IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 30 May 2022, accepted 28 June 2022, date of publication 1 July 2022, date of current version 11 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3187839

SURVEY

Load Forecasting Techniques for Power System: Research Challenges and Survey

NAQASH AHMAD[1], (Member, IEEE), YAZEED GHADI[2], (Senior Member, IEEE), MUHAMMAD ADNAN[1], (Member, IEEE), AND MANSOOR ALI[3], (Member, IEEE)

[1]Department of Electrical Engineering, National University of Computer and Emerging Sciences (FAST), Chiniot-Faisalabad Campus, Faisalabad 38000, Pakistan
[2]Department of Software Engineering, Al Ain University, Al Ain, United Arab Emirates
[3]Electrical Engineering Department, Ecole de Technologie Superieure (ETS), Montreal, QC H3C 1K3, Canada

Corresponding author: Muhammad Adnan (m.adnan@nu.edu.pk)

ABSTRACT The main and pivot part of electric companies is the load forecasting. Decision-makers and think tank of power sectors should forecast the future need of electricity with large accuracy and small error to give uninterrupted and free of load shedding power to consumers. The demand of electricity can be forecasted amicably by many Machine Learning (ML), Deep Learning (DL) and Artificial Intelligence (AI) techniques among which hybrid methods are most popular. The present technologies of load forecasting and present work regarding combination of various ML, DL and AI algorithms are reviewed in this paper. The comprehensive review of single and hybrid forecasting models with functions; advantages and disadvantages are discussed in this paper. The comparison between the performance of the models in terms of Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) values are compared and discussed with literature of different models to support the researchers to select the best model for load prediction. This comparison validates the fact that the hybrid forecasting models will provide a more optimal solution.

INDEX TERMS Load forecasting, machine learning, load shedding, root mean squared error, mean absolute percentage error.

I. INTRODUCTION

There is a need of uninterrupted provision of current to load system in modern power house. This need a suitable prediction about present and future load demands with very little errors. To achieve this goal, scholars and scientist tried to develop an optional and most efficient method called load forecasting, in which demand of future consumption of electricity is predicted. Many decisions such as unit commitment, off line network, dispatch and fuel allocation and several operations are controlled by load forecasting [1]. This provides a concept to power companies about consumption of electricity in future and time to reduce the difference between demand of load and generation capacity. The prediction of demand reduces the cost of power generation and helps to develop an organized power utility system. Several techniques based on machine Learning (ML) are used by energy and power utility companies to balance the demand and generation by predicting the need of energy and power. Load forecasting is a technique to manage supply and demand. However, the analysis of different affecting direct and indirect factors is required in this very difficult task. Although, there are a lot of benefits of using techniques of load forecasting, but there are some challenges in the way of accuracy of the methods. The convoluted and stochastic process is used in load forecasting. In the completion of forecasting, the data is influenced by weather related factors. So, the load of present hour depends on previous hour load, previous day load, demographic data, weather conditions, number of devices in forecasting area, economic data, number of customer and type of customer [2]. It is essential to maintain low load forecasting error in such complex situation. The measures such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) are determined in percentage and are the measures of prediction accuracy. To evaluate the certain algorithm, the values

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

are to be kept in range of few percentage points. The key component of the model of load forecasting is the load data. In order to train the model, it is needed to know the pattern of consumption of load data. So, the load data should be prepared for the purpose of training. The errors in data and missing values are corrected.

After that, the data about other factors such as historical event and historical weather is collaborated with electrical data. The accuracy in load forecasting depends on such factors. All the collected data is analyzed and different models are chosen for the process of load forecasting. The best accurate model is selected among all the models. The accuracy of the model also depends on other several factors. The variation in most of the factors largely depends on equipment and locations. These are also should be considered during the development of precise and accurate model. For the development of model, these factors are taken as input variables. Sometimes, all such factors are not considered due to their unavailability. Respective sources are used to collect the data about factors to develop a model in certain region. These respective sources may be weather office for weather data, and calendar for time factor. The accurate and precise load forecasting can be proved a great profit for electric utility companies but unnecessary errors can cause a loss of infrastructure and large amount of finance.

Haida and Muto *et al.* [3] described that negative and positive errors in forecasting may enhance the cost of electricity. Only decrease of 1% in mean absolute percentage error leave a impact of 3-5 % on the production side by minimizing generation cost from 0.1% to 0.3% [4]. Recently, renewable energy sources are gaining attention of electric companies and government. Some of the great challenges can be influenced by enhancing the progress in renewable energy sources. Time and location dependent nature of wind and solar energy is a challenge for electric company to create balance between load and variable production from new infrequent sides. Research is going on to replace traditional dispatch techniques with newly dispatch techniques which can be processed with wind and solar like energy sources [5]–[8]. The focus of this review paper is on Machine Learning based single and Hybrid methods.

Smart grid is a modern, accurate and fast electricity distribution and transmission network. The security, reliability, efficiency and precision of the grid are developed by using control, technologies, communication network and modern information. The planning and operation of electric power system strongly depends on important process of load forecasting. In terms of planning horizon's time, the load forecasting is divided into short term (up to 1 day/ week) medium term (1 day/ week to year) and long term (more than 1 year) [9].

The single predictive models are designed by Artificial Neural Network (ANN) and Support Vector Machine (SVM) learning methods. Before the description of ANN and SVM based hybrid algorithms, several known methods are presented in this paper. Such two algorithms have been used to optimize these single methods. The accuracy in forecasting

Abstract
I. Introduction
II. Aspects of load forecasting
III. Categories of load forecasting
IV. Classical load forecasting models
V. Time horizon's based types of load forecasting
VI. Evaluation criteria
VII. Data pre-processing techniques
VIII. Single methods for load forecasting
IX. Load forecasting in power and wind system
X. Statistical methods
XI. Probabilistic methods
XII. Probabilistic deep learning
XIII. Hybrid methods
XIV. Modern techniques
XV. Conclusions and future trends

FIGURE 1. Flow diagram of single and hybrid models for load forecasting.

improved by experimenting the three or more single methods on being the success of models based on two methods. The previous and present work on single models for short term forecasting have been reviewed in this paper. The significance and efficacy of the contingent factors based models have been evaluated statistically. The flow of useful information is shown in figure 1. The main contributions of this review paper are described next.

A. NEED OF FORECASTING

Since years, researchers are in search to improve the revenues and efficiency for the distribution and generating companies by developing an accurate load forecasting model. So far, this has invented many states of art methods. The cost effective supply and transmission can be made accessible by forecasting. Over and under consumption of production capacity can be avoided by well forecasting strategies. Forecasting can help to utilize the best possible capacity of electricity. The installation of future plants can be related accurately with future demand of consumption.

B. RESEARCH QUESTIONS

Recently, researchers had to face different challenges while improving the accuracy and precision of such models. Few of them had been described below:

1. To increase the forecast value by using single or integrating single methods.
2. The sudden change in prices and corresponding price based demand creates difficulty in getting accurate data.
3. To find best models for the investigation of static and dynamic time series forecasting.
4. Electric load forecasting relies on weather condition. Sometimes, unpredictable and sudden changes in weather condition may cause large error in load forecasting. This leaves a negative impact on the efficiency and revenue.

5. Further, it is difficult to get an accurate and precise demand forecast which is influenced by variable temperature and humidity.
6. Regular load forecasting models also affected by the sudden disturbances in power system. Unexpected faults result the development of unreliable and poor forecasting system.
7. Electrical distribution companies may face a big loss if they did not understand and take a decision about an acceptable error in short term load forecasting.

The other challenges for today researchers and utility companies are as Capacity planning • production and transmission capital investment • Financial forecasting • Effective Power Procurement Network Planning • Selling and saving of Excess Power • Planning and strategy of fuel ordering • Optimum Supply Schedule • Renewable Planning.

According to our best knowledge, there has been no review published which covers all the possible single and hybrid methods based on different load forecasting techniques. Also there is no extensive literature on the applications at the level of distribution grid for the comprehensive analysis including technical and non-technical losses, monitoring and operation, forecasting, predictive maintenance, flexible planning and interaction and relationship among them. Much of the identified applications lead to outputs which can be used as inputs for other applications. The important contributions of this review are listed next.

C. CONTRIBUTIONS

A holistic analysis of factors affecting the load forecasting along with benefits have been described in section II. Section III consists of categories of load forecasting. The classical load forecasting models such as economic and time series model have been provided in section IV. The time scale based types of forecasting techniques are detailed in section V. Also the recent published literature on the time scale based techniques is presented in the form of graph in this section.

The evaluation and comparison of different output of the techniques and models have been made by statistical formulas. Such formulas are detailed in section VI. Section VII describes the data pre-processing techniques. The single method for load forecasting including learning based methods; rule based methods have been detailed in section VIII. Also, energy management and applications of deep learning for wind forecasting is described in this section. Table 3 summarizes the uses of deep learning (DL) schemes in electric load forecasting. The comparison between the DL and other methods for different grid systems has been made in table 3. The comparison has been made in term of quantitative values.

The applications of deep learning methods in power load forecasting are described in section IX. The detailed section of statistical methods, probabilistic methods and probabilistic deep learning are described in next three sections. The section X consists of hybrid methods based on artificial neural network and support vector machine with sub section which are described in detail. The structure of extreme learning machine for single cluster was already published. In this review, structure of extreme learning machine for multiple clusters has been described. Modern techniques with smart grid and super smart grid are introduced and discussed in the section XI and XII describes the conclusions with way forward and future trends.

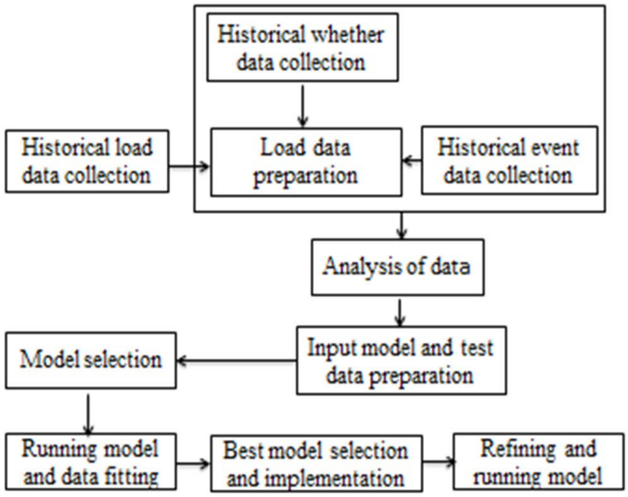


FIGURE 2. Flow chart of the development of load forecasting model.

II. ASPECTS OF LOAD FORECASTING

A. FACTORS AFFECTING LOAD FORECASTING

The experimental process of load forecasting depends on different agents which in result influence its precision. To obtain correct prediction, the dependent factors are needed to be selected carefully. The minor and major factors influence forecasting at each step but economic, time and weather are more considerable factors. The development of forecasting model is depicted in figure 2. A significant attention is needed in the selection of model because different algorithms are different variable parameters. The factors affect the selection of model and collection of data are described as follows.

1) WEATHER FACTOR

In the domain of load forecasting, the independent variable is the weather which influences the agricultural and domestic consumer. The consumer's behavior is affected by the weather. For example, in cold and hot seasons, the consumption of electricity increases due to turning ON and OFF of cooling and heating devices. This increases the demand of electricity in coolest and warmest seasons as compared to average temperature days. Also, sudden decrease in temperature can lead less consumption of temperature and so there is a probability of over-estimated load forecasting. The demand of future load is predicted by the results of weather forecast in different models. Temperature, dew factor and humidity are the weather factors. Also, electric utility companies use wind chill index (WCI) and temperature humidity index (THI). WCI and THI measure the winter's cold stress and heat discomfort in summer, respectively.

2) TIME FACTOR
The important and key factor in load forecasting is the time. Ruzic *et al.* [10] observed the load curve of many grid stations and concluded that load curve has ''day time'', ''week day'', ''week of month'', and also ''month of season'' properties. This also mentions that the accuracy and precision of the prediction relies on current data as well as previous day data. Further, the detection of load by timeframe is very important as it defines the quantity of data needed to run the process.

3) ECONOMIC FACTOR
Degree of industrialization, load management and price of electricity are the economic factors which influence the maximum demand and average load system [11]. The perfection of load forecasting is determined by influential level such as description of devices, behavior of customer, local population, compatibility of equipment and levels of employment. Such factors should be calculated for the prediction of long term forecasting in particular areas because the demand and extra generation of load affect the selection of model as well as acquisition of data [12].

B. BENEFITS
There are many advantages of load forecasting which make it interesting field for researchers. Since, from the early time of generation of electricity, there was a blazing question for the electric companies, how to balance the ever-enhancing demand of load and limited resources by determining the demand of load for the next hour, day and years. Although, the renewable energy resources have reduced the challenges of management but the process of energy harvesting from renewable resources is still cumbersome and expensive. The beforehand estimate of the load demand can optimize the dispatch of electricity. Apart from above mentioned advantages, there are many other load forecasting advantages as follows:

I. The emission of carbon and use of fossil fuels can be reduced by eradicating the over generation and under generation through the maximum utilization of power plants.
II. The maintenance of plants can be decided and well planned by understanding the load demand in load forecasting.
III. Load forecasting can be helpful in building a future generation plant. It can help in designing the size of plant, type of future plant, capacity, plant size and location. So, the cost of infrastructure for distribution and transmission can be estimated clearly.

III. CATEGORIES OF LOAD FORECASTING
There are two main groups of electric load forecasting's methods. (1) Classified methods (2) Artificial Intelligence based soft computing technologies. Different features such as type of customers, population, economic, indicators, time factors, new technologies, weather conditions and price of electricity affect the load forecasting [13]. The future value of single variable and statistical methods are the base of classical approach [14]. Repeated re-weighted least squares technologies; exponential smoothing, linear and multiple regressions are the classical load forecasting technologies [15]. Support vector machines, wavelet network, neural networks (NNS), genetic algorithms (GAS) and fuzzy logic (FL) are soft computing techniques [16]. Dependency of forecasting model on mathematical analysis, divide them into two categories: (1) Quantitative and (2) Qualitative methods. In quantitative techniques, the future forecasting is a function of past data, so these are considerable when past data is available. But the qualitative techniques depend on judgment and opinion of expert and consumers and such techniques are considerable when past data is not available [15].

IV. CLASSICAL LOAD FORECASTING MODELS
Econometric models and time series models are mostly used classical load forecasting models. Such technologies are deducted from the load relating observations of past, while the econometric methods are the combination of statistical techniques and economic theory [17].

A. ECONOMIC METHOD
In order to meet the demand of electric forecasting, the statistical techniques and economic theory are combined in econometric approach. So, the relationship between consumption's influencing factors and energy consumption can be estimated through this approach. Further, time series and least square method are used to estimate the relationships. The latest historical data is used to assemble the estimates, when consumption of electricity in industrial, residential, commercial sectors is measured as a function of economic, weather, and other factors [16], [18], [19]. The econometric is beneficial because it furnishes comprehensive information about future demands, also why future demand is increasing and how it is affected by different factors [16], [18]. Despite of such benefits, the modification in electricity, does not change in the forecast duration as in the past [18]. Thus its unsuccessful is to admit the interdependence between quantity and prices.

B. TIME SERIES MODELS
The load evolution is analyzed by time series methodologies in order to discover the dynamic attribute of load and to deduce them for future through mathematical tools. So, it is advantageous to break the load into constituent and separately treated them [17]. The prediction of time series can be considered as an issue of model formation that give rise to mapping between input and output values. After the model formation, the future values can be forecasted which depends on past and recent values [9]. So, the structural simplicity is the benefit of time series models. Further, under study variable's observations are absolute sufficient. Instead

of such advantages, the cause and effect relationship is not described by them [20]. So, the changes occur in variable are not provided by the argument of the time series model.

V. TIME HORIZON'S BASED TYPES OF LOAD FORECASTING

On the basis of time horizon, there are four classes of load forecasting. The implementation of different machine learning algorithms based on four classes: VSTLF, STLF, MTLF and VSTLF. The popular class of load forecasting is VSTLF. The future load is forecasted by past load in VSTLF [21]. So, many factors including information of used land, economics, and temperature can be optional. Time scale or time horizon is the time period required to generate the forecasts. It is the important parameter to classify the forecasting techniques in smart grid systems [22]. Energy forecasting is divided into following four types according to time scale.

A. VERY SHORT TERM FORECASTING (VSTLF)

This type of forecasting consists of time scale from minutes to hour (0-3h) [23]. It can help to deal with random variations in renewable energy production which can be predicted before very short duration of time. It has many uses relating renewable energy sources (RES) including solar and wind production forecasting [24], [25]. A hybrid approach consisting of fuzzy logic and artificial neural network (ANN) was used by potter *et al.* [26] to forecast the Tasmanian wind system before 2.5min.

B. SHORT TERM FORECASTING (STF)

It is the technique to forecast the energy forecasting ahead of few minutes to few days. It has a key role in different grid operations involving reliability analysis and dispatch analysis [27]. Further, it helps to avoid over estimation and under estimation of the energy demand and thus contribute substantially in the reliability of grid [28].

C. MEDIUM TERM FORECASTING (MTF)

It is used for time scale expands from few days to months ahead during a year [29]. It helps maintenance, adequacy assessment and fuel supply in smart grid systems. Further, it plays an important role to evaluate the financial attributes of energy system by contributing to risk management [30].

D. LONG TERM FORECASTING (LTF)

This type of forecasting involves time scale ranging from months to even years. LTF is very important for every production and load growth planning operations for long duration of time [31], [32]. The big advantage of LTF is that it can remove the effects of random fluctuations occur in short term and make the prediction of long term trends. Azad *et al.* [33] used neural network to predict the Malaysian meteorological station's winds speed swing for a year in order to control the challenges caused by irregular nature of wind production. The figure 3 shows the number of publications about energy forecasting systems with respect to time scale with the time interval of five-year duration since last two decades. During this period, STF stands first, while second most numbers of publications are made for LTF. The previous research made STF the most widely used forecasting technique for grid operations and planning in recent times. The suitable choice of the forecasting time scale is important to time the hyper parameter of the forecasting method. Relating this content, the author [34], [35] focused on selection of relevant time scale to build the machine learning model.

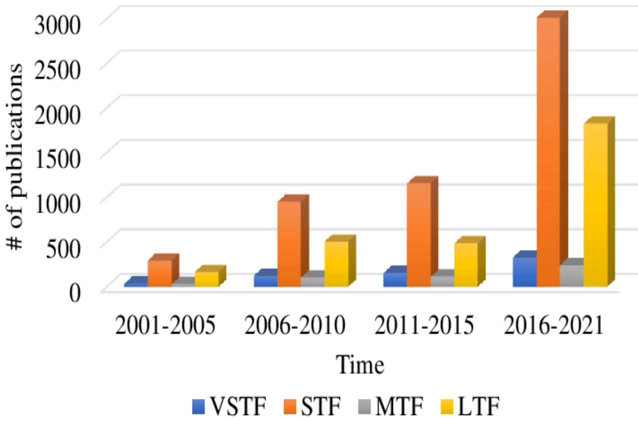


FIGURE 3. Number of publications with respect to time scale.

The modeling relationships among time, weather conditions and load are not used but the estimate of recent discovered load to nearly future is used. There are very few methods for VSTLF including genetic algorithm, autoregressive moving average models and artificial neural network. STLF is utilized for time hardly from minutes to hours. STLF is the important source of information for daily operations and it is important for system operations [36]. Researchers are taking more interest to design predictive models because STLF can be used to approximate the long time load.

It is essential to have accurate predict knowledge of affecting factors to improve short term model. The relation between demand of load and factors is the basic purpose to look for because instantaneous demand may be different. For duration of days to months, usually MTLF is used for load forecasting [37]. It becomes popular in peak summer or winter. For load duration from few weeks to many years, LTLF is considered [38]. The factors including weather data, characteristics of install devices at areas of interest, history of load and numbers of customers are accounted in it. The factors of economic are taken into account for long period methods of load forecasting. Load based companies among the factors, time period and applications has been described in table 1.

The most popular method among all four methods is STLF. It plays a crucial role in making power system's operating strategies due to its intrinsic connection with other class of forecasts. The addition of economic factor in STLF changed it into MTLF and LTLF. Also, STLF can be converted into VSTLF by the addition of previous hours loads as input factor

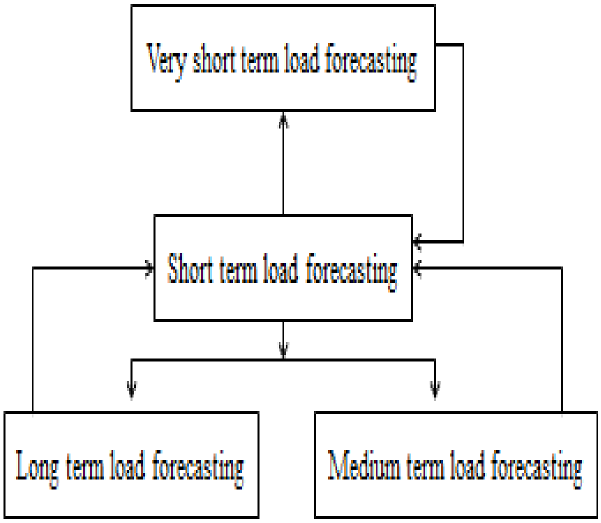


FIGURE 4. Conversion flow process between STLF and LTLF, MTLF, VSTLF.

in STLF model. STLF can capture the auto correlation of load of present hour and loads of previous hour. By taking STLF as a base, new series can be prepared and the residuals of previous loads can be obtained. By the addition of future residuals, into STLF, the VSTLF can be obtained. Figure 4 shows the conversion process flow between STLF and LTLF, MTLF and VSTLF. As it is shown in figure 5, load and weather history are taken as input in the processing of STLF method to model the extrapolating process with the summation of weather forecast data. The minute to hours' load prediction can be performed by the forecasting data. A lot of STLF techniques are designed for model. Few of them are regression analysis, fuzzy logic (FL), hybrid methods, time series, artificial neural network (ANN), genetic algorithms (GA's) and support vector machine (SVN) as shown in figure 6.

In current research, the methods of Computation Intelligence (CI) are mostly used. CI is considered as most potential computer algorithms to intuitively learn a specific task from past data. The unique property of CI methods is their ability for independent operation with no need of quantitative correlations or difficult mathematical formulations between inputs and output. The hybridized type of CI models has become more efficient as compare to their counterpart's single model.

VI. EVALUATION CRITERIA

To check the correctness of the methods used for the prediction of real values of load, different criteria are utilized to evaluate the techniques of load forecasting. The research of many researchers based on statistical metrics to optimize the precision of their model, newly developed statistical metrics such as probabilistic load forecasting metrics. Due to wide adaptation and extraordinary academic values in industry, literature on probabilistic forecasting is still in developing phase. The most important static metrics used by researchers are shown in table 2.

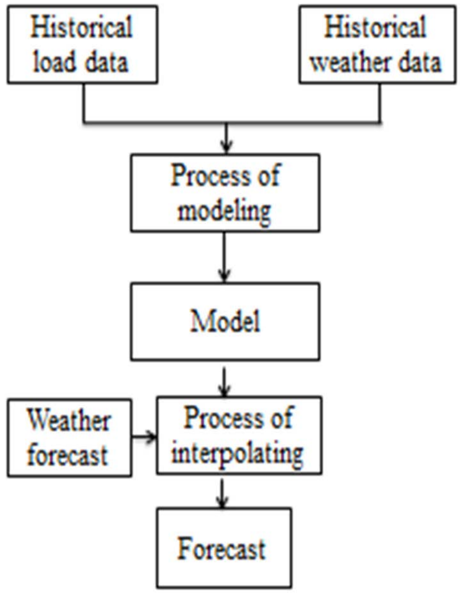


FIGURE 5. Flow chart of STLF method.

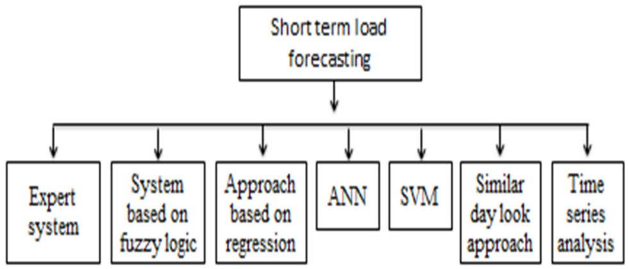


FIGURE 6. Mostly used single methods for STLF.

Here, **n** corresponds the number of samples, $\mathbf{y_i}'$ **and** $\mathbf{y_i}$ are the predicted and actual values of the model. Every metrics have disadvantages as well as advantages. Two-degree loss function is provided by RMSE but it gives an extra weight to large errors as compare to small ones. Naturally, the average error can be calculated by MAE. MAPE can be applied to low and high volumes products and it does not depend on scale. It may lead to biased forecasting due to differential penalty. Difficulty in controlling zero and small denominators are the weak points of MAPE. Such weak points are not related to problems of traditional load forecasting because of zero or very low level load at aggregated level [39], [40]. Also, the values of such metrics vary for different parameters and datasets. So, the comparison among the outcomes of different techniques is absolute difficult. For comparison, there is no such task to experiment the methods in a dataset.

VII. DATA PRE -PROCESSING TECHNIQUES FOR ENERGY FORECASTING

The early important phases in data investigation are the pre -processing and data representation. The incorrect and misleading forecasting outcomes can be produced by the

TABLE 1. Different types of load forecasting methods.

Methods	Time duration	Factors			Uses
		Temperature	Economics	Use of land	
VSTLF	Few minutes	Not compulsory	Not compulsory	Not compulsory	To generate forecasting. Distribution schedule.
STLF	Few hours	Compulsory	Not compulsory	Not compulsory	Maintenance schedule. Allocation of spinning reserve.
MTLF	Days to months	Simulated	Compulsory	Not compulsory	Seasonal forecasting
LTLF	More than year	Simulated	Simulated	compulsory	To plan about growth of generation

processing of data with redundant and irrelevant information. The pre- processing of raw data can be carried out by using different methods including cleaning feature engineering, normalization, transformation and dimensionality reduction. These, methods reshaped the data into usable training data set that can be used as an input for processing techniques [41]. Many authors have used various pre- processing techniques to increase the accuracy of forecasting for energy systems. Such pre- processing techniques are described in the following subsection.

