zhang, and J. Zhang, "Short-term electrical load forecasting based on error correction using dynamic mode decomposition," Applied Energy, vol. 261, p. 114368, 2020.

[29] P. P. Phyo, C. Jeenanunta, and K. Hashimoto, "Electricity load forecasting in thailand using deep learning models," International Journal of Electrical and Electronic Engineering and Telecommunications, vol. 8, no. 4, p. 221 – 225, 2019.

[30] K. Zheng, P. Li, S. Zhou, W. Zhang, S. Li, L. Zeng, and Y. Zhang, "A multi-scale electricity consumption prediction algorithm based on time-frequency variational autoencoder," IEEE Access, vol. 9, pp. 90 937–90 946, 2021.

[31] R. Houimli, M. Zmami, and O. Ben-Salha, "Short-term electric load forecasting in Tunisia using artificial neural networks," Energy Systems, vol. 11, no. 2, pp. 357–375, May 2020.

[32] Q. Cai, B. Yan, B. Su, S. Liu, M. Xiang, Y. Wen, Y. Cheng, and N. Feng, "Short-term load forecasting method based on deep neural network with sample weights," International Transactions on Electrical Energy Systems, vol. 30, no. 5, p. e12340, May 2020.

[33] T. N. Tran, B. M. Lam, A. T. Nguyen, and Q. B. Le, "Load forecasting with support vector regression: influence of data normalization on grid search algorithm," International Journal of Electrical and Computer Engineering, vol. 12, no. 4, p. 3410 – 3420, 2022.

[34] X. Fang, W. Zhang, Y. Guo, J. Wang, M. Wang, and S. Li, "A novel reinforced deep rnn-lstm algorithm: Energy

management forecasting case study," IEEE Transactions on Industrial Informatics, vol. 18, no. 8, pp. 5698–5704, Aug 2022.

[35] P. Zeng and M. Jin, "Peak load forecasting based on multi-source data and day-to-day topological network," IET Generation, Transmission and Distribution, vol. 12, no. 6, pp. 1374–1381, Mar. 2018.

[36] W. Jia, M. Sun, J. Lian, and S. Hou, "Feature dimensionality reduction: a review," Complex & Intelligent Systems, vol. 8, no. 3, pp. 2663–2693, Jun. 2022.

[37] Z. Hu, Y. Bao, T. Xiong, and R. Chiong, "Hybrid filter–wrapper feature selection for short-term load forecasting," Engineering Applications of Artificial Intelligence, vol. 40, pp. 17–27, 2015.

[38] M. Sheikhan and N. Mohammadi, "Neural-based electricity load forecasting using hybrid of ga and aco for feature selection," Neural Computing and Applications, vol. 21, no. 8, pp. 1961–1970, Nov. 2012.

[39] D. Sarkar, T. Ao, and S. K. Gunturi, "Bootstrap aggregating approach to short-term load forecasting using meteorological parameters for demand side management in the northeastern region of india," Theoretical and Applied Climatology, vol. 148, no. 3-4, pp. 1111–1125, May 2022.

[40] W.-J. Niu, Z.-K. Feng, S.-S. Li, H.-J. Wu, and J.-Y. Wang, "Short-term electricity load time series prediction by machine learning model via feature selection and parameter optimization using hybrid cooperation search algorithm," Environmental Research Letters, vol. 16, no. 5, 2021.

[41] A. Fallahpour, K. Barri, K. Y. Wong, P. Jiao, and A. H. Alavi, "An integrated data mining approach to predict electrical energy consumption," International Journal of Bio-Inspired Computation, vol. 17, no. 3, p. 142 – 153, 2021.

[42] A. Kavousi-Fard, T. Niknam, and M. Golmaryami, "Short term load forecasting of distribution systems by a new hybrid modified fa-backpropagation method," Journal of Intelligent and Fuzzy Systems, vol. 26, no. 1, p. 517 – 522, 2014.

