

Received September 3, 2021, accepted October 12, 2021, date of publication October 15, 2021, date of current version October 26, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3120731

Electrical Load Forecasting Models for Different Generation Modalities: A Review

ABDUL AZEEM^{ID1}, IDRIS ISMAIL¹, SYED MUSLIM JAMEEL^{ID2}, AND V. R. HARINDRAN^{ID3}

¹Electrical and Electronics Engineering Department, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Malaysia

²Postdoctoral Scientist at Structure Laboratory, School of Engineering, National University of Ireland Galway (NUIG), H91 TK33 Galway, Ireland

³Engineering and Maintenance Building, Pengerang Refining Company Sdn. Bhd., Pengerang 81600, Malaysia

Corresponding authors: Abdul Azeem (abdulazeem14974@gmail.com) and Idris Ismail (idrisim@utp.edu.my)

This work was supported by the Universiti Teknologi PETRONAS (UTP) and Yayasan Universiti Teknologi Petronas (YUTP) through the Grant Cost Center under Grant YUTP 015LC0-360.

ABSTRACT The intelligent management of power in electrical utilities depends on the high significance of load forecasting models. Since the industries are digitalized, power generation is supported by a variety of resources. Therefore, the forecasting accuracy of different models varies. The power utilities with different generation modalities (DGM) experience complexities and a noticeable amount of error in predicting future electrical consumption. To effectively manage the power flow with negligible power interruptions, a utility must utilize the forecasting tools to predict the future electricity demand with minimum error. Since the current literature supports individual and limited power sources involved in generation for load forecasting, thus the utilities with multiple power sources or DGM remain unexplored. Therefore, exploration of existing literature is required relating to analyzing the existing models which could be considered in load forecasting for DGM. This paper explores state-of-art methods recently utilized for electrical load forecasting highlighting the common practices, recent advances, and exposure of areas available for improvement. The review investigates the methods, parameters, and respective sectors considered for load forecasting. It performs in-depth analysis and discusses the strengths, weaknesses, and error percentages of models. It also highlights the peculiarities of methods used in residential, commercial, industrial, grid, and off-grid sectors aiming to help the researchers to appraise the common practices. Moreover, trends and research gaps are also discussed.

INDEX TERMS Electrical generation modality, grid data-analytics, isolated-grid load forecasting, load forecasting, machine learning, power plant load forecasting, review, smart grid, survey.

I. INTRODUCTION

The uninterrupted and stable supply of power is the fundamental requirement of an electrical power generation system. Such systems generate the necessity of meeting the load demand; by analyzing the history, behavior, trend, and factors affecting the load and predicting the future demand of load, termed as load forecasting (L.F). It is considered crucial to determine the supply and demand gaps of electrical power generation. It provides benefits to the management to understand the different variables influencing energy consumption, augmentation, and waste. Also, it is equally crucial for power control and exchange in different interconnected power systems since it plays a fundamental part in overlaying the disparity and concerns for energy demand [1].

The associate editor coordinating the review of this manuscript and approving it for publication was Bilal Alatas^{ID}.

The recent developments in electrical system by integrating several power resources (including renewables) for generation of electricity has made load forecasting more complex. This significant integration of renewable energy (solar, wind, thermal, etc.) has advanced the load forecasting phenomenon towards a more challenging phase. It has created tough environment for power professionals to interpret the electrical generation, supply, reserve, and manage the power demand [2]. The occurrence of this complexity is due to non-linearity and non-stationary load behavior; therefore, it involves detailed analysis and consideration of several factors affecting the forecasting procedure, directly or indirectly. Alongside the benefits of forecasting techniques, challenges of prediction accuracy still exist.

Due to rapid change in electrical load consumption patterns, it is important for forecasting models to learn and interpret the load consumption from historical data concerning the

data changes with respect to the associated factors. It helps to prepare the data for the persistence of training, identifying the error about required changes, and predict verification in the model. However, the model validation for several applications depends on the circumstances or limitations defined whilst the model development. The data validation has several objectives that include but are not limited to data loss, false data, data with error, and the number of more anomalies. Therefore, the existing models solely depend upon case-to-case replacement or modifications due to one or more than the above-defined factors.

The load consumption profoundly depends on meteorological variables, making the hourly, daily, weekly, and monthly forecasting more challenging. The accuracy of the model is considered critical, even the conditions verily unstable. A forecasting method with great accuracy and precision can help in cost-saving for the utilities. Similarly, a model lacking in such properties will result in a loss that could be either relating to finance, infrastructure, operation, or maintenance [3].

A load prediction with 1% error (lack of precision or accuracy) could lead to an operating cost of millions of dollars per year [4]. To counter the issue, numerous electrical load forecasting techniques were developed by scientists and researchers. However, the studies reveal that the existing forecasting methods carry a certain degree of error and inaccuracy, which differs from model to model, location to location, application to application, and region to region.

These variations are either due to the consideration of different variables of meteorological parameters, applied applications, different load behaviors, or model limitations. In literature, the model prediction is examined by its percentage, usually following the accuracy criteria of mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) defined in detail in section VII.

Load forecasting has a rigid and dependent relationship with meteorological parameters such as irradiance, rain, temperature, humidity, and wind speed, which should be considered with respect to the geographical region while developing any forecast model. The in-depth study of the electrical load has revealed that it follows patterns that are not always cyclic and have seasonal effects with a few uncertainties [5]. Therefore, the conventional approaches used in regular load forecasting models cannot handle the uncertainties experienced in a power utility with different generation modalities (DGM) or grid-isolated generations, due to their model limitations.

Since the power utilities and industries are digitalizing their operations and use of multi-source power generation has also increased. This also includes the contribution of residential and commercial sectors by means of renewables for eco-friendly environment. Therefore, the performance of existing load forecast models does not meet the model criteria of DGM. The DGM includes combination of two or more of these power sources such as hydro, gas, wind, solar, coal, and nuclear. Since the existing models have their limitations.

Therefore, this paper focuses on models that could be considered for tailoring a solution for DGM in short-term load forecasting (STLF) for power utilities in residential, commercial, industrial, grid and off-grid sectors. It also presents recent advancements and limitations of L.F models and introduces the existing research gaps for future research directions.

A. OTHER REVIEW

Several other reviews have been presented on electrical load forecasting models that emphasizes on single sector with different criteria and target different approaches and methods to conduct the literature review. However, they lack to include the review on industrial and grid-isolated sectors. Also, to highlight the models relating to develop DGM L.F-model. Since this addition is found rare, therefore we expect the review to add knowledge for researchers working on power sectors with different generation modalities. A comparative analysis of these reviews has been presented in Table 1.

The exploration domain of this paper is kept broader and includes the models used in multiple sectors i.e., residential, commercial, industrial, grid-connected, and grid-isolated load forecasting. The proposed review discusses diverse methods and techniques used for load forecasting by numerous researchers and present relevant technical recommendations for improvement. In this review, we consider discussing the methods that are most suitable to be tailored for multiple applications of load forecasting, which is currently needed other than the conventional approaches, particularly targeting the DGM; based on their strengths and weaknesses, which differ from model to model.

Organization: This study is managed so that the types of load forecasting with importance of dealing in STLF is discussed in section II. Importance of load forecasting model for DGM in power utilities including the benefits, affecting factors, and confrontations in achieving the desired model in conventional and DGM models are presented in section III. Section IV presents the clustering and optimization techniques. Moreover, section V introduces the recent advances in STLF with detailed discussion on residential, commercial, industrial, grid-connected, and grid-isolated load forecasting models. Furthermore, section VI examines an in-depth analysis of some load forecasting models and presents technical recommendations based on the studied algorithms. Section VII, VIII, and IX discuss the performance measurement criteria, outcomes, and conclusion, respectively.

II. TYPES OF ELECTRICAL LOAD FORECASTING

The electrical load forecasting has been distributed in four categories, concerning the forecasting time domain. Long-term load forecasting (LTLF), medium-term load forecasting (MTLF), short-term load forecasting (STLF), and very short-term load forecasting (VSTLF) [6], also shown in Fig. 1 with time stamp distribution. Several forecasting methods have been employed for these time domains, including knowledge base expert system and some statistical techniques, artificial

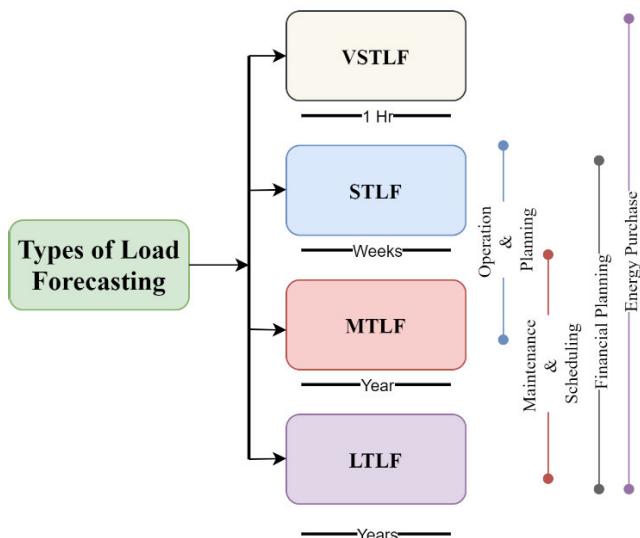


FIGURE 1. Electrical load forecasting categories concerning different time zones and defining respective operational domains.

intelligence (AI) techniques, including machine learning models, hybrid techniques and deep learning models.

A brief overview of the different types of load forecasting is presented in this section. Later, we provide a detailed literature review on the existing STLF methods, which could be considered to adapt further in load forecasting for different generation modalities.

A. LONG-TERM LOAD FORECASTING

LTLF targets the forecasting horizon of electrical demand for more than one year, including planning demand for several power systems. It provides an understanding of the system and adequacy for reliable operations of power supply and meeting future demands. The power system's infrastructure modifications and upgrades performed by power utilities can be adequately forecasted with the help of LTLF. It has significant effects on future-grid reliability, customer-utility satisfaction, investments, and utility companies [11].

Several forecasting techniques used in LTLF include multiple linear regression (MLR) to observe the relationship of load and external variables using the hourly informative approach to improve predictive modeling, perform analysis on different scenarios, and normalize the weather data [12].

Specific methods have relied upon sub-load profiling through individual or cluster aggregation, which helps short-term load forecasting but is unreliable for LTLF due to external parameters driving load variations. A study performed on Australia's electricity market discussed the changes in population, technology, weather and economic conditions, and individual electricity consumption. The load relationship with factors affecting load forecast was analyzed using semi-parametric models, also bootstrapping with the variable block method was utilized. The study predicted the annual and weekly demand peak for South Australia [13].

Reference [14] employs decomposition and bootstrap aggregation methods and auto-regressive moving average (ARIMA) for monthly electricity demand in the univariate forecast. Reference [15] has employed particle swarm optimization (PSO) for minimizing the error of the associated parameter while forecasting annual peak load utilizing the data from Egypt and Kuwait networks. Reference [16] forecasted the electricity consumption of China per capita utilizing an expert prediction and fuzzy Bayesian model to resolve the functional LTLF problems.

B. MEDIUM-TERM LOAD FORECASTING

MTLF deals in time horizons ranging from a month to a year, which usually covers the planning for the maintenance of the grid, consideration of prices for electricity and organization for energy sharing, and fuel scheduling. Several models have been employed for MTLF tasks. Reference [17] considered to eliminate irrelevant features from pre-processing data model comprising of feature selection, based on informative knowledge using Jaya (optimization algorithm) along with CRBM (conditional restricted Boltzmann machine) to forecast month electrical load for smart homes. Reference [18] presents a comprehensive survey regarding different employed techniques at Nigerian electric utility, and three regression techniques, linear, compound growth, and quadratic, were employed independently, helping the Ibadan electricity distribution company to make efficient energy management planning and management steps.

Reference [19] and [20] discuss some more models for MTLF with different time range utilizing time-dependent convolutional neural network (TD-CNN) and cycle-based long short-term memory (C-LSTM) for MTLF improvement and variants of models of machine learning predicting the district level forecast employing ANN, linear regression (LR) and adaptive boosting model classified into monthly, seasonal, and yearly forecast, respectively.

C. SHORT-TERM LOAD FORECASTING

STLF which deals in time zone ranging from a few minutes, hours, or days is a significant factor of day-to-day operation and planning of a power utility and critical component of an energy management system. The STLF approach is efficient in reducing financial costs and operational risks, impacting savings directly. Therefore, it is given much prominence and treated as a critical problem in the competitive energy market. Several STLF techniques have been proposed, some of which have been classified as statistical methods, including LR, ARIMA, exponential smoothing (ES) [21]. Later in STLF, scientists introduced specific new methods, including ANN, FL, SVM, recurrent neural network (RNN), and LSTM [22].

One of the major driving factors of a power utility is its operation and planning which depends on STLF. Since STLF ranges from minutes to days. Therefore, it is considered to be a crucial factor in decision making of a power utility for load control, dispatch, and secure operations of a power system. The digitalization of power utilities has commended towards

TABLE 1. List of related literature reviews covering different aspects of load forecasting with different review types and targets.

Reference	Year	Review Type	Review Target	Remarks
[7]	2016	Tutorial	Probabilistic approaches in load forecasting	The study presents the tutorial review on probabilistic load forecasting. It discusses the forecasting techniques, methodologies, and provided a research direction in probabilistic load forecasting.
[8]	2017	Systematic	Identification of scenario-based model for load forecasting	The study presents a systematic review of load forecasting methods. It focused on identifying the particular scenarios for models suited for their application. It also presents a taxonomy for model selection in accordance with their problem.
[3]	2017	Comparative	Architecture of RNNs for real-valued time series	The study performed a comparative study on STLF using different RNN networks and presented their important architectures and properties.
[9]	2019	Methodological	Computational intelligence based STLF	The study provides a framework for different computational intelligence methodologies considering several case studies as an example. It also presents an overview of difficulties experienced in STLF with possible strategies to tackle them.
[10]	2020	Comprehensive	Single and hybrid methods based on ML	The study covered the single, hybrid, and combined methods of load forecasting in STLF. However, in hybrid models prime target was SVM and ANN. It also presents the advantages, disadvantages, and functions of these models.
Proposed	2021	Narrative	ML approaches used for STLF for conventional and different generation modalities	The study presents a combined review differentiating in residential, commercial, industrial, grid and off-grid sectors. It highlights the limitations of existing load forecasting methods and discuss possible strategies to overcome these issues. It presents the possibilities to tailor existing load forecasting model in accordance with the requirement of different generation modalities. Also, it highlights the possible research gaps to work as future-work recommendations.

multiple power resources. Thus, the power utilities dealing in different generation modalities need to have an efficient and adaptable model for load forecasting. Since DGM involves multiple power resources, therefore STLF plays a critical role to direct the operations for shutting down or starting up the resources. Additionally, it will help the management to efficiently manage the distribution of power among DGM. Section V discusses various STLF methods utilized in their respective applications in detail.