A. SINGULAR VALUE DECOMPOSITION (SVD)

Feature engineering and dimensionality reduction are the important considered pre -processing techniques. The authors in [42] obtained the lower dimension's data by decomposing the high dimensional data. They used SVD prior to energy prediction. The authors obtained dimensionality reduction by using SVD decomposition matrix and tensors. Authors in [43] proposed compression and data decomposition method and k-means SVD based lead profiles. Initially, the sparsity of load profiles was exploited by compressing the data through coding technique. Then partial usage patterns (PUPS) were decomposed and extracted by using decomposed k-SVD method. The results of the proposed technique were better as compared to k-means clustering and discrete wavelet transform (DWT).

B. PRINCIPLE COMPONENT ANALYSIS (PCA)

PCA is traditional reduction technique applied to transform the high dimensionality data into lower dimension orthogonal matrix's form called principal components. The features present in PV generation data can be eliminated by using the proposed method which has no importance in forecasting output [44]. The predictions can be made more precise and accurate by proposed method. The results were compared with particle swarm optimization and differential evolution and found least RMSE values in their results.

TABLE 2. Formulas for evaluation criteria.

Name of criteria	Formula
Mean Absolute Error (MAE)	$\frac{\sum_{i=1}^{n}\lvert y_i' - y_i\rvert}{n}$
Root Mean Square Error (RMSE)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i' - y_i)^2}$
Mean Average Percentage Error (MAPE)	$\frac{1}{n}\sum_{1}^{n}\frac{(y_i - y_i')}{y_i}$

C. AUTO ENCODERS (AE)

AEs are the category of neural network, which use encoding layer to encode dimensional data. The reconstruction ratio in monitored and decoding layer is used to analyze the efficiency of AEs. Further, there are many versions of AEs including sparse auto-encoders (SAE). Chen *et al.* [45] used the unique type of SAE to classify the errors by recognizing the extra voltage and determined the disturbances in power quality. For changing weather conditions, the accuracy of PV estimation was improved by the combination of AEs and LSTM [46]. For multiple sites, authors used encode-decoder to reduce the influence of seasonal uncertainty for day ahead prediction.

D. CONVOLUTIONAL AUTO-ENCODERS (CAE)

Auto-encoders integrated with different layers of CNN are called CAEs. Ryu *et al.* [47] highlighted the role of data dimensionality reduction and data compression using CAES. They suggested a CAES based extraction technique to capture the seasonal and variations by showing many dimensional space into few dimensional vector. The authors used their technique and claimed 19-40% reduction in reconstruction error. They also claimed that compression ratio increased by 130% as compared to other standard methods. Shao *et al.* [48] combined CAE and LSTM to execute STLF involving time and energy. The authors claimed that they obtained a 10% improvement in prediction efficiency.

E. VARIATIONAL AUTO-ENCODERS (VAE)

VAE is a new type of AEs which performs encoding decoding process by using the idea of Bayesian optimization and variational inference [49], [50]. VAE combined with neural layers and its versions were explained for the uses of anomaly detection. Also, the process of forecasting became more efficient [51]–[54].

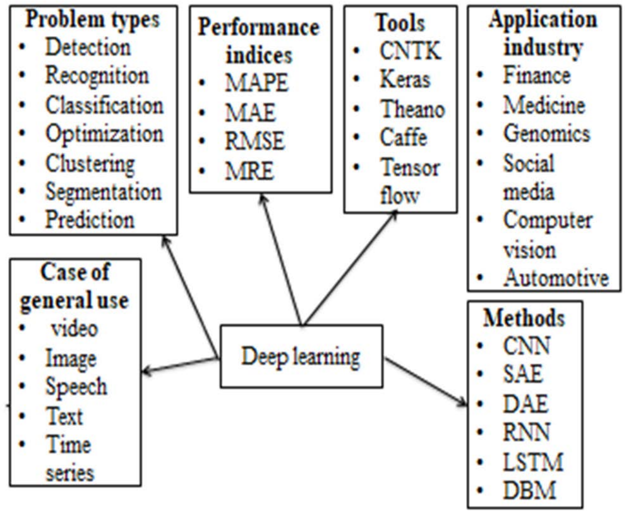


FIGURE 7. Aspects of DL.

VIII. SINGLE METHODS FOR LOAD FORECASTING

By using single methods which based on artificial intelligence, nonparametric and parametric, load forecasting is executed for short duration. Local fuzzy reconstruction and ANN method, statistical models and specific regression are employed in order to forecast from seconds to minutes [55]. Figure 6 shows the methods for STLF such as SVM, Time series analysis, FL augments, expert systems, similar day look up approach. Such single methods are also reliable for MTLF. Methods being used for MTLF with the addition of Adaptive Neuro Fuzzy Interference System (ANFIS), Grey model and wavelet transform are used for LTSM [56]. A clear picture of practical terminologies is crucial for the use of above mentioned methods, which are compared and provided in the following subsections.

A. LEARNING BASED METHODS

1) DEEP LEARNING

Dechter introduced the term of deep learning in 1986 [57]. After small modification in contents and algorithm, it was called Deep Neural Network (DNN). It is a shallow structure with hidden layer, input and output layer. However, the architecture of deep learning comprises more layers in compare to three-layer multilayer perception (MLP). Deep learning is a type of machine learning with non-linear aggregation of multi-layer and sophisticated algorithm. DL provides the solution of non-linear problems such as classification, recognition and detection. The working of deep neural network (DNN) is slow due to some technical restriction such as training methods and computing resources, tough training, unavailability of enough data, local minima and problems optimization [58], [59]. The codes and difficult mathematical algorithms made it to take much time to train. The scalable uses with large computing capacity can be developed by using graphical processing unit (GPU). Geoff *et al.* [60] and Yann *et al.* [61] used a newly algorithm called greedy layer-wise to create a breakthrough. According to these advancements, DL can be used in field of genomics [62], computer vision [63], robotics [64] and automatic vehicles [65]. The top companies with excellent technologies such as IBM, Facebook, Google, and Microsoft invested much in research and development of project including Watson platforms, Big sur, and Tensor flow. The problem categories, performance indices, model parameters, used methods, case of general use and performance indices are the important aspect of DL. Such classifications are shown in figure 7. Different fields such as social media, genomics computer, finance and automotive have used DL. Further, time series, any video, speech signal can be used by DL to detect, classify or predict the data set [66]. Training algorithm based architectures of DL can be divided into different groups.

Deep learning's network structure reflects the simulation of cerebral cortex system of human where it copies the function of human brain [67]. Deep learning has many applications in different forecasting due to its property of nonlinearity modeling [68]. It can exhibit hyper variable and high dimensional functions, which make worse the computational problem in forecasting modeling. In order to get precise output, many layers are needed which cause over fitting problem. The complex algorithms of deep learning need extra runtime. There are many deep learning methods such as deep belief networks (DBN), Convolution neural network (CNN) and deep auto encoder. Two distinct DNN models was proposed by (Ryu *et al.* [68]) in order to learn the difficult relations in current and past consumptions, weather changes for customers.

The load profile of 24 hours from the observations of past data can be produced by this DNN model. The model for forecasting of hourly load was suggested by He *et al.* [69]. The co-movement observation of Capula model is combined with deep belief network which based on layer wise pertaining. The comparison between classical DBN, Neural

Network, Extreme Learning Machine (ELM) and Support Vector Regression (SVR) in day week ahead forecasting represents excellent results in proposed data driven method. The important problem of over fitting in DL is described by Shi *et al.* [70] by enhancing the volume and diversity of data. The pooling based deep recurrent neural network (PDRNN) was proposed to batch the profiles of customer's load into the pool of inputs. There are two stages of proposed model: 1) Load profiles pooling, 2) STLF with deep Recurrent Neural Network (D-RNN).

The uncertainties in pertaining factors and weather conditions can be learned by this method. In Ireland, 920 smart meters were tested by this proposed method and found outperforms SVR (13.1%), classical D-RNN (6.5%) and ARIMA (19.5%) in terms of RMSE. Resident behavior based framework was designed with DL based on long short term memory (LSTM) [71]. The high volatility and variability is the steady nature of power system. The Capula Model is combined with DBN to reduce this challenge and errors of forecasting [72]. DP is a type of machine learning which depends on deep architectures. Many processing layers are used to arrange architectures in neural network to adjust the non-linear relationships between response and independent variable. Although DL has attained much attention of forecasting community but its major disadvantage is that it suffers from the issue of over fitting due to many layers. Shi *et al.* [70] addressed the over fitting issue of DL and tried to enhance the volume fed and data diversity into pool based deep recurrent neural networks. According to this proposed methodology, the addition of neighbor's historical load data can increase the input volume. In Ireland, 920 smart meters were tested by this method and the results were satisfactory. The DL architectures have been divided into five structures including auto-encoders (AEs), convolutional neural network (CNN), restricted Boltzmann machine (RBM), RNN and other integrated approaches.

2) CONVOLUTIONAL NEURAL NETWORK (CNN)

CNN is deliberated to train high level attributes through convolution [73]. CNN is widely used in sound processing, image, speech and video [74]. Such networks have become very popular due to their better attributes in object recognition via competition and image [75]. CNN has many applications in different fields such as individual recognition and traffic signal recognition [76]. CNN has billions of interconnected neurons and millions of weight values. The structure of CNN via image classification is shown in figure 8A. The structure consists of pooling layer, completely connected layer, input image and convolutional layer [63].

3) AUTOENCODERS AND VARIANTS

An AE is the individually learning algorithm based on back propagation (BP) to bring out a feature by using input data the three layered structure consists of input, hidden and output segment [61]. The encoder function $F_{\emptyset}$ consists of exhibition of inputs and decoder function (r) rebuilds the input into old representation. The encoder function, decoder function and small dimensional space are combined to minimize the reconstruction error to build the fundamental structure of AE. The function of $F_{\emptyset}$ is to transform the input 'x' into hidden representation $h^{(t)}$ through effective calculation, while 'r' performs the function of transformation of mapping from hidden layer to output layer. Both functions are described mathematically by equation (1) and (2). There are some restrictions on the AE due to the hidden units so AE is used in low dimensional representation of data.

$$h^{(t)} = F_{\varphi}\left(x^{(t)}\right) \rightarrow h = s\left(W * x + b\right) \tag{1}$$

$$r = g_{\emptyset'}\left(h\right) \rightarrow r = s\left(W' * h + b'\right) \tag{2}$$

The set parameter for a model is Q = {W, b, W', b'}. The parameters of data are variable such as bias vectors (b) and (b') with encoder and decoder weight function (W) and (W'), respectively. The model variables are optimized by minimizing the mean reconstruction function as indicated in equation (3). The error value of this auto encoder minimizes the reconstruction error of L (x; r) [77].

$$J_{DAE}\left(\emptyset\right) = \sum_{t} L(x^{(t)}, g_{\emptyset}(F_{\varphi}(x^{(t)}))) \tag{3}$$

The use of AEs is preferred in case of high dimensional and unlabeled data. The basic structure of AE provides base to derive various learning algorithms such as denoising auto encoders (DAEs) and stacked auto encoders. There are many AEs in the structure of learning algorithm (SAE) as shown in figure 8B [78]. While, the artificially reconstructed corrupted data ($\tilde{x}$) taken from clean data (x) is used in DAE algorithm. Then it is mapped to (h) by AE and is (x) is reconstructed. Resultantly, the output is brought up by the uncorrupted data and useful features are extracted.

4) RESTRICTED BOLTZMANN MACHINE (RBM)

RBM is a kind of ANN which consists of decision making units and uniformly connected neurons. It is a non-linear graphical developed model which represents probabilistic distribution made of observational, hidden or visible vectors [79], [80]. RBM can model the binary numbers into two layers. Here, features are detected by binary pixels through weighted links. The "visible" unit of RBM is represented by the pixels while feature detectors are related to "hidden" units. The equation (4) represents the values of energy for the common structures consists of visible and hidden units (h, v).

$$E\left(v, h\right) = v^{T} Wh - v^{T} b^{v} - h^{T} b^{h} \tag{4}$$

Here, W represents the weight matrix between the visible (v) layers and hidden (h) layers, b^h and b^v corresponds to the biases hidden and visible variables, respectively. The structures of designed Boltzmann machines are shown in figure 8C. The same layer's neurons are independent but dependent on next layer. Such properties of RBM make it faster as compared to classical Boltzmann machines [81]. The stack of RBMs forms the deep Boltzmann machine [82], [83].

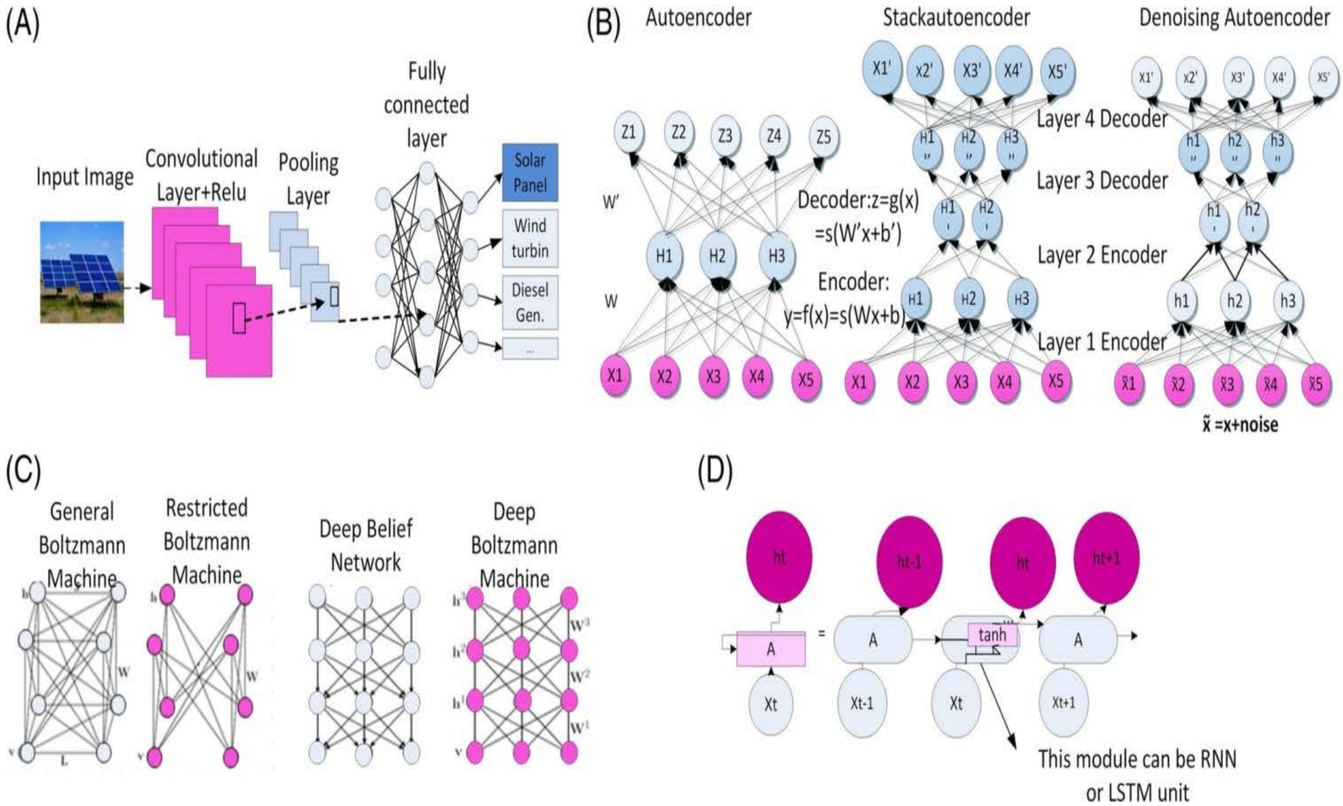


FIGURE 8. Structures of DL algorithms (A) CNN, (B) AEs, (C) RBM, (D) RNN [87].

5) RECURRENT NEURAL NETWORK (RNN)

RNN is a class of artificial neural network (ANN) which develop a directed graph and sequence unit by using network of units. This grants the representation of dynamic temporal conduct. Entire inputs are interconnected to each other in RNN and foregoing hidden value is given as input to the current state. So, the sequential data with changing length can be handled by RNN which makes it useful for DL [84], [85]. Usually, it is utilized in circumstances such as language processing, market machine translation and text to speech. The intrinsic structure of RNN is very deep since unfolded time of RNN is described by the combination of many non-linear layers. Different deep RNN structures with several approaches have been suggested to perform better as compared to classical RNN [81]. Among them long short term memory (LSTM) is widely used solution [86]. The structure of RNN and LSTM is shown in figure 8D.

IX. LOAD FORECASTING IN POWER AND WIND SYSTEMS

The adaptation and energy generation planning in power systems played an important role to increase the importance of energy demand forecasting. The power system should have the ability to handle changes in energy demand and respond dynamically in distribution of energy optimally and efficiently. Further, the optimum use of renewable sources for smart grids should be taken. For maximum efficiency, the smart power grids use adaptable and intelligent elements imperatively. Such elements need modern techniques for precise and accurate future energy generation and energy demands estimates. The use of different models in energy load forecasting depends on category of load forecasting. Energy consumption is a problem of time series forecasting various linear methods of forecasting including auto-regressive integrated moving average (ARMA), linear regression (LR), auto-regressive integrated moving average (ARIMA) and non-linear methods including artificial neural network (ANN), General regression neural network (GRNN). Multi-layer perception (MLP) and support vector machine (SVM) have been described in literature [88], [89]. Table (3) presents a comprehensive review of the applications of different DL methods in electrical load forecasting. The comparative methods also have been presented in table (3). As, it is clear from the table (3), that the contents of table are mostly about short term load forecasting. The most commonly used algorithms are of LSTM-RNN [90] followed by RBM based [91] and DBN deep architectures in this frame work of this topic. Different features of past energy consumption data set have been derived by using parallel components of CNN by [90] unlike relevant studies of Bouktif *et al.* [92] and Jian *et al.* [93]. Bouktif *et al.* [59] used the same data for MTLF and STLF studies to get maximum time lags and quantity of layers to predict performance by making a comparison with other studies through LSTM methods.

TABLE 3. Summary of uses of deep learning schemes in electric load forecasting.

DL methods	Uses	Comparative methods	Outputs	References
LSTM-RNN; Parallel CNN	AL, STLF	SVR, DNN, LR and CNN-RNN	MAPE 1.34	[44]
Integrated SAE and ELM	M-LTLF, BL	MLR and SVR	MRE 2.92	[45]
LSTM-RNN	S-MTLF, AL	ML benchmark	RMSE 341.4	[46]
LSTM-RNN	STLF, AL	NARX, NNETAR, SARIMA, SVR	MAPE 0.0535	[47]
EMD-DL Ensemble	M-LTLF, BL	ANN, SVR, DBN, EMD-SVR, EMD-ANN	MAPE 3.00	[48]
DBN	AL, STLF	Conventional NN	MAPE improvement .21%	[49]
FCRBM	AL, STLF	SVM	MAPE 1.43	[50]
DNN Deep energy	STLF, AL	RF, MLP, SVM, DT, LSTM	MAPE 9.77	[51]
LSH Deep auto encoder	STLF, AL, VSTLF	-	RMSE 6.99	[52]
SDA	STLF AL	SVM, LAS, SO, Conventional NNs	MAPE 2.47	[53]
CNN	STLF, AL	SVM, TSAM	DBI 8.57	[54]
IoT based DL	STLF, AL	DCEN, SDNN, HW	MAPE 1.00	[55]
CNN-LSTM	STLF, BL	SVR, LSTM, ARIMA	MAPE 10.16	[56]
LSTM-RNN	STLF, BL	ELM, KNN-R, BPNN	MAPE 8.18	[57]
D-RNN	STLF, BL	Various forecasting approaches and unit types	RMSE 111.9	[58]
LSTM based RNN	STLF, BL	KNN-R, FFNN	Mean MAPE 22	[59]
D-RNN	STLF, BL	SVR, ARIMA	RMSE 0.45 KWh	[60]
LSTM	S-LTLF, BL	Q-RNN, Q-LSTM	Error 0.18 KWh	[61]
HFM with GPU	STLF, BL	-	-	[62]
DNN	STLF, BL	ARIMA, DSHW	MAPE 8.84	[63]
CNN	STLF, BL	ANN, SVM, FCRBM	RMSE 0.732 %	[64]
FCRBM, CRBM	STLF, BL	RNN, SVM	RMSE 0.17 KW	[65]
DBN, SARSA with DBN	LTLF, BL	Q-learning without DBN, SARSA without DBN	RMSE 0.02	[66]
CPU-GPU	STLF, BL	CPU	-	[67]
DBN Ensemble	M-LTLF, AL	DBN, SVR, FF-NN	MAPE 5.93	[68]
LSTM-DRNN	M-LTLF, AL+BL	3 layer MLP	Mean RMSE7.10 KWh	[69]
CNN, LSTM-NN	M-LTLF, BL	RR, LR, PLS	RE 17.29 %	[70]
CNN, RNN	M-LTLF, BL	SARIMAX	Mean RMSE 17.84 KW	[71]

AL, aggregated level; **BL**, building level; **DBI**, Davies-Bouldin index; **GBRT**, gradient boosting regression tree; **MAPE**, mean absolute percentage error; **Q**, quantile; **RE**, relative error; **RMS**, root mean square; **SARSA**, state action reward state action; **TD**, time domain

A. ENERGY MANAGEMENT BY DEEP LEARNING

Energy management by deep learning is the process in which energy generation and consumption is observed, planned and controlled. The cost of electric bills of consumers can be reduced by better energy management [116], [117]. Energy management integrates the renewable energy source (RES) and energy storage system (ESS) in power systems [116]. The use of RES can be made optimal through proper management strategy. For example, cost can be reduced by shifting all the ESS and loads to solar energy in day instead of purchasing from utility companies. The life of ESS can be enhanced by proper management strategy. Proper charging and discharging of ESS up to safe specific limit can increase the life of batteries. For optimum life of batteries, the minimum and maximum storage levels are 10% and 90% respectively [118]–[120]. The efficient energy management can be attained by improving the energy production due to sporadic power generation from renewable energy sources. Various RESs and forecasting methods have been developed by the researchers which have important properties including temperature, solar irradiance and wind speed. There are three primary steps in load forecast of wind and solar energy following deep learning as shown in fig 9. Firstly, the input data is cleaned and normalized by data pre-processing step. Also,

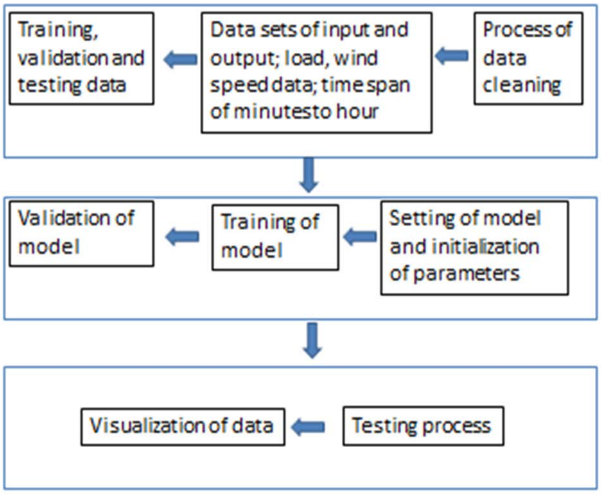


FIGURE 9. Flow chart of DL based wind energy forecasting.

the input data is cleaved into training, validation and testing data sets. Secondly, the valid and appropriate model is created by performing model training [103], [121]–[123]. Finally, the trained model is used to perform forecasting process. Wind energy forecasting based on deep learning techniques is given in the following section.

B. APPLICATIONS OF DEEP LEARNING IN WIND POWER SYSTEM

The concept of smart grid came out with the evolution of presently developed power systems that can combine production, demand of energy and storage areas. Flexibility is the requirement of these grids to generate and distribute energy at high level to reduce energy consumption and to minimize its usage. Further, the planning and operation of power system along with control evolution has become necessary due to latest improvements in super and smart grid systems including insistent requirement in quality and efficiency of power grids, deregulation of electric markets, divergence of distributed generation and exchange of power between utilities. In this study, the scope of DL methods in load forecasting has been discussed.

C. WIND ENERGY FORECASTING

Recently, wind energy has attained much attention due to pollution free energy source, wind turbines emit the lowest carbon [124]. The severe issues of fluctuations and uncertainty in the generation of wind energy obstruct the economic operations of power system. So, accuracy is necessary in wind energy management system for residential sections. Without, effective management, no optimal benefits from wind system can be obtained. Most of the data sets about wind speed were picked up from Asia with span up to three years and consists of wind direction, humidity wind speed, temperature and pressure. (Lin *et al.*, 2019 and Mosavi *et al.*, 2019) [125], [126] suggested a method for the forecasting of wind speed for effective energy management, where they used genetic algorithm (GA) along with deep learning. The parameters of deep belief network (DBN) were determined by using GA. The data of weather of different cities of Taiwan was used [104], [127], [128]. The wind speed was forecasted by exploiting the datasets of time series and multivariate regression. GA and DBN based model was validated by performing simulations [124]. The results showed the productiveness of design model over opposite. Cheng *et al.*, 2018 suggested a model consisted of RNN, wavelet threshold denoising (WTD) and an adaptive neuro-fuzzy interference system (ANFIS) was developed by [129]. The model was developed for wind energy forecast in residential sectors. Wind speed was made smooth by WTD to capture the fluctuation trends and data sets were used to train RNN. The final prediction of wind speed was performed by upper layer of ensemble model which was used further to predict the wind power production. The outcome of the model ensures its superiority over other models.