[43] U. Javed, K. Ijaz, M. Jawad, E. A. Ansari, N. Shabbir, L. Kütt, and O. Husev, "Exploratory data analysis based short-term electrical load forecasting: A comprehensive analysis," Energies, vol. 14, no. 17, 2021.

[44] W. Sun and C. Zhang, "A hybrid ba-elm model based on factor analysis and similar-day approach for short-term load forecasting," Energies, vol. 11, no. 5, 2018.

[45] Y. Gao, Y. Fang, H. Dong, and Y. Kong, "A multifactorial framework for short-term load forecasting system as well as the jinan's case study," IEEE Access, vol. 8, pp. 203 086–203 096, 2020.

[46] Y. Xuan, W. Si, J. Zhu, Z. Sun, J. Zhao, M. Xu, and S. Xu, "Multi-model fusion short-term load forecasting based on random forest feature selection and hybrid neural network," IEEE Access, vol. 9, pp. 69 002–69 009, 2021.

[47] S. Saleh, S. Mohammadi, M.-A. Rostami, and M.-R. Askari, "A hybrid artificial-based model for accurate short term electric load prediction," Journal of Intelligent & Fuzzy Systems, vol. 27, no. 6, pp. 3103–3110, 2014.

[48] A. Laouafi, M. Mordjaoui, F. Laouafi, and T. E. Boukelia, "Daily peak electricity demand forecasting based on an adaptive hybrid two-stage methodology," International Journal of Electrical Power & Energy Systems, vol. 77, pp. 136–144, 2016.

[49] Z. Shao, F. Gao, S.-L. Yang, and B. gong Yu, "A new semi-parametric and eemd based framework for mid-term electricity demand forecasting in china: Hidden characteristic extraction and probability density prediction," Renewable and Sustainable Energy Reviews, vol. 52, pp. 876–889, 2015.

[50] Y. Yang, Z. Wang, Y. Gao, J. Wu, S. Zhao, and Z. Ding, "An effective dimensionality reduction approach for short-term load forecasting," Electric Power Systems Research, vol. 210, p. 108150, 2022.

[51] Y. Liang, D. Niu, M. Ye, and W.-C. Hong, "Short-term load forecasting based on wavelet transform and least squares support vector machine optimized by improved cuckoo search," Energies, vol. 9, no. 10, 2016.

[52] X. Guo, Q. Zhao, D. Zheng, Y. Ning, and Y. Gao, "A short-term load forecasting model of multi-scale cnn-lstm hybrid neural network considering the real-time electricity price," Energy Reports, vol. 6, pp. 1046–1053, 2020.

[53] V. Shrivastava, R. Misra, and R. Bansal, "An assessment of electrical load forecasting using artificial neural network," International Journal of Computer Aided Engineering and Technology, vol. 4, no. 1, pp. 80–89, 2012.

[54] C. Hamzaçebi, H. A. Es, and R. Çakmak, "Forecasting of turkey's monthly electricity demand by seasonal artificial neural network," Neural Computing and Applications, vol. 31, no. 7, pp. 2217–2231, Jul. 2019.

[55] N. SINGH, A. Singh, and M. Tripathy, "Selection of hidden layer neurons and best training method for ffnn in application of long term load forecasting," International Journal of Electrical Engineering, vol. 63, no. 3, pp. 153–161, 2012.

[56] P. Liu, P. Zheng, and Z. Chen, "Deep learning with stacked denoising auto-encoder for short-term electric load forecasting," Energies, vol. 12, no. 12, 2019.

[57] R. Wang, S. Chen, and J. Lu, "Electric short-term load forecast integrated method based on time-segment and improved mdsc-bp," Systems Science & Control Engineering, vol. 9, no. S1, pp. 80–86, 2021.