D. VERY SHORT-TERM LOAD FORECASTING

VSTLF is famous for dealing load forecasts in a time frame ranging from minutes to an hour. As the VSTLF deals in near future forecasting, certain load forecasting factors consideration is conditional, but mostly time and temperature are considered, with current values exploration to forecast future values. However, the forecasting in VSTLF is primarily dependent on the recent load pattern compared to the rest (STLF, MTLF, and LTLF), which depends on multivariable relationships. Some of the methods that include VSTLF include classical-statistical and computational intelligence techniques, out of which some were exploited for forecasting in district buildings [23].

III. DIFFERENT GENERATION MODALITIES

The accumulative generation of electrical energy through a combination of two or more power resources from

conventional and renewable energy sources is termed as different generation modalities. In the DGM environment, the generation of an individual source is managed through load demand forecast produced in collective efforts of these modalities. They do have individual proportions to contribute and keep enough reserve to support in any event of equipment failure. The conventional energy sources include coal, nuclear, hydro, oil, and gas. However, renewable energy sources include thermal, wind, or solar energy, etc. The DGM has multiple applications such as smart grid (SG), power utilities, and independent power producers (IPPs). The DGM could be the combination of any two or more of the above-mentioned power resources for energy generation (For e.g., coal, gas, solar, and wind).

Since the amount of energy consumed by a person has increased, therefore it has enormously increased the energy demand affecting the environment along with the problem of energy generation and shortage [24]. Such kinds of issues highlighted the importance of DGMs in SG, which has the capability of managing the energy more efficiently [25]. The inclusion of multiple power resources strengthens the power system sustainability, improves system demand, and helps in cost reduction [26]. In conjunction with benefits, it also raises challenges for power professionals, as it increases the system complexity. Energy consumption does get affected by several factors, similarly, generation is also affected [27]. This increases the complexities in the correlational

analysis [28], as multiple factors get involved due to DGM.

In electrical load forecasting for conventional and hybrid renewable generations, the models having capabilities to deal with non-linear and multivariate data have outperformed the competition [10]. Most of these models are deep learning models that include ANN, DNN, CNN, and LSTM [5], [9], [10], [28], [36]. The types of modalities forecasted in recent papers along with suitable algorithms have been listed in Table 7. However, there exist different hybridizations of these methods along with different optimization algorithms (discussed in section IV) to improve the prediction accuracy. But these models still need modifications, expansion, and alteration to cope with existing complexities such as adaptivity, the introduction of new parameters, sudden sag or lag, stress on a single generative system, etc. Still, with such an improved performance of these algorithms in the existing state, they are not suitable for the environment like DGMs. Since DGMs have multiple sources and each source has its unique set of parameters to deal with. Therefore, an individual contribution forecast is better to anticipate the generation requirement at this stage in comparison to a combined forecast.

In hybrid power plants, the load demand is anticipated based on previously accumulated load patterns (patterns from a conventional and renewable source) with single point accumulated forecast with relevant parameters, which variate the generation. This variation of power generation from sources is one of the major causes to trigger the distress in electrical systems. Thus, one of the power sources experiences high volumes of load and encounters the surges and put risk to meet the demand [29]. This risk is usually confronted by conventional power sources since renewables have a large range of parameters that affect their power generation.

However, in DGMs the contribution of each power source is forecasted individually based on its previous patterns and relevant parameters which variate the amount of generation. This encompasses the sole contribution of the source and later accumulated in the total forecast, but each forecast is on an individual basis. This leads to a multipoint forecast with a final accumulated forecast. Due to this equity of treatment of power sources, the other power resources are not distressed. In case of power variation from any of DGMs, the requirement is equally distributed on other power sources to effectively manage the demand.

Since the electrical load forecast is crucial for operations and planning therefore if the management is unable to foresee the deviations of contributing resources precisely and accurately than it impacts the energy management [30]. The modern power systems face several challenges including peak demand which is another cause for exhibiting the non-linear, noisy, and distinct pattern existence in load anticipation. This peak demand could be a result of meteorological variables, huge consumer loads, large industrial operations, and an inclined acceptance of electric vehicles (EVs) [31]. These peak demands are more likely to be satisfied through

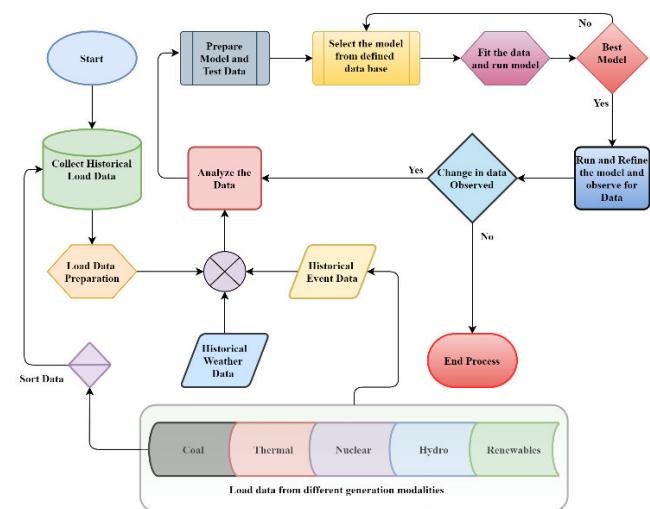


FIGURE 2. Sample flowchart for electrical load forecasting model with respect to DGM comprising of different historical values (Source, load, weather, and event).

conventional means or calculated DES. Table. 2 provides some insight differences on the hybrid power plants and DGMs in the light of certain standards to clarify the concept and brief its vitality in the smart grid environment with respect to the DGM's importance.

A facility's power plant is one of the most crucial departments that facilitates the rest departments to ensure their operations are timely. For this purpose, a felicitous approach of anticipating the future requirement should be computed to prevent any downtime at an economical cost. The load forecast stands as its main deriving factor, which plays a vital role in regulating how and when to expand the power facility. Fig. 2 gives a sample flowchart of the electrical load forecasting model in the DGM environment. Every aspect has its benefits, affecting factors, and some confrontations which should be addressed, accordingly. This section describes these aspects in detail, which could lead to the improvement of conventional and DGM load forecasting models. The limitations of existing models with respect to the DGM environment are presented in section VI.

A. IMPORTANCE OF LOAD FORECASTING

Beside the importance of anticipating the future load demand, load forecasting helps in numerous other sectors as well. It directly helps the organizations relating to power for their required calculations and non-power sector to plan their operations in accordance. It also plays an important role to lay foundation of SG and smart city.

The benefits of load forecasting in dealing with operations, planning, and investment, including market bidding, controllable loads scheduling, generation, fuel purchases, load-switching, and infrastructure maintenance schedule, have highlighted its importance among utilities to contemplate its effective harvesting. Table 3 lists some benefits relating to

TABLE 2. Standard differences between hybrid power plants and DGMs to clarify the concept and brief its vitality in the smart grid environment with respect to the DGM's importance.

Standards	Hybrid Power Plants	DGMs
Type	The hybrid power plants are a combination of generation or storage resources put together (solar and wind) or single conventional power source with single or double renewable energy sources.	The power plants with DGMs have multiple conventional and multiple renewable energy sources.
Load Balancing	The load is usually balanced between single conventional and single or double renewable power resources.	The load is usually balanced between multiple conventional and multiple renewable power resources.
Bad Weather	In bad weather scenarios, the conventional power resource is distressed with more load.	In such scenarios the load is balanced between two or more conventional power resources.
Worst Weather	In such scenarios additional power resource is required or load shedding is performed to keep the system running with maximum output.	In such scenarios no additional power resource is required since the load is balanced with multiple conventional power resources.
Dependency	High dependency on renewable power resource.	Low dependency on single power resource.
Risk	Moderate risk of spikes and load surges on electrical system.	Low risk of load surges on electrical system.
Forecast	Single point forecasting	Multiple point forecasting
Error	There exists moderate error rate due to single point forecast.	Due to multiple point forecasting the error rate is minimum.
Back up	The backup is supported by either batteries or single conventional source.	The backup is supported by multiple power sources.
Source Failure	In such scenario, one of the systems will have great stress. Also, it could cause a cascaded disaster.	In such scenario, the system will not be stressed. Thus, no cascading impact.
Smart Grid	Such utilities face complexities to manage two-way power flow and smart power management.	It effectively manages the two-way power flow, making it a fundamental part of SG's.
Smart Cities	Such utilities cannot be part of smart cities due to their complex behavior in dealing with power failure.	Such utilities are considered essential for implementing smart cities due to their multi-source power generation having minimal impact on consumers in power failure situations.

conventional and DGM load forecasting with the effects on their beneficiaries.

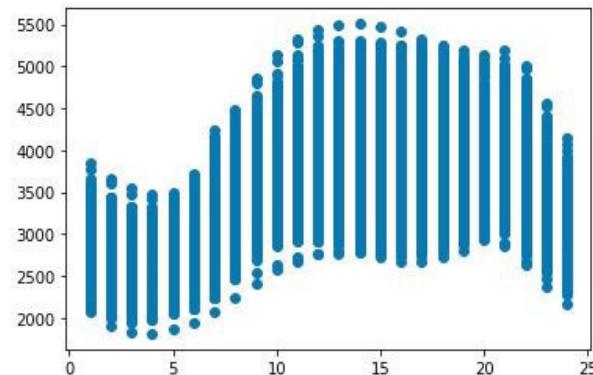
B. IMPACT OF INFLUENTIAL FACTORS

The electrical load has several direct and indirect relationships with certain factors which affect load forecasting. These factors directly contribute towards achieving the prediction's precision and accuracy, if considered wisely while developing the model. Researchers need to explore the most pertinent factors for their model and develop a relevant scenario about the electrical consumption pattern, due to the nature of dependency of electrical load and the forecast model on these factors. A correlation of time and temperature is presented in Fig. 3 and Fig. 4 using the dataset from GEFCOM 2014.

Some of the significant factors having impact on load forecasting models in conventional and DGMs approaches are explained briefly as follows:

1. Time: The relationship of load forecasting with time is beneficial for pattern recognition and depends on the hour to an hour (less or more) period, for which it should be considered during model development. After performing several simulations referring to every hour of the day, the hourly load consumption changes were observed, varied responses for every day of the week [29].

Reference [30] has extensively described the load behavior that varies from hour to days to weeks to

**FIGURE 3.** Load to time correlation depicting the continuous change in load behavior relating to time (X-axis representing hours and Y-axis load in MW's).

weekends and months. It states that the data interpretation depends on past hours following their day and time, contributing to an accurate model. The time frame of activities performed also plays an essential role in determining the future load. A load to time correlation is presented in Fig. 3, which depicts the load changes concerning the time domain.

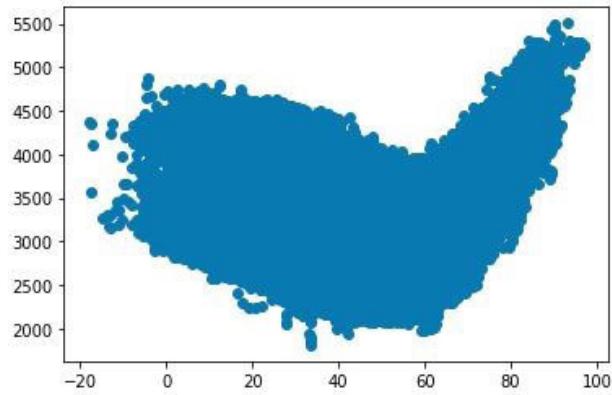
2. Meteorology: One of the most dominant load forecasting factors consists of meteorology, which consists of multiple variables. A correlational data interpretation

TABLE 3. Benefits of electrical load forecasting in conventional and DGMs environment.

S. No	List of Benefits	Direct Beneficiaries						Remarks on Benefits of Load Forecasting
		Government	Environment	Power Utility	Industrial	Commercial	Residential	
1.	Load demand anticipation	✓	-	✓	✓	-	-	It guides the beneficiaries in decisions, anticipating load demand for the utilities to perform future demand planning considering their generation and transmission accordingly, at economical costs by a valid forecast.
2.	Operations and planning	✓	-	✓	✓	-	-	It guides the beneficiaries to conduct their operations and planning in accordance with the accurate predictions to avoid the hassles. It helps to decide about managing the modalities to maintain economical operations.
3.	Infrastructure and maintenance	✓	-	✓	✓	-	-	It guides the beneficiaries to understand the increasing demand and perform related actions timely to avoid the downtime. Also, it helps to recognize the improvement in infrastructure and highlight the scheduling of maintenance required for smooth operations.
4.	Purchase and saving of fuel	-	-	✓	✓	-	-	An accurate forecast can help to pre-bargain and prepare the warehouse in accordance with the required demand of electricity which eventually results in savings in terms of anticipating the requirements earlier.
5.	Resource and reserve calculation	✓	-	✓	✓	-	-	Earlier demand anticipation helps the utilities to manage the spinning and non-spinning reserve of the system for sudden demand requirement.
6.	Over and Under energy preparation	✓	-	✓	✓	-	-	It projects the future demands which helps to effectively manage the uncertainties and production facilities.
7.	Facility development	-	-	✓	✓	-	-	It leads the utilities to observe the trend of consumption to generate a guideline to enhance the generation facility or include a new resource to bridge the forthcoming gap.
8.	Uninterrupted power supply	✓	✓	✓	✓	✓	✓	An accurate forecast can help to intelligently manage the power supply guiding the utility to perform smart operations to avoid power surges, spikes, sags, dips, and outages.
9.	Loss, waste, and theft calculation	✓	-	✓	-	-	-	It provides an understanding to disclose the issues causing system to overload.
10.	Renewable energy contribution	✓	✓	✓	✓	✓	✓	It enables the effective integration of renewable energy after understanding the load behaviors in different time intervals.
11.	IPP's contribution	✓	-	✓	✓	✓	✓	It enables the beneficiaries and independent power producers to understand their generation segment to avoid cascading collapse affect.
12.	Industrial Digitalization	✓	✓	✓	✓	✓	-	It empowers the industries to smartly manage their operations based on intelligent forecasting.

TABLE 3. (Continued.) Benefits of electrical load forecasting in conventional and DGMs environment.

13.	Smart grid implementation	✓	✓	✓	✓	✓	✓	It helps the utilities to understand the consumption pattern and act accordingly to design the power flow to benefit and smartly manage the distributed energy resources.
14.	Smart city implementation	✓	✓	✓	✓	✓	✓	It is considered a founding pillar in implementing the smart city by foreseeing the exponentially increased demand and mapping it with relevant factors to efficiently manage the sources and intelligently produce the power.
15.	Investment and Market bidding	✓	-	✓	-	-	-	It enlightens the beneficiaries to manage the investment and bidding smartly with preprocessed information provided by forecasting.
16.	Reduction in CO ₂ emission	✓	✓	✓	✓	-	-	It enables to integrate renewables in densely to utilize different energy resources at different intervals of time contributing towards reduction in CO ₂ emission.
17.	Preservation of fossils	✓	✓	✓	✓	-	-	It provides the consumption pattern revealing the demand peak and stability hours which enables to utilize respective energy resources which could help to recover the depletion rate of fossils.