Large quantity of simulated scenarios based technique known as probabilistic wind energy ramp forecasting (P-WPRF) was proposed by [130]. The efficiency of the model was verified by exploiting the public place data set for area of Dallas, Texas, USA. The results were verified by performing the simulations studies which confirmed the effectiveness of their research work with higher stability and accuracy. A model for the prediction of wind speed was proposed by [131] under the cost based loss function. Various effective forecasting for the wind speed was formulated by developing a cost-oriented boosted regression tree (BRT). The productivity of this model was verified by different case studies with different data sets. Also the comparison between the conventional unbiased and proposed method was made. A hybrid approach called EWT-LSTML-Elman was proposed for the prediction of wind speed [132], [133]. This approach is the integration of two RNNs and empirical wavelet transformation (EWT). The data about wind speed was decomposed by EWT into many sub-layers, and the low frequency of sub layers was forecasted by LSTM network. Finally, high frequency of sub-layer was predicted by Elman neural network (ENN). Eleven forecasting algorithms were benchmarked to verify the performance of this proposed model. The experimental results showed the high precision. The summary of different approaches for wind speed and energy prediction is given in table 4.

Here, $RMSE_p$ means RMSE of proposed model and $RMSE_c$ means RMSE of compared model and similarly others.

The applications of DL methods in different fields have been summarized in table 5.

D. EXTREME LEARNING MACHINE (ELM)

ELM is a kind of learning scheme which is utilized for single hidden layer networks. ELM is considered as a quick learning algorithm due to its remarkable capability of generalization [197]. Ertugrul *et al.* [198] incorporated ELM into recurrent

TABLE 4. Summary of approaches for wind speed and energy prediction.

Location	Time scale	Methods	Compared methods	Outcomes	References
Taiwan	Hourly	BRT	Conventional unbiased	$RMSE_p$=0.1389 $RMSE_c$=0.1734	[92]
Taiwan	15 min	WTD-RNN-ANFIS	WTD-ANN, WTD-SVM, ANN,RNN	$RMSE_p$=0.9678, $RMSE_c$=1.0045 MAE_p= 0.6516, MAE_c= 0.6989	[87]
China	Hourly	EWT-LSTM-Elman	ARIMA, Elman, LSTM	$MAPE_p$= 10.93, $MAPE_c$= 24.95	[93]
China	10 min	Ensem LSTM	ARIMA, ANN, SVR	MAE_p= 1.1416, MAE_c= 1.3753 $RMSE_p$=1.5335, $RMSE_c$=1.8337	[94]
China	15 min	WT-CNN	SVM and back propagation	$PINC_p$ 99%=-0.78 $PINC_c$ 99%=-3.11	[95-97]

TABLE 5. Summary of different deep learning based methods in electric load forecasting system.

Types of problem	Methods	Application discipline	Input data in training	References
Load forecasting	LSTM-RNN, RBM, DBN, Deep CNN	Load consumption	Resident behaviors, load characterization	[98-125]
Wind forecasting	LSTM-RNN	Wind power and speed	Power, direction, atmosphere pressure, wind speed, temperature, humidity	[126-140]
Solar forecasting	CNN and K-means clustering, D-RNN, FF-DNN	Solar irradiance and power	Power, weather data, the sun, solar radiation	[141-147]
Power detection and classification	SAE, WT-SAE, CNN, SSAE, SAE-SVM, FF-DNN	Island recognition in DG systems, power quality	Current and voltage signals, evaluation of current, reactive power, active power, voltage	[148-164]
Detection of fault in power system and instrument	SAE, WT-SAE, CNN, SSAE, SAE-SVM,	Types of Error in power system. Fault detection in fuel cell, gearbox, transformer,	current signal, lubricant pressure data in gearbox, oil data of transformer, data of current and voltage signals	[165-172]
Applications of power system	BM, WT-DNN, CNN, Deep reinforcement	Security of power system, Detection of electricity theft in grid, Optimization of decentralized power system	Electricity production, energy consumption, electric devices, Data of energy interruption, fire, environmental pollution, natural disaster , fire	[173-182]

neural networks (RNN) and named it as recurrent extreme learning machine (RELM). This proposed RELM was used to get high accuracy and small error rate. It was noticed that RNN provide satisfactory results dynamic forecasting system. So, the incorporation of ELM into RNN can be very effective solution for forecasting model in actual time dynamical systems.

Huang *et al.* [197] described the following reasons behind the slow learning speed of neural network as compared to the required speed: (1) use of learning algorithms based on slow gradient to train the neural network, (2) the repeated tuning of parameters of networks. A newly algorithm known as extreme learning machine was proposed by the author to overcome the above mentioned problem. The proposed algorithm for single hidden layer feed forward neural networks (SLFNs) which determine the output weight of SLFNs analytically. ELM are feed forward neural network and consists of single to multiple layers of hidden nodes. Such nodes are never updated and are randomly assigned. There is no need to tune the hidden layers due to interesting property of this model. The structure of ELM for multiple structure is shown in figure 10, where **m** denotes the inputs as x_1, x_2, x_3,….x_m. The **n** hidden layers are traversed by these inputs. From input to these hidden layers, different weight (w_1, w_2,……w_n) are assigned to each values of input neurons. The linear weights β_1, β_2, β_3……β_n are obtained from hidden layer to get predicted output $\mathbf{O_j}$. It is thousand times faster than learning algorithms of traditional feed-forward neural network [197].

The previous machine learning methods for load forecasting were reviewed by Ma *et al.* [199]. It was emphasized

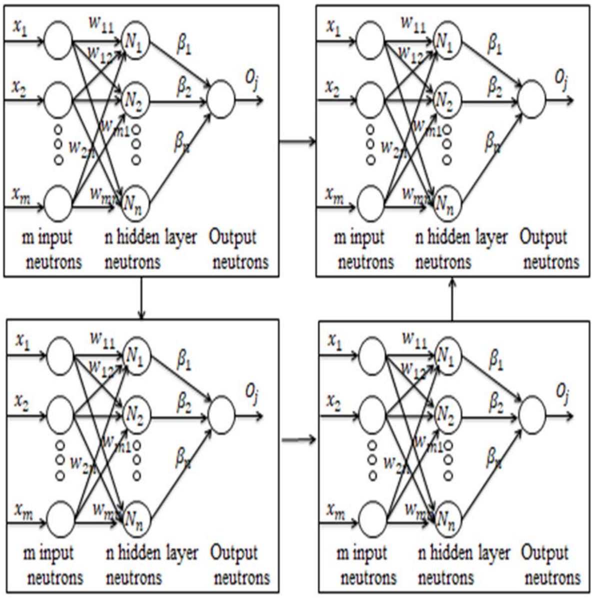


FIGURE 10. Structure of ELM for multiple cluster.

by authors that ML methods are appropriate for STLF in scalability and accuracy compared to conventional load forecasting methods. In [200], the photovoltaic generation was forecasted through support vector machine by the authors. They used the weather related data such as cloudy, rainy, clear and foggy. Support vector regression was used for STLF in energy system by [201]. The results of the proposed scheme were compared with other techniques such as ANN and ARMA.

Also, without facing the problems like over fitting, improper rate of training and local minima, ELM approaches solutions. A newly ensemble methodology was proposed for STLF by Li *et al.* [202], where least squares regression, wavelet transform and ELM are integrated.

Individual forecasters can be derived from the several combinations of decompositions levels and mother wavelet. Next day hourly load can be predicted by 24 ELMS consisted parallel model. Partial least squares regression methodology is used to compose ensemble forecast by combining the individual forecasts. The data obtained from two electric companies for one hour and one day ahead load forecasting was used to test the proposed method and improved accuracy was found in outcome as compared to other models. Recurrent extreme learning machine (RELM) was proposed by Ertugrul *et al.* [203] by integrating RNN with ELM. In forecasting dynamic system, the outcome of RNN is better as compared to feed forward AAANN model. RELM model is better due to low training time and its use in real-time dynamical system

Zhang *et al.* [204] dealt with STLF which came out from ELM under the management on Improved Gravitational Search Algorithm (IGSA), which is the collection of Gravitational Search Algorithm and partial Swarm optimization. Improved extreme learning based STLF method was proposed by Li *et al.* [205] which can choose automatically the number of hidden neurons according to number of input samples, which reduces the training error up to zero and as possible the test error. ELM and Empirical mode decomposition (EMD) based STLF was proposed by Chen *et al.* [206]. Data of instantaneous frequency can be obtained from EMD empirical approach through non-linear and non-stationary data sets. Load series is decomposed by EMD to capture the difficult features of electric load [207]. This method was tested for half hour electric load in Queens land and Victoria in Australia and the outcome was too much improved.

E. MULTILAYER PERCEPTION (MLP)

Park *et al.* [208] developed the architecture of MLP to predict the 24 h and one **h** ahead load. The outcomes were totally consistent with the real load with error value of 2.06% and 1.40 % for daily and hourly prediction data, respectively. However, the temperature was the only weather data taken in this study. According to the value of error, the holidays had various pattern as compared to the initial week days like Mondays. In that work, it was recommended to add sophisticated topology in neural networks to capture these data features. Alireza *et al.* [209] developed an MLP model by classifying the interaction between temperature and load into three kinds of weekly, days and hours trend. By clustering these three modules and adaptive weight new strategy, one to k-days ahead hourly load might be predicted dynamically. For daily and hourly forecasts, the MAPE value of 2.34% and 1.67 % was recorded, respectively.

Perception is that algorithm which uses straight line to divide the input. Further, it is a linear divider. The equation (5) describes the single outcome **y** which is generated by perception by using linear combination based on different actual real value input.

$$\mathrm{Y} = \emptyset\Big(\sum_{i=1}^{n} w_i x_i + b\Big)$$
$$\mathrm{Y} = \emptyset\left(w^T x + b\right) \tag{5}$$

where **x** and **w** corresponds to vectors of input and weights respectively. **Ø** is function of non-linear activation and **b** being bias.

Basically, MLP is artificial neural network (ANN). Three layered-structure of MLP is shown in figure 11. The signal is received by input layer and prediction about input is made by output layer. Two hidden layer in between of input and output layer corresponds to the computation core machine of MLP. It can manipulate input by continuously modifying the weight metrics until the error between predicted and target value becomes minimum. The issues of input selection and NN structure were solved by developing a two nonparametric method by Ferreira and Da Silva [210]. The outcome of RBFs and MLP can be improved by this proposed model. Ding *et al.* [211] made a comparison between NN and Naïve

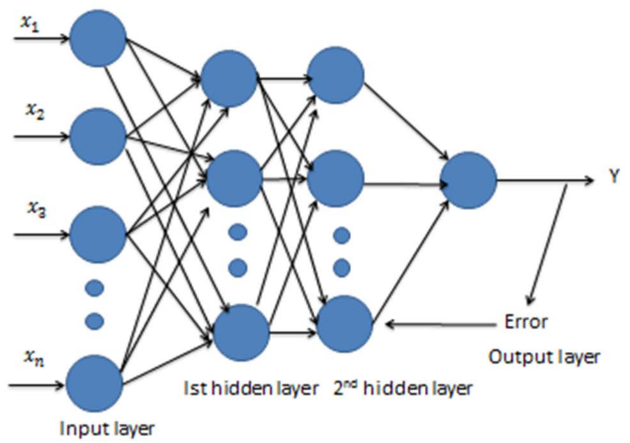


FIGURE 11. Structure of an ideal MLP.

model by using MLP. The analysis revealed that MLP model is more accurate than Naïve model by 4.7%.

F. SELF ORGANIZING MAP (SOM)

SOM neural network is a method of unsupervised cluster with unique property of producing local representation of the input data [212]. Lamedica *et al.* [213] used the Kohonen *et al.* [215] model to develop double stage load forecasting for networking the load data into different load profiles by deducing many features. In a second stage, the forecasting was performed through the arrangement of supervised MLP. The proposed method decreased the error rate as compared to outcome of MLP method. Lopez *et al.* [214] proposed the algorithm of SOM for STLF. The historical load and meteorological data was used to train the map. Further, the influence of data frame and various input selection on training the map was studied.

Xu *et al.* [31] developed the self-organizing map. It is also called kohonen network. The visualization and analysis of large dimensional data can be performed by this computational method. Unsupervised learning train this ANN type to produce discretized and small dimensional exhibition of training samples known a map. Competitive learning is applied in SOM and topological properties of input space are preserved by using neighborhood function. Suppose data items of n-dimensional Euclidean Vector Space

$$x(t) = [\xi_1(t), \xi_2(t), \xi_3(t), \ldots\ldots \xi_n(t)] \tag{6}$$

where, **t** represents the index of data observations of the sequence in equation (6).

Suppose the equation (7) represents the ith model

$$m_i(t) = [\mu_{i1}(t), \mu_{i2}(t), \mu_{i3}(t), \mu_{i4}(t), \ldots\ldots\ldots \mu_{in}(t)] \tag{7}$$

where **t** represents the index of generation of models in sequence.

By using new item **x(t)** of data and previous value $\boldsymbol{m_i}(\boldsymbol{t})$, the new value $\boldsymbol{m_i}(\boldsymbol{t}+1)$ can be calculated iteratively as equation (8):

$$m_i(t+1) = m_i(t) + \alpha(t)\, hc_i(t)\, [x(t) - m_i(t)] \tag{8}$$

where the correction size is defined by scalar factor $\boldsymbol{\alpha}(\mathbf{t})$ whose value reduces by increasing step index **(t).** The under process model is described by index **I** and **c** corresponds to the model with lowest distance from **x(t)** in Euclidean space. The smoothing Kernel type factor $\boldsymbol{hc_i}(\boldsymbol{t})$ have value **1** when **i =c** and its value reduces with the increase in grid distance between the models $\mathbf{m_c}$ **and** $\mathbf{m_i}$. Also, increase in step index **t** decreases the spatial width of kernel. Such convergence determined step index functions must be selected very delightfully.

Fan *et al.* [216] improved the prediction accuracy and learned time series load data by using conventional SOM and parameters of weather information. The data of weather and daily peak consumption of electricity of long Island and New York was used in this study. The time duration of taking values was from July1, 2001 to September 31, 2004. The extension of algorithm of SOM was used, which was based on error correction rule. The average of outcome of entire neurons was taken to generate the peak load. The smallest number of MAPE (1.93%) value was obtained by this model using (15∗15) neurons. The model was validated by using data of Spain energy consumption since 2001 to 2010. The model forecasts daily market load with 2.32% MAPE value.

For the prediction of load, Aprillia *et al.* [217] made a system as follows; An optimization algorithm to find and select the suitable level of the wavelet breakdown, the transform of discrete wavelet to decompose data. They used multiple linear regression to predict the outcome of the load. The proposed system was tested for holidays and weekdays of all season and builds a small forecasting error in compare to different models. For load forecast, the multiple linear regression was designed by Amral *et al.* [218]. The data of rainy and dry seasons were taken in experiments. For dry days, the MAPE error between forecasted and actual values was 3.52% and for rainy days, it was 4.34% [178]. The relation between the demand and weather condition was found through multiple linear regression by Saber and Razaul [219]. The large data was processed by multi core parallel processing. The MAPE error value was about 3.99%.

G. RULE BASED METHODS

1) FUZZY C-MEANS (FCM)

Dunn *et al.* [220] developed the Fuzzy C-means method and later Bezdek *et al.* [221] developed it. FCM is a clustering method which connects one sample of data with two or more clusters. Clustering is a widely used and popular numerical tool that can discover certain patterns or structures in dataset and there exist a certain similarity between the objects of each cluster. Fuzzy set theory and one of the popular Fuzzy clustering algorithm was proposed by Zedeh *et al.* [222]. He described the idea of uncertainty through the membership function. Minimization of objective function is the base of

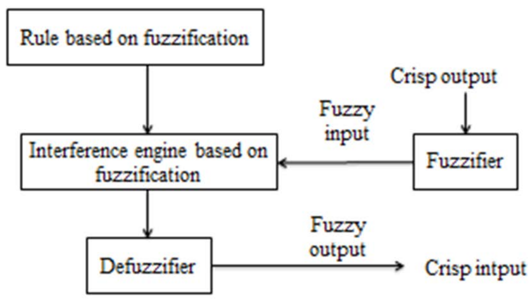


FIGURE 12. Algorithm structure of FRBS.

FCM algorithm, which is described as equation (9):

$$\mathrm{I_m} = \sum_{i=1}^{N}\sum_{j=1}^{C}\mu_{ij}^{m}\mathrm{m} > 1 \tag{9}$$

where **m** is real and $\boldsymbol{\mu_{ij}}$ corresponds to the degree of membership function of $\mathbf{x_i}$, $\mathbf{x_i}$ corresponds to ith of d-dimensional calculated data and $\mathbf{C_j}$ is the center of d-dimension in cluster. The approaching ability for nonlinear function and convergence speed of Radial Basis function (RBF) made it higher than BP network [223]. Zhu and He [224] introduced the use of RBF model based on FCM to overcome the BP network' issues such as local minima and low convergence speed. The load data of actual power was used to test the proposed model and the outcome revealed better mean percentage error (4.04%).

2) FUZZY RULE BASE SYSTEM (FRBS)

The important application of fuzzy logic or fuzzy set is the fuzzy rule base system. Fuzzy logic methods defined the fuzzy sets and fuzzy numbers that can be described in linguistic parameters. The inaccuracy of input and output variables can be addressed by fuzzy logic methods. The verbally prepared rules based fuzzy approach is overlapped in entire space of parameter. The difficult non-linear relation is controlled by using numerical interpolation. The general form of the fuzzy rules is **"IF A Then B",** where **A** and **B** are linguistic variables containing propositions.

A and **B** are known as the premise and consequence (result) of rule, respectively. The tolerance for uncertainty and inaccuracy is exploited by using the fuzzy rules of **IF Then** and linguistic variables. Human brain's ability to aggregate the data and decision making information is copied by fuzzy logic. The fuzzy input is generated by a **"Fuzzyfier"** to **"Fuzzify"** process by inserting crisp output or raw output in it as shown in figure 12. The user determined the fuzzy rule base to meet the need of forecasting system. The fuzzy output is generated by the inference engine by following the conditions given by Fuzzy input. For real applications, **"Defuzzifier"** convert the Fuzzy output into crisp input. Knowledge and reasoning are combined for the FRBS. Reasoning is well known fuzzy set with more value logic system that was developed by Lukasiwicz in 1930 and later on changed by Zadeh in 1960 [222]. Ranaweera *et al.* [225] comprehensively investigated the use of fuzzy logic system for the problem of STLF. MAPE value (<2.3%) was shown by this proposed model. The finding of maximum fuzzy rule base was the important issue that was faced during the designing of fuzzy model. Kang *et al.* [226] proposed an approach for controlling and modeling the evolutionary design where parameter and structure for fuzzy rule are simultaneously evolved by using evolutionary programming. Khosravi *et al.* [227] discussed the ability of method to handle qualitative and quantitative information and uncertainties'. Also, the use of IT2FLS for the STLF was proposed by them. Subsequently, they utilized the Takagi Sugeno Kang Fuzzy Inference system for the development of IT2-TSK_FLS hybrid algorithm. Hassan *et al.* [228] suggested a fuzzy logic model of type-2 (IT2FLS) where ELM was applied to tune the variables of IT2FLS. Ali *et al.* [229] proposed historical and weather parameters based fuzzy logic system for LTLF. The efficiency of the model was 93.1% with MAPE value of 6.9%. Advantages and limitations of single methods are stated in table 6.

X. STATISTICAL METHODS

These methods are traditional in nature and mostly used in time series forecasting such as STF by employing the historical data of smart grid [230]. To fit a regression model on classical data this is the base of these methods. Then model is validated by measuring the difference between the predicted and actual values. The existing statistical methods are described in this subsection.

A. AUTOREGRESSIVE MOVING AVERAGE MODEL (ARMA)

ARMA is the basic method which is commonly used in time series investigation. It is made by combining auto regressive (AR) and moving average (MA) methods. If (P) represents the order of AR model, then AR part can forecast the value at time stamp (t) as a function of its foregoing value (t-p). The observed data is formed by combining the previous values with error term by using the MA part. Sansa *et al* [231] used ARMA to forecast the winter day's solar irradiation with optimum 10% changes in the production. However, ARMA is applicable only for static time series.

B. AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)

ARIMA is the generalized form of ARMA to deal dynamic time series. In this context, the values of data are replaced with the difference of previous and current values. Amini *et al.* [232] used historical data to charge the electric vehicle (EV). Similarly, it used to control and optimizes the residential micro grid by [233], [234]. However, in large data set, ARIMA provides large RMSE values and shows large execution time for non-linear data. So, ARIMA is very suitable for linear data and time series [235]. An extra technique such as long transformation is used in ARIMA models to manage the non-linearity.

TABLE 6. Advantages and limitations of different single load forecasting methods.

Methods	Advantages	Limitations
DL	Human engineers do not design the layers of features. Computational complexities are reduced. Better provision of network's initial values.	Low probability of empirical coverage. Not scalable. Convergence is very slow. Bad optima may trap the process of training.
ELM	Very quick training. Capability of universal classification. Better generation.	The problem of generalization degradation and uncertainty.
Fuzzy C Means	Useful for pattern recognition. Better analysis for error index as compared to other methods of load forecasting. Practicable and accurate.	Require to consider weather, temperature and humidity type. Large computational time. Sensitivity to early guess (local minima, speed).
Fuzzy Rule Base	Capability of universal approximation. Having the ability of controlling uncertainties and qualitative/quantitative information.	Difficult to recognize optimal fuzzy rule base. Large uses in practical applications are hindered by computational complexities.
MLP	Excellent in patter recognition and modeling. Provision of better accuracy in non-linear problem. Have insignificant errors and classifications accuracy.	Slow convergence. Training process may be affected by local minima.
SOM	Ability to recognize the difficult sets of pattern. Need less training samples. Good interpretation of outcomes.	Inputs required being vectors. Problems of scalability.

C. VECTOR AUTOREGRESSION

It is defined as an addition in univariate auto regression for time series with (n) number. VAR has been used as an statistical approach is energy forecast to get linear dependencies among the many time series. The solar irradiation, temperature and speed of wind of 61 location of America were observed by using VAR [236]. The results showed lower values of RMSE as compared to simple persistence method.

XI. PROBABILISTIC METHODS

Forecasting tendency is shifting from point to probabilistic approach to meet the future generation and demand at disjoint levels [237]. The importance of probabilistic methods was reviewed over point prediction with changing requirements of electric power industry [40], [238]. The past literature regarding probabilistic forecasting has been identified in this subsection. The literature is divided into non-parametric and parametric approaches. These two approaches have been briefly reviewed by authors in [8t]. Dowell *et al.* [239] suggested a Bayesian probability and sparse VAR based parametric probabilistic approach to predict the VST wind power production with time interval of 5min in southeastern Australian. The results of the proposed model showed least RMSE value as compared to the standard VAR and AR methods. Hong *et al.* [32] used parametric approach to predict the probabilistic horizons energy consumption. Various authors [240], [241] used probabilistic methods for price forecasting. Weron *et al.* [242] highlighted the importance of future demand forecasting in probabilistic shape to assist the grid planning and operations in regard of energy production and distribution. They forecasted the energy consumption of around 3700 residencies in Ireland with time interval of 30 min by using additional quantile regression (QR). Liu *et al.* [243] assist the probabilistic forecasting by proposing quantile regression averaging (QRA) with 90% percent get PIs for consumption. However, the existing developed probabilistic scheme often suffers in complexities of computation, thus further need is to develop more efficient methods by considering the scenarios of energy market [244], [245].

XII. PROBABILISTIC DEEP LEARNING (PDL)

The Bayesian probability integrated with deep learning methods to supply forecasting outcomes in the shape of PIs. The classical deep neural systems are inevitable in nature and produce point forecasts. The model parameters are described as a function of probability distributions. For precise results, the future PIs with various percentiles can be predicted by PDL models. The percentiles can describe the uncertain and certain factors in the data set which make it finer decision making. This section defines the important role in the discipline of PDL to help the applications of generation, needs and price forecasting in current power systems.

A. BAYESIAN NEURAL NETWORK (BNN)

The idea of Bayesian probability integrated with artificial neural network is outlined as BNN. Yang *et al.* [246] quantified the contributed uncertainties between various group of customers and proposed BNN technique to predict the demands of energy at residential levels. Further, a data system based on cluster was presented to handle the problem of over fitting by enhancing the diversity and volume of data. The authors showed pinball scores and lower winkles for probabilistic approaches.

B. BAYESIAN LSTM

Sun *et al.* [247] presented a BNN hybridized LSTM scheme to curb the issues caused by seasonal uncertainties in distributed PV producers and then accurately generate the total load forecasts in shape of PIs. Further, they clustered the discrete sub profiles to improve the forecasting performance. The patterns of energy consumption were the base of sub profiles prior the use of Bayesian approach. The methods were implemented on actual data set of AUS GRID for three years with time interval of half an hour.

C. BAYESIAN BIDIRECTIONAL LSTM

The authors of [248] suggested a PDL scheme to solve the issue of uncertainty in energy systems. The authors enabled the propagating to make the training sequence backwards and forwards by integrating the bidirectional RNN with LSTM. Then authors suggested their method known as bidirectional-LSTM (BLSTM). The proposed network was trained to create non- parametric and Gaussian predictive distribution of non- independent parameters of data set. Further, sampling based on copula was used to create predictive scenarios. However, the more sample space made the probabilistic methods computationally expensive. In this content, [249] proposed potential solutions in the form of dropout or reduce the complexity and computation of Bayesian interference. However, the issue of computational complexity still remains as big concern, so more effective and generalized solutions are needed to be developed in future.

XIII. HYBRID MODELS

Higher percentage of error, computational complexity and computing efficiency are the different kinds of disadvantages of single methods in load forecasting. From many years, researchers have worked to build hybrid methods and models with large precision and accuracy but minimum rate of error. In hybrid models, two or more than two single methods are combined to get more efficiency and accuracy. In hybrid model, single methods are selected according to their needs, where they can contribute beneficially in load forecasting. There are two important and popular methods such as SVM and ANN. Such methods are hybridized with other single methods to get minimum error rate and best load forecasting model. SVM is suitable for semi structured and unstructured data to get maximum output. The actual power of the SVM's model is the kernel trick where any complex issue can be solved by any suitable kernel function. A unique solution can be produced through SVM by rounding the optimality. This is the basic difference between SVM and Neural Networks, which generate local minima based multiple solutions due to which they are not trustable for several samples. This section makes a comparison between the different hybrid models based on ANN and SVM to optimize the output of load forecasting.