[58] O. Bozkurt, G. Biricik, and Z. Taysi, "Artificial neural network and sarima based models for power load forecasting in turkish electricity market," PLoS ONE, vol. 12, no. 4, pp. 1–24, 2017.

[59] N. Jha, D. Prashar, M. Rashid, S. K. Gupta, and R. Saket, "Electricity load forecasting and feature extraction in smart grid using neural networks," Computers and Electrical Engineering, vol. 96, p. 107479, 2021.

[60] Z. Shafiei Chafi and H. Afrakhte, "Short-term load forecasting using neural network and particle swarm optimization (pso) algorithm," Mathematical Problems in Engineering, vol. 2021, p. 5598267, 2021.

[61] L. Alhmoud, R. Abu Khurma, A. Al-Zoubi, and I. Aljarah, "A real-time electrical load forecasting in jordan using an enhanced evolutionary feedforward neural network," Sensors, vol. 21, 09 2021.

[62] R. M. Rizk-Allah, I. Eldesoky, and A. Ayad, "Long term load forecasting based on hybrid model of feed forward net and modified grey wolf optimization," WSEAS Transactions on Power Systems, vol. 15, pp. 133–142, 2020.

[63] E. Ofori-Ntow Jnr, Y. Y. Ziggah, and S. Relvas, "Hybrid ensemble intelligent model based on wavelet transform, swarm intelligence and artificial neural network for electricity demand forecasting," Sustainable Cities and Society, vol. 66, p. 102679, 2021.

[64] M. Moazzami, A. Khodabakhshian, and R. Hooshmand, "A new hybrid day-ahead peak load forecasting method for iran's national grid," Applied Energy, vol. 101, pp. 489–501, 2013, sustainable Development of Energy, Water and Environment Systems.

[65] K. Abeyrathna and C. Jeenanunta, "Hybrid particle swarm optimization with genetic algorithm to train artificial neural networks for short-term load forecasting," International Journal of Swarm Intelligence Research, vol. 10, no. 1, pp. 1–14, 2019.

[66] X. Wu, J. Mao, Z. Du, and Y. Chang, "Online training algorithms based single multiplicative neuron model for energy consumption forecasting," Energy, vol. 59, pp. 126–132, 2013.

[67] K. D. Ünlü, "A data-driven model to forecast multi-step ahead time series of turkish daily electricity load," Electronics (Switzerland), vol. 11, no. 10, 2022.

[68] W. Zhang, J. Qin, F. Mei, J. Fu, B. Dai, and W. Yu, "Short-term power load forecasting using integrated methods based on long short-term memory," Science China Technological Sciences, vol. 63, no. 4, pp. 614–624, 2020.

[69] K. Chreng, H. Lee, and S. Tuy, "Electricity demand prediction for sustainable development in cambodia using recurrent neural networks with era5 reanalysis climate variables," Energy Reports, vol. 8, pp. 76–81, 2022.

[70] D. P. Ananthu and K. Neelashetty, "Electrical load forecasting using arima, prophet and lstm networks," International Journal of Electrical and Electronics Research, vol. 9, no. 4, p. 114 – 119, 2021.

[71] S. H. Rafi, Nahid-Al-Masood, S. R. Deeba, and E. Hossain, "A short-term load forecasting method using integrated cnn and lstm network," IEEE Access, vol. 9, pp. 32 436–32 448, 2021.

[72] K. Miao, Q. Hua, and H. Shi, "Short-term load forecasting based on cnn-bilstm with bayesian optimization and attention mechanism," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 12606 LNCS, pp. 116–128, 2021.

[73] U. Javed, K. Ijaz, M. Jawad, I. Khosa, E. Ahmad Ansari, K. Shabih Zaidi, M. Nadeem Rafiq, and N. Shabbir, "A novel short receptive field based dilated causal convolutional network integrated with bidirectional lstm for short-term load forecasting," Expert Systems with Applications, vol. 205, p. 117689, 2022.