**FIGURE 4.** Load to temperature correlation showing the continuous change in load behavior with respect to change in temperature (X-axis representing temperature in °C and Y-axis load in MWs).

of load and temperature is presented in Fig. 4, presenting the changes observed in load consumption concerning the temperature increment or decrement. Temperature, humidity, wind speed, rain, cloud cover, and solar radiation influence electricity consumption. The load consumption could be less, equal, or more significant in the summer and winter seasons depending on the meteorological factors of different regions explained in part 3 (region). However, the increase or decrease in temperature from normal levels causes the increment in load consumption as air-conditioning increases in summer, whereas heating increases in winter. A moderate temperature level at any of these seasons results in less electricity consumption, thus

causes the over or underestimation of load forecast, challenging the forecasters to find an adaptive model for these deviations. Researchers have proposed different variables out of the above mentioned to reach the maximum accuracy, but still lack to achieve the desired results. Reference [31] discusses several studies that include the humidex index, wind chill index, and different weather uncertainties showing the importance of including the meteorological inputs for achieving better forecasts.

- Region: The world comprises of very different climates having remarkable variations, which differ from region to region. The regional consumption of electricity differs as different regions have distinct meteorological constants, affecting their consumption level. Therefore, to achieve a higher degree of forecasting accuracy, regional climatic variables need to be incorporated in the model development. Every region has its weather and climate; some countries cover all four seasons, some have two to three, and some exist with only one climate with minor changes. Due to this issue, the load demand varies, particularly the load demand of tropical countries does not match with the rest. A model designed only for tropical countries would not be valid for other countries until it consists of regional adaptability. Therefore, the region also retains importance as an influential forecasting factor. Fig. 5 highlights the different countries have a different climate, which keeps variating their energy consumption level; the color level indicates the max, medium, or moderate level of temperature in the region [32]. Reference [33] discussed the approach of weather priority indexing

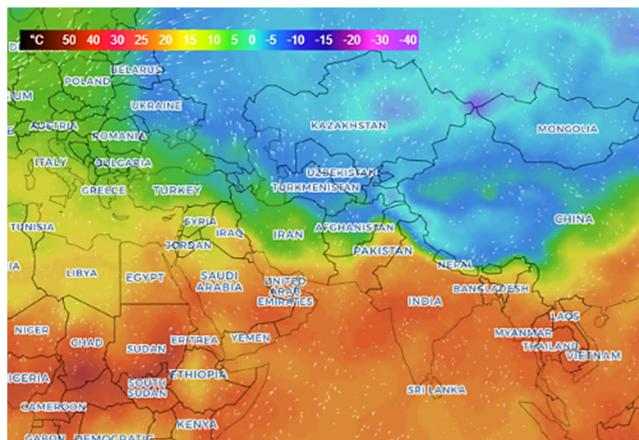


FIGURE 5. Climatic variations representation of different countries (taken from world weather online) showing different climate phases, affecting the energy consumption [13].

and explained the effects of different regions within Iran holding different weather categories throughout the year with sixteen (16) different regional electrical companies (RECs).

- Application: The application factor is quite interesting as it does affect the load forecasting if the designed model was for any other application. If designed for the small grid, it will not be efficient for large scale consumers.

Similarly, if it is designed for a grid isolated network, its efficiency for grid-connected would differ. Therefore, the application also plays a role in model development [7].

- Economic: Electricity has become a basic necessity of life, resulting in an economic impact on its usage. The economic factor does include industrialization, load management, electricity pricing, and consumer behavior. The price of electricity directly relates to buying capability and is directly proportional to the defined tariffs with peak and non-peak hours. The appliances also play their roles in economic factors as developed and modernized countries have more advanced power consumption devices than under-developed countries; their load is also different. Thus, the country's economy plays a role in long-term load forecasting as it advances industrial development [34].

- Events: Since November 2019, the pandemic situations are going to be counted as one of the influential factors in meeting the demand and supply as such situations put load increment in the residential sector, decrement in commercial and load variations concerning country policies can be observed in industrial sectors, as some completely and some partially closed their operations. On the other hand, the week data differ with weekends and other events or particular holiday data. Reference [30] explains the importance of these events for

consideration and discusses their significance in detail by developing daily load profiles.

- Historical information: The historical data for training of models is of much importance since it decides the features to be adapted by the model as it sets the trend of model. Hence, the data should be considered to have a variety in the learning process so that the model can learn at full capabilities to perform the predictions [35].
- Data quality: The quality of data plays an important role in electrical load forecasting. If the data has errors, it will result in bad forecast. Ultimately, it also affects the performance of the model. Therefore, the data quality should be maintained either while using the data as input for model or before collecting. However smart meters do provide sufficient data to use directly [36].
- Technology: The technological changes have advanced towards more sustainable power systems, but it has also contributed to errors of load forecasting as the increased renewables penetration and electric vehicles (EV's) have added more uncertainties and complexities towards load prediction [36].
- Distributed power resources: The contribution of electrical power by commercial and residential sector which is on the other side of the meter increases the errors in load forecasting for which meter level load forecasting should be considered [37].

C. CONFRONTATIONS

To modernize the conventional grid and transform it into a SG, the researchers and scientists have worked in many electrical engineering subjects and developed many state-of-art methods in load forecasting to improve electrical load forecasting efficiency make an economic impact in electrical generation, transmission, and distribution sector. Several methods have been developed on which scientists have kept improvisations under consideration so that the model accuracy and precision could be enhanced, and an adaptable model could be established to deal with different generation modalities. During these improvement and enhancement periods, several confrontations were observed and held responsible for achieving the accurate model.

Out of many confrontations, some are listed as below:

- Lack of consideration of influential meteorological factors, causing an abrupt error in load forecasting.
- Lack of adaptable model considering regional adaptability of an electrical load forecasting model; causing models to behave inaccurately in load forecasting and generating large errors, including economic impact in different regions.
- Consumer behavior analysis, including digital data collection of consumers.
- Power surge, maintenance, loss, waste, and flawed data analysis, along with data validation.
- Extreme penetration of distributed energy resources.

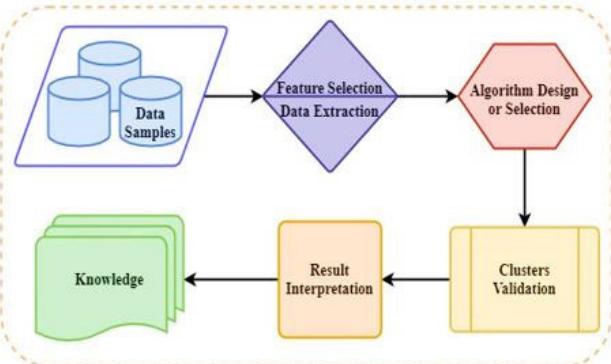


FIGURE 6. Flowchart representation for developing a clustering technique method to analyze the data insights for adequate knowledge extraction.

6. Model limitations to withstand the requirements of DGM.
7. Grid isolated power plants load forecasting.

The above mentioned are some of the significantly observed confrontations that inspired the researchers to continuously improve and develop new models focusing on creative challenges in data acquisition, data clustering, parameter defining, segregation of days based on their data consumption structure, special events, and holidays.

The current pandemic situation has also shuffled all models' bases, as it has converted the traditional working model into work from the home model, which has created a significant impact on power industries as the official load has been distributed to homes, hospitals, and quarantine centers due to high consumption demand at residential, medical and hospitality consumers. Simultaneously, the demand for the rest of the industries has fallen from the calculated point. An adaptable model would outperform all the rest models by transforming the current online data into live feed streaming.

IV. AN OVERVIEW OF CLUSTERING AND OPTIMIZATION TECHNIQUES

The clustering and optimization techniques have widely been used to gain information from data and optimize the parameters to have adequate inputs. Therefore, we have summarized some of the clustering and optimized techniques with their models, features, and specifications used in electrical load forecasting to improve the prediction results and minimize the error expectancy efficiently.

A. CLUSTERING TECHNIQUES

The clustering techniques have been widely used in electrical load forecasting to sort and properly arrange the raw data in the proper format. It is used to perform analysis and remove the unwanted data for forecasting (a sample of data clustering procedure is also highlighted in Fig. 6, which describes the typical working of a data clustering algorithm.), as the issue of clustering is highlighted in several different areas like mining of data, recognition of patterns, statistical analysis of data

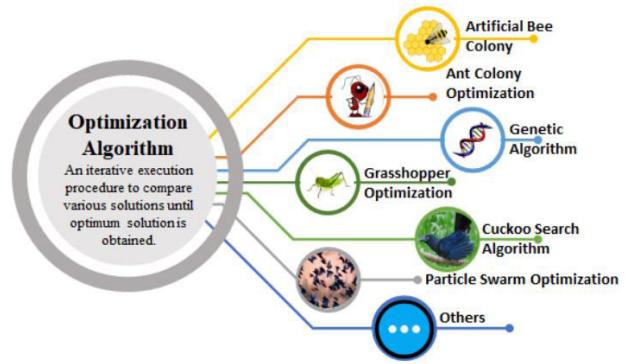


FIGURE 7. Types of Nature-Inspired optimization algorithms used for different load forecasting models for parameters optimization to improve prediction results.

etc. The clustering (supervised or unsupervised) also serves as a tool to analyze the data insights and gain data to detect their characteristics, distribution, outliers, and noise. Data clustering requirements deal with scalability, different nature of data, cluster and attribute discovery, high dimensionality, data with noise, and provide an interpretable, coherent, and functional data format.

Clustering is forming data groups, recognized as clusters, having a similarity between the same cluster elements and high dissimilarity between elements of different clusters [38]. There exist numerous clustering techniques, out of which some are presented in Fig. 8.

The distance-based methods comprise two algorithms: partitioning, in which datasets single clustering is created, and hierarchical, in which dataset's nested clustering sequence is created. They are further breakdown into k-means, k-median and k-medoids, and agglomerative and divisive methods, respectively. Also, some density-based methods which cluster the objects' dense region and highlight the noise include density-based spatial clustering of applications with noise (DBSCAN).

The recently improved DBSCAN by hierarchical clustering converting it into HDBSCAN has also received immense acceptance; DENCLUE and OPTICS are also some of the extended versions of density-based clustering [38]- [39]. Multi-resolution grid data structure with a sub-category of STING, Wave-Cluster, and CLIQUE; the expectation-maximization algorithm for underlying data-points estimation probability and subspace clustering for dimensionality reduction is part of grid-based, probabilistic-generative, and high dimensional clustering methods, respectively.

B. OPTIMIZATION ALGORITHMS

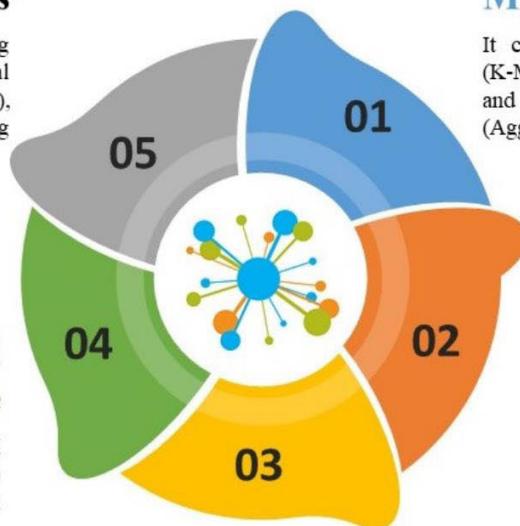
In recent times the optimization algorithms have been widely used, mostly inspired by nature as nature is still unexplored, yet the best mathematical calculations are performed by nature. Some of the nature-inspired optimization algorithms explored by scientists and researchers are shown in Fig. 7, which includes the artificial bee colony (ABC)

High Dimensional Clustering Methods

It comprises of subspace clustering (bottom up, top down, correlational based and delta cluster methods), dimensionality reduction (co-clustering of PLSI and NMF).

Probabilistic and Generative

It estimates model parameters with EM algorithm (Expectation-Maximization) for generative probability of estimation for underlying data-points.



Grid Based Methods

It uses a multi-resolution grid data structure. Further sub-categorizing into STING, WaveCluster and CLIQUE.

Distance Based Methods

It comprises of partitioning algorithms (K-Means, K-Medians and K-Medoids) and hierarchical algorithms (Agglomerative and Divisive methods).

Density Based Methods

It comprises of DBSCAN and recently introduced HDBSCAN algorithms. Also includes OPTICS and DENCLUE.

FIGURE 8. Typical data clustering techniques widely used in electrical load forecasting defining their extents of operational capabilities.

algorithm based on the foraging behavior of honeybees and how their colonies are structured functions. The bees are divided into groups, namely, employed (food source exploitation), onlooker (exploration decisions), and scout (exploring new food sources, when needed) structures in their social organization.

Similarly, the ant colony optimization (ACO) is based on ants' real foraging behavior (development of the shortest possible path from source to nest). The ant communicates about food sources through pheromone trails left by the ants taking food back to the nest, making it observable by others to travel on provided trails towards food sources [40].

Moreover, the genetic algorithm (GA) belongs to the evolutionary algorithms in inspiration to the theory of Charles Darwin's natural evolution. GA comprises crossover, mutation, and selection operators for diversity and particle swarm optimization (PSO) was an inspiration work observed by birds' flocks and schooling fish searching best locations collectively for optimum fitness [41].

Furthermore, the grasshopper optimization algorithm (GOA) is based on teeming grasshopper's behavior working in teams of nymph (without wings) exploiting the neighborhood and adults (with wings) responsible for locating better regions of food sources [42]. Also, the cuckoo search algorithm (CSA) is one of the many natures inspired optimization algorithms to solve different engineering problems based on family parasitism. All these and other optimization algo-

rithms work iteratively until reaching satisfactory or optimum solutions.

V. LITERATURE REVIEW ON STL

There are various databases with numerous papers in electrical load forecasting; however, we have deduced some literature from Science Direct, Web of Science, and Scopus databases. These papers comprise of majorly past 7 years, taking the most contribution from journal papers in residential, commercial, power plant/ industrial, grid, and off-grid sectors. Various contribution from different databases is also shown in Fig. 9. Load forecasting has become the energy sector's spinal column in any decision-making process as the complete operations are enormously affected by it. The electrical load pattern and its behavior are intrinsically transient. Therefore, affected by several parameters causing model to vary efficacy.

As discussed in section II C, that power sector organization holds STL accountable to generation sources, power economics, start-up and shutdown forecast, hourly operations, management and scheduling for cost savings, security assessment of power systems, and vulnerability occurrence analysis. Therefore, STL has been in exploration (highlighted in Fig. 10) with different methods that changed, upgraded, modified, or newly developed from time to time. Consequently, this section presents an in-depth analysis of different sectors considered for exploration of STL.