A. ARTIFICIAL NEURAL NETWORK (ANN)

The conventional models including regression can lead to undesirable results and so are limited. The basic cause is computational complexity which lead to many solution times and difficulty in non-linear data design. ANN provides an attractive, efficient and promising analytical alternative of conventional techniques which are limited by assumptions including linearity, normality and variable independence. ANNs are applied to estimate cooling/heating loads, consumption of electricity and optimization of output. ANNs are real hardware or processing devices. ANN is a machine learning method based on the human brain's structure. A big ANN may contain thousands of units of neurons and interactions of multiple layers. The main unit of ANN is the neuron which sends and receives normalised signals to and from the other neurons of the network. The interaction wires between the neurons are known as ''weight''. The ANN is determined by three main features; the structure of the network (recurrent or feed-forward), the weights defining learning rules (Hebbian, perception), the activation function between the input and output neurons. Mostly used ANN is the Multi-layer-Perception. Backpropagation rule is the base of multilayer network which determine the error in output and minimize it. ANNs are very suitable for the energy load forecasting [250]. Further, it is a self-adaptive model that consists of (a) captures subtle and pattern recognition relationship, (b) related with noise, (c) independent of knowledge of programmers about rules, (d) independent and same operations can be performed simultaneously. However, the outcome of ANNs cannot be evaluated easily because (a) there is no concept of mathematic, (b) it consume time in computation, (c) optimisation of training process is difficult, (d) large data is needed and (e) non-convergence of model in few cases [250]. Baek *et al.* [251] proposed a recurrent type of ANN with MAPE value (1.57%) in South Korea. In this work, the data of actual time temperature, load consumption of day to day, weather and month of July, 2011 was used. At the interval of 15 min, the value of power consumption along with temperature was taken. Whereas, the weather and day type was recorded once a day. Multilayer perception (MLP) is architecture of Artificial Neural Network. Park *et al.* [208] used MLP to STLF. The major drawback of this methodology is that it consumes large MLP structure to control the actual dataset and produce issues of redundancy.

In cognitive neurosciences and machine learning, ANN is considered as an intelligent method because the functional aspects of bio-neural network motivate it. Different tasks such as classification, forecasting, data mining, pattern recognition and process modelling can be implemented by organizing ANN in different arrangements. Its excellent features such as generalization, parallel processing, learning ability and error tolerance provide solution in case of nonlinear and linear mapping.

Before the process of prediction, there is no need of specific relationship between input and output variable which is the biggest advantage of ANN. So, for non-linear regression applied in load forecasting, ANN is the most popular learning tool. ANN model's architecture can be classified into DL, MLP, SOM and ELM. MLP is the design of feed forward ANN which consists of multiple layers of nodes, where every node is linked with next layer [250]. Neto and Fiorelli [252] used the data of hourly load of university administration block to make a comparison between thermal model and ANN model in order to forecast hourly load. For thermal model, the daily MAPE value was smaller than 13% but daily MAPE value of ANN model was 10%. Further, the study compared the complex ANN model (solar radiations, temperature and relative humidity was considered as effective parameters) and simple ANN model (only temperature was considered as influence parameter). The mean MAPE value of complex ANN was 9.5% which was smaller than simple ANN.

Din and Marnerides [253] used Recurrent Neural Network (RNN) and Feed forward Neural (FNN) along with deep learning for (STLF). They applied these networks on the data set taken from New England for the duration of 2007 to 2012. The model was checked against two cases. In first case, frequency domains and time features were used and in the second case, the attributes of time domain were used. The system was evaluated by RMSE, MAPE and MAE errors, which furnished small rates in first case as compared to second case. Accuracy of model was boosted in the first case [253]. Reddy and Jung used ANN with wavelet decomposition. The results of the experiments revealed the efficiency of proposed system which surpassed the ANN [254]. In order to forecast the electric load, the dynamic neural network was proposed by Mardjaoui *et al.* [255]. The data of French transmission system was used to design and test the proposed system. The result of simulation proved the validation of proposed system. The following are the most important hybrid methods used to integrate the ANN with other single methods.

Usually, the mean age of western high voltage grid is about 40 years. The average life period of installed equipment at high voltage grid is about 50 years. So, the western electricity distribution system needs to be upgraded and massively refurnished in future. According to the DG TREN [256], the life time of few distribution system will end in 10 years. The hierarchical based grid was built to interconnect the small voltage network and to cope with the breakdown of a power system. From the decades, the grid system was not reconstructed despite of the increase in number of customers and load demand. The basic reason of this failure is the liberalization and deregulation of large power markets since last 15 years. The grid runs at their full capacity by weak load area's wind power. Further, the remote areas are occupied by the excellent power resources, where electric grids are not too much strong. This could be applied to mid voltage (up to 110kv) as well as high voltage level, where decent rally generated electricity is brought to high voltage grid. The stability of grid is a limitation for the use of wind power and renewable sources of energy. Therefore, the un-stability of renewable energies is the cause of strain in the infrastructure of grid. Some of the well-known hybrid methods combined with ANN and single methods are discussed in this subsection.

1) NEURO-FUZZY (NF)

Neuro Fuzzy was introduced by J.S.R Jang [257]. It was discovered to have memories and to cover the lack capabilities of Fuzzy system by combining the learning arrangement of ANN and reasoning mechanism of fuzzy sets. In fact, the classification of fuzzy technique help ANN to calculate the optimize parameters by the arrangement of rule base. This is very helpful in forecasting task with raw datasets [258]. Fuzzy logic (FL) is very influencive method to characterize the load uncertainty which is caused by various behavioral and environmental factors. The effect of human behavior on hourly load profile was considered by the neuro-fuzzy model [259]. In this research, the price dependent load data was created by using FL without the information of price. It has been found that the efficiency of approach based on NF was better than single neural network.

2) ANN AND WAVLET TRANSFORM (ANN-WT)

WT is a strong tool to analyze the non-stationary signals the domains of time frequency [260]. The data of time series can be decomposed into various levels with low and high frequency components with the help of WT. The stationary data can be improved by the transformation approach and it can be integrated with ANN to select the input. Guan *et al.* [24] decomposed the load into many components of frequency by wavelet approach. The transformed normalized dataset was given as an input to ANN to learn the properties of individual components. Results of data from ISO New England showed that wavelet neural technique is very good.

The classical FOURIER decomposition approach can be completed by wavelet. The constituent of non-stationary signal can be examined by WT and non-stationary and stationary signals can be filtered. WT has been dictated in different works to disintegrate the load into many frequency components. Jawerth and Sweldens [260] concentrated on multistage analysis of WT. Zhang *et al.* [261] captured the important information on different time scales by using WT. The load of electricity was predicted by the combination of ANN and proposed method. Australian markete's data was used to validate the method. Guan *et al.* [24] combined the data pre-filtering with wavlet neural networks (WNN).

3) ANN AND FRUIT FLY OPTIMIZATION ALGORITHM (ANN-FOA)

FOA algorithm helps to measure the minimal and maximal value and the basic part of the FOA algorithm was discussed by Pan *et al* [262]. The searching possibility of FOA algorithm is very helpful to select the parameter of neural network. For the annual electric load forecasting, FOA and GRNN was combined by Li *et al.* [263], where the value of spread parameter of GRNN model is determined with FOA.

According to proposed model, the MSE and MAPE values were 1.421 and 1.149% respectively. The proposed model was compared with GRNN, FOA-GRNN and PSO-GRNN. As compared to other methods, FOA-GRNN showed best performance. FOA is a kind of optimization and computation technique. It is easy to code and understand it as compared to different algorithms and it is based on swarm intelligence.

4) ANN AND FIREFLY ALGORITHM (ANN-FA)

Yang *et al.* [264] introduced FA to state the problem of optimization. It was developed to solve the drawback of local optimal of neural network. Liye *et al.* [265] developed the new change of FA to minimize the weight coefficient of integrated FA neural network for the purpose of load forecast. The data set of the state of victoria, New South wales and Queensland of Australia was used to evaluate the ability of forecasting. The results of different combined algorithms such as WT-ANN and GA-ANN were compared with the algorithm of proposed model. The proposed model showed better results as compared to other models. In ANN-FA methods, the non-linear mapping is created by using FA and its learning ability is achieved by the works of ANN. The accurate and efficient forecasting model can be developed through this model, but its performance is not good with respect to RMSE value. Kavousi-Fard *et al.* [266] developed an accurate and efficient forecasting model by combining ANN and FA. The mathematical expression (10) of the proposed hybrid ANN and MFA model is

$$X_k = [w_{i,1}, w_{i,2}, w_{i,3} \ldots\ldots w_{i,Mw}, b_{i,1}, b_{i,2}, \ldots . b_{i,Mb}]_{(1,M)}$$
$$\mathrm{M} = \mathrm{M_w} + \mathrm{M_b} \tag{10}$$

Where, X_k represent the firefly which was selected randomly from the set of fireflies, w_i represents the adjusting coefficient **M** dimensional row vector, and M_b, M_w represents the biasing and weighting factors, respectively.

5) ANN AND PARTICLE SWARM OPTIMIZATION (ANN-PSO)

The network weights as well as optimal arrangement were measured by adapting PSO algorithm to the neural network. Telbany [267] assessed the feasibility of forecasting of MLP recurrent network which was developed by PSO for the forecasting of daily load. PSO assisted to get rid of few of the neuron weights by reducing their values through global searching. The algorithm of back propagation was outperformed by the results of PSO. BP displayed poor performance due to complexity and over training. Liu *et al.* [268] also used the combination of ANN-PSO. They optimized the parameter of ANN in order to predict the load in a small grid with high randomness and small capacity. The high quality of PSO is that it can find effective element from many possible alternative and also it can be implemented easily and computationally in- expensive. It only needs the values of objective function rather than its gradient information. It was suggested that the proposed method is suitable for small grids with large load fluctuations. Zhang and Ma [269] combined the RBF and PSO algorithm where RBF neural network was used to learn accuracy and weight were optimized by PSO. This proposed model performed very well.

6) ANN AND ARTIFICIAL IMMUNE SYSTEM (ANN-AIS)

The immune system is very beneficial due to which the system can be made parallel because the ANN-AIS has the information by processing capabilities. This system is invigorated by the immunology of human. The complex patterns can be identified by both ANN and AIS. The immune algorithm can overcome the defect of premature phenomenon. So, the precision and speed of searching can be improved by the special trait of AIS. Yong *et al.* [270] used immune algorithm (IA) to design the BP neural network (BPNN) for STLF. It was given the name of Artificial immune network (AIN). The proposed method found the optimize parameters by taking the benefits of quick searching ability of AIS algorithm. The AIN MAPE value and MAPE value of proposed model was 2.038% and 2.52%, respectively.

7) ANN AND OPTIMIZATION ALGORITHMS

ANN is very popular forecasting technique, which is used for the prediction of load forecasting. The fitting of model to actual data decides the performance of prediction. A neural network with back propagation algorithm have problem of convergence to local minima and parameters initial values sensitivity. Many layered neural network based on gradient has low possibility to find the optimal solution, due to which many local minima results in cost function.

8) ANN AND GENETIC ALGORITHM (GA)

The GA is an optimization engine which provides a technique of global search. The non-linear issues of neural network can be solved by this suggested technique. This integration was used by Ling *et al.* [271] to design a forecasting model. The proposed GA algorithm based neural network. GA with non-uniform mutation and arithmetic crossover was used to help in adjusting the proposed network's parameter. The proposed technique reduced the number of parameters. There are two main function of this model called dynamic and static activation function. If v_{ij} represents the connection weight from node x_i to j_{th} neuron, n_h corresponds nodes of hidden layer, net_a^j (.) is the j_{th} dynamic activation function, net_s^j (.) represents the static activation function, then the mathematical equation (11) for the daily based load forecasting is:

$$y_l(t) = net_0^l\left(\sum_{j=1}^{n_h} net_d^j\left(net_s^j\left(\sum_{i=1}^{24} x_i v_{ij}\right), m_d^j, \sigma_d^j\right) w_{jl}\right) \tag{11}$$

where

net_o^l (o) is the 'l' output neuron's activation function m_d^j represents dynamic mean

σ_d^j corresponds to dynamic standard deviation for the j_{th} DAF.

The MAPE value of this proposed model is less than the conventional neural models. Azadeh *et al.* [272] integrated

TABLE 7. Comparison of MAPE, MAE and RMSE values for different ANN based models.

Models	Input	Mean MAPE value (%)	Mean MAE value (MW)	Mean RMSE value (MW)	Reference
Combination of neural networks and wavelet decomposition	Similar's day load	With air temperature (1.58), with wind chill temperature (1.55)	51.45	-	[267]
SD-EMD-LSTM,	24-h forecast	1.10,	-	-	[268]
SD'-EMD-LSTM	24-h forecast	1.44	-		
BNN1,[1]	hourly load forecasting	0.5071,	-	12..71,	[269]
BNN2,		12.91,	-	210.63,	
ANN		13.08	-	213.73	
DT,	One day ahead with 5 clusters	440.61,	-	-	[270]
ANN		47.95	-	-	
BNN[2]	For 1500 samples and 10 iterations, hourly load	1.81	-	-	[271]
IEMD,	Whole Year of 2012	0.82,	69.03,	2.15,	[272]
ELM,		0.91,	80.23,	2.88,	
SVM,		1.12,	99.26,	3.59,	
ENN		2.34	204.19	7.05	
MI+ANN,	Whole Year of 2012	3.78,	-	-	[273]
Bi-Level,		2.31,	-	-	
MI+ANN+mEDE		1.23	-	-	
EMD-mRMR-FOA-GRNN	October 31, 2017	4	38	39	[274]
Cluster 1,	January, February, March, April, November, December and October (from 15th to 31^{st},	9.63,	-	-	[275]
Cluster 2,	: Saturdays, Sundays and Holidays,	7.58,	-	-	
Cluster 3	May, June, July, August, September and October (from 1st to 14th)	4.50	-	-	
ANN-BKP,	Winter season	-	-	0.1416,	[276]
GNN-BKP,		-	-	0.1329,	
GNN-GAF,		-	-	0.1032,	
GNN-W-BKP,		-	-	0.0610,	
GNN-W-GAF		-	-	0.0486	

Similar days (**SD**) selection +empirical mode decomposition **(EMD)** + long short-term memory (**LSTM**) (**SD-EMD-LSTM**), Bayesian neural network (**BNN**[1]), artificial neural network (**ANN**), decision tree (**DT**), Bagged neural networks (BNNs[2]), Improved empirical mode decomposition (**IEMD**), extreme learning machine (ELM), support vector machine (**SVM**), Elman neural network (**ENN**), mutual information-based *artificial neural network* (***MI-ANN***), mutual information+modified enhanced differential evolution + artificial neural network (**ANN**) based model (**MI-mEDE-ANN**), empirical mode decomposition **(EMD)**+ minimal redundancy maximal relevance (**mRMR**),+general regression neural network (**GRNN**) with fruit fly optimization algorithm (**FOA**) (**EMD-mRMR-FOA-GRNN**), Generalized Neural Network (**GNN**)+back-propagation (**BKP**) (**GNN-BKP**), artificial neural network (**ANN**)+ back-propagation (**BKP**) (**GANN-BKP**), Generalized Neural Network (**GNN**)+ genetic algorithm and fuzzy system (**GNN-GAF**), Integration of wavelet and **GNN-BKP** systems (**GNN-W-BKP**), Integration of wavelet and **GNN-GAF** systems (**GNN-W-GAF**).

the ANN and GA algorithm to access its performance. They designed a logarithmic linear model for the forecasting of energy. Different variables including number of customers, price, and value of electricity consumption were used to the GNN. The affective coefficient with less error rate has been recognized by changing the parameters through GA. The performance of this proposed model was excellent as compared to series model. Table 7 shows the comparative study of MAE, MAPE and RMSE values of various ANN based models.

B. SUPPORT VECTOR MACHINES (SVM)

In machine learning, SVMs are much effective models with ability to solve non-linear problems with less training data [283]. Vapnik *et al.* [284] developed this popular ML algorithm in 1995. They can be employed for regression problem and for classification. In first case, it is called support vector regression (SVR). The big advantage of SVMs over ANN is that they have the ability to search global minima [285]. Borges *et al.* [286] used three different buildings of eastern slovakian to compare the different ML methods. The author concluded that day ahead hourly forecasting results of SVR model were more accurate as compared to ANN for these three buildings. SVM has been remained popular choice among ML methods since 1990. The convex optimized problem defined the SVM. There are effective methods for optimization problem such as sequential minimal optimization SVM provides a unique solution. For the prediction of load, SVM is a bright learning tool. It theoretically guarantees to get special global minima. Slow running and computationally expensive are the drawbacks of SVM. The unique solution of SVM depends upon important parameters for the used-selected kernel function. The SVM parameter values and classes of Kernel function can be optimized by different optimization algorithms. The well-known hybrid methods that combine single methods with SVM are discussed in this subsection.

1) SVM AND PARTICLE SWARM OPTIMIZATION (SVM-PSO)

Wang *et al.* [287] combined weighted LS-SVM with PSO to predict the load. It was revealed that the performance of this proposed model was better than the other methods because the mean MAPE value was reduced to 3.095% from 3.22%. Yu *et al.* [288] employed the algorithm of k-nearest neighbors (KNN) to preprocess the load data and integrated SVM with PSO to predict the load. The big advantage of the PSO is that it has the memory space to store the solution of all particles. Also, same parameters of PSO can be adjusted due to which it become an optimization method to find the solution of non-linear problems [289]. PSO was combined with adaptive ANN to adjust the weights of network [290]. The proposed model showed better computational results as compared to conventional BP algorithm. Wang *et al.* [291] proposed a hybrid model based on SVM. They integrated empirical mode decomposition (EMD), SVM and PSO. The effect of holidays, weekend and temperature were also considered. The residential load data was decomposed into intrinsic mode function by EMD. These functions were forecasted by SVM. The selections of parameter were performed by PSO. The results showed that (EMD-PSO-SVM) model was significant tool and effective for residential STLF.

2) LEAST SQUARES SUPPORT VECTOR MACHINE (LSSVM)

Chen *et al.* [292] proposed a new forecasting model by combining EMD, PSO and least squares support vector machine. They used EMD based filtering method to decrease to effect of noise signals. The seasonal components of de-noised resulting series were eliminated by ESPLSSVM and the resulting series were modeled by LSSVM. The proposed ESPLSSVM model reduced the mean MAPE, MAE and RMSE values by 26.16%, 24.70% and 34.89%, respectively. So the proposed model improved the accuracy of load forecasting. Jin *et al.* [293] proposed an LSSVM and sperm whale algorithm based load forecasting model to improve the accuracy of the forecasting. The redundancy of input vector was reduced by selecting the optimal feature by discrete wavelet transform and inconsistency rate model. The mapping ability of LSSVM was improved by replacing the kernel function of LSSVM by wavelet kernel function. The sperm whale algorithm was used to optimize the parameter of W-LSSVM and finally the W-LSSVM-SWA method was established. The proposed method was feasible and effective for STLF in energy system. Wei *et al.* [294] proposed on LSSVM and wavelet transform model to improve the accuracy of load forecasting by reducing the effect of external factors. Fruit fly algorithm (FOA) was used to optimize the proposed model for STLF. To increase stability of data and remove errors points, wavelet transform was used.

3) SVM AND GENETIC ALGORITHM (SVM-GA)

The increasing complexity and importance of STLF demands an accurate forecast model. Wei *et al.* [295] proposed genetic algorithm (GA) based SVM model with deterministic annealing (DA). DA was adopted to cluster the load in order to solve the issues then GA-SVM model was established. The performance of GA-SVM was compared with conventional BP forecasting model. The MAPE and RMSE values were 1.66% and 386.42 for GA-SVM model. But for the BP forecasting model, these values were 4.03% and 877.94. The values of errors parameters for GA-SVM were smaller than the BP forecasting model. Arash *et al.* [296] developed a modified support vector regression based long short term memory (SVR-LSTM) model. The proposed model was applied on the data set from micro-grid MG in Africa. The value of co-relation coefficient was 0.9901 for SVR-LSTM and 0.9809, 0.9770 were for the SVR and LSTM respectively.

4) SVM AND FRUITFLY OPTIMIZATION ALGORITHM (SVM-FOA)

Food process based FOA was proposed by Pan [262]. It is popular due to short program code as compared to other optimization algorithm. FOA can search quickly to the global optimal solution. So it had been used for various forecasting applications.

A hybrid SVM-FOA model was proposed by Li *et al.* [297] in his research work. The result of SVM-FOA was compared with heuristic optimization algorithms including GA and stimulated Annealing (SA). The proposed model showed better results by less researching time to find global optimum with 3% forecasting error for annual load. Also, the accuracy of FOA over PSO for SVM parameter was proposed by Cao *et al.* [298]. Comparative study for advantages and disadvantages of different state-of-art forecasting methods is given in table 8.

XIV. MODERN TECHNIQUES

A. FUZZY LOGIC (FL)

Fuzzy logic is based on Boolean theory, but instead of taking a value of 1 or 0 as an input, it considers particular qualitative ranges. The input is related to qualities based comparison. For example, the temperature of some quantity may be ''low'', ''medium'' or ''high'' however, the outputs can be deduced from fuzzy or noisy inputs in fuzzy logic and there is no need to specify the mapping of inputs to outputs [340]. The uncertainties can be controlled by fuzzy methods. There is a membership function for every fuzzy set, which shows a continuous fitting curve which change from 0 to 1. To get better prediction results, often the other method are combined with fuzzy theory. The main benefit of fuzzy logic is that there is no existence of mathematical model to design map between inputs and outputs. Also, there is no existence of noise free inputs properly designed and general rules based fuzzy logic systems are powerful in the forecast of electric load. In whole processing through fuzzy logic, there is a need of ''defuzzification'' for precise outputs [341].

B. EXPERT SYSTEMS

The expert system used computer programming to understand, explain and expand knowledge base information with the access of new information. Expert system integrates the procedure developed by human expert and rules. There should be convenience in the knowledge of expert system so that its code could be developed in the form of software. In other words, the decisions of experts must be understandable to the program developer. The information is translated in coding form by using facts and IF THEN conditions. The code develops relationships between the effecting factors and load of systems. With the passage of time, some conditions of code have to change continuously but some rules do not have to change [340].

Interference engine is the part of expert system which is used to search the solution or reasoning related to the conclusions. The expert system should be able to trace reasoning if asked by the customers. Interference component is used to built this facility. Rehman and Baba demonstrated this rule based algorithm in their research work. This function based algorithm developed to forecast the logical based model in the form of rules in rule base approach. This approach consists of relationships between variations influencing factors and variations in system load. Statistical criteria were used to accept the possible relationship [342].

C. GENETIC ALGORITHM

In 20th century, computer scientist studied effectively the evolutionary system to solve the engineering problems by using evolution as an optimization tool. The idea was to develop a population of solution for a given engineering problem by using operators, which copy the natural selection and genetic variations [343]. The idea of genetic algorithms was introduced by John Holland and he published his idea in a book. Also, GA was popularized by David Goldberg in 1989. Genetic Algorithms is the most commonly used evolutionary computational techniques [344]. GAs represents a search technique which based on natural selection and principles of genetics. They combine the individuals under selection rules to reduce the cost function and to enhance the efficiency of solution. The principle ''survival of the best'' based GA methaheuristic method was presented by Holland. This popularized GAs as a useful tool to solve intense optimization problems [344]. In summary, GA represents the computer programming based field which can find the optimal solution to minimize or maximize the criterion function. Genetic Algorithm consists of the following important components [345]–[347].

1) Fitness function which is important component of algorithm. It is the function to be optimized by the algorithm.
2) Population of values which shows a solution of solvable problem. Random sample of chromosomes (values) is selected as an initial population. Then the problem solved ability of each chromosome is tested by the fitness function.
3) A reproducible chromosome is selected which based on the probability distribution.
4) Next level of chromosomes generation is produced which resemble the division of chromosomes in biological cell meiosis process.
5) A random mutation to flips the individual bits in the newly produce chromosomes.

In load forecasting system, GAs are well suited for non-linear systems. GAs can conduct a specific optimization which based on principles of natural selection found in candidates' population [348]. GAs had been used to measure the optimal parameters of ARIMA model [349]. Gupta and Savangi [350] used back propagation methodology based on GA for electric load forecasting.

Newly developed approaches regarding learning and architecture are required in order to handle the variations in data, to make model to adopt itself promptly and capture the patterns with new revealing.

D. ONLINE ADAPTIVE RNN

This load forecasting system works with continuously reaching data and adjusts to the new design. These approaches employ batch-normalized RNN (BNRNN) as base learner and integrate performance monitoring, Bayesian optimization and buffering to line the BNRNN framework on the fly. Figure 13 depicts the newly online adaptive model. Data

TABLE 8. Summary of advantages and disadvantages of different forecasting methods.