[74] S. Fattaheian-Dehkordi, A. Fereidunian, H. Gholami-Dehkordi, and H. Lesani, "Hour-ahead demand forecasting in smart grid using support vector regression (svr)," International Transactions on Electrical Energy Systems, vol. 24, no. 12, pp. 1650–1663, Dec. 2014.

[75] Z. Zhang and S. Ye, "Long term load forecasting and recommendations for china based on support vector regression," in 2011 International Conference on Information Management, Innovation Management and Industrial Engineering, vol. 3, 2011, pp. 597–602.

[76] G. Zhang and J. Guo, "A novel method for hourly electricity demand forecasting," IEEE Transactions on Power Systems, vol. 35, no. 2, pp. 1351–1363, March 2020.

[77] A. Kavousi-Fard, H. Samet, and F. Marzbani, "A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting," Expert Systems with Applications, vol. 41, no. 13, pp. 6047–6056, 2014.

[78] Y. Chen, W.-C. Hong, W. Shen, and N. Huang, "Electric load forecasting based on a least squares support vector machine with fuzzy time series and global harmony search algorithm," Energies, vol. 9, no. 2, 2016.

[79] H. Li, S. Guo, H. Zhao, C. Su, and B. Wang, "Annual electric load forecasting by a least squares support vector machine with a fruit fly optimization algorithm," Energies, vol. 5, no. 11, pp. 4430–4445, 2012.

[80] J. Wang, L. Li, D. Niu, and Z. Tan, "An annual load forecasting model based on support vector regression with differential evolution algorithm," Applied Energy, vol. 94, pp. 65–70, 2012.

[81] F. Liu, T. Dong, T. Hou, and Y. Liu, "A hybrid short-term load forecasting model based on improved fuzzy c-means clustering, random forest and deep neural networks," IEEE Access, vol. 9, pp. 59 754–59 765, 2021.

[82] M.-W. Tian, K. Alattas, F. El-Sousy, A. Alanazi, A. Mohammadzadeh, J. Tavoosi, S. Mobayen, and P. Skruch, "A new short term electrical load forecasting by type-2 fuzzy neural networks," Energies, vol. 15, no. 9, p. 3034, 2022.

[83] R. Jain, N. Jain, Y. Gupta, T. Chugh, T. Chugh, and D. J. Hemanth, "A modified fuzzy logic relation-based approach for electricity consumption forecasting in india," International Journal of Fuzzy Systems, vol. 22, no. 2, pp. 461–475, 2020.

[84] R. Rathor and A. Bharagava, "Day ahead regional electrical load forecasting using anfis techniques," Journal of The Institution of Engineers (India): Series B, vol. 101, no. 5, pp. 475–495, 2020.

[85] S. Mollaiy-Berneti, "Optimal design of adaptive neuro-fuzzy inference system using genetic algorithm for electricity demand forecasting in iranian industry," Soft Computing, vol. 20, no. 12, pp. 4897–4906, 2016.

[86] S. K. Panda and P. Ray, "An effect of machine learning techniques in electrical load forecasting and optimization of renewable energy sources," Journal of The Institution of Engineers (India): Series B, vol. 103, no. 3, p. 721–736, 2022.

[87] A. Kavousi-Fard and F. Kavousi-Fard, "A new hybrid correction method for short-term load forecasting based on arima, svr and csa," Journal of Experimental and Theoretical Artificial Intelligence, vol. 25, no. 4, pp. 559–574, 2013.

[88] P. Mokilane, P. Debba, V. Yadavalli, and C. Sigauke, "Bayesian structural time-series approach to a long-term electricity demand forecasting," Applied Mathematics and Information Sciences, vol. 13, no. 2, pp. 189–199, 2019.

[89] P. Nop and Z. Qin, "Cambodia mid-term transmission system load forecasting with the combination of seasonal arima and gaussian process regression," in 2021 3rd Asia Energy and Electrical Engineering Symposium (AEEES), 2021, pp. 700–707.