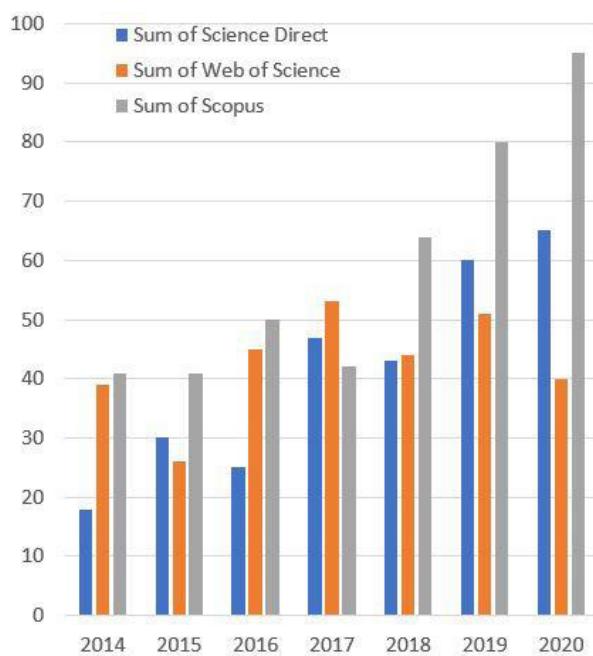


FIGURE 9. Representation of selected literature contributions from different databases for load forecasting.

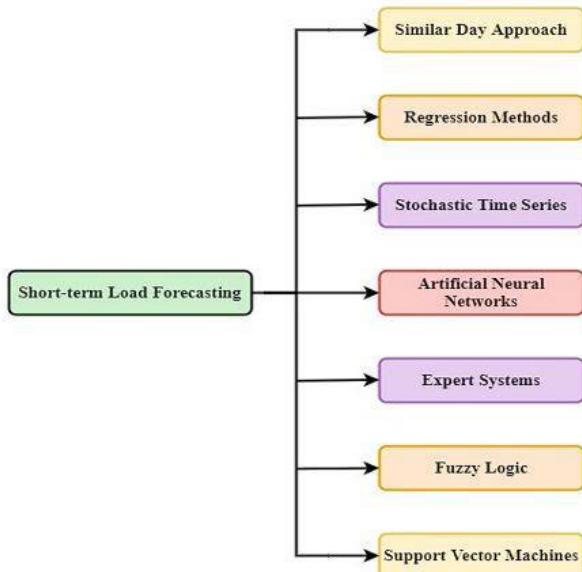


FIGURE 10. List of different methods utilized for STLF exploration in different periods for various applications.

A. RESIDENTIAL STUDIES

Several studies on residential electrical consumers have been carried out due to their intrinsic power consumption nature. The electrical consumption pattern in the residential sector differs from studies carried out in different countries or cities. Some of the studies have been discussed as below:

Barman Mayur *et al.* [35] carried out four different case studies investigating the impact of seasonal changes on res-

idential load consumption; the study proposed incorporating seasonality effects for STLF in Assam, India. It utilized the Firefly Algorithm (FA), SVM, and season-specific similarity concept (SSSC) under different meteorological variables exploration after plotting the correlations.

Season-specific and traditional approaches were compared in two different parts: considering the varying variables of input factors and the approach. However, it was determined that the inclusion of multiple meteorological inputs had a certain degree of positive impact in improving the STLF. A similar study by Barman Mayur *et al.* [29] presented the regional approach's consideration and included the regional factors affecting STLF to improve the load forecasting model's efficacy while comparing the study with GA-SVM and PSO-SVM.

Weicong Kong *et al.* [36] discussed the issues on STLF for residential customers by performing the data analytics on individual households and system load. A density-based clustering is also utilized to evaluate the load data concerning consumption and errors. LSTM has been employed for the prediction task for its pattern recognition ability which outperforms other machine learning algorithms. Forecasting topology is based on individual basis load forecasting from the provided data set, which is later aggregated together due to Australia's SG's fashionable society. The smart society data enabled the gathering of residential consumer behavior towards consumption analysis, which is further sub-sectioned under different data interpretation categories. The observations included customers' different consumption consistency, creating a challenge for forecasting accuracy; also, the forecasting accuracy tends to drop significantly as the aggregation level decreases. However, the LSTM can contribute more towards the improvement of forecasting accuracy irrespective of the consumption inconsistency.

Weicong Kong *et al.* [37] presented the residential load forecasting that depends on the resident behavior analysis. It learns via appliance consumption with a data frequency of every 30 minutes interpreting the consumer to load analysis for effective STLF in residential forecasting through LSTM and compared with feed-forward neural network (FNN). The test scenarios were differentiated concerning time intervals and appliances selected, claiming to show a significant improvement in STLF through load consumption by appliances.

Wang *et al.* [43], to deal with (variability and uncertainty) futuristic load profiles, presented a probabilistic approach for individual consumer load forecasting. To achieve the task, the LSTM is supported by pinball loss which guides LSTM for parameter training. It presents a multioutput LSTM model to forecast all quantiles in one network instead of multiple individuals on the Ireland dataset (comprising of 6000 residential profiles out of which selection of 100 random profiles were selected) with different lead times (1/2, 1, 2, and 4 hours). The model was compared with the quantile regression neural network (QRNN) and quantile gradient boosting regression tree (QGBRT). The dataset was exclusive of weather data,

which raises a significant concern if the weather variables change.

Wang and Ahn [44] presented a framework for residential electrical load anomaly detection (RELAD) by utilizing a rule-engine-based load-anomaly detector (RE-AD) along with one step ahead load predictor (OSA-LP) to detect a load of about 44 residential customers in a Tanzanian rural area. OSA-LP is a combinational framework of ARIMA and ANN, employing Bayesian information criterion (BIC) for data fitting problems and optimization of model. Whereas RE-AD is a combination of SVM and kNN, the forecasting is performed by utilizing the past 24 hours.

Amara et al. [45] examines the relationship between temperature and electricity in Montreal, a city of Quebec province, Canada, where electronic devices are used as a primary source of heating in residences, increasing the electric consumption and adding complexity in forecasting as duration and number of houses varies. This study considered temperature as the major influential factor and utilized an adaptive conditional density estimation (ACDE) model and compared it with multi-layer perceptron (MLP) and recursive least squares (RLS) to forecast the heating and air-conditioning consumption and data analysis of thermal behavior of the building.

Bruno et al. [46] on an R&D project of Italy, performed on public buildings which targeted to design an energy management system (EMS) device responsible for optimization and control for distributed energy sources. The work majorly contributes towards developing control modules for EMS applications, consisting of flexibility in adopting different forecasting techniques (ES, ARIMA, and NN) and relevant impact variables. Model predictive control is used in a hierarchical structure for limiting the error in the forecast.

B. COMMERCIAL STUDIES

This section provides a detailed review for state-of-art methods for the commercial sector, including buildings, offices, institutes, etc. considering their indoor, outdoor, design, layout, social and economic factors influencing the electrical load forecasting:

Ma et al. [47] discussed distributed energy systems (DES's) while targeting the district load forecasting (DLF) and presented different facets affecting the district load. Classification and comparison of methods (top-down and bottom-up) were performed to observe the current lack of forecasting, claiming to forecast accuracy or workload burden on the system. The study also discusses the indoor, design, layout, social and economic factors that have an underlying impact on load forecasting associated with buildings.

Nepal et al. [48] proposed a method comprising of k-means clustering and auto-ARIMA for electrical load peak forecasting for buildings taking the data of Chubu University (East Campus) from 2017-2018, to take respective actions for designing the strategy that helps in peak reduction of load and curtail the electricity bills.

Chen et al. [49] have focused on the demand response (DR), the difference of baseline and actual load, and being inspired by support vector regression (SVR) and its capabilities, have presented an approach to improve the load forecasting for office buildings utilizing the historical load and weather data also stating that the dry bulb temperature which is taken "pre-two hours" can contribute more in terms of accuracy in forecasting. The working dataset was based on four buildings' load data for 2 months for a specific time (9 AM to 5 PM), including the working factor schedule.

Duan et al. [50] presented a model comprising feature selection for selecting the best inputs and features. The proposed model was constructed using SVM with finely tuned parameters to forecast for an aggregated load of the selected buildings to consider electric vehicles' impacts.

Dagdougui et al. [23] have examined their proposed approach of ANN with Bayesian regularization (BR) and Levenberg–Marquardt (LM) for STLF for district buildings in Montreal (downtown campus). The study evaluates the ANN performance and its learning algorithm of backpropagation, analyzing model performance for an hour and day ahead forecasting for buildings of different types and investigating internal architecture impacts on models forecasting ability and different input features, weather conditions, and previous load data. The study claims that better performance was observed for one-hour ahead load forecasting.

Liu et al. [51] outlined a new prediction technique for building level consumers load forecasting using a sliding window empirical mode decomposition (SWEMD) for maximizing the relevancy and minimizing the redundancy depending on Pearson's correlation (MRMRPC). The forecasting engine is an improved Elman neural network (IENN) and novel shark smell optimization (NSSO) algorithm.

Skomski et al. [52] explores the implementation of sequence-to-sequence RNN for STLF considering the case study of four commercial buildings; comprising offices located in eastern Washington. The study also deduced that a long forecast window increases the probability of low predictions. The study investigates the performance of neural networks with varying time resolutions (1, 15, or 60 mint data frequency), the date range's dependency for data training, decay effect, and short-term context. The sensitivity of sequence-to-sequence models for hyperparameters and the generalized effect of one model application to others (one building model for others) were also part of the performed investigation.

Chitsaz et al. [53] explored a new prediction model for an educational building supplied by a micro-grid, comprising a self-recurrent wavelet neural network (SRWNN) for its training and implementation. LM algorithm is used to forecast the volatile and disrupt time series of the electrical load of British Columbia Institute of Technology (BCIT), Vancouver, which has very high volatility (load fluctuates severely on a day-to-day basis compared to normal grids).

Zulfiquar A.K et al. [54] have developed a hybrid method combining CNN and LSTM-AE (autoencoder) for energy

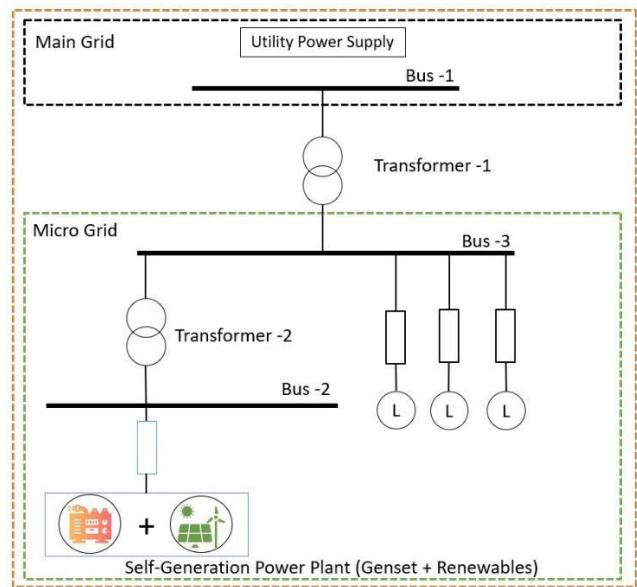


FIGURE 11. Diagram illustrating the coupling of the primary grid and industrial micro-grid for industrial energy consumption.

forecasting in residential and commercial buildings (UCI France and Korean commercial building, respectively) utilizing the data obtained from smart meters and claiming to be the first to incorporate the framework with utility pre-processing. The model works in a manner where CNN performs feature extraction from data input after pre-processing and normalizing the data (removing the outliers and missing values but collected data was from smart meters. Therefore, no such values were found and transferred for encoded sequence generation to one of the LSTM modules and later decoded by another LSTM module advancing to predict energy.

C. INDUSTRIAL STUDIES

The worldwide electricity consumption is enormously contributed by the industrial sector, which either has its generation through power plants or manages the supplied power through the national or contractual grid and personal generation. Industrial consumption varies and increases with time as the region, state, or country progress in industrialization for economic growth. Therefore, this sector needs explicit exploration as a separate field in terms of short-term electrical load forecasting.

The power plant of a facility is one of the most crucial and essential departments, which facilitates the rest of the departments to ensure their operations are timely achieved; representing industrial power plants is presented in Fig. 11. For this purpose, a felicitous approach of anticipating the future requirement should be computed to prevent any downtime at an economical cost. The load forecast stands as its main deriving factor, which plays a vital role in regulating how and when to expand the power facility.

The STLF problem can be observed as a non-linear forecasting problem. Therefore, it is highly necessitated for power

plants dealing with the industrial or commercial sector to consider the influential parameters affecting the forecast, such as process, meteorological factors, and its associates, day, time, and event. Models not considering these constraints will generate low predictions resulting in the loss of millions of dollars in operations, maintenance, and generation process. Some of the studies focusing on industries are discussed in this section to highlight the sector's importance.

Bracale *et al.* [55] provided the industrial load modeling insights considering a set of MLR models while observing transformer manufacturing Italian factory and characterizing the production and working shifts utilizing qualitative variables for a valid industrial load forecast. The load measurement was performed effectively by taking the total, each feeder, and single loads of machine, buildings, etc. The study analyzed the importance of considering the inclusion and exclusion of industrial load information for efficient modeling. Also, Bracale *et al.* [56] presents the active and reactive power forecasting for a probabilistic industrial load forecast investigated by univariate methods (quantile regression forests - QRF's) focusing on suitable variables and multivariate (vector autoregressive exogenous model - VARX) methods focusing on multivariate approach; by exploring the same factory dataset; characteristically a heterogeneous load.

Kim *et al.* [57] performed a study on two small scale industries from Silicon Valley (CA, USA) for industrial load forecast focusing on peak-load prediction for more accurate results using the different combination of regression models (bagging, random forest, extra trees, ada boost, and gradient boosting regressors). The data collection frequency was continued for 2 years with the time interval of 5-minutes. The study contributes towards understanding and depicting the peak-load with features selection capability and evaluation of underestimation. The model was designed as a generalized process and had added compensation values in the predicted results.

Bracale *et al.* [58] present an approach of probabilistic industrial load forecasting which includes the active power and reactive power of an Italian factory. The method comprises two stages out of which first consists of individual univariate forecast generation for active and reactive power, and second amalgamates these forecasts into a multivariate based regression model (comprised of QRF or URF).

Zhang and Chiang [59] have presented a framework of two-layered neural network for STLF, namely Enhanced ELITE (based on optimal structure, accurate and diverse ensemble of neural networks) and consensus-based mixed-integer PSO trust-tech (CMPSOATT) for high-quality solution optimization, computation, and accuracy of STLF for industrial applications. The E-ELITE model constructs several individuals and internal NN models and then selects the best out of it based on accuracy. The model's network architecture comprises 3 stages: exploration and exploitation of data in the bottom layer and ensemble in the top layer.

Park *et al.* [60] investigates the combined cooling, heat, and power (CCHP) in Korea (the energy markets in Korea are

working towards SG transformation) to curtail energy costs in the process of power generation utilizing useful and accurate STLF. It proposes a two-stage model in which executions first stage involves models of extreme gradient boosting (XGB) and random forest (RF), and the second stage works on deep neural networks (DNN's) for calculating optimal operation schedule, electrical charge minimization, and electric rate. The performance measurement of this model was observed by conducting Wilcoxon and Friedman tests.

Hu *et al.* [61] have presented a case study in which a detailed discussion of STLF for the process industry (paper-making process) has been briefed. The study presents the model of “GA-PSO-BPNN” in which GA-PSO has been considered for parameter optimization (weights and thresholds) of BPNN (an MFFN backpropagation trained, for error optimization of ANN). The study considers the data from two different process industries to verify the proposed model and generate a comparison. The Chinese process industrial enterprises are accounted for around 70% of electricity consumption in comparison to society. The study is subdivided into four sections of obtaining the historical data, pre-processing it by outlier’s removal using 3 sigma method and gap filling of results utilizing interpolation methods and noise filtration utilizing moving average (MA) and KKF methods: training and forecasting, and evaluation. The input variables affecting the load forecasts were considered from external and internal industrial environments.