Forecasting methods	Advantages	Disadvantages	References
Statistical methods	Uncomplicated and less computationally expensive. Use to determine relationship between predicted power and weather features.	Less reliability for large and non-linear data in dataset of energy. Difficult to control complex weather conditions.	[294-295]
Machine learning	Simple and can deal with large datasets	Less reliability for large and heterogeneous data issues, produce point prediction	[296-301]
Deep learning	Efficient for large and non-linear datasets. Show better results for multilayer neural networks. Do not require data pre-processing or additional neural networks.	Create point prediction, over fitting, need suitable hyper-parameter tuning. Performance depends on input data quality and architecture design.	[302-307]
Probabilistic	Provide prediction intervals (PIs)	Highly computationally expensive	[308-313]
Probabilistic deep learning	Efficient, reliable, provide PIs	Highly computationally expensive	[313-317]
Hybrid methods	Highly accurate and scalable. Integrate different model's weight to improve the performance of model by conserving the advantages of approach.	Specific uses, Highly computationally expensive. Face issues of stable prediction due to their complex learning structure. May have long duration of training, under fitting problem and low efficiency.	[318-323]
Pre-processing	Better forecasting performance	Specific uses	[4-7]
ARIMA	Usable for non-linear model. Forecast a regression model and develop a fit.	Data linearization is needed. Suited to stationary data. Need of data preprocessing. Chance of loss of information. Less accurate for time series data.	[324-325]
ANN	No needs of extra expertise for statistical training. Ability to observe the interaction between independent and dependent variables. Ability to detect all relationships among predictor variables. Quantity of training algorithms can be accessed. Can control non-linear interaction in load consumption by adjustment of weight in training process.	Black box in nature. Computationally expensive. Have a tendency of overfitting. Model development with empirical nature. Need of large quantity of datastes to train the model and its complexity. Need data pre-processing and neural network.	[326-328]
SVM	Overfitting can be nullify through regularization. Kernel trick can be utilized to built expert knowledge about problem. No local minima and many effective methods to solve problems.	Dependency on kernel. At testing and training levels, offers limitation in size and speed. Have slow testing process	[329-331]
Time series analysis	Have ability to adopt seasonal effects	Numerical instability	[332-333]
Fuzzy inference system	Quick and accurate in execution	Trail and error based selection of membership function in the formation of rule	[334-335]
Regression	Functional interaction between the influencing factors (temperature, weather and	Could not control consumption of non-linear load.	[336-337]

TABLE 8. *(Continued.)* **Summary of advantages and disadvantages of different forecasting methods.**

	time of day) and former forecast load. Beneficial in imaginary time forecasting.	Addition of parameters reduces its stability. Have less accuracy for actual time load.	
ANFIS	Integration of both ANN and fuzzy system can solve the basic problems of fuzzy system. The basic issues of fuzzy system can be eliminated through the learning capability of ANN and parameter optimization.	Sensible to first number of fuzzy rules. Enhancement in fuzzy rules can enhance the computational complexity.	[338]
GA	Selection of high quality of chromosomes can reduce the computational complexity.	Random generation and selection of number of genes of chromosomes. Genes quality is not high. Higher complexity in selection of initial population. Degeneracy in clusters. Need to rearrange dataset with high dimension. Need to reduce time complexity.	[339]

becomes available from sensors and other smart meters with the passage of time. The pre-processing module transforms this data into processing form for RNNs. The pre-processing module comprises of sliding window and online normalization. Batch normalization is used to control the covariate shift and to minimize the training time. The sensitivity to variations in the learning rate is also reduced by batch normalization. Consequently, it support the turning module to adjust the rate of learning in order to capture the data in better way. Batch normalization consists of prediction, online normalization and model training [351].

E. GENERALIZED ADDITIVE MODELS (GAM)

The additive models decompose the response variable y_t as described by equation (12):

$$y_t = \beta_0 + \sum_{j=1}^{d} f_j\left(x_{t,j}\right) + \varepsilon_t \tag{12}$$

where ε_t is random noise and $x_t = (x_{t,1}, \ldots\ldots x_{t,d})$ represents the explanatory variable and f_j depicts the non-linear effect which is decayed on a spline basis $(B_{j,k})$ with β_j coefficients. The function of non-linear effect is described by equation (13) as:

$$f_j(x) = \sum_{k=1}^{m_j} \beta_{j,k} B_{j,k}(x) \tag{13}$$

where m_j is the function of dimension of spline basis [352].

F. SINGLE MULTIPLICATIVE NEURON MODEL (SMN)

Figure 14 shows the architecture of SMN model having single neuron for I inputs, where ω_i represents weights, b_i is biases and u_i is input of the framework. The multiplicative operator Z can be written as by equation (14):

$$z = \prod_{i=1}^{I} (\omega_i u_i + b_i) \tag{14}$$

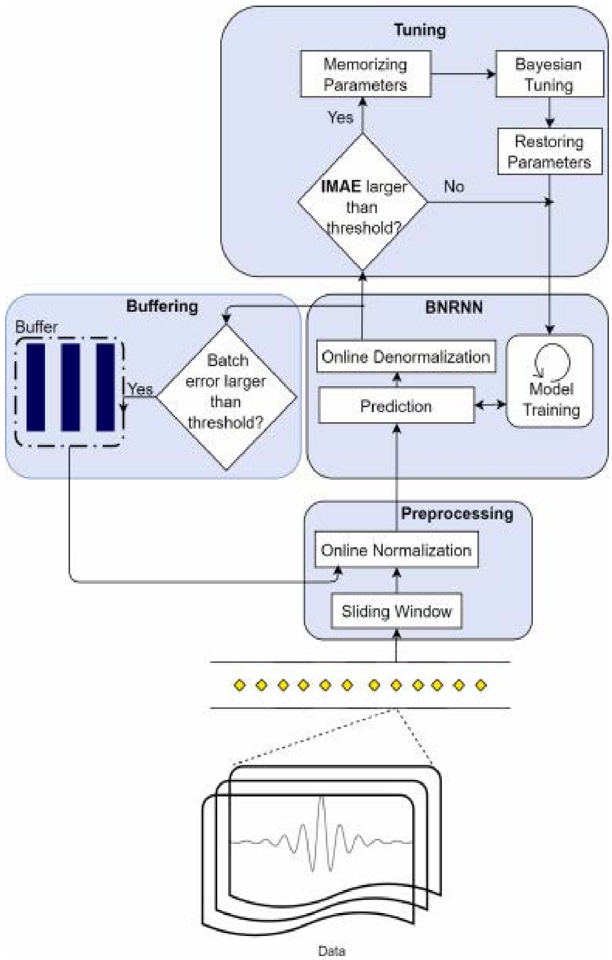


FIGURE 13. **Structure of online adaptive RNN [351].**

If the activation function's nature is logistic, then equation (15) represents the outcome function of the SMN model as [353].

$$y = \frac{1}{1 + e^{-z}} \tag{15}$$

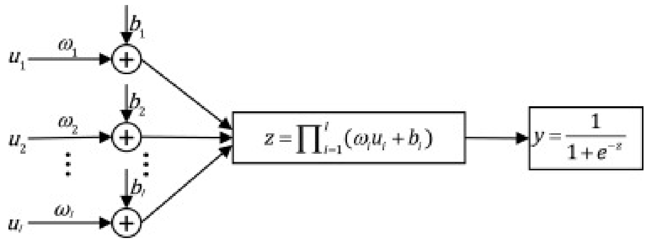


FIGURE 14. Architecture of SMN model [353].

G. FUTURE GRID

In field of non-conventional energy, there are two disputed approaches:

i. Distribution to electricity to wide area
ii. Decentralized production

1). Super grid is the idea of transmission of non-conventional renewable electricity of large scale and over large distances. In world, there is non-uniform distribution of energy sources [354]. The prerequisite for energy source based electric system is an expansion of effective and large distance transmission grid. The solar and wind energy potential in the wordly deserts would be able to fulfill the demand of energy across the world. The supply of electricity up to large area depends on structure which transmits the electricity from generation places to consumption sites. High voltage direct current lines (HVDC) technology has been developed to transmit the electricity over large distances. There are many advantages of direct current system over alternating current system. There is no limit on underground or sea cables and aerial lines for transmission over long distance. Also, the impact of environment is small. DESERTEC idea of Rome's club is the vision of a super grid. In it, the desert having massive potentials for non-conventional with assets of technology [355], [356].

2). The decentralized access based on cluster of distributed production installations (biogas digester, mean wind turbines, gas turbines, and fuels cells) which are linked through a smart grid. The practical power plants or clusters are jointly derived by a control entity that deals with the output of full cluster by governing the power plants. Some of the renewable technologies are able to generate electricity on need like biomass or hydro plants, provide controlled feed-in of photovoltaic or wind. The entire feed-in of virtual plant is handled and kept at a fixed level. The cluster of renewable energy power plant is considered to be reliable and controllable by this approach [357], [358]. Smart grid consist of hardware tools and software with ability of routing power in more effective way, thus minimize the requirement of extra capacity. The basic difference between smart grid and current grid is that smart grid is a modified electricity and supply network. In smart grid, modern intelligent technologies and two-way communications are used to ameliorate the reliability and efficiency of electric transmission and supply system. Smart grid are equipped with optimize technology based on Information and Communication Technologies, which are able to communicate with demanded loads. The production and grid load can be made more predicted by the option provided by the smart grids.

H. SUPER SMART GRID

The lower scale decentralized smart grid and large or prominent scale approach of super grid are sensed as alternate of each other. However, we indicate that these ideas are interchangeable and can co-exist to assure a transition to a carbon free economy. What is needed is, hence, a super smart grid (SSG). The infrastructure to supply non-conventional renewable production of electricity from large and small generating places across wide areas with the quality of managing loads and fluctuating supply. The important issue of fluctuating transmission renewable production is addressed comprehensively by super smart grid. From the idea of Super grid, mostly fluctuations will finally cancel out each other in a supply across wide area [359], [360]. There is a high probability that wind will blow and sun will shine somewhere in the domain of large region. Super smart grid also uses the concept that hydro power storage plants and pumps are not expensive technology with enough capacity. Further, it also depends on flexible and quick biomass power stations to complement the hydro storage. The concept of smart grid assures the possible economic of an appreciable share decentralized production and also show important for the transitional phase. Basically, such technologies are related to load management and control the demand in proper way to assure that demand and supply are in balance at any time. New developments of wind energy and mass electro-mobility are potential candidates in keeping the entire costs of demand and matching supply within reasonable range.

The comparison of related reviews and novelty of our review work is summarized in table 9. The aforementioned review work did not presents a comprehensive detail of types of load forecasting techniques along with their advantages and limitations. Also, none of review work presented datasets for forecasting. Our review is intrinsically different due to presentation of detailed types of load forecasting techniques. DP, ML and AI based single and hybrid methods to forecast load and energy consumption and production, presentation of datasets, current challenges, way forwards and future directions.

XV. CONCLUSION AND FUTURE TRENDS

The prediction of electrical load need accuracy, precision and often checks for changing parameters. The researchers had researched on different ML, DL and AI models to increase the efficiency of developed forecast technologies. This review paper has addressed the study of single methods and hybrid methods. Time horizon regarding various load forecasting technologies has been presented based on which different prediction models enlisted for comparative analysis. SVM, ANN and related models have proved fruitful because these schemes showcased better opportunities in getting a popular

TABLE 9. Summary of comparison of existing review and our work. Note: PY: published year; LF: load forecasting; DP: deep learning; HM: hybrid method; WSF: wind speed forecasting; PPT: pre-processing data techniques; Ref: references.

PY	Dura-tion	LF	DP	HM	WSF	PPT	Review	Ref
2017	1981-2017	×	×	√	√	×	Artificial intelligence to improve the efficiency of renewable energy sources and energy generation from them	[361]
2017	1992-2016	√	×	×	×	×	Intrusive load manipulating techniques to reduce energy cost and power consumption and forecasting of energy consumption to make a balance in demand and supply	[362]
2018	2002-2017	√	×	×	×	√	Forecasting of energy consumption, data pre-processing techniques and data properties	[363]
2018	1986-2017	√	×	×	×	×	Management of energy generation and consumption by data driven techniques	[364]
2018	2000-2018	×	×	×	√	×	Forecasting of wind speed by ANN based methods, use of ANN in system designing and error detection	[365]
2019	2008-2018	×	×	√	√	×	Forecasting of wind speed using DL based method, and concludes that hybrid methods are more effective than single	[366]
2020	2002-2019	√	√	×	×	×	Forecasting of energy generation and consumption through DL based approaches	[367]
2021	2000-2020	√	√	×	√	×	Forecasting of load and energy consumption from wind turbines by DL based approaches	[368]
-	Up to 2022	√	√	√	√	√	Detailed types of load forecasting techniques. DP, ML and AI based methods to forecast load and energy consumption, presentation of datasets, current challenges , way forwards and future directions.	Our work

power system where the prediction about demand load have little error percentage.

The percentage error has been compared by using statistical measurements in order to select the best method for certain forecasting routine.

In electric load forecasting, the MAPE value of Deep learning method STLF-RNN is 0.0535. The MRE and RMSE values are 2.92 and 6.99 for DL based Integrated SAE and ELM method and LSH Deep auto encoder method, respectively. In case of wind speed prediction, RMSE and MRE values for DL based boosted regression tree (BRT) method and wavelet threshold denoising- recurrent neural network - adaptive neuro-fuzzy interference system (WTD-RNN-ANFIS) method are 0.1389 and 0.6516. For hybrid methods, the mean MAPE (%) and RMSE (MW) values are 0.5071 and 0.0486 for ANN based BNN method and Integration of wavelet and GNN-GAF systems (GNN-W-GAF) method, respectively. Table 4 also represent the outcomes of the performance comparison in terms of MAE, RMSE and MAPE of different forecasting techniques in this review work.

A. WAY FORWARD

There is a significance influence of weather, forecasting time and economic system on the accuracy of design of specific load forecaster. The previous load bank of data set has a key scale in the augmentation of precision and accuracy. The processing of statistical operations and addition of load data or economic factors can approximate the STLF values to MTLF or LTLF.

The accuracy of single and hybrid predictive model can be analyzed by the evaluation criteria such as MAE, MAPE, and RMSE.

- The forecast value can be improved and made accurate by single predictive models as well as by the integration of single models.
- In single method described above, the MLP shows the best classification precision and accuracy while ELM gives quick training.
- The error analysis index of FCM is better while the universal accuracy is of FRBS.
- The forecasting of time series can be performed by statistical methods.
- ARM and ARIMA are best for the investigation of static and dynamic time series, respectively.
- The outcomes of hybrid predictive models such as SVM and ANN are much better then single forecasting predictive models.

The readers can obtain deep knowledge from the outcomes of the review about the comprehensive detailed predictive models. The new hybrid model can be designed to forecast the load and energy. The disadvantages of the enlisted methods based on single or two models can be addressed by the future research about hybrid models consisted of more than two methods.

1) FUTURE TRENDS

In order to increase the forecasting accuracy, the following strategies could be considered.

a: DEVELOPMENT OF NEW DATA PRE-PROCESSING METHODS

There is a uncertainty in power and wind data which has not been evaluated and analyzed sufficiently. It means existing data based analysis are not effective for power and wind data which turn to biased conclusions. So, there is a deep need to develop new data pre-processing methods to reduce the uncertainty and to handle problem with complex uncertainty. This problem can be solved by integrating various types of pre-processing approaches.

b: INCREMENT IN NUMBER OF EFFECTING FEATURES

Fluctuations in data are the results of many factors. So, the use of more meteorological and topographical factors at study site may precisely describe the variations in the datasets. The problem can be solved by recording the data through the installment of modern suitable sensors.

c: INCREASE IN ABILITY FOR NON-LINEAR FITTING

Usually, the difficult terrain and climate variability make it tough to characterize and evaluate the fluctuations in data. Although, the non-linear fitting tendency of different model is better and their outcome is constrained due to their limited ability of learning and fluctuations in training data. So, it is need to combine different model to enhance the non-linear fitting ability.

d: DEVELOPMENT OF HYBRID MODEL CONSISTING OF MORE THAN TWO MODELS

The readers can obtain deep knowledge from the outcomes of the review about the comprehensive detailed predictive models. The new hybrid model can be designed to forecast the load and energy. The disadvantages of the enlisted methods based on single or two models can be addressed by the future research about hybrid models consisted of more than two methods.

REFERENCES

[1] D. W. Bunn, "Forecasting loads and prices in competitive power markets," *Proc. IEEE*, vol. 88, no. 2, pp. 163–169, Feb. 2000, doi: 10.1109/5.823996.

[2] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 44–55, Feb. 2001, doi: 10.1109/59.910780.

[3] T. Haida and S. Muto, "Regression based peak load forecasting using a transformation technique," *IEEE Trans. Power Syst.*, vol. 9, no. 4, pp. 1788–1794, Nov. 1994, doi: 10.1109/59.331433.

[4] W. Gao, A. Darvishan, M. Toghani, M. Mohammadi, O. Abedinia, and N. Ghadimi, "Different states of multi-block based forecast engine for price and load prediction," *Int. J. Elect. Power Energy Syst.*, vol. 104, pp. 423–435, Jan. 2019, doi: 10.1016/j.ijepes.2018.07.014.

[5] D. Sáez, F. Ávila, D. Olivares, C. Cañizares, and L. Marín, "Fuzzy prediction interval models for forecasting renewable resources and loads in microgrids," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 548–556, Mar. 2015, doi: 10.1109/TSG.2014.2377178.

[6] F. Bouffard and F. D. Galiana, "Stochastic security for operations planning with significant wind power generation," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 306–316, May 2008, doi: 10.1109/TPWRS.2008.919318.

[7] Y. S. Manjili, R. Vega, and M. M. Jamshidi, "Data-analytic-based adaptive solar energy forecasting framework," *IEEE Syst. J.*, vol. 12, no. 1, pp. 285–296, Mar. 2018, doi: 10.1109/JSYST.2017.2769483.

[8] G. Li and H.-D. Chiang, "Toward cost-oriented forecasting of wind power generation," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2508–2517, Jul. 2018, doi: 10.1109/TSG.2016.2614341.

[9] M. Stojanović, M. Bozić, and M. Stanković, "Mid-term load forecasting using recursive time series prediction strategy with support vector machines," *Facta Universitatis-Ser., Electron. Energetics*, vol. 23, no. 3, pp. 287–298, 2010.

[10] S. Ruzic, A. Vuckovic, and N. Nikolic, "Weather sensitive method for short term load forecasting in electric power utility of Serbia," *IEEE Trans. Power Syst.*, vol. 18, no. 4, pp. 1581–1586, Nov. 2003, doi: 10.1109/TPWRS.2003.811172.

[11] E. A. Feinberg and D. Genethliou, "Load forecasting," in *Applied Mathematics for Restructured Electric Power Systems: Optimization, Control, and Computational Intelligence*, J. H. Chow, F. F. Wu, and J. Momoh, Eds. Boston, MA, USA: Springer, 2005, pp. 269–285.

[12] G. Singh, D. S. Chauhan, A. Chandel, D. Parashar, and G. Sharma, "Factor affecting elements and short term load forecasting based on multiple linear regression method," *Int. J. Eng. Res. Technol.*, vol. 3, no. 12, pp. 736–740, Dec. 2014.

[13] H. Seifi and M. Sadegh, *Electric Power System Planning Issues, Algorithms and Solutions* (Power Systems). Springer, 2011.

[14] N. Amjady, "Short-term hourly load forecasting using time-series modeling with peak load estimation capability," *IEEE Trans. Power Syst.*, vol. 16, no. 4, pp. 798–805, Aug. 2001.

[15] A. Kumar, S. Khatoon, and M. Muazzam, "An overview electricity demand forecasting techniques," in *Proc. Nat. Conf. Emerg. Trends Electr., Instrum. Commun. Eng.*, vol. 3, no. 3, pp. 38–48, 2013.

[16] L. Ghods and M. Kalantar, "Different methods of long-term electric load demand forecasting; a comprehensive review," *Iranian J. Electr. Electron. Eng.*, vol. 7, no. 4, pp. 249–259, 2011.

[17] M. Kandil, S. El-Debeiky, and N. Hasanien, "Overview and comparison of long-term forecasting techniques for a fast developing utility: Part I," *Fuel Energy Abstr.*, vol. 43, no. 4, p. 294, Jul. 2002.

[18] L. Ghods and M. Kalantar, "Methods for long-term electric load demand forecasting; a comprehensive investigation," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Apr. 2008, pp. 1–4.

[19] E. Badar, "Comparison of conventional and modern load forecasting techniques based on artificial intelligence and expert systems," *Int. J. Comput. Sci. Issues*, vol. 8, no. 3, p. 504, 2011.

[20] *Long-Term Load Forecasting Options in Georgia*, Usaid Hydropower Investment Promotion Project (HIPP), Deloitte Consulting, United States Agency, Washington, DC, USA, 2013.

[21] L. Katzir, "The effect of system characteristics on very-short-term load forecasting," *Przegląd Elektrotechniczny*, vol. 1, no. 11, pp. 121–125, Nov. 2015, doi: 10.15199/48.2015.11.31.

[22] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," in *Proc. North Amer. Power Symp.*, Sep. 2010, pp. 1–8.

[23] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice* (Business Forecasting), 3rd ed. Australia: OTexts, 2018.

[24] C. Guan, P. B. Luh, L. D. Michel, Y. Wang, and P. B. Friedland, "Very short-term load forecasting: Wavelet neural networks with data pre-filtering," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 30–41, Feb. 2013, doi: 10.1109/TPWRS.2012.2197639.

[25] C. Guan, P. B. Luh, L. D. Michel, and Z. Chi, "Hybrid Kalman filters for very short-term load forecasting and prediction interval estimation," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3806–3817, Nov. 2013.

[26] C. W. Potter and M. Negnevitsky, "Very short-term wind forecasting for Tasmanian power generation," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 965–972, May 2006.

[27] J. C. López, M. J. Rider, and Q. Wu, "Parsimonious short-term load forecasting for optimal operation planning of electrical distribution systems," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1427–1437, Mar. 2019.

[28] S. Fan and R. J. Hyndman, "Short-term load forecasting based on a semi-parametric additive model," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 134–141, Feb. 2012.

[29] N. Amjady and A. Daraeepour, ''Midterm demand prediction of electrical power systems using a new hybrid forecast technique,'' *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 755–765, May 2011.
[30] Y. Wang, Z. Zhou, A. Botterud, and K. Zhang, ''Optimal wind power uncertainty intervals for electricity market operation,'' *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 199–210, Jan. 2018.
[31] W.-T. Xu, Y.-L. Wang, T. Li, Q. Tang, J.-F. Zhang, L. Shen, and M. Zhu, ''Long term intelligent load forecasting method considering the expectation of power market transaction,'' in *Proc. 29th Chin. Control Decis. Conf. (CCDC)*, May 2017, pp. 2310–2315.
[32] T. Hong, J. Wilson, and J. Xie, ''Long term probabilistic load forecasting and normalization with hourly information,'' *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 456–462, Jan. 2014.
[33] H. B. Azad, S. Mekhilef, and V. G. Ganapathy, ''Long-term wind speed forecasting and general pattern recognition using neural networks,'' *IEEE Trans. Sustain. Energy*, vol. 5, no. 2, pp. 546–553, Apr. 2014.
[34] C. Voyant, G. Notton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, and A. Fouilloy, ''Machine learning methods for solar radiation forecasting: A review,'' *Renew. Energy*, vol. 105, pp. 569–582, May 2017.
[35] M. N. Akhter, S. Mekhilef, H. Mokhlis, and N. M. Shah, ''Review on forecasting of photovoltaic power generation based on machine learning and Metaheuristic techniques,'' *IET Renew. Power Gener.*, vol. 13, no. 7, pp. 1009–1023, May 2019.
[36] M. R. Haq and Z. Ni, ''A new hybrid model for short-term electricity load forecasting,'' *IEEE Access*, vol. 7, pp. 125413–125423, 2019, doi: 10.1109/ACCESS.2019.2937222.
[37] N. Abu-Shikhah, F. Elkarmi, and O. M. Aloquili, ''Medium-term electric load forecasting using multivariable linear and non-linear regression,'' *Smart Grid Renew. Energy*, vol. 2, no. 2, pp. 126–135, 2011, doi: 10.4236/sgre.2011.22015.
[38] S. Ye, G. Zhu, and Z. Xiao, ''Long term load forecasting and recommendations for China based on support vector regression,'' *Energy Power Eng.*, vol. 4, no. 5, pp. 380–385, 2012, doi: 10.4236/epe.2012.45050.
[39] S. Fallah, R. Deo, M. Shojafar, M. Conti, and S. Shamshirband, ''Computational intelligence approaches for energy load forecasting in smart energy management grids: State of the art, future challenges, and research directions,'' *Energies*, vol. 11, no. 3, p. 596, Mar. 2018, doi: 10.3390/en11030596.
[40] T. Hong and S. Fan, ''Probabilistic electric load forecasting: A tutorial review,'' *Int. J. Forecasting*, vol. 32, no. 3, pp. 914–938, Jul. 2016, doi: 10.1016/j.ijforecast.2015.11.011.
[41] Z. Zhu, D. Zhou, and Z. Fan, ''Short term forecast of wind power generation based on SVM with pattern matching,'' in *Proc. IEEE Int. Energy Conf. (ENERGYCON)*, Apr. 2016, pp. 1–6.
[42] D. Kaur, R. Kumar, N. Kumar, and M. Guizani, ''Smart grid energy management using RNN-LSTM: A deep learning-based approach,'' in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.
[43] Y. Wang, Q. Chen, C. Kang, Q. Xia, and M. Luo, ''Sparse and redundant representation-based smart meter data compression and pattern extraction,'' *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 2142–2151, May 2017.
[44] Y.-C. Huang, C.-M. Huang, S.-J. Chen, and S.-P. Yang, ''Optimization of module parameters for PV power estimation using a hybrid algorithm,'' *IEEE Trans. Sustain. Energy*, vol. 11, no. 4, pp. 2210–2219, Oct. 2020.
[45] K. Chen, J. Hu, and J. He, ''A framework for automatically extracting overvoltage features based on sparse autoencoder,'' *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 594–604, Mar. 2018.
[46] Y. Zhang, C. Qin, A. K. Srivastava, C. Jin, and R. Sharma, ''Data-driven day-ahead PV estimation using autoencoder-LSTM and persistence model,'' *IEEE Trans. Ind. Appl.*, vol. 56, no. 6, pp. 7185–7192, Nov./Dec. 2020.
[47] S. Ryu, H. Choi, H. Lee, and H. Kim, ''Convolutional autoencoder based feature extraction and clustering for customer load analysis,'' *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1048–1060, Mar. 2020.
[48] X. Shao and C. S. Kim, ''Multi-step short-term power consumption forecasting using multi-channel LSTM with time location considering customer behavior,'' *IEEE Access*, vol. 8, pp. 125263–125273, 2020.
[49] C. Zhang, J. Bütepage, H. Kjellström, and S. Mandt, ''Advances in variational inference,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 8, pp. 2008–2026, Aug. 2019.
[50] M. R. Ibrahim, J. Haworth, A. Lipani, N. Aslam, T. Cheng, and N. Christie, ''Variational-LSTM autoencoder to forecast the spread of coronavirus across the globe,'' *PLoS ONE*, vol. 16, no. 1, Jan. 2021, Art. no. e0246120.
[51] A. Dairi, F. Harrou, Y. Sun, and S. Khadraoui, ''Short-term forecasting of photovoltaic solar power production using variational auto-encoder driven deep learning approach,'' *Appl. Sci.*, vol. 10, no. 23, p. 8400, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/23/8400
[52] J. Pereira and M. Silveira, ''Unsupervised anomaly detection in energy time series data using variational recurrent autoencoders with attention,'' in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2018, pp. 1275–1282.
[53] P. Bachhav, M. Todisco, and N. Evans, ''Latent representation learning for artificial bandwidth extension using a conditional variational autoencoder,'' in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 7010–7014.
[54] X. Wang, P. Cui, Y. Du, and Y. Yang, ''Variational autoencoder based fault detection and location method for power distribution network,'' in *Proc. 8th Int. Conf. Condition Monitor. Diagnosis (CMD)*, Oct. 2020, pp. 282–285.
[55] W. Charytoniuk and M.-S. Chen, ''Very short-term load forecasting using artificial neural networks,'' *IEEE Trans. Power Syst.*, vol. 15, no. 1, pp. 263–268, Feb. 2000, doi: 10.1109/59.852131.
[56] L. Ghods and M. Kalantar, ''Different methods of long-term electric load demand forecasting a comprehensive review,'' *Electron. Eng.*, vol. 7, no. 4, 2011, Art. no. 4.
[57] R. Dechter, ''Learning while searching in constraint-satisfaction problems,'' in *Proc. 5th AAAI Nat. Conf. Artif. Intell.*, Philadelphia, PA, USA, Aug. 1986, pp. 178–183. Accessed: Apr. 25, 2020.
[58] G. Tesauro, ''Practical issues in temporal difference learning,'' *Mach. Learn.*, vol. 277, no. 4, pp. 257–277, 1992.
[59] Y. Bengio and Y. Lecun, ''Scaling learning algorithms towards AI,'' in *Large-Scale Kernel Machines*, vol. 1. Cambridge, MA, USA: MIT Press, 2007, pp. 1–41.
[60] J. Schmidhuber, ''Deep learning in neural networks: An overview,'' *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015, doi: 10.1016/j.neunet.2014.09.003.
[61] Y. Lecunn, L. Bottou, Y. Bengiu, and P. Haffner, ''Reducing the dimensionality of data with neural networks,'' *Science*, vol. 313, no. 5786, pp. 504–507, 2006, doi: 10.1126/science.1127647.
[62] T. Yue and H. Wang, ''Deep learning for genomics: A concise overview,'' in *Handbook of Deep Learning Applications*. New York, NY, USA: Springer, 2018, doi: 10.1021/acs.molpharmaceut.5b00982.
[63] A. Krizhevsky, I. Sutskever, and H. GE, ''ImageNet classification with deep convolutional neural networks,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 1, 2012, pp. 1–9, doi: 10.1016/j.protcy.2014.09.007.
[64] H. A. Pierson and M. S. Gashler, ''Deep learning in robotics: A review of recent research,'' 2017, pp. 1–41, *arXiv:1707.07217*.
[65] NVIDIA. *Deep Learnıng for Self-Drıvıng Cars*. Accessed: May 10, 2018. [Online]. Available: https://www.nvidia.com/en-us/deep-learning-ai/industries/automotive/
[66] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. Accessed: Oct. 9, 2019. [Online]. Available: http://www.deeplearningbook.org/
[67] P. Kuang, W.-N. Cao, and Q. Wu, ''Preview on structures and algorithms of deep learning,'' in *Proc. 11th Int. Comput. Conf. Wavelet Actiev Media Technol. Inf. Processing(ICCWAMTIP)*, Dec. 2014, pp. 176–179, doi: 10.1109/ICCWAMTIP.2014.7073385.
[68] S. Ryu, J. Noh, and H. Kim, ''Deep neural network based demand side short term load forecasting,'' *Energies*, vol. 10, no. 1, p. 3, Dec. 2016.
[69] Y. He, J. Deng, and H. Li, ''Short-term power load forecasting with deep belief network and copula models,'' in *Proc. 9th Int. Conf. Intell. Hum.-Mach. Syst. Cybern. (IHMSC)*, Aug. 2017, pp. 191–194, doi: 10.1109/IHMSC.2017.50.
[70] H. Shi, M. Xu, and R. Li, ''Deep learning for household load forecasting—A novel pooling deep RNN,'' *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5271–5280, Sep. 2018, doi: 10.1109/TSG.2017.2686012.
[71] W. Kong, Z. Y. Dong, D. J. Hill, F. Luo, and Y. Xu, ''Short-term residential load forecasting based on resident behaviour learning,'' *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1087–1088, Jan. 2018, doi: 10.1109/TPWRS.2017.2688178.
[72] T. Ouyang, Y. He, H. Li, Z. Sun, and S. Baek, ''Modeling and forecasting short-term power load with copula model and deep belief network,'' *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 3, no. 2, pp. 127–136, Apr. 2019, doi: 10.1109/TETCI.2018.2880511.