[90] B. Zeng, M. Zhou, and J. Zhang, "Forecasting the energy consumption of china's manufacturing using a homologous grey prediction model," Sustainability (Switzerland), vol. 9, p. 1975, 10 2017.

[91] C. Hamzacebi and H. Es, "Forecasting the annual electricity consumption of turkey using an optimized grey model," Energy, vol. 70, pp. 165–171, 2014.

[92] Y.-S. Lee and L.-I. Tong, "Forecasting nonlinear time series of energy consumption using a hybrid dynamic model," Applied Energy, vol. 94, pp. 251–256, 2012.

[93] E. Yukseltan, A. Yucekaya, and A. Bilge, "Hourly electricity demand forecasting using fourier analysis with feedback," Energy Strategy Reviews, vol. 31, 2020.

[94] J. R. Cedeño González, J. J. Flores, C. R. Fuerte-Esquivel, and B. A. Moreno-Alcaide, "Nearest neighbors time series forecaster based on phase space reconstruction for short-term load forecasting," Energies, vol. 13, no. 20, 2020.

[95] A. Gafar Abdullah, W. Sopian, W. Arasid, A. Nandiyanto, A. Danuwijaya, and C. Abdullah, "Short-term peak load forecasting using pso-ann methods: The case of indonesia," Journal of Engineering Science and Technology, vol. 13, pp. 2395–2404, 08 2018.

[96] N. Q. Dat, N. T. Ngoc Anh, N. Nhat Anh, and V. K. Solanki, "Hybrid online model based multi seasonal decompose for short-term electricity load forecasting using arima and on-line rnn," Journal of Intelligent & Fuzzy Systems, vol. 41, no. 5, pp. 5639–5652, 2021.

[97] M.-R. Kazemzadeh, A. Amjadian, and T. Amraee, "A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting," Energy, vol. 204, p. 117948, 2020.

[98] S. Essallah and A. Khedher, "A comparative study of long-term load forecasting techniques applied to tunisian grid case," Electrical Engineering, vol. 101, no. 4, pp. 1235–1247, 2019.

[99] W. Zhang, W.-C. Hong, Y. Dong, G. Tsai, J.-T. Sung, and G.-F. Fan, "Application of svr with chaotic gasa algorithm in cyclic electric load forecasting," Energy, vol. 45, no. 1, pp. 850–858, 2012.

[100] C. Li, S. Li, and Y. Liu, "A least squares support vector machine model optimized by moth-flame optimization algorithm for annual power load forecasting," Applied Intelligence, vol. 45, no. 4, pp. 1166–1178, Dec. 2016.

[101] N. A. Mohammed and A. Al-Bazi, "An adaptive backpropagation algorithm for long-term electricity load forecasting," Neural Computing and Applications, vol. 34, no. 1, pp. 477–491, Jan. 2022.

[102] M. Li, X. Xie, and D. Zhang, "Improved deep learning model based on self-paced learning for multiscale short-term electricity load forecasting," Sustainability (Switzerland), vol. 14, no. 1, 2022.

[103] K. E. Farfar and M. T. Khadir, "A two-stage short-term load forecasting approach using temperature daily profiles estimation," Neural Computing and Applications, vol. 31, no. 8, pp. 3909–3919, Aug. 2019.

[104] P. P. Phyo and C. Jeenanunta, "Daily load forecasting based on a combination of classification and regression tree and deep belief network," IEEE Access, vol. 9, pp. 152 226–152 242, 2021.

IEEE Access

[105] A. Azadeh, R. Babazadeh, and S. Asadzadeh, "Optimum estimation and forecasting of renewable energy consumption by artificial neural networks," Renewable and Sustainable Energy Reviews, vol. 27, pp. 605–612, 2013.