Son *et al.* [62] present demand forecasting technique for the industrial sector using time-series clustering of different methods combined (double seasonal Holt-Winters-DSHW, Trigonometric transform, Box cos transform, ARMA errors, Trend, and seasonal components - TBATS, ARIMA Variants, and NN-AR) by utilizing the data obtained through advanced metering infrastructure (AMI) with the interval of an hour. It presents the power prediction method using the bottom-up approach after analyzing electricity consumption clusters (based on autocorrelation and normalized periodogram distances).

Wang *et al.* [63] carried out an experimental study on a company of Hunan (a province in China, where industries have to purchase the load consumption based on pre-calculations) which presented an ensemble Hidden Markov model (e-HMM) for STLF to understand the characteristics and learn the dynamics of the different pattern formed during power consumption of an industrial consumer concerning time considering the date, meteorological and electrical features obtained by GRA for redundancy and efficiency of the model. The “log-likelihood” strategy is designed with time windows to improvise HMM results from past generated patterns to improve the accuracy and prediction results. The bagging ensemble algorithm’s framework was also utilized to minimize the prediction errors in the single model. The study mainly contributes towards the combined approach of data mining and prediction and multiple HMM’s integrated frameworks. The model comparison was made with state-of-art methods of SVR, RF, CNN, and LSTM.

Hoori *et al.* [64] utilized the ISO New England input features and hourly load data as a dataset to test and train the model validation of Multicolumn radial basis neural network (MCRN) and Radial base function neural network (RBFN), and further investigate the renewables high penetration. The study parted the input-features of large datasets into multiple sub-datasets and trained them in individual RBFN in a similar manner, reducing the computation interval. The study also provided insights regarding the reliability of renewables distributed generation potentials for considering the best features that affect the load forecasting for either shortage or surplus of energy in planning.

Ahmed *et al.* [65] claims to present the model with great accuracy and fast convergence (AFC) rate of STLF for industrial applications but cannot forecast for two or more days. The study also discusses the improvement in STLF models’ accuracy and convergence on existing ANN-based methods and modifications in two techniques: “mutual info-based feature selection and enhanced differential evaluation (EDE) algorithm.” The proposed AFC-STLF model comprises a feature selector which removes the redundant and irrelevant inputs and transfers the data to the next module termed as a forecaster and made up of ANN where training and validation are performed; later, the data is feed into an optimizer module where the error is calculated and optimized to a minimum using mEDE iterative algorithm.

Kou and Gao [66] focused on probabilistic day-ahead load forecasting for energy-intensive enterprises (EIE’s) considering a model based on Gaussian process and named it as a “sparse heteroscedastic” model using the data generated at the power plant of steel manufacturing in China, collected through SCADA system for an averaged frequency of 5 minutes interval for 5 months period. The autocorrelation was plotted to observe the nature of load consumption. HGP encounters high-computational complexity, which makes the case complex for practical implications. Therefore, $l_{1/2}$ is used to overcome the issue.

Ungureanu *et al.* [67] aims for the machine learning model’s efficiency concerning industrial load forecasting. The dataset utilized is of meat processing facility (capacity of almost 35 tons meat/day) comprised of hourly consumption values. The study implemented RF and LSTM algorithms utilizing python-based programming language utilizing TensorFlow and Keras libraries for machine learning. The study also investigated the electricity dependencies, including meat production and storage warehouse, which is directly linked to the temperature; sales are also essential factors.

Tan *et al.* [68] proposed the ensemble deep learning model for load forecasting and controlling the ultra-short-term power demand. The study contributes towards the accuracy and robustness of forecasting by the hybrid strategy of models and model ability enhancement for peak demand forecast. The proposed model takes input from demand and external data to conduct an exploratory data analysis (EDA) on raw data, later normalizing it and treating several LSTM networks groups (LSTM consideration is made because of

its strong processing capability for time-series) off-line with subsets and subspace of data. The model performance was assessed with an open dataset from the Australian energy market operator (AEMO) with the same data testing and training. However, long-time forecasting for peak demand needs to be improved.

Petrosanu [69] provides a month ahead hourly forecasting of electrical consumption for medium industrial consumers utilizing the customized NARX, ANN, and LSTM on a timestamped dataset of hourly values consisting of 8760 records. The data was collected from the smart meters database; therefore, no missing or abnormal values were observed, making the forecast more reliable. The method claims to harness the benefits from custom-designed combinational models based on NARX, LSTM, and ANN, which generates the daily forecast at the initial stage through NARX-ANN and then refined with LSTM-ANN for achieving better precision and accuracy, also parallel computing architecture was utilized.

D. GRID STUDIES

This section provides insight into the proposed electrical load forecasting methods focusing on the electrical grids considering different conditions, regions, and applications of different types. The exemplary smart power grid architecture with a two-way flow of information is shown in Fig. 12. The review for grid studies is provided as below:

Ahmed I. Saleh *et al.* [70] have presented a load forecasting strategy using data mining techniques and feature-selection and outlier-rejection methodologies. The prominent feature selection has enabled genetic-based feature selector and rough set-based feature selector (GBFS and RBFS). The obtained results are treated with kNN and NB classifier for load predictions.

Neethu Mohan *et al.* [71] presented the work relating to developing the STL model based on data-driven strategy and employing the method famous for its information extraction, namely, dynamic mode decomposition (DMD), a methodology for the underlying system that captures their spatial-temporal-dynamics. The model claims to be efficient in identifying the load data's characteristics and the factors affecting it (time, calendar, meteorological, social, and economic) and improvising the forecasting accuracy and complexity in grid regions. The test observations were performed on two energy market operators that belong to Australia and North America and compared with other state-of-art methods, and the results were satisfactorily acceptable compared to others.

Xiangyu Kong *et al.* [72] have presented a method developed based on error correction methodology that predicts the error in the load data utilizing DMD, corrected in data selection and forecasting error. GRA is employed to input the model in the data selection stage, which performs required data collection (previous day data, same day data in the previous week, and similar day data). DMD's load data is further refined for stable predictions involving the extreme value constraint method (EVCM). However, the prediction

accuracy differs concerning areas (lower accuracy for small areas and vice versa).

Nian Liu *et al.* [73] have presented a hybrid forecasting model for micro-grids developed with Empirical mode decomposition (EMD), which decomposes the load data into different scale trends and scales components, extended KF and extreme machine learning with the kernel (KELM) algorithms which create a forecast model of dual prediction, and PSO for parameter optimization. Four micro-grids having different load characteristics were considered for method verification, out of which the load of two belonged to small residential areas and two to commercial buildings.

Ying Nie *et al.* [74] have proposed a hybrid model (RBF-GRNN-ELM) with combinational data pre-processing. The datasets from Queensland and Victoria (Australia) are considered for the study. Complimentary ensemble EMD is used for data decomposition. In contrast, singular spectrum analysis (SSA) is used for information extraction and forecasting. The combinational model of RBF-GRNN-ELM and further improvisation of results are performed with a multi-objective grey wolf optimizer (MOGWO) algorithm. After data pre-processing, the standalone forecasting is performed, and later an aggregated model is constructed.

Yeming Dai and Pei Zhao [75], considering the dataset from a SG from Singapore; have provided some improvements for SVM and have made a proposition of a hybrid model consisting of intelligence of feature selection and parameter optimization method (second-order oscillation and repulsion PSO) for improvement of forecasting accuracy. The study considers pricing as one of the major influential factors for consumption pattern variations. Therefore, it is considered one of the inputs with other redundancy and relevancy features and holidays projection; a weighted grey relation is employed.

A. Ghasemi *et al.* [76] proposes a price and demand forecast method using useful tools in pre-processing, forecasting, and algorithm tuning. The method is divided into specific parts, which first involves signal decomposition into multiples at different intervals utilizing flexible wavelet packet transform (FWPT) and feature selection from input data employing conditional mutual information (CMI). The second stage behaves like a multiple inputs and multiple-output (MIMO) model comprised of a Non-Linear Square SVM (NLS-SVM) and ARIMA for correlational analysis. The last stage involves a time-varying ABC algorithm that tunes the parameters.

Youlong Yang *et al.* [77] forecasted short-term power load using the sequential-grid-approach (SGA) based SVR, which subsamples the parameter and fine-tune regions it. The study contributes towards an efficient SVR framework for accuracy improvement of STL, improving SVR calculation intervals, introducing sequential inference for the selection process, and validating the proposed framework over two real cases from Jiangxi (province of China) and California electric utility.

Feifei He *et al.* [78] proposed a day-ahead QRF load forecasting model based on decomposition strategy. The study

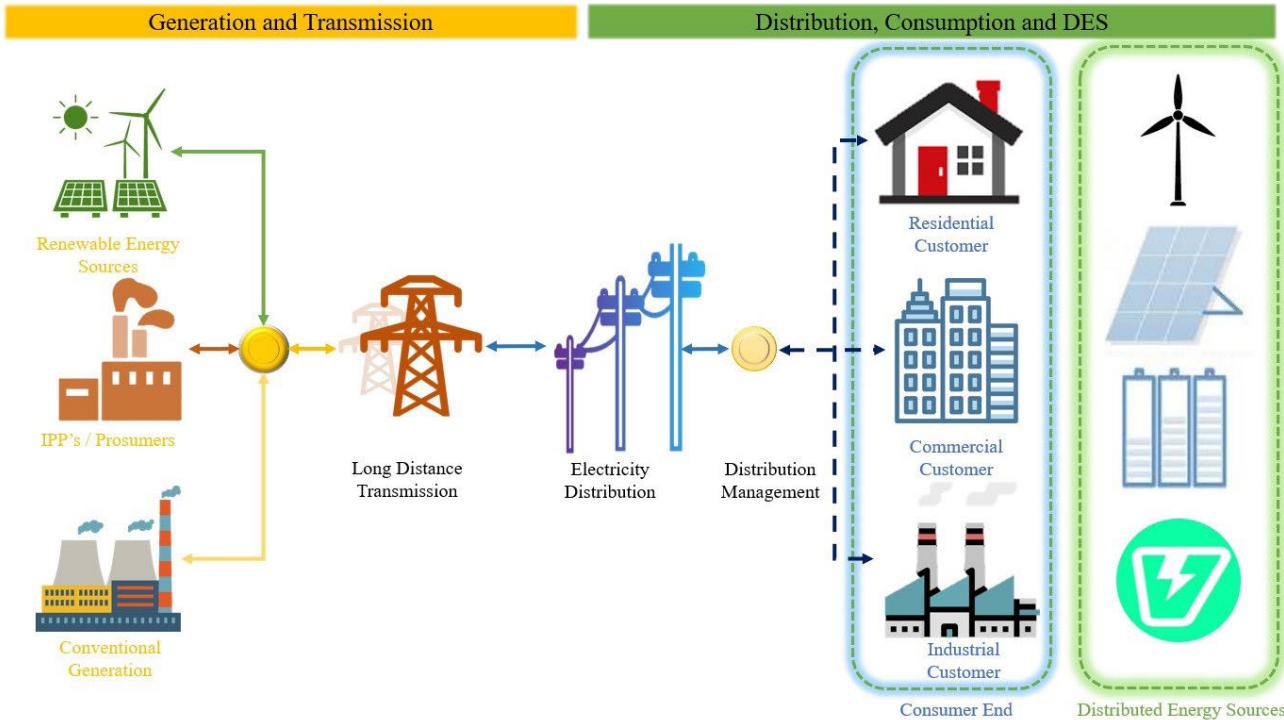


FIGURE 12. Architecture of an ideal SG with bi-directional information flow, including distributed energy sources.

decomposes the load data into sub-models with variational mode decomposition (VMD), the influential factors WTHI (i.e., weighted temperature and humidity index), and type of day are taken into consideration for each decomposed sub-model. The prediction model is then involved with a multi-step strategy to forecast the load for each sub-model using QRF, and the decomposed segments are reconstructed back for a compete for prediction model probability density with the help of kernel density estimation (KDE). The parameter optimization for the dataset obtained from real load data of Henan (province of China) was dependent on the BOA-based Tree-structured of Parzen Estimator (TPE).

Muhammad Qamar Raza *et al.* [79] considered a three-layered feed-forward ANN (FF-ANN) for a week ahead load forecasting over a year and utilized global best PSO (GPSO is used for NN weight bias values optimization) as a training technique for performance enhancement of the forecasting model. The model input comprises meteorological, exogenous, and load data tested and validated over the New England grid (ISO) for performance measurement. NN's designed architecture consisted of 8 input, 20 hidden, and 1 output layer and tuning parameters included no. of particles, time interval, G and P best components.

I. P. Panapakidis [80] have directed the attention towards buses of transmission and distribution systems of urban, suburban and industrial load values of 10 buses (Thessaloniki, North-Greece) and have presented the busload more stochastic in nature in comparison to system load and have considered the busload attributes in his day and hour ahead

load prediction model based on ANN and supported by the clustering technique for the hybrid formation of forecasting model resulting in efficient results.

Salah Bouktif *et al.* [81] have proposed an enhanced-LSTM model considering the periodic electrical characteristics using input time lag of multiple sequences to capture the sequence features for forecasting accuracy of aggregated load for achieving targeted predictions against time. The dataset and meteorological variables have been processed for null values or outlier detection for LSTM and gated recurrent unit (GRU) models. Also, Salah Bouktif *et al.* [41] have used LSTM-RNN deep learning method for STLF utilizing GA and PSO for hyperparameter learning and tuning. These selected parameters then are later used for final LSTM model construction and its training. The model is then compared with benchmarks of RF, SVR and other LSTM's. The dataset utilized was from RTE Corporation (French electric transmission network) consisting of a data frequency of 30 minutes over nine years.

Jian Zheng *et al.* [22] have used a long-term electricity consumption dataset and an airline data in comparison to verify the compatibility of LSTM-RNN models over other benchmark models (SARIMA, NARX, and SVR) to explore the dependencies of electrical load for a long time for accurate load forecasting.

E. OFF-GRID STUDIES

The grid-isolated power plants such as islands, remote communities and industrial operations are typically cut off from

the infrastructure of the normal grid. Due to which they experience unique challenges. Such power plants have different operations due to their size; characteristically smaller and not as diverse compared to most market grid. Due to which, stability cannot be mostly managed with generations alone, reacting in load disconnections.

Therefore, this section presents the studies, including off-grid power plants along with some SG studies for review as below:

Yuanzhang-Sun *et al.* [82] have discussed the wind power efficiency for transmission over long distance and locally consumption presenting their challenges and solutions. They compared the two approaches for their utilization in the context of security, reliability, capability, and cost along with economic and technical feasibility in one of China's large isolated-power industrial power-system. The study also claimed the wind power to be an efficient utilization around the globe for isolated-super-micro-grid.

Danilo P. e Silva *et al.* [83] presents a hybrid model of optimization for isolated-microgrid. Considering forecast of meteorology model for ancillary services of purchasing or selling the power in a situation of requirement or access, respectively. The presented optimization algorithm follows specific rules for connection and disconnection with the primary grid (minimum time connection) considering the battery bank's operations.