[73] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, ''Gradient-based learning applied to document recognition,'' *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
[74] W. Rawat, ''Deep convolutional neural networks for image classification: A comprehensive review,'' *Neural Comput.*, vol. 2449, pp. 2352–2449, Aug. 2017, doi: 10.1162/neco_a_00990.
[75] C. Robinson, B. Dilkina, J. Hubbs, W. Zhang, S. Guhathakurta, M. A. Brown, and R. M. Pendyala, ''Machine learning approaches for estimating commercial building energy consumption,'' *Appl. Energy*, vol. 208, pp. 889–904, Dec. 2017.
[76] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, ''A survey on deep learning in medical image analysis,'' *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.
[77] Y. Bengio, A. Courville, and P. Vincent, ''Representation learning: A review and new perspectives,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013, doi: 10.1109/TPAMI.2013.50.
[78] G. E. Hinton, ''Learning multiple layers of representation,'' *Trends Cogn. Sci.*, vol. 11, no. 10, pp. 428–434, 2007, doi: 10.1016/j.tics.2007.09.004.
[79] D. E. Rumelhart, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA, USA: MIT Press, 1986.
[80] D. Erhan, A. Courville, and P. Vincent, ''Why does unsupervised pre-training help deep learning?'' *J. Mach. Learn. Res.*, vol. 11, no. 2007, pp. 625–660, 2010.
[81] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA, USA: Morgan Kaufmann Publishers, 1988.
[82] R. Salakhutdinov and G. Hinton, ''Deep Boltzmann machines,'' in *Proc. Int. Conf. Artif. Intell. Statist.*, D. van Dyk and M. Welling, Eds., vol. 5, 2009, pp. 448–455.
[83] R. Salakhutdinov and H. Larochelle, ''Efficient learning of deep Boltzmann machines,'' in *Proc. Int. Conf. Artif. Intell. Statist.*, vol. 9, 2010, pp. 693–700.
[84] C.-C. Ku and K. Y. Lee, ''Diagonal recurrent neural networks for dynamic systems control,'' *IEEE Trans. Neural Netw.*, vol. 6, no. 1, pp. 144–156, Jan. 1995.
[85] R. J. Williams and D. Zipser, ''A learning algorithm for continually running fully recurrent neural networks,'' *Neural Comput.*, vol. 1, no. 2, pp. 270–280, Jun. 1989, doi: 10.1162/neco.1989.1.2.270.
[86] S. Hochrejter and J. Schmidhuber, ''Long short term memory,'' *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[87] A. K. Ozcanli, F. Yaprakdal, and M. Baysal, ''Deep learning methods and applications for electrical power systems: A comprehensive review,'' *Int. J. Energy Res.*, vol. 44, no. 9, pp. 7136–7157, Jul. 2020.
[88] M. Saviozzi, S. Massucco, and F. Silvestro, ''Implementation of advanced functionalities for distribution management systems: Load forecasting and modeling through artificial neural networks ensembles,'' *Electr. Power Syst. Res.*, vol. 167, pp. 230–239, Feb. 2019, doi: 10.1016/j.epsr.2018.10.036.
[89] S. S. Pappas, L. Ekonomou, P. Karampelas, D. C. Karamousantas, S. K. Katsikas, G. E. Chatzarakis, and P. D. Skafidas, ''Electricity demand load forecasting of the hellenic power system using an ARMA model,'' *Electr. Power Syst. Res.*, vol. 80, no. 3, pp. 256–264, Mar. 2010, doi: 10.1016/j.epsr.2009.09.006.
[90] W. He, ''Load forecasting via deep neural networks,'' *Proc. Comput. Sci.*, vol. 122, pp. 308–314, Jan. 2017, doi: 10.1016/j.procs.2017.11.374.
[91] C. Li, Z. Ding, D. Zhao, J. Yi, and G. Zhang, ''Building energy consumption prediction: An extreme deep learning approach,'' *Energies*, vol. 10, no. 10, p. 1525, Oct. 2017, doi: 10.3390/en10101525.
[92] S. Bouktif, A. Fiaz, A. Ouni, and M. Serhani, ''Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches,'' *Energies*, vol. 11, no. 7, p. 1636, Jun. 2018, doi: 10.3390/en11071636.
[93] J. Zheng, C. Xu, Z. Zhang, and X. Li, ''Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network,'' in *Proc. 51st Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2017, pp. 1–6, doi: 10.1109/CISS.2017.7926112.
[94] X. Qiu, Y. Ren, P. N. Suganthan, and G. A. J. Amaratunga, ''Empirical mode decomposition based ensemble deep learning for load demand time series forecasting,'' *Appl. Soft Comput.*, vol. 54, pp. 246–255, May 2017, doi: 10.1016/j.asoc.2017.01.015.
[95] A. Dedinec, S. Filiposka, A. Dedinec, and L. Kocarev, ''Deep belief network based electricity load forecasting: An analysis of Macedonian case,'' *Energy*, vol. 115, pp. 1688–1700, Nov. 2016, doi: 10.1016/j.energy.2016.07.090.
[96] P.-H. Kuo and C.-J. Huang, ''A high precision artificial neural networks model for short-term energy load forecasting,'' *Energies*, vol. 11, no. 1, p. 213, Jan. 2018, doi: 10.3390/en11010213.
[97] E. Mocanu, P. H. Nguyen, M. Gibescu, E. M. Larsen, and P. Pinson, ''Demand forecasting at low aggregation levels using factored conditional restricted Boltzmann machine,'' presented at the 19th Power Syst. Comput. Conf., Jun. 2016, doi: 10.1109/PSCC.2016.7540994.
[98] E. D. Varga, S. F. Beretka, C. Noce, and G. Sapienza, ''Robust real-time load profile encoding and classification framework for efficient power systems operation,'' *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1897–1904, Jul. 2015, doi: 10.1109/TPWRS.2014.2354552.
[99] L. Wang, Z. Zhang, and J. Chen, ''Short-term electricity price forecasting with stacked denoising autoencoders,'' *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2673–2681, Nov. 2017, doi: 10.1109/TPWRS.2016.2628873.
[100] L. Li, K. Ota, and M. Dong, ''Everything is image: CNN-based short-term electrical load forecasting for smart grid,'' in *Proc. 11th Int. Conf. Frontier Comput. Sci. Technol.*, Exeter, U.K., Jun. 2017, pp. 344–351, doi: 10.1109/ISPAN-FCST-ISCC.2017.78.
[101] L. Li, K. Ota, and M. Dong, ''When weather matters: IoT-based electrical load forecasting for smart grid,'' *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 46–51, Oct. 2017, doi: 10.1109/MCOM.2017.1700168.
[102] C. Fan, J. Wang, W. Gang, and S. Li, ''Assessment of deep recurrent neural network-based strategies for short-term building energy predictions,'' *Appl. Energy*, vol. 236, pp. 700–710, Feb. 2019, doi: 10.1016/j.apenergy.2018.12.004.
[103] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, ''Short-term residential load forecasting based on LSTM recurrent neural network,'' *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 841–851, Jan. 2019, doi: 10.1109/TSG.2017.2753802.
[104] Y. Wang, D. Gan, M. Sun, N. Zhang, Z. Lu, and C. Kang, ''Probabilistic individual load forecasting using pinball loss guided LSTM,'' *Appl Energy*, vol. 235, pp. 10–20, Feb. 2019, doi: 10.1016/j.apenergy.2018.10.078.
[105] K. Yan, X. Wang, Y. Du, N. Jin, H. Huang, and H. Zhou, ''Multi-step short-term power consumption forecasting with a hybrid deep learning strategy,'' *Energies*, vol. 11, no. 11, p. 3089, Nov. 2018, doi: 10.3390/en11113089.
[106] V. N. Coelho, I. M. Coelho, E. Rios, A. S. T. Filho, A. J. R. Reis, B. N. Coelho, A. Alves, G. G. Netto, M. J. F. Souza, and F. G. Guimarães, ''A hybrid deep learning forecasting model using GPU disaggregated function evaluations applied for household electricity demand forecasting,'' *Energy Proc.*, vol. 103, pp. 280–285, Dec. 2016, doi: 10.1016/j.egypro.2016.11.286.
[107] S. Ryu, J. Noh, and H. Kim, ''Deep neural network based demand side short term load forecasting,'' in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Sydney, NSW, Australia, Nov. 2016, pp. 308–313, doi: 10.1109/SmartGridComm.2016.7778779.
[108] K. Amarasinghe, D. L. Marino, and M. Manic, ''Deep neural networks for energy load forecasting,'' in *Proc. IEEE 26th Int. Symp. Ind. Electron. (ISIE)*, Jun. 2017, pp. 1483–1488, doi: 10.1109/ISIE.2017.8001465.
[109] E. Mocanu, P. H. Nguyen, M. Gibescu, and W. L. Kling, ''Deep learning for estimating building energy consumption,'' *Sustain. Energy, Grids Netw.*, vol. 6, pp. 91–99, Jun. 2016, doi: 10.1016/j.segan.2016.02.005.
[110] E. Mocanu, P. H. Nguyen, W. L. Kling, and M. Gibescu, ''Unsupervised energy prediction in a smart grid context using reinforcement cross-building transfer learning,'' *Energy Buildings*, vol. 116, pp. 646–655, Mar. 2016, doi: 10.1016/j.enbuild.2016.01.030.
[111] I. M. Coelho, V. N. Coelho, E. J. D. S. Luz, L. S. Ochi, F. G. Guimarães, and E. Rios, ''A GPU deep learning Metaheuristic based model for time series forecasting,'' *Appl. Energy*, vol. 201, pp. 412–418, Sep. 2017, doi: 10.1016/j.apenergy.2017.01.003.
[112] X. Qiu, L. Zhang, Y. Ren, P. Suganthan, and G. Amaratunga, ''Ensemble deep learning for regression and time series forecasting,'' presented at the IEEE Symp. Comput. Intell. Ensemble Learn. (CIEL), Dec. 2014, doi: 10.1109/CIEL.2014.7015739.
[113] A. Rahman, V. Srikumar, and A. D. Smith, ''Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks,'' *Appl. Energy*, vol. 212, pp. 372–385, Feb. 2018, doi: 10.1016/j.apenergy.2017.12.051.

[114] R. F. Berriel, A. T. Lopes, A. Rodrigues, F. M. Varejao, and T. Oliveira-Santos, "Monthly energy consumption forecast: A deep learning approach," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Anchorage, AK, USA, May 2017, pp. 4283–4290, doi: 10.1109/IJCNN.2017.7966398.

[115] M. Cai, M. Pipattanasomporn, and S. Rahman, "Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques," *Appl. Energy*, vol. 236, pp. 1078–1088, Feb. 2019, doi: 10.1016/J.APENERGY.2018.12.042.

[116] J. Kim, J. Moon, E. Hwang, and P. Kang, "Recurrent inception convolution neural network for multi short-term load forecasting," *Energy Buildings*, vol. 194, pp. 328–341, Jul. 2019.

[117] S. Aslam, "An optimal home energy management scheme considering grid connected microgrids with day-ahead weather forecasting using artificial neural network," M.S. thesis, COMSATS Univ. Islamabad, Islamabad, Pakistan, 2018.

[118] W. Hu, Z. Chen, and B. Bak-Jensen, "Optimal operation strategy of battery energy storage system to real-time electricity price in Denmark," in *Proc. IEEE PES Gen. Meeting*, Jul. 2010, pp. 1–7.

[119] S. Aslam, N. Javaid, F. Khan, A. Alamri, A. Almogren, and W. Abdul, "Towards efficient energy management and power trading in a residential area via integrating a grid-connected microgrid," *Sustainability*, vol. 10, no. 4, p. 1245, Apr. 2018.

[120] S. Aslam, Z. Iqbal, N. Javaid, Z. Khan, K. Aurangzeb, and S. Haider, "Towards efficient energy management of smart buildings exploiting heuristic optimization with real time and critical peak pricing schemes," *Energies*, vol. 10, no. 12, p. 2065, Dec. 2017.

[121] (Jun. 2019). *Recurrent Neural Networks Tutorial*. [Online]. Available: http://www.wildml.com/2015/09/recurr entneural-networks-tutorial-part-1-introduction-to-rnns/

[122] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, 8, pp. 1735–1780, 1997.

[123] G. FA, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," in *Proc. 9th Int. Conf. Artif. Neural Networks (ICANN)*, 1999, pp. 850–855.

[124] K.-P. Lin, P.-F. Pai, and Y.-J. Ting, "Deep belief networks with genetic algorithms in forecasting wind speed," *IEEE Access*, vol. 7, pp. 99244–99253, 2019.

[125] A. Mosavi, M. Salimi, S. F. Ardabili, T. Rabczuk, S. Shamshirband, and A. Varkonyi-Koczy, "State of the art of machine learning models in energy systems, a systematic review," *Energies*, vol. 12, no. 7, p. 1301, Apr. 2019.

[126] A. Mostafaeipour, M. B. Fakhrzad, S. Gharaat, M. Jahangiri, J. A. Dhanraj, S. S. Band, A. Issakhov, and A. Mosavi, "Machine learning for prediction of energy in wheat production," *Agriculture*, vol. 10, no. 11, p. 517, Oct. 2020.

[127] Y. Li, H. Shi, F. Han, Z. Duan, and H. Liu, "Smart wind speed forecasting approach using various boosting algorithms, big multi-step forecasting strategy," *Renew. Energy*, vol. 135, pp. 540–553, May 2019.

[128] X. Mi, H. Liu, and Y. Li, "Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine," *Energy Convers. Manage.*, vol. 180, pp. 196–205, Jan. 2019.

[129] L. Cheng, H. Zang, T. Ding, R. Sun, M. Wang, Z. Wei, and G. Sun, "Ensemble recurrent neural network based probabilistic wind speed forecasting approach," *Energies*, vol. 11, no. 8, p. 1958, Jul. 2018.

[130] M. Cui, J. Zhang, Q. Wang, V. Krishnan, and B.-M. Hodge, "A data-driven methodology for probabilistic wind power ramp forecasting," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1326–1338, Mar. 2019.

[131] C. Wan, J. Lin, J. Wang, Y. Song, and Z. Y. Dong, "Direct quantile regression for nonparametric probabilistic forecasting of wind power generation," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2767–2778, Jul. 2017.

[132] J. Wang and S. Xiong, "A hybrid forecasting model based on outlier detection and fuzzy time series—A case study on Hainan wind farm of China," *Energy*, vol. 76, pp. 526–541, Nov. 2014.

[133] J. Wang, Y. Song, F. Liu, and R. Hou, "Analysis and application of forecasting models in wind power integration: A review of multi-step-ahead wind speed forecasting models," *Renew. Sustain. Energy Rev.*, vol. 60, pp. 960–981, Jul. 2016.

[134] H. Liu, X.-W. Mi, and Y.-F. Li, "Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network," *Energy Convers. Manage.*, vol. 156, pp. 498–514, Jan. 2018.

[135] J. Chen, G.-Q. Zeng, W. Zhou, W. Du, and K.-D. Lu, "Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization," *Energy Convers. Manage.*, vol. 165, pp. 681–695, Jun. 2018.

[136] H. Z. Wang, G.-Q. Li, G.-B. Wang, J.-C. Peng, H. Jiang, and Y.-T. Liu, "Deep learning based ensemble approach for probabilistic wind power forecasting," *Appl. Energy*, vol. 188, pp. 56–70, Feb. 2017, doi: 10.1016/j.apenergy.2016.11.111.

[137] R. Yu, J. Gao, M. Yu, W. Lu, T. Xu, M. Zhao, J. Zhang, R. Zhang, and Z. Zhang, "LSTM-EFG for wind power forecasting based on sequential correlation features," *Future Gener. Comput. Syst.*, vol. 93, pp. 33–42, Apr. 2019.

[138] Y.-L. Hu and L. Chen, "A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and differential evolution algorithm," *Energy Convers. Manage.*, vol. 173, pp. 123–142, Oct. 2018, doi: 10.1016/j.enconman.2018.07.070.

[139] M. Abdel-Nasser and K. Mahmoud, "Accurate photovoltaic power forecasting models using deep LSTM-RNN," *Neural Comput. Appl.*, vol. 31, no. 7, pp. 2727–2740, Jul. 2019, doi: 10.1007/s00521-017-3225-z.

[140] A. Alzahrani, P. Shamsi, C. Dagli, and M. Ferdowsi, "Solar irradiance forecasting using deep recurrent neural networks," in *Proc. ICRERA*, vol. 114, 2017, pp. 304–313, doi: 10.1016/j.procs.2017.09.045.

[141] J. Ma, J. Zhang, L. Xiao, K. Chen, and J. Wu, "Classification of power quality disturbances via deep learning," *IETE Tech. Rev.*, vol. 34, no. 4, pp. 408–415, Jul. 2017, doi: 10.1080/02564602.2016.1196620.

[142] X. Kong, X. Xu, Z. Yan, S. Chen, H. Yang, and D. Han, "Deep learning hybrid method for islanding detection in distributed generation," *Appl Energy*, vol. 210, pp. 776–785, Jan. 2018, doi: 10.1016/j.apenergy.2017.08.014.

[143] Y. Wang, M. Liu, and Z. Bao, "Deep learning neural network for power system fault diagnosis," in *Proc. 35th Chin. Control Conf. (CCC)*, Jul. 2016, pp. 6678–6683.

[144] Z. Yizhe, M. Wang, D. Gang, G. Jun, and W. Pai, "A cable fault recognition method based on a deep belief network," *Comput. Electr. Eng.*, vol. 71, pp. 452–464, Oct. 2018, doi: 10.1016/j.compeleceng.2018.07.043.

[145] Y. He, G. J. Mendis, and J. Wei, "Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2505–2516, Sep. 2017.

[146] X. Z. Wang, J. Zhou, Z. L. Huang, X. L. Bi, Z. Q. Ge, and L. Li, "A multilevel deep learning method for big data analysis and emergency management of power system," in *Proc. IEEE Int. Conf. Big Data Anal. (ICBDA)*, Mar. 2016, pp. 1–5, doi: 10.1109/ICBDA.2016.7509811.

[147] J. Zheng, C. Xu, Z. Zhang, and X. Li, "Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network," in *Proc. 51st Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2017, pp. 1–6, doi: 10.1109/CISS.2017.7926112.

[148] S. Bouktif, A. Fiaz, A. Ouni, and M. Serhani, "Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches," *Energies*, vol. 11, no. 7, p. 1636, Jun. 2018, doi: 10.3390/en11071636.

[149] E. Mocanu, P. H. Nguyen, M. Gibescu, E. M. Larsen, and P. Pinson, "Demand forecasting at low aggregation levels using factored conditional RBM," presented at the 19th Power Syst. Comput. Conf., 2016, doi: 10.1109/PSCC.2016.7540994.

[150] P.-H. Kuo and C.-J. Huang, "A high precision artificial neural networks model for short-term energy load forecasting," *Energies*, vol. 11, no. 1, pp. 1–13, Jan. 2018, doi: 10.3390/en11010213.

[151] K. Yan, X. Wang, Y. Du, N. Jin, H. Huang, and H. Zhou, "Multi-step short-term power consumption forecasting with a hybrid deep learning strategy," *Energies*, vol. 11, no. 11, p. 3089, Nov. 2018, doi: 10.3390/en11113089.

[152] C. Li, Z. Ding, D. Zhao, J. Yi, and G. Zhang, "Building energy consumption prediction: An extreme deep learning approach," *Energies*, vol. 10, no. 10, p. 1525, Oct. 2017, doi: 10.3390/en10101525.

[153] A. S. Qureshi, A. Khan, A. Zameer, and A. Usman, "Wind power prediction using deep neural network based meta regression and transfer learning," *Appl. Soft Comput.*, vol. 58, pp. 742–755, Sep. 2017, doi: 10.1016/j.asoc.2017.05.031.

[154] M. Khodayar, O. Kaynak, and M. E. Khodayar, "Rough deep neural architecture for short-term wind speed forecasting," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 2770–2779, Dec. 2017.

[155] M. Dalto, J. Matuško, and M. Vašak, "Deep neural networks for ultra-short-term wind forecasting," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Seville, Spain, Mar. 2015, pp. 1657–1663.

[156] Y. Tao, H. Chen, and C. Qiu, "Wind power prediction and pattern feature based on deep learning method," in *Proc. IEEE PES Asia–Pacific Power Energy Eng. Conf. (APPEEC)*, Hong Kong, Dec. 2014, pp. 1–4.

[157] C.-Y. Zhang, C. L. P. Chen, M. Gan, and L. Chen, "Predictive deep Boltzmann machine for multiperiod wind speed forecasting," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1416–1425, Oct. 2015.

[158] K. Wang, X. Qi, H. Liu, and J. Song, "Deep belief network based k-means cluster approach for short-term wind power forecasting," *Energy*, vol. 165, pp. 840–852, Dec. 2018, doi: 10.1016/j.energy.2018.09.118.

[159] W. Wu, K. Chen, Y. Qiao, and Z. Lu, "Probabilistic short-term wind power forecasting based on deep neural networks," in *Proc. Int. Conf. Probabilistic Methods Appl. Power Syst. (PMAPS)*, Beijing, China, 2016, pp. 1–8, doi: 10.1109/PMAPS.2016.7764155.