[106] Z. Guo, K. Zhou, X. Zhang, and S. Yang, "A deep learning model for short-term power load and probability density forecasting," Energy, vol. 160, pp. 1186–1200, 2018.

[107] R. Salgado and R. Lemes, "A hybrid approach to the load forecasting based on decision trees," Journal of Control, Automation and Electrical Systems, vol. 24, no. 6, pp. 854–862, 2013.

[108] M. Ma, B. Jin, S. Luo, S. Guo, and H. Huang, "A novel dynamic integration approach for multiple load forecasts based on q-learning algorithm," International Transactions on Electrical Energy Systems, vol. 30, no. 7, p. e12146, Jul. 2020.

[109] F. Yu and X. Xu, "A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved bp neural network," Applied Energy, vol. 134, pp. 102–113, 2014.

[110] N. Bendaoud, N. Farah, and S. Ben Ahmed, "Comparing generative adversarial networks architectures for electricity demand forecasting," Energy and Buildings, vol. 247, p. 111152, 2021.

[111] Y. Hu, C. Ye, Y. Ding, and C. Xu, "Short-term load forecasting utilizing wavelet transform and time series considering accuracy feedback," International Transactions on Electrical Energy Systems, vol. 30, no. 7, p. e12455, 2020.

[112] J. Sun, H. Dong, Y. Gao, Y. Fang, and Y. Kong, "The short-term load forecasting using an artificial neural network approach with periodic and nonperiodic factors: A case study of tai'an, shandong province, china," Computational Intelligence and Neuroscience, vol. 2021, 2021.

[113] C. Velasquez, M. Zocatelli, F. Estanislau, and V. Castro, "Analysis of time series models for brazilian electricity demand forecasting," Energy, vol. 247, p. 123483, 2022.

[114] A. Asrari, D. Javan, M. Javidi, and M. Monfared, "Application of gray-fuzzy-markov chain method for day-ahead electric load forecasting [zastosowanie kombinowanej metody: Gray-fuzzy-markov do prognozowania obciążenia sieci elektrycznej]," Przeglad Elektrotechniczny-Electrical Review, vol. 88, no. 3 B, pp. 228–237, 2012.

[115] L. Xie, J. Xiao, Y. Hu, H. Zhao, and Y. Xiao, "China's energy consumption forecasting by gmdh based auto-regressive model," Journal of Systems Science and Complexity, vol. 30, no. 6, pp. 1332–1349, Dec. 2017.

[116] H. Zhao and S. Guo, "An optimized grey model for annual power load forecasting," Energy, vol. 107, pp. 272–286, 2016.

[117] X. Han and J. Chang, "A hybrid prediction model based on improved multivariable grey model for long-term electricity consumption," Electrical Engineering, vol. 103, no. 2, pp. 1031–1043, Apr. 2021.

[118] P. Li, J. Zhang, C. Li, B. Zhou, Y. Zhang, M. Zhu, and N. Li, "Dynamic similar sub-series selection method for time series forecasting," IEEE Access, vol. 6, pp. 32 532–32 542, 2018.

[119] X. Serrano-Guerrero, M. Briceño-León, J.-M. Clairand, and G. Escrivá-Escrivá, "A new interval prediction methodology for short-term electric load forecasting based on pattern recognition," Applied Energy, vol. 297, 2021.

[120] H. Migon and L. Alves, "Multivariate dynamic regression: Modeling and forecasting for intraday electricity load," Applied Stochastic Models in Business and Industry, vol. 29, no. 6, pp. 579–598, 2013.

[121] H. Shi, K. Miao, and X. Ren, "Short-term load forecasting based on cnn-bilstm with bayesian optimization and attention mechanism," Concurrency and Computation: Practice and Experience, vol. 12606 LNCS, pp. 116–128, Oct. 2021.

[122] G. Nalcaci, A. Özmen, and G. W. Weber, "Long-term load forecasting: models based on mars, ann and lr methods," Central European Journal of Operations Research, vol. 27, no. 4, pp. 1033–1049, Dec. 2019.