Life Lozano *et al.* [84] have presented a case study for Gilutongan-Island (one of many off-grid Islands with limited electricity) in the Philippines, where a 194-kVA DG powers this off-grid island. This study presents a hybrid optimization model for electric renewables (HOMER) model, providing the analysis for adequate power generation on the island for 24/7 electricity provision.

Arslan Ahmed Bashir *et al.* [85] have presented an approach for economically reliable operations of islanded and grid-connected micro-grids (MG) utilizing a framework of energy scheduling and management of associated uncertainties with an islanded grid (IG) and connected micro-grid (C-MG). MG comprises various renewable energy sources (RES's) like solar, wind, tidal, biomass, etc., including Diesel Generators (DG). The proposed mixed-integer-linear-programming (MILP) model predicts a 24 hour ahead by input data updating considering minimizing MG's cost keeping the consumers comfort priorities (HVAC and EWH) while investigating the parametric impacts on cost and risk indices of MG over a year in IG and C-MG.

Asif Islam *et al.* [86] have performed an analysis for electrical load forecasting to accommodate an isolated island (Kutubdia, Bangladesh) with no historical load demand patterns. However, the analysis comprised driving factors considering the load growth of an area involving two methods: inverse matrix and regression analysis.

Shengxi Yuan *et al.* [87] have performed a case study on Sao Vicente (a small island of Cape Verde with high wind resources in the Atlantic Ocean), having an isolated electrical grid-primarily dependent on DG and have an integrated

wind power (capacity of 5.950 MW). To achieve conservative forecasting, the study presents a wind-speed forecast utilizing ARIMA and energy storage. The study also discusses the advantages of renewable integration resulting in cost savings. The forecast was done on an hourly basis rolling mechanism and claims to have more noticeable results regarding reliability, fuel savings, and wind capacity.

Tanveer Ahmad and Huan Xin Chen [88] have examined the utilities, IPP's and Industrial consumers taking historical and climate data as inputs for energy usage estimation for the month, season, and year energy predictions on a large scale utilizing the non-linear autoregressive-model (NARM), linear-model stepwise-regression (LMSR), and RF-least-square-boosting (RF-LSBoost) methods on actual energy data and employing clustering analysis for outlier removals and detection. The model presented a city-wide energy demand in the form of energy prediction.

Azim Heydari *et al.* [89] have adopted the composite neural network and gravitational search optimization algorithm (GSA) for price forecasting and STLF in isolated power-grids. The constructed model has various internal sub-modeling (variational model decomposition, the mixture of data-modeling, selection of features, generalized regression NN and GSA) for price and load forecast. Power markets of Pennsylvania and Spanish electricity have been put to the test for the proposed model to observe the forecasting accuracy and later the Isolated Power-Grid data of Favignana's Island have been taken into consideration for their testing of the proposed model for winter, spring, summer, and fall.

Ibrahim S. Jahan *et al.* [90] have presented a load forecasting model and data clustering techniques to improve electrical loads' forecasting accuracy. The collected dataset belonged to an off-grid platform of VSB-Technical University (Ostrava, Czech Republic), comprising meteorological and power data variables. The forecasting setup consisted of steps, out of which 1st involved the data clustering using k-means and k-medoids. Secondly, utilizing the square Euclidean equation, distances were measured, which improvised the dataset's features, and in the last step, the training and testing of the forecasting model were performed utilizing (DT) ANN.

Tolga Turai *et al.* [91] have taken a remotely located oil and gas facility (data of 160 days with different time interval) for a case study of electrical load forecasting in off-grid platforms considering the ANN (multi-layer NN's - Nonlinear Autoregressive NN model) as a prediction model taking 60 and 15-min data of electrical load. The two cascaded non-linear autoregressive NN models (CNNARX) were utilized with different hidden layers topology to improve forecasting accuracy. However, compared with the 60 and 15-min, the 15 min interval produced better results.

Spyridon Chapalaglou *et al.* [92] have presented an algorithm for managing the energy smartly of an islanded-power-system with different power sources (DG, PV, and battery storage). The algorithm schedules comprise load forecasting, pattern recognition, and optimal power-flow customization to efficiently manage the system's power-flow. The L.F module

TABLE 4. Summary of the comparative study of selected literature comprising prominent methods based on their strengths (defined in section V part F).

References	Methods	Strengths					Remarks
		Acc	Dur	Adap	Pro	Com	
[38, 44 - 47, 66, 67, 74 – 77]	Regression	**	*	x	*	*	Often misses vital information and consists of the problem of collinearity and multicollinearity. Also, not very effective on cross-sectional data
[28, 32, 39, 64, 65]	SVM	**	**	x	**	**	Provides effective results in well separated and clear dimensional spaces but provides ineffective performance in noisy data.
[23, 35, 40, 42, 49, 50, 54, 63, 68, 69, 78 – 81]	ANN	***	***	x	***	**	ANNs are very good with non-linear data with fast predictions but have performance degradation relating to overfitting and memorization, dealing in noisy datasets, and inadequate learning rate.
[30, 41, 63, 22, 27,]	RNN	***	***	x	***	**	RNNs are good with performance in sequential data due to internal memory. It also generates impactful outputs but is prone to the issues of vanishing gradient and exploding problems
[22, 27, 29, 31, 43, 56 - 58, 70]	LSTM	***	***	x	***	***	Since LSTM is an enhanced form of RNN, therefore it caters to the RNNs highlighted issues. However, high memory bandwidth requirement, small weight initializations, and overfitting are still some challenges in LSTMs.

*	Good	x	Feature does not exist	Adap	Adaptability
**	Better	Acc	Accuracy	Pro	Processing
***	Best	Dur	Durability	Com	Complexity

depends on clustering algorithms and FF-ANN for the day ahead STLF. The results are treated as inputs to algorithm for pattern recognition for its classification according to the load curve. The study presented a methodology to improvise the smart flow of energy for efficient energy management and maximize the utilization of RES in peak demands and extend DG's life and save fuel costs.

F. COMPARATIVE STUDY TABLE

The summary of the works reviewed for this paper are presented in Table 4. The table emphasizes the methods utilized in residential, commercial, industrial, grid, and off-grid electrical load forecasting sectors. It presents the approach, discusses its strengths, and provides a review of advantages and disadvantages in the remarks section. In the table, the term "Acc" denotes accuracy, "Dur" denotes durability, and "Adapt" denotes adaptability. Additionally, the term "Pro" denotes processing time, and "Com" defines the complexity. Notice that we have utilized the scale in terms of "*", "**", "***", and "x" which symbolizes good, better, best, and no existence of the feature.

VI. LIMITATIONS OF EXISTING L. F MODELS

Different prediction models have different pros and cons. Some models' performance is limited to stationary time series, and some have good performance in non-stationary time series but can work with limited parameters. Whereas the data in DGM environment has totally different nature

as it is produced with the generation of different sources affected by different parameters. The data in result has large variations in comparison to existing available data. Therefore, the existing prediction models without modifications stand obsolete in such type of environment.

Considering the global digitalization and clean energy production uprising the need of SG and implementation of smart city, such models are required which can deal in load forecasting with multiple features and can withstand the changes occurring in real time. The limitations of some of the existing models with suggestions to tailor them in accordance with DGM environment are presented in Table 5. However, these changes are not limited to presented remarks but could increase or decrease in accordance with the requirement.

VII. PERFORMANCE MEASUREMENT CRITERIA

Numerous researchers' overtime has used different criteria to evaluate their methods, models, and techniques in predicting load forecasting and in time new are on way to discovery; to quantify their results compared with actual load values. These indices justify the correction or precision of predicted values for their evaluation. The most famous performance measurement criteria for load forecasting among researchers are combined in Table 6. Where N is the number of samples, Y_i' and Y_i represent the models' predicted and actual desired values, respectively.

The metric values of these criteria vary concerning the difference in defined parameters and provided datasets. Each

TABLE 5. Limitation of conventional load forecasting models with respect to DGM environment.

References	Methods	Limitations	Remarks		
[11]	rGk-TS	Not witnessed a separate model for weekends which would be based on the previous weekend and not on the previous day.	Observed no automatic feature selection for minimal-redundancy-maximal-relevance. However, this feature can help to transform the model and curtail it for DGM.		
[33]	GA-PSO-LSTM	Inappropriate choice of mutation and cross-over operators. Not feasible for non-stationary. Premature convergence and low accuracy are observed. Not feasible for handling non-stationary input data assumptions (which exists due to online time series). Observed performance degradation due to complexity of online time series, which causes non-stationary and non-linearity of load patterns.	Since the DGM are affected by several factors therefore it is required to work on multi-scale property of input sequences, multiple features, and multiple variables. However, with improvement of these features LSTM's have great capabilities to be explored with DGM.		
[34]	GOA-SVM	Limited correlational analysis on climatic factor is presented. FA experiences the issue of premature convergence. However, GOA unbalances the exploration and exploitation factors by means of convergence parameters, have unstable speed of convergence, and falls easily into local-optimum. Moreover, SVM has limitations to work with large and noisy datasets. It faces issues of overlapping and underperformance.	DGM contains huge datasets where balanced convergence of parameters is required. It consists variations in data samples along with noise. This leaves SVM out of the league of DGM. However, improved versions managing these issues could be considered.		
[38]	FA-SVM	The study comprised of scenario-based load forecasting where the time period was limited to reduce the peak consumption. However, the limitations of ARIMA include its linear nature and order selection process.	DGM have a complex structure of load demand and requires the exploitation of leading indicator along with explanatory variables which is limitation for such models.		
[43]	ARIMA	The study removed the low-quality data and had missing data to re-fill. The study was conducted making the environment comfortable for model. However, the model needs validation on other datasets as well.	Owing to the fact of sensitivity, BPNN are sensitive to noisy data and performance depends upon input data. However, in DGM the data have marginal noise and input data keeps variating which can cause this model to perform poor.		
[56]	GA-PSO-BPNN	The collected data was from smart meters resulting in no data irregularities and was a custom-design model. However, the load data was pre-maintained, and architecture was observed to be complex. Autocorrelation is not present.	The nature of DGM data is highly complex in comparison to industrial data where it can be separated. However, the introduction of multiple variables could imbalance the system. However, the modified model could be considered for DGM environment with inclusion of multivariable and autocorrelations.		
[64]	NARX-ANN-LSTM	DMD corrects the load time series which is limited by EVCM putting the requirement of higher accuracy in historical data. Prediction accuracy varies inversely with the areas and accumulated errors have higher values.	The model experiences accuracy issues in small segments. DGM has multiple small segments which need to be dealt independently but the accumulated error accuracy will be large.		
[67]	DMD-EVCM	GRU troubles with convergence rate, efficiency of learning and fitting problems. Whereas LSTM were considered to remove gradient vanishing problem, which still persists, it faces issues with model hyperparameters selection and is prone to overfitting.	Same model is not suitable for multiple type of dataset, similarly same parameters are not benchmark for other type of models, therefore an auto-tuning of parameters is required to tailor the model for DGM. Since specific regions have their dedicated inputs, therefore autocorrelation for optimized of inputs would also be required for DGM with appropriate settings to avoid try-and-see based configurations.		
[76]	LSTM-GRU	BPTT has contributed towards efficient training for RNN, but BPTT is only suitable for offline training and RNN. However, this approach introduces gradient vanishing and exploding problems.	BPTT limitation with noisy data and online training will be greatly affected in DGM environment. RNN with modified version with reduced issues of gradient vanishing and exploding can be tailored for DGM but in combination with other algorithms.		
[93]	BPTT-RNN				
ANN	Artificial neural network	FA	Firefly algorithm	PSO	Particle swarm optimization
ARIMA	Autoregressive integrated moving avg.	GA	Genetic algorithm	rGk	recursive Gustafson-Kessel
BPNN	Back propagation neural network	GOA	Grasshopper optimization algorithm	RNN	Recurrent neural network
BPTT	Back propagation through time	GRU	Gated recurrent unit	SVM	Support vector machine
DMD	Dynamic Mode Decomposition	LSTM	Long short-term memory	TS	Takagi-Sugeno
EVCM	Extreme Value Constraint Method	NARX	Non-linear autoregressive network with exogenous inputs.		

of the performance metrics has its pros and cons in terms of evaluation. Like RMSE focuses on more significant errors concerning small but has a second-degree loss function, MAE

provides average error measurement, and MAPE can lead to biasness. Also, it observes denominator complications. Therefore, comparing the results of different methods with

TABLE 6. Performance measurement criteria for load forecasting models.

S. No	Performance Metrics	Equations
i.	Root mean square error (RMSE)	$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y'_i - Y_i)^2}$
ii.	Mean absolute error (MAE)	$MAE = \frac{1}{N} \sum_{j=1}^N Y'_j - Y_j $
iii.	Mean absolute percentage error (MAPE)	$MAPE = \frac{1}{N} \sum_{j=1}^N \frac{ Y_j - Y'_j }{Y_j} * 100$

different metrics will not be as just as required. However, Table 7 is compiled with the available methods forecasting different modalities based on the MAPE % evaluation of different models from surveyed papers of different sectors.

VIII. OUTCOMES

The critical discoveries were drawn from this survey after scrutinizing, reviewing recent studies, and comparing state-of-art electrical load forecasting models. The outcomes are listed in this section under trends (presenting the findings) and research gaps (highlighting the possibilities of unexplored areas) in STLF.

A. TRENDS

The trends identified from the reviewed papers are listed below:

- The forecasting accuracy of model variate with respect to the application, the change can cause inconsiderable amount of error. The change could be related to different application, sector, and region.
- The correlational analysis of factors affecting the load forecasting have significant impact on model accuracy. However, these correlations are also dependent on regions, change in region will adversely affect the model accuracy.
- Different papers have considered different meteorological variables relating to their application and region. Thus, there are no defined standard variables to be considered in load forecasting.
- The STLF is considered more reliable for operations and planning of utilities. It is also considered to further expand the operations and contribute in MTLF and LTLF.
- Historical load data contributes towards better model designing, resulting in better accuracy and model performance.
- The missing data points, noise, and bad data cause error in predictions. However, they can be improvised by the help of respective algorithms.
- Better accuracy results were observed from the datapoint received from smart meters having no bad data and noise.

- Every model has its own merits and de-merits; however, hybrid methods were observed to have more performance efficiency in comparison to standalone methods. The hybridization of models with fine-tuned parameters had better performance.
- Due to the sector differences of power utility, the model, choice of parameters and algorithms, and the load pattern were observed to be different and followed a different cycle making every sector unidentical to be replicated.
- The utilization of different clustering techniques was employed in many papers however the results of k-means outperformed others. The recent introduction of HDBSCAN will also have improved results.
- Extensive use of optimization algorithms was observed in load forecasting models. However, the GA and PSO optimization algorithms were mostly used due to their high-performance capability of crossover and mutation, and optimum fitness, respectively.
- The existing models have limitations in adaptivity and in dealing with utilities having different generation modalities.
- The load patterns of grid-isolated industries and with DGM, SG, and smart cities have a different nature due to their intrinsic nature of operations.