[160] Y. Qin, K. Li, Z. Liang, B. Lee, F. Zhang, Y. Gu, L. Zhang, F. Wu, and D. Rodriguez, "Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal," *Appl. Energy*, vol. 236, pp. 262–272, Feb. 2019, doi: 10.1016/j.apenergy.2018.11.063.

[161] H. Liu, X. Mi, and Y. Li, "Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network," *Energy Convers. Manage.*, vol. 166, pp. 120–131, Jun. 2018, doi: 10.1016/j.enconman.2018.04.021.

[162] Q. Xiaoyun, K. Xiaoning, Z. Chao, J. Shuai, and M. Xiuda, "Short-term prediction of wind power based on deep long short-term memory," presented at the IEEE PES Asia–Pacific Power Energy Eng. Conf. (APPEEC), Xi'an, China, Oct. 2016, doi: 10.1109/APPEEC.2016.7779672.

[163] A. Gensler, J. Henze, B. Sick, and N. Raabe, "Deep learning for solar power forecasting—An approach using AutoEncoder and LSTM neural networks," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Budapest, Hungary, Oct. 2016, pp. 2858–2865, doi: 10.1109/SMC.2016.7844673.

[164] L.-L. Li, P. Cheng, H.-C. Lin, and H. Dong, "Short-term output power forecasting of photovoltaic systems based on the deep belief net," *Adv. Mech. Eng.*, vol. 9, no. 9, pp. 1–13, Sep. 2017, doi: 10.1177/1687814017715983.

[165] Y. Q. Neo, T. T. Teo, W. L. Woo, T. Logenthiran, and A. Sharma, "Forecasting of photovoltaic power using deep belief network," in *Proc. IEEE Region Conf. (TENCON)*, Penang, Malaysia, Nov. 2017, pp. 1189–1194, doi: 10.1109/TENCON.2017.8228038.

[166] H. Wang, H. Yi, J. Peng, G. Wang, Y. Liu, H. Jiang, and W. Liu, "Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network," *Energy Convers. Manage.*, vol. 153, pp. 409–422, Dec. 2017, doi: 10.1016/j.enconman.2017.10.008.

[167] A. Alzahrani, P. Shamsi, C. Dagli, and M. Ferdowsi, "Solar irradiance forecasting using deep neural networks," *Proc. Comput. Sci.*, vol. 114, pp. 304–313, Jan. 2017, doi: 10.1016/j.procs.2017.09.045.

[168] M. Mishra, "Power quality disturbance detection and classification using signal processing and soft computing techniques: A comprehensive review," *Int. Trans. Electr. Energy Syst.*, vol. 29, no. 8, p. e12008, Aug. 2019, doi: 10.1002/2050-7038.12008.

[169] M. Mahdi and V. M. I. Genc, "Post-fault prediction of transient instabilities using stacked sparse autoencoder," *Electr. Power Syst. Res.*, vol. 164, pp. 243–252, Nov. 2018, doi: 10.1016/j.epsr.2018.08.009.

[170] X. Shi, H. Yang, Z. Xu, X. Zhang, and M. R. Farahani, "An independent component analysis classification for complex power quality disturbances with sparse auto encoder features," *IEEE Access*, vol. 7, pp. 20961–20966, 2019, doi: 10.1109/ACCESS.2019.2898211.

[171] Y. Deng, L. Wang, H. Jia, X. Tong, and F. Li, "A sequence-to-sequence deep learning architecture based on bidirectional GRU for type recognition and time location of combined power quality disturbance," *IEEE Trans. Ind. Informat.*, vol. 15, no. 8, pp. 4481–4493, Aug. 2019, doi: 10.1109/TII.2019.2895054.

[172] W. Qiu, Q. Tang, J. Liu, and W. Yao, "An automatic identification framework for complex power quality disturbances based on multifusion convolutional neural network," *IEEE Trans. Ind. Informat.*, vol. 16, no. 5, pp. 3233–3241, May 2020, doi: 10.1109/TII.2019.2920689.

[173] E. Balouji, I. Y. H. Gu, M. H. J. Bollen, A. Bagheri, and M. Nazari, "A LSTM-based deep learning method with application to voltage dip classification," in *Proc. 18th Int. Conf. Harmon. Quality Power (ICHQP)*, May 2018, pp. 1–5, doi: 10.1109/ICHQP.2018.8378893.

[174] C.-M. Li, Z.-X. Li, N. Jia, Z.-L. Qi, and J.-H. Wu, "Classification of power-quality disturbances using deep belief network," in *Proc. Int. Conf. Wavelet Anal. Pattern Recognit. (ICWAPR)*, Chengdu, China, Jul. 2018, pp. 231–237.

[175] F. Mei, Y. Ren, Q. Wu, C. Zhang, Y. Pan, H. Sha, and J. Zheng, "Online recognition method for voltage sags based on a deep belief network," *Energies*, vol. 12, no. 1, p. 43, Dec. 2018, doi: 10.3390/en12010043.

[176] S. Wang and H. Chen, "A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network," *Appl. Energy*, vol. 235, pp. 1126–1140, Feb. 2019, doi: 10.1016/j.apenergy.2018.09.160.

[177] H. Liu, F. Hussain, Y. Shen, S. Arif, A. Nazir, and M. Abubakar, "Complex power quality disturbances classification via curvelet transform and deep learning," *Electr. Power Syst. Res.*, vol. 163, pp. 1–9, Oct. 2018, doi: 10.1016/j.epsr.2018.05.018.

[178] Y. Shen, M. Abubakar, H. Liu, and F. Hussain, "Power quality disturbance monitoring and classification based on improved PCA and convolution neural network for wind-grid distribution systems," *Energies*, vol. 12, no. 7, p. 1280, Apr. 2019, doi: 10.3390/en12071280.

[179] N. Mohan, K. P. Soman, and R. Vinayakumar, "Deep power: Deep learning architectures for power quality disturbances classification," in *Proc. Int. Conf. Technol. Advancements Power Energy (TAP Energy)*, Dec. 2017, pp. 1–6, doi: 10.1109/TAPENERGY.2017.8397249.

[180] J. Wang, Z. Xu, and Y. Che, "Power quality disturbance classification based on compressed sensing and deep convolution neural networks," *IEEE Access*, vol. 7, pp. 78336–78346, 2019, doi: 10.1109/ACCESS.2019.2922367.

[181] H. Liu, F. Hussain, S. Yue, O. Yildirim, and S. J. Yawar, "Classification of multiple power quality events via compressed deep learning," *Int. Trans. Electr. Energy Syst.*, vol. 29, no. 6, pp. 1–14, Jun. 2019, doi: 10.1002/2050-7038.12010.

[182] J. J. Q. Yu, Y. Hou, A. Y. S. Lam, and V. O. K. Li, "Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1694–1703, Mar. 2019, doi: 10.1109/TSG.2017.2776310.

[183] S. Zhang, Y. Wang, M. Liu, and Z. Bao, "Data-based line trip fault prediction in power systems using LSTM networks and SVM," *IEEE Access*, vol. 6, pp. 7675–7686, 2018, doi: 10.1109/ACCESS.2017.2785763.

[184] L. Wang, Z. Zhang, H. Long, J. Xu, and R. Liu, "Wind turbine gearbox failure identification with deep neural networks," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1360–1368, Jun. 2017.

[185] F. Cheng, J. Wang, L. Qu, and W. Qiao, "Rotor-current-based fault diagnosis for DFIG wind turbine drivetrain gearboxes using frequency analysis and a deep classifier," *IEEE Trans. Ind. Appl.*, vol. 54, no. 2, pp. 1062–1071, Mar. 2018, doi: 10.1109/TIA.2017.2773426.

[186] Z. Zhang, S. Li, Y. Xiao, and Y. Yang, "Intelligent simultaneous fault diagnosis for solid oxide fuel cell system based on deep learning," *Appl. Energy*, vols. 233–234, pp. 930–942, Jan. 2019, doi: 10.1016/j.apenergy.2018.10.113.

[187] L. Wang, X. Zhao, J. Pei, and G. Tang, "Transformer fault diagnosis using continuous sparse autoencoder," *SpringerPlus*, vol. 5, no. 1, p. 448, Dec. 2016, doi: 10.1186/s40064-016-2107-7.

[188] J. Dai, H. Song, G. Sheng, and X. Jiang, "Dissolved gas analysis of insulating oil for power transformer fault diagnosis with deep belief network," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 5, pp. 2828–2835, Oct. 2017, doi: 10.1109/TDEI.2017.006727.

[189] T.-E. Huang, Q. Guo, H. Sun, C.-W. Tan, and T. Hu, "A deep spatial-temporal data-driven approach considering microclimates for power system security assessment," *Appl. Energy*, vol. 237, pp. 36–48, Mar. 2019, doi: 10.1016/j.apenergy.2019.01.013.

[190] Z. Zheng, Y. Yang, X. Niu, H.-N. Dai, and Y. Zhou, "Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids," *IEEE Trans. Ind. Informat.*, vol. 14, no. 4, pp. 1606–1615, Apr. 2018, doi: 10.1109/TII.2017.2785963.

[191] J. M. Malof, L. M. Collins, K. Bradbury, and R. G. Newell, "A deep convolutional neural network and a random forest classifier for solar photovoltaic array detection in aerial imagery," in *Proc. IEEE Int. Conf. Renew. Energy Res. Appl. (ICRERA)*, vol. 5, Nov. 2016, pp. 5–9.

[192] S. Shen, M. Sadoughi, X. Chen, M. Hong, and C. Hu, "A deep learning method for online capacity estimation of lithium-ion batteries," *J. Energy Storage*, vol. 25, Oct. 2019, Art. no. 100817, doi: 10.1016/j.est.2019.100817.
[193] E. Mocanu, D. Constantin Mocanu, P. H. Nguyen, A. Liotta, M. E. Webber, M. Gibescu, and J. G. Slootweg, "On-line building energy optimization using deep reinforcement learning," 2017, pp. 1–9, *arXiv:1707.05878*.
[194] V. François-lavet, D. Taralla, E. Damien, and R. Fonteneau, "Deep reinforcement learning solutions for energy microgrids management," in *Proc. Eur. Workshop Reinforcement Learn.*, 2016, pp. 1–7.
[195] N. Wu and H. Wang, "Deep learning adaptive dynamic programming for real time energy management and control strategy of microgrid," *J. Cleaner Prod.*, vol. 204, pp. 1169–1177, Dec. 2018, doi: 10.1016/j.jclepro.2018.09.052.
[196] T. Sogabe, I. Haruhisa, K. Sakamoto, K. Yamaguchi, M. Sogabe, and T. Sato, "Optimization of decentralized renewable energy system by weather forecasting and deep machine learning techniques," in *Proc. IEEE Innov. Smart Grid Technol.*, Nov./Dec. 2016, pp. 1–5.
[197] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006, doi: 10.1016/j.neucom.2005.12.126.
[198] Ö. F. Ertugrul, "Forecasting electricity load by a novel recurrent extreme learning machines approach," *Int. J. Electr. Power Energy Syst.*, vol. 78, pp. 429–435, Jun. 2016.
[199] Z. Ma, C. Zhang, and C. Qian, "Review of machine learning in power system," in *Proc. IEEE Innov. Smart Grid Technol.-Asia (ISGT Asia)*, May 2019, pp. 3401–3406.
[200] J. Shi, W.-J. Lee, Y. Liu, Y. Yang, and P. Wang, "Forecasting power output of photovoltaic systems based on weather classification and support vector machines," *IEEE Trans. Ind. Appl.*, vol. 48, no. 3, pp. 1064–1069, May/Jun. 2012.
[201] 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 Syst. Appl.*, vol. 41, no. 13, pp. 6047–6056, 2014.
[202] S. Li, L. Goel, and P. Wang, "An ensemble approach for short-term load forecasting by extreme learning machine," *Appl. Energy*, vol. 170, pp. 22–29, May 2016, doi: 10.1016/j.apenergy.2016.02.114.
[203] Ö. F. Ertugrul, "Forecasting electricity load by a novel recurrent extreme learning machines approach," *Int. J. Elect. Power Energy Syst.*, vol. 78, pp. 429–435, Jun. 2016, doi: 10.1016/j.ijepes.2015.12.006.
[204] W. Zhang, H. Hua, and J. Cao, "Short term load forecasting based on IGSA-ELM algorithm," in *Proc. IEEE Int. Conf. Energy Internet (ICEI)*, Apr. 2017, pp. 296–301, doi: 10.1109/ICEI.2017.59.
[205] J. Li, Z. Song, Y. Zhong, Z. Zhang, and J. Li, "Short-term load forecasting based on improved extreme learning machine," in *Proc. IEEE 2nd Int. Conf. Big Data Anal. (ICBDA)*, Mar. 2017, pp. 584–588, doi: 10.1109/ICBDA.2017.8078701.
[206] Y. Chen, M. Kloft, Y. Yang, C. Li, and L. Li, "Mixed kernel based extreme learning machine for electric load forecasting," *Neurocomputing*, vol. 312, pp. 90–106, Oct. 2018, doi: 10.1016/j.neucom.2018.05.068.
[207] B. Jatin and D. Toshniwal, "Empirical mode decomposition based deep learning for electricity demand forecasting," *IEEE Access*, vol. 6, pp. 49144–49156, 2018, doi: 10.1109/ACCESS.2018.2867681.
[208] D. C. Park, M. A. El-Sharkawi, R. J. Marks, L. E. Atlas, and M. J. Damborg, "Electric load forecasting using an artificial neural network," *IEEE Trans. Power Syst.*, vol. 6, no. 2, pp. 442–449, May 1991.
[209] A. Khotanzad, R.-C. Hwang, A. Abaye, and D. Maratukulam, "An adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities," *IEEE Trans. Power Syst.*, vol. 10, no. 3, pp. 1716–1722, Aug. 1995.
[210] V. H. Ferreira and A. P. A. da Silva, "Toward estimating autonomous neural network-based electric load forecasters," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1554–1562, Nov. 2007, doi: 10.1109/TPWRS.2007.908438.
[211] N. Ding, C. Benoit, G. Foggia, Y. Bésanger, and F. Wurtz, "Neural network-based model design for short-term load forecast in distribution systems," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 72–81, Jan. 2016, doi: 10.1109/TPWRS.2015.2390132.
[212] T. Kohonen, "The self-organizing map," *Neurocomputing*, vol. 21, pp. 1–6, Nov. 1998.
[213] R. Lamedica, A. Prudenzi, M. Sforna, M. Caciotta, and V. O. Cencellli, "A neural network based technique for short-term forecasting of anomalous load periods," *IEEE Trans. Power Syst.*, vol. 11, no. 4, pp. 1749–1756, Nov. 1996.
[214] M. López, S. Valero, C. Senabre, J. Aparicio, and A. Gabaldon, "Application of SOM neural networks to short-term load forecasting: The Spanish electricity market case study," *Electr. Power Syst. Res.*, vol. 91, pp. 18–27, Oct. 2012.
[215] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biol. Cybern.*, vol. 43, no. 1, pp. 59–69, 1982, doi: 10.1007/BF00337288.
[216] S. Fan, C. Mao, and L. Chen, "Peak load forecasting using the self-organizing map," in *Advances in Neural Networks—ISNN*. Berlin, Germany: Springer, 2005, pp. 640–647, doi: 10.1007/11427469_102.
[217] H. Aprillia, H.-T. Yang, and C.-M. Huang, "Optimal decomposition and reconstruction of discrete wavelet transformation for short-term load forecasting," *Energies*, vol. 12, no. 24, p. 4654, 2019.
[218] N. Amral, C. S. Ozveren, and D. King, "Short term load forecasting using multiple linear regression," in *Proc. 42nd Int. Univ. Power Eng. Conf.*, Brighton, U.K., Sep. 2007, pp. 1192–1198.
[219] A. Y. Saber and A. K. M. R. Alam, "Short term load forecasting using multiple linear regression for big data," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Honolulu, HI, USA, Nov./Dec. 2017, pp. 1–6.
[220] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *J. Cybern.*, vol. 3, no. 3, pp. 32–57, Jan. 1973, doi: 10.1080/01969727308546046.
[221] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Comput. Geosci.*, vol. 10, no. 2, Art. no. 2, Jan. 1984, doi: 10.1016/0098-3004(84)90020-7.
[222] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, Jun. 1965, Art. no. 3, doi: 10.1016/S0019-9958(65)90241-X.
[223] Z. Xin and C. Tian-Lun, "Nonlinear time series forecast using radial basis function neural networks," *Commun. Theor. Phys.*, vol. 40, no. 2, Aug. 2003, Art. no. 2, doi: 10.1088/0253-6102/40/2/165.
[224] Y. Zhu and Y. He, "Short-term load forecasting model using fuzzy c means based radial basis function network," in *Proc. 6th Int. Conf. Intell. Syst. Design Appl.*, vol. 1, Oct. 2006, pp. 579–582, doi: 10.1109/ISDA.2006.235.
[225] D. K. Ranaweera, N. F. Hubele, and G. G. Karady, "Fuzzy logic for short term load forecasting," *Int. J. Electr. Power Energy Syst.*, vol. 18, no. 4, May 1996, Art. no. 4, doi: 10.1016/0142-0615(95)00060-7.
[226] S.-J. Kang, C.-H. Woo, H.-S. Hwang, and K. B. Woo, "Evolutionary design of fuzzy rule base for nonlinear system modeling and control," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 1, pp. 37–45, Feb. 2000, doi: 10.1109/91.824766.
[227] A. Khosravi, S. Nahavandi, D. Creighton, and D. Srinivasan, "Interval type-2 fuzzy logic systems for load forecasting: A comparative study," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1274–1282, Aug. 2012, doi: 10.1109/TPWRS.2011.2181981.
[228] S. Hassan, A. Khosravi, J. Jaafar, and M. A. Khanesar, "A systematic design of interval type-2 fuzzy logic system using extreme learning machine for electricity load demand forecasting," *Int. J. Electr. Power Energy Syst.*, vol. 82, pp. 1–10, Nov. 2016, doi: 10.1016/j.ijepes.2016.03.001.
[229] D. Ali, M. Yohanna, M. I. Puwu, and B. M. Garkida, "Long-term load forecast modelling using a fuzzy logic approach," *Pacific Sci. Rev. A, Natural Sci. Eng.*, vol. 18, no. 2, pp. 123–127, Jul. 2016, doi: 10.1016/j.psra.2016.09.011.
[230] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah, "A review on time series forecasting techniques for building energy consumption," *Renew. Sustain. Energy Rev.*, vol. 74, pp. 902–924, Jul. 2017.
[231] I. Sansa, Z. Boussaada, M. Mazigh, and N. M. Bellaaj, "Solar radiation prediction for a winter day using ARMA model," in *Proc. 6th IEEE Int. Energy Conf. (ENERGYCon)*, Sep. 2020, pp. 326–330.
[232] M. H. Amini, O. Karabasoglu, M. D. Ilic, K. G. Boroojeni, and S. S. Iyengar, "ARIMA-based demand forecasting method considering probabilistic model of 'electric vehicles' parking lots," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2015, pp. 1–5.
[233] D. Z. Amirhosseini, R. Sabzehgar, and M. Rasouli, "Power management and optimization for a residential smart microgrid using stochastic methods," in *Proc. 9th IEEE Int. Symp. Power Electron. Distrib. Gener. Syst. (PEDG)*, Jun. 2018, pp. 1–6.

[234] A. Elmouatamid, R. Ouladsine, M. Bakhouya, N. E. Kamoun, K. Zine-Dine, and M. Khaidar, "A control strategy based on power forecasting for micro-grid systems," in *Proc. IEEE Int. Smart Cities Conf. (ISC)*, Oct. 2019, pp. 735–740.

[235] R. Kedrowski, J. Nelson, A. S. Nair, and P. Ranganathan, "Short-term seasonal energy forecasting," in *Proc. IEEE Int. Conf. Electro/Inf. Technol. (EIT)*, May 2018, pp. 696–700.

[236] Q. Hu, S. Zhang, M. Yu, and Z. Xie, "Short-term wind speed or power forecasting with heteroscedastic support vector regression," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 241–249, Jan. 2016.

[237] T. Gneiting and M. Katzfuss, "Probabilistic forecasting," *Annu. Rev. Statist. Appl.*, vol. 1, pp. 125–151, Jan. 2014.

[238] H. Tao, "Short term electric load forecasting," Pro-Quest Disseretation, North Caroline Univ., NC, USA, 2010.

[239] J. Dowell and P. Pinson, "Very-short-term probabilistic wind power forecasts by sparse vector autoregression," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 763–770, Mar. 2016.

[240] J. Nowotarski and R. Weron, "Computing electricity spot price prediction intervals using quantile regression and forecast averaging," *Comput. Statist.*, vol. 30, no. 3, pp. 791–803, Sep. 2015.

[241] K. Maciejowska, J. Nowotarski, and R. Weron, "Probabilistic forecasting of electricity spot prices using factor quantile regression averaging," *Int. J. Forecasting*, vol. 32, no. 3, pp. 957–965, Jul./Sep. 2016.

[242] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *Int. J. Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014.

[243] B. Liu, J. Nowotarski, T. Hong, and R. Weron, "Probabilistic load forecasting via quantile regression averaging on sister forecasts," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 730–737, Mar. 2017.

[244] Y. Wang, N. Zhang, Y. Tan, T. Hong, D. S. Kirschen, and C. Kang, "Combining probabilistic load forecasts," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3664–3674, Jul. 2019.

[245] H. Aprillia, H.-T. Yang, and C.-M. Huang, "Statistical load forecasting using optimal quantile regression random forest and risk assessment index," *IEEE Trans. Smart Grid*, vol. 12, no. 2, pp. 1467–1480, Mar. 2021.

[246] Y. Yang, W. Li, T. A. Gulliver, and S. Li, "Bayesian deep learning-based probabilistic load forecasting in smart grids," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4703–4713, Jul. 2020.

[247] M. Sun, T. Zhang, Y. Wang, G. Strbac, and C. Kang, "Using Bayesian deep learning to capture uncertainty for residential net load forecasting," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 188–201, Jan. 2020.

[248] J.-F. Toubeau, J. Bottieau, F. Vallée, and Z. De Greve, "Deep learning-based multivariate probabilistic forecasting for short-term scheduling in power markets," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1203–1215, Mar. 2019.

[249] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1050–1059.

[250] C. Kuster, Y. Rezgui, and M. Mourshed, "Electrical load forecasting models: A critical systematic review," *Sustain. Cities Soc.*, vol. 35, pp. 257–270, Nov. 2017.

[251] S.-M. Baek, "Mid-term load pattern forecasting with recurrent artificial neural network," *IEEE Access*, vol. 7, pp. 172830–172838, 2019, doi: 10.1109/ACCESS.2019.2957072.

[252] A. H. Neto and F. A. S. Fiorelli, "Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption," *Energy Buildings*, vol. 40, no. 12, pp. 2169–2176, Jan. 2008, doi: 10.1016/j.enbuild.2008.06.013.

[253] G. M. U. Din and A. K. Marnerides, "Short term power load forecasting using deep neural networks," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Santa Clara, CA, USA, Jan. 2017, pp. 594–598.

[254] S. S. Reddy and C. M. Jung, "Short-term load forecasting using artificial neural networks and wavelet transform," *Int. J. Appl. Eng. Res.*, vol. 11, pp. 9831–9836, Oct. 2016.

[255] M. Mordjaoui, S. Haddad, A. Medoued, and A. Laouafi, "Electric load forecasting by using dynamic neural network," *Int. J. Hydrogen Energy*, vol. 42, no. 28, pp. 17655–17663, Jul. 2017.

[256] *Energy Infrastructure Costs and Investments Between 1996 and 2013 (Medium-Term) and Further to 2023 (Long-Term) on the Trans-European Energy Network and to Neighbouring Regions With Emphasis on Investments on Renewable Energy Sources and Their Integration Into the Trans-European Energy Networks, Including an Inventory of the Technical Status of the European Energy Network for the Year 2003*, Eur. Commission, Directorate-Gen. Transp. Energy (DG-TREN), CESI spa, Instituto de Investigacion Tecnologica, Mercados Energeticos, Brussels, Belgium, 2005.

[257] J.-S. Jang and C.-T. Sun, "Neuro-fuzzy modeling and control," *Proc. IEEE*, vol. 83, no. 3, pp. 378–406, Mar. 1995.

[258] K. Li, H. Su, and J. Chu, "Forecasting building energy consumption using neural networks and hybrid neuro-fuzzy system: A comparative study," *Energy Buildings*, vol. 43, no. 10, pp. 2893–2899, Oct. 2011.

[259] A. Khotanzad, E. Zhou, and H. Elragal, "A neuro-fuzzy approach to short-term load forecasting in a price-sensitive environment," *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 1273–1282, Nov. 2002.

[260] B. Jawerth and W. Sweldens, "An overview of wavelet based multiresolution analyses," *SIAM Rev.*, vol. 36, no. 3, pp. 377–412, Sep. 1994, doi: 10.1137/1036095.

[261] B.-L. Zhang and Z.-Y. Dong, "An adaptive neural-wavelet model for short term load forecasting," *Electr. Power Syst. Res.*, vol. 59, no. 2, pp. 121–129, Sep. 2001.

[262] W.-T. Pan, "A new fruit fly optimization algorithm: Taking the financial distress model as an example," *Knowl.-Based Syst.*, vol. 26, pp. 69–74, Feb. 2012, doi: 10.1016/j.knosys.2011.07.001.

[263] H.-Z. Li, S. Guo, C.-J. Li, and J.-Q. Sun, "A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm," *Knowl.-Based Syst.*, vol. 37, pp. 378–387, Jan. 2013, doi: 10.1016/j.knosys.2012.08.015.

[264] Z. W. Geem, J. H. Kim, and G. V. Loganathan, "A new heuristic optimization algorithm: Harmony search," *Simulation*, vol. 76, no. 2, pp. 60–68, Feb. 2001.

[265] L. Xiao, W. Shao, T. Liang, and C. Wang, "A combined model based on multiple seasonal patterns and modified firefly algorithm for electrical load forecasting," *Appl. Energy*, vol. 167, pp. 135–153, Apr. 2016.