SAADMAN S. ARNOB received the B.Sc. degree from the Islamic University of Technology, Bangladesh. After completing his bachelor's degree, he is working as a Research Engineer in the Advanced Intelligent Multidisciplinary Systems Laboratory (AIMS Lab), United International University (UIU). His research interests include Forecasting, Optimization, AI, IoT, Big Data, MLOps, Renewable energy, Smartgrid, and AI in Healthcare sectors.

ABU ISHA MD. SADOT AREFIN received the Bachelor of Science in Computer Science & Engineering from Rajshahi University of Engineering & Technology, Bangladesh. He is currently the Chief Software Architect at an MNC, headquartered in Bangladesh. He engineers large scale software systems in multiple platforms leveraging cloud computing systems including VR/AR. He won the Mujib 100 Idea Contest in 2021 and received various rewards in his academic and professional career. He has designed the national database of National Load Dispatch Center of Bangladesh and contributed to many machine learning projects for Japanese and Saudi Arabian clients. His research interests include forecasting, smart-grid, intelligent systems, IoT, surveillance system, and smart city research.

AHMED Y. SABER received the Ph.D. degree from University of the Ryukyus, Japan, in 2007. He is currently VP Optimization and AI, ETAP, USA. He develops power systems simulation tools in ETAP using both deterministic and intelligent methods. His timely researches have been funded nationally and internationally including DOE. He won the IEEE Outstanding Engineer Award in Southern California, USA among more than 12,000 engineers for his contributions on smart-grid in 2012. He has published over 75 technical papers and holds 3 patents on power system optimization and AI. His research interests include forecasting, optimization, smart-grid, vehicle-to-grid, renewables, micro-grid, cyber-physical systems, intelligent systems and operations research.

KHONDAKER A. MAMUN received his Ph.D. in Computer and Biomedical Engineering from University of Southampton, UK. After that he worked as a Postdoctoral Research Scientist in the University of Toronto, Toronto, Canada. Currently he is a Professor of Data Science and Digital Health at the United International University (UIU), Dhaka, Bangladesh. Prof. Mamun is the Founder and Director of AIMS Lab (www.aimsl.uiu.ac.bd) at UIU, where his team actively performs R&D on the applications of 4IR technologies like, artificial intelligence, big data and internet of things in large scale intelligent system, healthcare, disabilities, education and brain computer interface (BCI). He has published more than 160 articles and holds patients. He has invented and implemented a number of intelligent systems, digital healthcare and rehabilitations service delivery model and mobile based solutions for Bangladesh. His research and innovation have also been featured in a local and international news, also recognized internationally and by the government of Bangladesh. He received National Award from Honorable Prime Minister Sheikh Hasina for his contribution to Autism. Recently, Prof. Mamun presented "Digital Bangladesh Award 2022" for his significant contribution for building Digital Bangladesh by Honorable Prime Minister Sheikh Hasina. Prof. Mamun is the founder of CMED Health (www.cmed.com.bd), a university research to commercialization that are working to create health inclusion for UHC through digital healthcare platform. CMED Health is currently serving over 3.9 million people in Bangladesh. He contributed to introduce first Deep Brain Stimulation surgery for Parkinson's patients in Bangladesh in collaboration with NINS and Prof. Tipu Z. Aziz, University of Oxford, UK. Prof. Mamun is the member WHO/ITU Focus Group on AI for Health, Digital Health Strategy 2020-2025 for the GoB, MTR of 4th HPNSP, MOHFW; Innovation Ecosystem Development, Hi-Tech Park, ICT Division; Industry Accademia Collaboration, UGC, MOE. He contributed to design and implement digital GP model with referrals for urban and rural areas of Bangladesh. He is also worked as a consultant for WHO, Unicef, ADB and World Bank on digital system design and implementation.

· · ·