B. RESEARCH GAPS

In this section, we present our observation leading to the available research gaps in electrical load forecasting.

1) CORRELATIONAL ANALYSIS

The existing models have set no standard for including the factors affecting load forecasting. Some studies have worked on 2 factors, and some have included more than 2. However, just few of the studies have considered relevancy and redundancy for inclusion of most relevant parameters. Therefore, inclusion of more parameters whether meteorological, economical, or regional with relevance to correlational analysis should be included in forecasting model.

2) MODEL ADAPTABILITY PROBLEMS

Since the existing load forecasting models are limited towards dedicated applications and cannot produce the desired results with the change of application. Therefore, it opens a direction to design a model with adaptability. No such models have been observed through-out the review of the papers to the best of our knowledge. Adaptive model is defined as a model which can evolve with changing conditions (input parameters, region, or application). Such models found the basis of load forecasting in different generation modalities environment.

3) ISOLATED-GRID FORECASTING

The load forecasting models have extensively explored the load phenomenon in grid connected sectors, the sectors isolated from grid have been neglected. This sector has different

TABLE 7. Summary of selected literature based on Modalities and MAPE % of prominent methods.

Reference	Modalities	Method	Parameters	Application	MAPE %	Comparison
B. Mayur et al. [35]	Gas	FA-SVM	L, t, T, CC, WS		1.66	GA-SVM, PSO-SVM
W. Kong et al. [37]	Hydro	LSTM	HD	Residential	4.24	FFNN
Y. Wang et al. [43]	Gas + Coal + RES	Pinball LSTM	L, t, w		2.19	QRNN and QGBRT
W. Ma et al. [47]	DES	LSSVM	L, t, T, BI		2.15	LSSVM, LSSVM- PSO, SVR
B. Nepal et al. [48]	Hydro + Nuclear + RES	k-means and auto-ARIMA	L, t, T		5.1	ARIMA
M. Duan et al. [50]	-	ESVM	L, t, d	Commercial	6.5	ARIMA, SVR, RBFNN
D. Hanane et al. [23]	Hydro + Wind	ANN-LM	L, t, T, H, w		4.18	ANN-BR
Y. Liu et al. [51]	Gas + Oil	IENN-NSSO	L, t, T		3.95	ARIMA, SVR, RBFNN
Y. F. Zhang et al. [59]	Nuclear + Gas + Coal	CMPSOATT	L, t, T		1.72	ANN, SIWNN, SVR
Sungwoo Park et al. [60]	Nuclear + Gas + Oil+ Hydro	XGB-RF-DNN	L, t, T, CC, H, WS		21.09	DNN, DT, MLR GOA-SVM, GPSO-ANN, DMD
Yusha hu et al. [61]	Coal + Hydro	GA-PSO-BPNN	L, t, T, P, H		0.77	
H. -g. Son et al. [62]	Hydro + Nuclear	NN-AR	L, t, T	Industrial	3.86	FARIMA, ANN
Y. Wang et al. [63]	Coal + Hydro	e-HMM	L, t, w		3.13	SVR, CNN, LSTM
A. O. Hoori et al. [64]	RES	MCRN - RBFN	L, t, T, d		4.59	SVM, ELM
S. Ungureanu et al. [67]	-	LSTM	L, t, d		17.10	RF, Regression, NN
M. Tan et al. [68]	Coal + Wind	EDA - LSTM	D, t, m, d		3.875	SVR, LSTM, XGBoost
X. Kong et al. [72]	Coal + Hydro + Wind	GRA-EVCM EMD-EKF- KELM	D, t, d, w		4.772	SVM, SSA-AR
N. Liu et al. [73]	-	RBF-GRNN- ELM	L, t, T, d		13.78	-
Y. Nie et al. [74]	Coal + Solar	mRMR-GA- SVM	L, t, T	Grid	2.89	CEEMD-SSA- ELM
Y. Dai et al. [75]	Gas	LBoost, LMSR, NARM	L, t, T		7.22	BPNN
M.Q. Raza et al. [79]	Nuclear + Gas + Coal	FF-ANN	d, w, t, T		1.87	BP-NN, LM-NN
S. Bouktif et al. [81]	Nuclear + Oil + Gas	e-LSTM	PL		0.55	MLR, MLP, DNN
T. Ahmad et al. [88]	Nuclear + Gas + Coal	GRNN-GSA	L, t, T		3.64, 4.15, 4.11	-
A. Heydari et al. [89]	DES	ANN	L, t, T		1.33	RBF, GRNN
I. S. Jahan et al. [90]	Solar	CNNARX	L, t, T, WS	Off-Grid	47.95	DT
T. Turai et al. [91]	Solar	FFNN	L, t, T		12.60	NAR, NARX
S. Chapaloglou et al. [92]	Diesel + Gas	FFNN	L, t, T, d, w		1.71	-

L	load	t	time	T	Temperature	D	Date
d	days	w	Week	WS	Windspeed	CC	Cloud cover
P	Process	H	Humidity	HD	Household data	BI	Building Layout
m	Month	PL	Past Loads	RES	Renewable energy sources	DES	Distributed energy sources

load behavior due to the change of structure and production capability. It faces heavy issues regarding load balancing and load surges causing the system to distress and produce randomly changing load patterns. This situation creates a challenging environment for power professionals to predict the future load demand. This random load behavior cannot be explored with existing developed models and need some more

custom-designed parameters. Thus, the sector requires more exploration for accurate load forecasting for this intrinsic nature of load.

4) INDEPENDENT POWER PRODUCERS (IPP'S)

The independent power producers (IPP's) are also key sources for power production, which are not a public utility

but contribute to power pool exchange. These utilities have a different pattern for power dispatch because it differs from the pool exchange position. Therefore, an accurate model for an exact node can help IPP's with their power management. This sector also needs immense exploration in terms of load forecast modeling because its nature is distinctive compared to others.

5) PROCESS BASED INDUSTRIAL LOAD FORECAST

The survey also revealed that the industrial load forecasting model based on processes could lead towards more accurate and better precise results. It can help to enhance the energy management, to supply the energy on pre-determined power consumption patterns based on model output and plan accordingly for facility enhancement. But the huge variations are observed with change in process, therefore the model needs to have parallel operational window to accommodate multiple forecasting operations for different processes at same time. Also, it can produce an accumulated results of these forecast at final stage resulting in better forecast.

6) ANOMALY BASED FORECASTING

Another fascinating area of further research includes load forecasting with anomaly detection. The transformation of the electrical grid from conventional to smart has opened a broad research area. Because a SG requires massive two-way communication, which is to be dealt with a large amount of data and integration of renewables gives it more complexity. Therefore, we firmly believe that the electrical modeling capable of forecasting with anomalies detection will further improvise forecasting. Since the load patterns in different generation modalities environment has multiple resources involved, it can create hinderances in accurate load prediction.

IX. CONCLUSION

Electrical load forecasting is the key component in power system operations. Different decisions related to energy transactions and management are based on it. Therefore, its accuracy and precision for decision-makers in the energy sector are utmost important. This paper has presented a detailed review of diverse load forecasting models relating to various periods and considered in distinct sectors.

The hybrid methods were observed to have better forecasting accuracy, but still they had limitations. We discussed the conventional methods of load forecasting along with their performance and limitations. We discussed the importance of different generation modalities in existing power systems and presented an analysis to tailor the existing load forecasting methods for DGM scenarios (SG and Smart city). Finally, highlighting the limitations, trends, and research gaps will aid the readers for selection and evaluation of models as per their requirement and get future working directions. Future work will focus to enhance the limitations of existing load forecasting models, DGM and adaptability of load forecasting models.

REFERENCES

- [1] 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.
- [2] C. Cecati, J. Kolbusz, P. Rózycki, P. Siano, and B. M. Wilamowski, "A novel RBF training algorithm for short-term electric load forecasting and comparative studies," *IEEE Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6519–6529, Oct. 2015.
- [3] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer, A. Rizzi, and R. Janssen, "An overview and comparative analysis of recurrent neural networks for short term load forecasting," 2017, *arXiv:1705.04378*. [Online]. Available: <http://arxiv.org/abs/1705.04378>
- [4] P. Singh, P. Dwivedi, and V. Kant, "A hybrid method based on neural network and improved environmental adaptation method using controlled Gaussian mutation with real parameter for short-term load forecasting," *Energy*, vol. 174, pp. 460–477, May 2019.
- [5] M. Q. Raza and A. Khosravi, "A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings," *Renew. Sustain. Energy Rev.*, vol. 50, pp. 1352–1372, Oct. 2015.
- [6] P. Singh and P. Dwivedi, "Integration of new evolutionary approach with artificial neural network for solving short term load forecast problem," *Appl. Energy*, vol. 217, pp. 537–549, May 2018.
- [7] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *Int. J. Forecasting*, vol. 32, no. 3, pp. 914–938, Jul. 2016.
- [8] 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.
- [9] S. Fallah, M. Ganjkhani, S. Shamshirband, and K.-W. Chau, "Computational intelligence on short-term load forecasting: A methodological overview," *Energies*, vol. 12, no. 3, p. 393, Jan. 2019.
- [10] 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.
- [11] F. Elkarmi, *Power System Planning Technologies and Applications: Concepts, Solutions and Management*. Hershey, PA, USA: IGI Global, 2012.
- [12] 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.
- [13] R. J. Hyndman and S. Fan, "Density forecasting for long-term peak electricity demand," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 1142–1153, May 2010.
- [14] E. M. de Oliveira and F. L. C. Oliveira, "Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods," *Energy*, vol. 144, pp. 776–788, Feb. 2018.
- [15] M. R. Al Rashidi and K. El-Naggar, "Long term electric load forecasting based on particle swarm optimization," *Appl. Energy*, vol. 87, no. 1, pp. 320–326, Jan. 2010.
- [16] L. Tang, X. Wang, X. Wang, C. Shao, S. Liu, and S. Tian, "Long-term electricity consumption forecasting based on expert prediction and fuzzy Bayesian theory," *Energy*, vol. 167, pp. 1144–1154, Jan. 2019.
- [17] O. Samuel, F. A. Alzahrani, R. J. U. H. Khan, H. Farooq, M. Shafiq, M. K. Afzal, and N. Javaid, "Towards modified entropy mutual information feature selection to forecast medium-term load using a deep learning model in smart homes," *Entropy*, vol. 22, no. 1, p. 68, Jan. 2020.
- [18] O. Olabode, O. Amole, T. Ajewole, and I. Okakwu, "Medium-term load forecasting in a Nigerian electricity distribution region using regression analysis techniques," in *Proc. Int. Conf. Math., Comput. Eng. Comput. Sci. (ICMCECS)*, Mar. 2020, pp. 1–5.
- [19] L. Han, Y. Peng, Y. Li, B. Yong, Q. Zhou, and L. Shu, "Enhanced deep networks for short-term and medium-term load forecasting," *IEEE Access*, vol. 7, pp. 4045–4055, 2019.
- [20] T. Ahmad and H. Chen, "Potential of three variant machine-learning models for forecasting district level medium-term and long-term energy demand in smart grid environment," *Energy*, vol. 160, pp. 1008–1020, Oct. 2018.
- [21] G. Dudek, "Pattern-based local linear regression models for short-term load forecasting," *Electr. Power Syst. Res.*, vol. 130, pp. 139–147, Jan. 2016.
- [22] 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.