[266] A. Kavousi-Fard, T. Niknam, and M. Golmaryami, "Short term load forecasting of distribution systems by a new hybrid modified FA-backpropagation method," *J. Intell. Fuzzy Syst.*, vol. 26, no. 1, pp. 517–522, 2014, doi: 10.3233/IFS-131025.

[267] M. El-Telbany and F. El-Karmi, "Short-term forecasting of Jordanian electricity demand using particle swarm optimization," *Electr. Power Syst. Res.*, vol. 78, no. 3, pp. 425–433, Mar. 2008.

[268] N. Liu, Q. Tang, J. Zhang, W. Fan, and J. Liu, "A hybrid forecasting model with parameter optimization for short-term load forecasting of micro-grids," *Appl. Energy*, vol. 129, pp. 336–345, Sep. 2014, doi: 10.1016/j.apenergy.2014.05.023.

[269] C. Zhang and P. Ma, "Short-term electricity price forecasting based on PSO algorithm and RBF neural network algorithm," in *Proc. Int. Conf. Measuring Technol. Mechatronics Autom.*, Mar. 2010, pp. 334–337, doi: 10.1109/ICMTMA.2010.22.

[270] Y. Yong, W. Sun'an, and S. Wanxing, "Short-term load forecasting using artificial immune network," in *Proc. Int. Conf. Power Syst. Technol.*, 2002, pp. 2322–2325, doi: 10.1109/ICPST.2002.1047199.

[271] S. H. Ling, F. H. F. Leung, H. K. Lam, Y.-S. Lee, and P. K. S. Tam, "A novel genetic-algorithm-based neural network for short-term load forecasting," *IEEE Trans. Ind. Electron.*, vol. 50, no. 4, pp. 793–799, Aug. 2003.

[272] A. Azadeh, S. F. Ghaderi, S. Tarverdian, and M. Saberi, "Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption," *Appl. Math. Comput.*, vol. 186, no. 2, pp. 1731–1741, Mar. 2007.

[273] Y. Chen, P. B. Luh, C. Guan, Y. Zhao, L. D. Michel, M. A. Coolbeth, P. B. Friedland, and S. J. Rourke, "Short-term load forecasting: Similar day-based wavelet neural networks," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 322–330, Feb. 2010.

[274] H. Zheng, J. Yuan, and L. Chen, "Short-term load forecasting using EMD-LSTM neural networks with a xgboost algorithm for feature importance evaluation," *Energies*, vol. 10, no. 8, p. 1168, Aug. 2017.

[275] D.-X. Niu, H.-F. Shi, and D. D. Wu, "Short-term load forecasting using Bayesian neural networks learned by hybrid Monte Carlo algorithm," *Appl. Soft Comput.*, vol. 12, no. 6, pp. 1822–1827, Jun. 2012.

[276] I. S. Jahan, M. Prilepok, S. Misak, and V. Snasel, "Intelligent system for power load forecasting in off-grid platform," in *Proc. 19th Int. Sci. Conf. Electr. Power Eng. (EPE)*, May 2018, pp. 1–5.

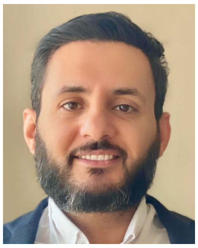
[277] A. S. Khwaja, M. Naeem, A. Anpalagan, A. Venetsanopoulos, and B. Venkatesh, "Improved short-term load forecasting using bagged neural networks," *Electr. Power Syst. Res.*, vol. 125, pp. 109–115, Aug. 2015.

[278] Y. Wang, D. Niu, and L. Ji, "Short-term power load forecasting based on IVL-BP neural network technology," *Syst. Eng. Proc.*, vol. 4, pp. 168–174, Jan. 2012.

[279] S. Ekici, "Electric load forecasting using regularized extreme learning machines," *Int. J. Ind. Electron. Elect. Eng.*, vol. 4, no. 6, pp. 119–122, 2016.
[280] J. Zhang, Y.-M. Wei, D. Li, Z. Tan, and J. Zhou, "Short term electricity load forecasting using a hybrid model," *Energy*, vol. 158, no. 1, pp. 774–781, Sep. 2018.
[281] A. Ahmad, N. Javaid, A. Mateen, M. Awais, and Z. A. Khan, "Short-term load forecasting in smart grids: An intelligent modular approach," *Energies*, vol. 12, no. 1, p. 164, Jan. 2019.
[282] Y. Liang, D. Niu, and W.-C. Hong, "Short term load forecasting based on feature extraction and improved general regression neural network model," *Energy*, vol. 166, pp. 653–663, Jan. 2019.
[283] H.-X. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renew. Sustain. Energy Rev.*, vol. 16, no. 6, pp. 3586–3592, Aug. 2012, doi: 10.1016/j.rser.2012.02.049.
[284] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, pp. 273–297, Sep. 1995, doi: 10.1023/A:1022627411411.
[285] S. M. Clarke, J. H. Griebsch, and T. W. Simpson, "Analysis of support vector regression for approximation of complex engineering analyses," *J. Mech. Des.*, vol. 127, no. 6, pp. 1077–1087, Nov. 2005, doi: 10.1115/1.1897403.
[286] C. E. Borges, Y. K. Penya, and I. Fernández, "Optimal combined short-term building load forecasting," in *Proc. IEEE PES Innov. Smart Grid Technol.*, Nov. 2011, pp. 1–7, doi: 10.1109/ISGT-Asia.2011.6167091.
[287] L. J. Wang, "Research on forecasting methods of natural gas load based on particle swarm optimization-least square support vector machine, China Shanghai," M.S. Dissertation, East China Univ. Sci. Technol., Shanghai, China, 2012.
[288] F. Yu and X. Xu, "A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network," *Appl. Energy*, vol. 134, pp. 102–113, Dec. 2014.
[289] Eberhart and Y. Shi, "Particle swarm optimization: Developments, applications and resources," in *Proc. Congr. Evol. Comput.*, Seoul, South Korea, May 2001, pp. 81–86.
[290] A. P. Engelbrecht, *Computational Intelligence: An Introduction*. Hoboken, NJ, USA: Wiley, 2007.
[291] X. Wang and Y. Wang, "A hybrid model of EMD and PSO-SVR for short-term load forecasting in residential quarters," *Math. Problems Eng.*, vol. 2016, pp. 1–10, Dec. 2016.
[292] Y. Chen, Y. Yang, C. Liu, C. Li, and L. Li, "A hybrid application algorithm based on the support vector machine and artificial intelligence: An example of electric load forecasting," *Appl. Math. Model.*, vol. 39, no. 9, pp. 2617–2632, May 2015.
[293] J.-P. Liu and C.-L. Li, "The short-term power load forecasting based on sperm whale algorithm and wavelet least square support vector machine with DWT-IR for feature selection," *Sustainability*, vol. 9, no. 7, p. 1188, Jul. 2017.
[294] W. Sun and M. Ye, "Short-term load forecasting based on wavelet transform and least squares support vector machine optimized by fruit fly optimization algorithm," *J. Electr. Comput. Eng.*, vol. 2015, pp. 1–9, Dec. 2015.
[295] W. Sun, "A novel hybrid GA based SVM short term load forecasting model," in *Proc. 2nd Int. Symp. Knowl. Acquisition Modeling*, Nov. 2009, pp. 227–229.
[296] A. Moradzadeh, S. Zakeri, M. Shoaran, B. Mohammadi-Ivatloo, and F. Mohammadi, "Short-term load forecasting of microgrid via hybrid support vector regression and long Short-term memory algorithms," *Sustainability*, vol. 12, no. 17, p. 7076, Aug. 2020.
[297] 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, Nov. 2012, doi: 10.3390/en5114430.
[298] G. Cao and L. Wu, "Support vector regression with fruit fly optimization algorithm for seasonal electricity consumption forecasting," *Energy*, vol. 115, pp. 734–745, Nov. 2016.
[299] Y. Liu, M. C. Roberts, and R. Sioshansi, "A vector autoregression weather model for electricity supply and demand modeling," *J. Mod. Power Syst. Clean Energy*, vol. 6, no. 4, pp. 763–776, Jul. 2018.
[300] J. W. Messner and P. Pinson, "Online adaptive lasso estimation in vector autoregressive models for high dimensional wind power forecasting," *Int. J. Forecasting*, vol. 35, no. 4, pp. 1485–1498, Oct. 2019.
[301] D. S. Kumar, G. M. Yagli, M. Kashyap, and D. Srinivasan, "Solar irradiance resource and forecasting: A comprehensive review," *IET Renew. Power Gener.*, vol. 14, no. 10, pp. 1641–1656, Jul. 2020.
[302] S. Fathi, R. Srinivasan, A. Fenner, and S. Fathi, "Machine learning applications in urban building energy performance forecasting: A systematic review," *Renew. Sustain. Energy Rev.*, vol. 133, Nov. 2020, Art. no. 110287.
[303] S. D. Percy, M. Aldeen, and A. Berry, "Residential demand forecasting with solar-battery systems: A survey-less approach," *IEEE Trans. Sustain. Energy*, vol. 9, no. 4, pp. 1499–1507, Oct. 2018.
[304] H.-T. Yang, C.-M. Huang, Y.-C. Huang, and Y.-S. Pai, "A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output," *IEEE Trans. Sustain. Energy*, vol. 5, no. 3, pp. 917–926, Jul. 2014.
[305] H. Jiang, Y. Zhang, E. Muljadi, J. J. Zhang, and D. W. Gao, "A short-term and high-resolution distribution system load forecasting approach using support vector regression with hybrid parameters optimization," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3341–3350, Jul. 2018.
[306] Y. Ren, P. N. Suganthan, and N. Srikanth, "A novel empirical mode decomposition with support vector regression for wind speed forecasting," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 8, pp. 1793–1798, Aug. 2016.
[307] R. Saeedi, S. K. Sadanandan, A. K. Srivastava, K. L. Davies, and A. H. Gebremedhin, "An adaptive machine learning framework for behind-the-meter load/PV disaggregation," *IEEE Trans. Ind. Informat.*, vol. 17, no. 10, pp. 7060–7069, Oct. 2021.
[308] 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. 59754–59765, 2021.
[309] H. Jahangir, H. Tayarani, S. S. Gougheri, M. A. Golkar, A. Ahmadian, and A. Elkamel, "Deep learning-based forecasting approach in smart grids with microclustering and bidirectional LSTM network," *IEEE Trans. Ind. Electron.*, vol. 68, no. 9, pp. 8298–8309, Sep. 2021.
[310] H. Wang, Z. Lei, X. Zhang, B. Zhou, and J. Peng, "A review of deep learning for renewable energy forecasting," *Energy Convers. Manage.*, vol. 198, Oct. 2019, Art. no. 111799.
[311] C. Guan, P. B. Luh, L. D. Michel, Y. Wang, and P. B. Friedland, "Very short-term load forecasting: Wavelet neural networks with data pre-filtering," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 30–41, Feb. 2013.
[312] Y. Wang, Q. Chen, D. Gan, J. Yang, D. S. Kirschen, and C. Kang, "Deep learning-based socio-demographic information identification from smart meter data," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2593–2602, May 2019.
[313] D. Zhang, X. Han, and C. Deng, "Review on the research and practice of deep learning and reinforcement learning in smart grids," *CSEE J. Power Energy Syst.*, vol. 4, no. 3, pp. 362–370, Sep. 2018.
[314] K. Bhaskar and S. N. Singh, "AWNN-assisted wind power forecasting using feed-forward neural network," *IEEE Trans. Sustain. Energy*, vol. 3, no. 2, pp. 306–315, Apr. 2012.
[315] N. M. Pindoriya, S. N. Singh, and S. K. Singh, "An adaptive wavelet neural network-based energy price forecasting in electricity markets," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1423–1432, Aug. 2008.
[316] S. Ben Taieb, R. Huser, R. Hyndman, and M. G. Genton, "Forecasting uncertainty in electricity smart meter data by boosting additive quantile regression," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2448–2455, Sep. 2016.
[317] W. Zhang, H. Quan, and D. Srinivasan, "An improved quantile regression neural network for probabilistic load forecasting," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4425–4434, Jul. 2019.
[318] M. Torabi, S. Hashemi, M. R. Saybani, S. Shamshirband, and A. Mosavi, "A hybrid clustering and classification technique for forecasting short-term energy consumption," *Environ. Prog. Sustain. Energy*, vol. 38, no. 1, pp. 66–76, Jan. 2019.
[319] H. Liu, H.-Q. Tian, D.-F. Pan, and Y.-F. Li, "Forecasting models for wind speed using wavelet, wavelet packet, time series and artificial neural networks," *Appl. Energy*, vol. 107, pp. 191–208, Jul. 2013.
[320] O. Triebe, N. Laptev, and R. Rajagopal, "AR-Net: A simple auto-regressive neural network for time-series," 2019, *arXiv:1911.12436*.
[321] Y.-K. Wu, C.-R. Chen, and H. Abdul Rahman, "A novel hybrid model for short-term forecasting in PV power generation," *Int. J. Photoenergy*, vol. 2014, pp. 1–9, Jun. 2014.
[322] Y. Zhang, Y. Xu, and Z. Y. Dong, "Robust ensemble data analytics for incomplete PMU measurements-based power system stability assessment," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1124–1126, Jan. 2018.
[323] R. Zhang, Z. Y. Dong, Y. Xu, K. Meng, and K. P. Wong, "Short-term load forecasting of Australian national electricity market by an ensemble model of extreme learning machine," *IET Gener., Transmiss. Distrib.*, vol. 7, no. 6, pp. 391–397, Apr. 2013.

[324] E. Cadenas, W. Rivera, R. Campos-Amezcua, and C. Heard, "Wind speed prediction using a univariate ARIMA model and a multivariate NARX model," *Energies*, vol. 9, no. 2, p. 109, Feb. 2016.
[325] M. A. Mohandes, T. O. Halawani, S. Rehman, and A. A. Hussain, "Support vector machines for wind speed prediction," *Renew. Energy*, vol. 29, no. 6, pp. 939–947, May 2004.
[326] D. Li, F. Jiang, M. Chen, and T. Qian, "Multi-step-ahead wind speed forecasting based on a hybrid decomposition method and temporal convolutional networks," *Energy*, vol. 238, Jan. 2022, Art. no. 121981.
[327] L. Donadio, J. Fang, and F. Porté-Agel, "Numerical weather prediction and artificial neural network coupling for wind energy forecast," *Energies*, vol. 14, no. 2, p. 338, Jan. 2021.
[328] Y.-Y. Hong and C. L. P. P. Rioflorido, "A hybrid deep learning-based neural network for 24-h ahead wind power forecasting," *Appl. Energy*, vol. 250, pp. 530–539, Sep. 2019.
[329] M. Mana, D. Astolfi, F. Castellani, and C. Meißner, "Day-ahead wind power forecast through high-resolution mesoscale model: Local computational fluid dynamics versus artificial neural network downscaling," *J. Sol. Energy Eng.*, vol. 142, no. 3, Jun. 2020.
[330] C. Emeksiz and M. Tan, "Multi-step wind speed forecasting and Hurst analysis using novel hybrid secondary decomposition approach," *Energy*, vol. 238, Jan. 2022, Art. no. 121764.
[331] P. Jiang, Z. Liu, X. Niu, and L. Zhang, "A combined forecasting system based on statistical method, artificial neural networks, and deep learning methods for short-term wind speed forecasting," *Energy*, vol. 217, Feb. 2021, Art. no. 119361.
[332] C. E. Velasquez, M. Zocatelli, F. B. G. L. Estanislau, and V. F. Castro, "Analysis of time series models for Brazilian electricity demand forecasting," *Energy*, vol. 247, May 2022, Art. no. 123483.
[333] R. Kumar, P. Kumar, and Y. Kumar, "Multi-step time series analysis and forecasting strategy using ARIMA and evolutionary algorithms," *Int. J. Inf. Technol.*, vol. 14, no. 1, pp. 359–373, Feb. 2022.
[334] P. A. Adedeji, S. Akinlabi, N. Madushele, and O. O. Olatunji, "Hybrid adaptive neuro-fuzzy inference system (ANFIS) for a multi-campus university energy consumption forecast," *Int. J. Ambient Energy*, vol. 43, no. 1, pp. 1685–1694, Dec. 2022.
[335] H. H. H. Aly, "A hybrid optimized model of adaptive neuro-fuzzy inference system, recurrent Kalman filter and neuro-wavelet for wind power forecasting driven by DFIG," *Energy*, vol. 239, Jan. 2022, Art. no. 122367.
[336] S. Dang, L. Peng, J. Zhao, J. Li, and Z. Kong, "A quantile regression random forest-based short-term load probabilistic forecasting method," *Energies*, vol. 15, no. 2, p. 663, Jan. 2022.
[337] C. Zhang, T. Peng, and M. S. Nazir, "A novel hybrid approach based on variational heteroscedastic Gaussian process regression for multi-step ahead wind speed forecasting," *Int. J. Electr. Power Energy Syst.*, vol. 136, Mar. 2022, Art. no. 107717.
[338] M. A. A. Al-Qaness, A. A. Ewees, H. Fan, L. Abualigah, and M. A. Elaziz, "Boosted ANFIS model using augmented marine predator algorithm with mutation operators for wind power forecasting," *Appl. Energy*, vol. 314, May 2022, Art. no. 118851.
[339] D. H. Nguyen, X. H. Le, D. T. Anh, S.-H. Kim, and D.-H. Bae, "Hourly streamflow forecasting using a Bayesian additive regression tree model hybridized with a genetic algorithm," *J. Hydrol.*, vol. 606, Mar. 2022, Art. no. 127445.
[340] R. Weron, *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Hoboken, NJ, USA: Wiley, 2006.
[341] B. U. Islam, "Comparison of conventional and modern load forecasting techniques based on artificial intelligence and expert systems," *Int. J. Comput. Sci. Issues*, vol. 8, no. 5, p. 504, 2011.
[342] A. P. Douglas, A. M. Breipohl, F. N. Lee, and R. Adapa, "Risk due to load forecast uncertainty in short term power system planning," *IEEE Trans. Power Syst.*, vol. 13, no. 4, pp. 1493–1499, Nov. 1998.
[343] M. Mitchell, *An Introduction to Genetic Algorithms: A Bradford Book*, 7th ed. Cambridge, MA, USA: MIT Press, 2001, p. 209.
[344] R. L. Haupt and S. E. Haupt, *Practical Genetic Algorithms*, 2nd ed. Hoboken, NJ, USA: Wiley, 2004.
[345] J. Carr, "An introduction to genetic algorithms," *Senior Project*, vol. 1, no. 40, p. 7, 2014.
[346] C.-C. Hsu, C.-H. Wu, S.-C. Chen, and K.-L. Peng, "Dynamically optimizing parameters in support vector regression: An application of electricity load forecasting," in presented at the 39th Annu. Hawaii Int. Conf. Syst. Sci. (HICSS), 2006.
[347] D. Beasley, D. R. Bull, and R. R. Martin, "An overview of genetic algorithms: Part 1, fundamentals," *Univ. Comput.*, vol. 15, no. 2, pp. 56–69, 1993.
[348] Y. K. Al-Douri, H. Al-Chalabi, and J. Lundberg, "Time series forecasting using genetic algorithm," in *Proc. 12th Int. Conf. Adv. Eng. Comput. Appl. Sci.*, 2018, pp. 1–12.
[349] R. R. B. de Aquino, O. N. Neto, M. M. S. Lira, A. A. Ferreira, and K. F. Santos, "Using genetic algorithm to develop a neural-network-based load forecasting," in *Artificial Neural Networks—ICANN*. Berlin, Germany: Springer, 2007, pp. 738–747.
[350] A. Gupta and P. K. Sarangi, "Electrical load forecasting using genetic algorithm based back-propagation method," *ARPN J. Eng. Appl. Sci.*, vol. 7, no. 8, pp. 1017–1020, 2012.
[351] M. N. Fekri, H. Patel, K. Grolinger, and V. Sharma, "Deep learning for load forecasting with smart meter data: Online adaptive recurrent neural network," *Appl. Energy*, vol. 282, Jan. 2021, Art. no. 116177.
[352] D. Obst, J. de Vilmarest, and Y. Goude, "Adaptive methods for short-term electricity load forecasting during COVID-19 lockdown in France," *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4754–4763, Sep. 2021.
[353] X. Wu, Y. Wang, Y. Bai, Z. Zhu, and A. Xia, "Online short-term load forecasting methods using hybrids of single multiplicative neuron model, particle swarm optimization variants and nonlinear filters," *Energy Rep.*, vol. 7, pp. 683–692, Nov. 2021.
[354] *DLR. Trans-Mediterranean Interconnection for Concentrating Solar Power*, Study Commissioned German Ministry Environ., German Aerosp. Center (DLR), Berlin, Germany, 2006.
[355] D. Michael, "Clean power from deserts," Club Rome, DESERTEC, Hamburg, Germany, White Paper, 2007, pp. 263–275.
[356] S. Erdle, "The DESERTEC initiative: Powering the development perspectives of Southern Mediterranean countries?" Discuss. Paper 12/2010, 2010.
[357] RWE. (2007). *Siemens und RWE Energy Entwickeln Gemeinsam Virtuelles Kraftwerk*. Accessed: Jun. 13, 2008. [Online]. Available: https://www.rwe.com/generator.aspx/presse/languagede/id76858?pmid4001753
[358] A. N. Andersen and H. Lund, "New CHP partnerships offering balancing of fluctuating renewable electricity productions," *J. Cleaner Prod.*, vol. 15, no. 3, pp. 288–293, Jan. 2007.
[359] G. Czisch, "Szenarien zur zukünftigen Stromversorgung. Kostenoptimierte Variationen zur Versorgung Europas und seiner Nachbarn mit Strom aus erneuerbaren Energien," Ph.D. dissertation, Universität Kassel, Kassel, Germany, 2006.
[360] A. A. Mamun, M. Sohel, N. Mohammad, M. S. H. 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. 134911–134939, 2020, doi: 10.1109/ACCESS.2020.3010702.
[361] S. K. Jha, J. Bilalovic, A. Jha, N. Patel, and H. Zhang, "Renewable energy: Present research and future scope of artificial intelligence," *Renew. Sustain. Energy Rev.*, vol. 77, pp. 297–317, Sep. 2017.
[362] I. Abubakar, S. N. Khalid, M. W. Mustafa, H. Shareef, and M. Mustapha, "Application of load monitoring in appliances' energy management—A review," *Renew. Sustain. Energy Rev.*, vol. 67, pp. 235–245, Jan. 2017.
[363] K. Amasyali and N. M. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 1192–1205, Jan. 2018.
[364] Y. Wei, X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, M. Han, and X. Zhao, "A review of data-driven approaches for prediction and classification of building energy consumption," *Renew. Sustain. Energy Rev.*, vol. 82, pp. 1027–1047, Feb. 2018.
[365] A. P. Marugán, F. P. G. Márquez, J. M. P. Perez, and D. Ruiz-Hernández, "A survey of artificial neural network in wind energy systems," *Appl. Energy*, vol. 228, pp. 1822–1836, Oct. 2018.
[366] S. Shamshirband, T. Rabczuk, and K.-W. Chau, "A survey of deep learning techniques: Application in wind and solar energy resources," *IEEE Access*, vol. 7, pp. 164650–164666, 2019.
[367] M. Aslam, J.-M. Lee, H.-S. Kim, S.-J. Lee, and S. Hong, "Deep learning models for long-term solar radiation forecasting considering microgrid installation: A comparative study," *Energies*, vol. 13, no. 1, p. 147, Dec. 2019.
[368] S. Aslam, H. Herodotou, S. M. Mohsin, N. Javaid, N. Ashraf, and S. Aslam, "A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids," *Renew. Sustain. Energy Rev.*, vol. 144, Jul. 2021, Art. no. 110992.

NAQASH AHMAD (Member, IEEE) received the B.S. degree in electrical engineering with the specialization in power from The University of Faisalabad, Faisalabad, in 2018. He is currently pursuing the master's degree in electrical engineering with FAST National University, Chiniot Campus.



YAZEED GHADI (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from The University of Queensland. His dissertation on developing novel hybrid plasmonicphotonic on-chip biochemical sensors received the Sigma Xi best Ph.D. thesis award. He is currently an Assistant Professor in software engineering at Al Ain University. He was a Postdoctoral Researcher with The University of Queensland, before joining Al Ain University. He has published more than 25 peer-reviewed journals and conference papers and he holds three pending patents. His current research interests include developing novel electro-acousto-optic neural interfaces for large-scale highresolution electrophysiology and distributed optogenetic stimulation. He was a recipient of a number of awards.

MUHAMMAD ADNAN (Member, IEEE) received the B.S. degree in electrical engineering from the National University of Computer and Emerging Sciences, Peshawar, Pakistan, in 2013, the M.S. degree in electrical engineering from the COMSATS Institute of Information and Technology, Islamabad, Pakistan, in 2015, and the Ph.D. degree in electrical engineering from the National University of Computer and Emerging Sciences. He worked as a Research Fellow with the Department of Electrical Power Engineering, National University of Computer and Emerging Sciences, from January 2017 to December 2019. Currently, he is working as an Assistant Professor with the Department of Electrical Engineering, FAST NUCES, CFD Campus. His research interests include energy management systems, load flow balancing, load forecasting, power systems dynamic analysis, protection, stability, and intelligent control in renewable energy resources using a fuzzy controller and unified power flow controller.

MANSOOR ALI (Member, IEEE) received the B.S. degree in electrical engineering from the National University of Computer and Emerging Sciences, Peshawar, Pakistan, in 2013, the M.S. degree in electrical engineering from the CECOS University of IT and Emerging Sciences, Peshawar, in 2016, and the Ph.D. degree in electrical engineering from the National University of Computer and Emerging Sciences. Since 2021, he has been working as a Postdoctoral Research Fellow with the Electrical Engineering Department, Ecole de Technologie Superieure (ETS), Montreal, Canada. His research interests include the load forecasting in a power system networks, fuzzy control, and power flow control under various disturbances.