- [23] H. Dagdougui, F. Bagheri, H. Le, and L. Dessaint, "Neural network model for short-term and very-short-term load forecasting in district buildings," *Energy Buildings*, vol. 203, Nov. 2019, Art. no. 109408.
- [24] H. Tabrizchi, M. M. Javidi, and V. Amirzadeh, "Estimates of residential building energy consumption using a multi-versus-optimizer-based support vector machine with k-fold cross-validation," *Evolving Syst.*, vol. 12, pp. 755–767, Sep. 2021, doi: [10.1007/s12530-019-09283-8](https://doi.org/10.1007/s12530-019-09283-8).
- [25] M. Morandin, S. Bolognani, and A. Faggion, "Active torque damping for an ICE-based domestic CHP system with an SPM machine drive," *IEEE Trans. Ind. Appl.*, vol. 51, no. 4, pp. 3137–3146, Jul. 2015.
- [26] R. R. J. Al Doury, S. Ozkan, and M. P. Mengüç, "Cascaded thermodynamic and environmental analyses of energy generation modalities of a high-performance building based on real-time measurements," *Entropy*, vol. 22, no. 4, p. 445, Apr. 2020.
- [27] Y.-T. Chen, "The factors affecting electricity consumption and the consumption characteristics in the residential sector—A case example of Taiwan," *Sustainability*, vol. 9, no. 8, p. 1484, Aug. 2017.
- [28] 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.
- [29] Y. Wang and J. M. Bielicki, "Acclimation and the response of hourly electricity loads to meteorological variables," *Energy*, vol. 142, pp. 473–485, Jan. 2018.
- [30] G. Černe, D. Dovžan, and I. Škrjanc, "Short-term load forecasting by separating daily profiles and using a single fuzzy model across the entire domain," *IEEE Trans. Ind. Electron.*, vol. 65, no. 9, pp. 7406–7415, Sep. 2018.
- [31] P. Lusis, K. R. Khalilpour, L. Andrew, and A. Liebman, "Short-term residential load forecasting: Impact of calendar effects and forecast granularity," *Appl. Energy*, vol. 205, pp. 654–669, Nov. 2017.
- [32] (Dec. 9, 2020). *World Weather Map*. Accessed: Aug. 14, 2021. [Online]. Available: <https://map.worldweatheronline.com/>
- [33] M. Karimi, H. Karami, M. Gholami, H. Khatibzadehazad, and N. Moslemi, "Priority index considering temperature and date proximity for selection of similar days in knowledge-based short term load forecasting method," *Energy*, vol. 144, pp. 928–940, Feb. 2018.
- [34] M. U. Fahad and N. Arbab, "Factor affecting short term load forecasting," *J. Clean Energy Technol.*, vol. 2, no. 4, pp. 305–309, 2014.
- [35] M. Barman and N. B. Dev Choudhury, "Season specific approach for short-term load forecasting based on hybrid FA-SVM and similarity concept," *Energy*, vol. 174, pp. 886–896, May 2019.
- [36] 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, Sep. 2017.
- [37] 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.
- [38] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Amsterdam, The Netherlands: Elsevier, 2011.
- [39] B. K. Patra, S. Nandi, and P. Viswanath, "A distance based clustering method for arbitrary shaped clusters in large datasets," *Pattern Recognit.*, vol. 44, no. 12, pp. 2862–2870, Dec. 2011.
- [40] M. Kefayat, A. L. Ara, and S. A. N. Niaki, "A hybrid of ant colony optimization and artificial bee colony algorithm for probabilistic optimal placement and sizing of distributed energy resources," *Energy Convers. Manage.*, vol. 92, pp. 149–161, Mar. 2015.
- [41] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, "Multi-sequence LSTM-RNN deep learning and metaheuristics for electric load forecasting," *Energies*, vol. 13, no. 2, p. 391, Jan. 2020.
- [42] M. Barman, N. B. D. Choudhury, and S. Sutradhar, "A regional hybrid GOA-SVM model based on similar day approach for short-term load forecasting in Assam, India," *Energy*, vol. 145, pp. 710–720, Feb. 2018.
- [43] 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.
- [44] X. Wang and S.-H. Ahn, "Real-time prediction and anomaly detection of electrical load in a residential community," *Appl. Energy*, vol. 259, Feb. 2020, Art. no. 114145.
- [45] F. Amara, K. Agbossou, Y. Dubé, S. Kelouwani, A. Cardenas, and J. Bouchard, "Household electricity demand forecasting using adaptive conditional density estimation," *Energy Buildings*, vol. 156, pp. 271–280, Dec. 2017.
- [46] S. Bruno, G. Dellino, M. La Scala, and C. Meloni, "A microforecasting module for energy management in residential and tertiary buildings," *Energies*, vol. 12, no. 6, p. 1006, Mar. 2019.
- [47] W. Ma, S. Fang, G. Liu, and R. Zhou, "Modeling of district load forecasting for distributed energy system," *Appl. Energy*, vol. 204, pp. 181–205, Oct. 2017.
- [48] B. Nepal, M. Yamaha, A. Yokoe, and T. Yamaji, "Electricity load forecasting using clustering and ARIMA model for energy management in buildings," *Jpn. Archit. Rev.*, vol. 3, no. 1, pp. 62–76, Jan. 2020.
- [49] Y. Chen, P. Xu, Y. Chu, W. Li, Y. Wu, L. Ni, Y. Bao, and K. Wang, "Short-term electrical load forecasting using the support vector regression (SVR) model to calculate the demand response baseline for office buildings," *Appl. Energy*, vol. 195, pp. 659–670, Jun. 2017.
- [50] M. Duan, A. Darvishan, R. Mohammaditab, K. Wakil, and O. Abedinia, "A novel hybrid prediction model for aggregated loads of buildings by considering the electric vehicles," *Sustain. Cities Soc.*, vol. 41, pp. 205–219, Aug. 2018.
- [51] Y. Liu, W. Wang, and N. Ghadimi, "Electricity load forecasting by an improved forecast engine for building level consumers," *Energy*, vol. 139, pp. 18–30, Nov. 2017.
- [52] E. Skomski, J.-Y. Lee, W. Kim, V. Chandan, S. Katipamula, and B. Hutchinson, "Sequence-to-sequence neural networks for short-term electrical load forecasting in commercial office buildings," *Energy Buildings*, vol. 226, Nov. 2020, Art. no. 110350.
- [53] H. Chitsaz, H. Shaker, H. Zareipour, D. Wood, and N. Amjadi, "Short-term electricity load forecasting of buildings in microgrids," *Energy Buildings*, vol. 99, pp. 50–60, Jul. 2015.
- [54] Z. Khan, T. Hussain, A. Ullah, S. Rho, M. Lee, and S. Baik, "Towards efficient electricity forecasting in residential and commercial buildings: A novel hybrid CNN with a LSTM-AE based framework," *Sensors*, vol. 20, no. 5, p. 1399, Mar. 2020.
- [55] A. Bracale, G. Carpinelli, P. De Falco, and T. Hong, "Short-term industrial load forecasting: A case study in an Italian factory," in *Proc. IEEE PES Innov. Smart Grid Technol. Conf. Eur. (ISGT-Europe)*, Sep. 2017, pp. 1–6.
- [56] A. Bracale, P. De Falco, and G. Carpinelli, "Comparing univariate and multivariate methods for probabilistic industrial load forecasting," in *Proc. 5th Int. Symp. Environ.-Friendly Energies Appl. (EFEA)*, Sep. 2018, pp. 1–6.
- [57] D.-H. Kim, E.-K. Lee, and N. B. S. Qureshi, "Peak-load forecasting for small industries: A machine learning approach," *Sustainability*, vol. 12, no. 16, p. 6539, Aug. 2020.
- [58] A. Bracale, P. Caramia, P. De Falco, and T. Hong, "A multivariate approach to probabilistic industrial load forecasting," *Electr. Power Syst. Res.*, vol. 187, Oct. 2020, Art. no. 106430.
- [59] Y.-F. Zhang and H.-D. Chiang, "Enhanced ELITE-load: A novel CMP-SOATT methodology constructing short-term load forecasting model for industrial applications," *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2325–2334, Apr. 2020.
- [60] S. Park, J. Moon, S. Jung, S. Rho, S. W. Baik, and E. Hwang, "A two-stage industrial load forecasting scheme for day-ahead combined cooling, heating and power scheduling," *Energies*, vol. 13, no. 2, p. 443, Jan. 2020.
- [61] Y. Hu, J. Li, M. Hong, J. Ren, R. Lin, Y. Liu, M. Liu, and Y. Man, "Short term electric load forecasting model and its verification for process industrial enterprises based on hybrid GA-PSO-BPNN algorithm—A case study of papermaking process," *Energy*, vol. 170, pp. 1215–1227, Mar. 2019.
- [62] H.-G. Son, Y. Kim, and S. Kim, "Time series clustering of electricity demand for industrial areas on smart grid," *Energies*, vol. 13, no. 9, p. 2377, May 2020.
- [63] Y. Wang, Y. Kong, X. Tang, X. Chen, Y. Xu, J. Chen, S. Sun, Y. Guo, and Y. Chen, "Short-term industrial load forecasting based on ensemble hidden Markov model," *IEEE Access*, vol. 8, pp. 160858–160870, 2020.
- [64] A. O. Hoori, A. A. Kazzaz, R. Khimani, Y. Motai, and A. J. Aved, "Electric load forecasting model using a multicolumn deep neural networks," *IEEE Trans. Ind. Electron.*, vol. 67, no. 8, pp. 6473–6482, Aug. 2020.
- [65] A. Ahmad, N. Javaid, M. Guizani, N. Alrajeh, and Z. A. Khan, "An accurate and fast converging short-term load forecasting model for industrial applications in a smart grid," *IEEE Trans. Ind. Informat.*, vol. 13, no. 5, pp. 2587–2596, Oct. 2017.
- [66] P. Kou and F. Gao, "A sparse heteroscedastic model for the probabilistic load forecasting in energy-intensive enterprises," *Int. J. Electr. Power Energy Syst.*, vol. 55, pp. 144–154, Feb. 2014.
- [67] S. Ungureanu, V. Topa, and A. Cziker, "Industrial load forecasting using machine learning in the context of smart grid," in *Proc. 54th Int. Univ. Power Eng. Conf. (UPEC)*, Sep. 2019, pp. 1–6.
- [68] M. Tan, S. Yuan, S. Li, Y. Su, H. Li, and F. H. He, "Ultra-short-term industrial power demand forecasting using LSTM based hybrid ensemble learning," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 2937–2948, Jul. 2020.

- [69] D.-M. Petroșanu, "Designing, developing and validating a forecasting method for the month ahead hourly electricity consumption in the case of medium industrial consumers," *Processes*, vol. 7, no. 5, p. 310, May 2019.
- [70] A. I. Saleh, A. H. Rabie, and K. M. Abo-Al-Ez, "A data mining based load forecasting strategy for smart electrical grids," *Adv. Eng. Informat.*, vol. 30, no. 3, pp. 422–448, Aug. 2016.
- [71] N. Mohan, K. Soman, and S. S. Kumar, "A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model," *Appl. Energy*, vol. 232, pp. 229–244, Dec. 2018.
- [72] X. Kong, C. Li, C. Wang, Y. Zhang, and J. Zhang, "Short-term electrical load forecasting based on error correction using dynamic mode decomposition," *Appl. Energy*, vol. 261, Mar. 2020, Art. no. 114368.
- [73] N. Liu, Q. Tang, J. Zhang, W. Fan, and J. Liu, "A hybrid forecasting model with parameter optimization for short-term load forecasting of microgrids," *Appl. Energy*, vol. 129, pp. 336–345, Sep. 2014.
- [74] Y. Nie, P. Jiang, and H. Zhang, "A novel hybrid model based on combined preprocessing method and advanced optimization algorithm for power load forecasting," *Appl. Soft Comput.*, vol. 97, Dec. 2020, Art. no. 106809.
- [75] Y. Dai and P. Zhao, "A hybrid load forecasting model based on support vector machine with intelligent methods for feature selection and parameter optimization," *Appl. Energy*, vol. 279, Dec. 2020, Art. no. 115332.
- [76] A. Ghasemi, H. Shayeghi, M. Moradzadeh, and M. Nooshyar, "A novel hybrid algorithm for electricity price and load forecasting in smart grids with demand-side management," *Appl. Energy*, vol. 177, pp. 40–59, Sep. 2016.
- [77] Y. Yang, J. Che, C. Deng, and L. Li, "Sequential grid approach based support vector regression for short-term electric load forecasting," *Appl. Energy*, vol. 238, pp. 1010–1021, Mar. 2019.
- [78] F. He, J. Zhou, L. Mo, K. Feng, G. Liu, and Z. He, "Day-ahead short-term load probability density forecasting method with a decomposition-based quantile regression forest," *Appl. Energy*, vol. 262, Mar. 2020, Art. no. 114396.
- [79] M. Q. Raza, M. Nadarajah, D. Q. Hung, and Z. Baharudin, "An intelligent hybrid short-term load forecasting model for smart power grids," *Sustain. Cities Soc.*, vol. 31, pp. 264–275, May 2017.
- [80] I. P. Panapakidis, "Clustering based day-ahead and hour-ahead bus load forecasting models," *Int. J. Electr. Power Energy Syst.*, vol. 80, pp. 171–178, Sep. 2016.
- [81] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, "Single and multi-sequence deep learning models for short and medium term electric load forecasting," *Energies*, vol. 12, no. 1, p. 149, Jan. 2019.
- [82] Y. Sun, X. Ma, J. Xu, Y. Bao, and S. Liao, "Efficient utilization of wind power: Long-distance transmission or local consumption?" *Frontiers Mech. Eng.*, vol. 12, no. 3, pp. 440–455, Sep. 2017.
- [83] D. P. E. Silva, J. L. F. Salles, J. F. Fardin, and M. M. R. Pereira, "Management of an island and grid-connected microgrid using hybrid economic model predictive control with weather data," *Appl. Energy*, vol. 278, Nov. 2020, Art. no. 115581.
- [84] L. Lozano, E. M. Querikiol, M. L. S. Abundo, and L. M. Bellotindos, "Techno-economic analysis of a cost-effective power generation system for off-grid island communities: A case study of Gilutongan Island, Cordova, Cebu, Philippines," *Renew. Energy*, vol. 140, pp. 905–911, Sep. 2019.
- [85] A. A. Bashir, M. Pourakbari-Kasmaei, J. Contreras, and M. Lehtonen, "A novel energy scheduling framework for reliable and economic operation of islanded and grid-connected microgrids," *Electr. Power Syst. Res.*, vol. 171, pp. 85–96, Jun. 2019.
- [86] A. Islam, S. R. Hasib, and M. S. Islam, "Short term electricity demand forecasting for an isolated area using two different approaches," *J. Power Technol.*, vol. 93, no. 4, pp. 185–193, 2013.
- [87] S. Yuan, A. S. Kocaman, and V. Modi, "Benefits of forecasting and energy storage in isolated grids with large wind penetration—The case of Sao Vicente," *Renew. Energy*, vol. 105, pp. 167–174, May 2017.
- [88] T. Ahmad and H. Chen, "Nonlinear autoregressive and random forest approaches to forecasting electricity load for utility energy management systems," *Sustain. Cities Soc.*, vol. 45, pp. 460–473, Feb. 2019.
- [89] A. Heydari, M. M. Nezhad, E. Pirshayan, D. A. Garcia, F. Keynia, and L. De Santoli, "Short-term electricity price and load forecasting in isolated power grids based on composite neural network and gravitational search optimization algorithm," *Appl. Energy*, vol. 277, Nov. 2020, Art. no. 115503.
- [90] I. S. Jahan, M. Prilepol, 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.
- [91] T. Turai, I. Ballard, and R. Rob, "Short-term electrical load demand forecasting using artificial neural networks for off-grid distributed generation applications," in *Proc. Saudi Arabia Smart Grid (SASG)*, Dec. 2017, pp. 1–7.
- [92] S. Chapaloglou, A. Nesiadis, P. Iliadis, K. Atsonios, N. Nikolopoulos, P. Grammelis, C. Yiakopoulos, I. Antoniadis, and E. Kakaras, "Smart energy management algorithm for load smoothing and peak shaving based on load forecasting of an island's power system," *Appl. Energy*, vol. 238, pp. 627–642, Mar. 2019.
- [93] T. Mikolov and G. Zweig, "Context dependent recurrent neural network language model," *SLT*, vol. 12, pp. 234–239, 2012.



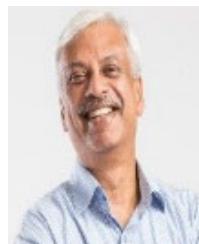
ABDUL AZEEM received the bachelor's degree in electronics engineering with majors in electrical power from IQRA University (IU), Karachi, Pakistan, in 2014, and the master's degree in engineering management with specialization in "energy management" from the Electrical Engineering Department, N. E. D. University of Engineering and Technology, Karachi, in 2018. He is currently pursuing the Ph.D. degree in electrical and electronics engineering with the Universiti Teknologi PETRONAS, Malaysia. He is a Registered Engineer at Pakistan Engineering Council (PEC). He has five years of professional work experience in energy sector. He is working on new algorithm development for electrical load forecasting for industrial and off-grid power plants. His research interests include artificial intelligence, load forecasting, energy management, renewable energy, smart grid, and smart cities.



IDRIS ISMAIL received the bachelor's degree in electrical engineering from Wichita State University, Wichita, KS, USA, in 1986, the master's degree in control system engineering from The University of Sheffield, U.K., in 2000, and the Ph.D. in electrical and electronics engineering from The University of Manchester, U.K., in 2009. He is currently an Associate Professor with the Universiti Teknologi PETRONAS, Malaysia, and a Registered Professional Engineer with the Board of Engineers, Malaysia. His current research interests include data analytics and model predictive control.



SYED MUSLIM JAMEEL is currently working as a Postdoctoral Scientist at the National University of Galway Ireland (NUIG), Ireland. He has vast experience and interest in intelligent algorithms design and development with proper mathematical modeling, statistics, and probability using deep learning. He particularly more focused on the computer vision applications for health informatics. He has several international publications in high reputed international journal and conferences. He is also an experienced Software Engineer and hands-on experience in a Python programming language with some specialized libraries, such as TensorFlow, Keras, OpenCV, Scikitlearn, YellowBrick, and others. He has established his knowledge of blockchain technology to attack the research issues of SMART CITIES with the amalgamation of deep learning and blockchain technologies.



V. R. HARINDRAN received the bachelor's degree in electrical and electronics engineering from the University of Huddersfield, U.K., in 1984, and the master's degree in technology management from the University of Staffordshire, U.K., in 2010. He is currently the Head Instrument at PETRONAS Refinery and Petrochemical Corporation Sdn. Bhd. (PRPC) and also an Adjunct Professor with the Universiti Teknologi PETRONAS, Malaysia. He is also a Registered Professional Engineer with the Board of Engineers, Malaysia.