Energy Load Forecasting Based on the Load Consumption Factors and Techniques Employed: A Review

Mrs. Swati Patil1, Dr. Mukund Kulkarni2, Ms. Swati Anil Patil3
1, 2 Bharti Vidyapeeth University, Kadamwadi, Kolhapur, Maharashtra, India.
3 Sanjay Ghodawat University, Aigre, Kolhapur, Maharashtra, India.

Emails: patilswati7922@gmail.com1, my_mukund@yahoo.com2, swati.pati1@sanjayghodawatuniversity.ac.in3

Abstract
Due to the contradiction between energy production and demand in the current energy crisis, power consumption (PC) is crucial to the world economy. Energy industry can improve power system control and energy usage by using machine learning (ML) models, which are widely acknowledged as a precise and computationally efficient prediction solution. For the purpose of projecting energy consumption and performance, machine learning (ML) techniques and artificial intelligence (AI) have been suggested recently. The various machine learning (ML) techniques i.e. artificial neural networks (ANN), support vector machines (SVM), Gaussian-based regressions and Fuzzy logic etc. that have been frequently used in predicting and enhancing energy performance are reviewed in-depth in this research.

Keywords: Artificial Intelligence; Energy Consumptions; Machine Learning

1. Introduction
A nation's economic development depends heavily on the electric power sector, which also significantly improves society through its effective functioning. The need for power has recently reached unfathomable heights as technology has advanced in cities all over the world. Global warming issues will worsen if national governments on a global scale are unable to meet this energy needs. Energy demand in India continues to rise at over 3% per year in the Stated Policies Scenario from 2021 to 2030. [1] According to the International Energy Agency's 2023 report, the average annual growth rate of total energy demand through 2030 is 0.7%. Demand is expected to rise until 2050. India is predicted to have the fastest-growing energy consumption globally over the next thirty years. By 2030, peak electricity demand has increased by almost 60% from 2022 [2]. Because of the growing population, the comfort levels of buildings, and people spending more time inside of them, the energy demand will therefore climb significantly in the near future. Peak consumption can present significant difficulties for electricity providers since they have to over-dimension the grid to accommodate the unusually high load of consumption. In order to prevent serious consequences like power outages, energy scarcity must be managed by the providers. An alternative strategy is to balance the grid by introducing intelligent methods for controlling these peaks, which will help prevent overreach and lower the costs associated with over-dimensioning and enormity. There are several ways to control the peaks, including load balancing and the development of dynamic and intelligent pricing strategies that take into account the fact that end users are price sensitive and may cut back on consumption when electricity prices are high. Therefore, it is necessary to forecast energy consumption in order to support national production and consumption balance. This review paper focuses on the different methods that are employed to forecast energy usage. There are three types of load forecasts: short-term forecasts, which are typically from one hour to one week; medium-term forecasts, which are typically from a week to a year; and long-term forecasts, which are typically more than a year. For the scheduling and control of power systems, short-term forecasting is crucial. [3] However, predicting daily load is a difficult process since it depends on
other factors, such as temperature and the calendar effect, in addition to the load from the previous days. As a result, anyone looking to save energy should focus on improving energy usage and the effectiveness of energy in buildings. We highlight a machine learning-based methodology for predicting energy usage in this review. [4] To make the most accurate prediction, it is crucial to identify the optimal strategy and model. Finding the answers to the following questions is the basis of this research:

1. What are the key elements that have a significant impact on how much energy is consumed?
2. How do the daily patterns of electric energy use in winter, rainy and summer differ?
3. Which machine learning techniques have the highest forecast accuracy limitations?
4. What advantages do various models have for accurately estimating how much electricity is used?

Accordingly, Figure 1 depicts the information flow. Apart from above information, the rest of the paper is divided into the following categories: [7-9] Second section discusses the factors affecting for load forecasting, and third section the performance metrics of the models, last section discusses shows the comparative study of the load forecasting techniques. Prediction category and duration shown in Table 1.

![Figure 1 Flow of Information in the Review Paper for Electricity Load Forecasting](image)

### 1.1. Categories of Electricity Load Forecasting

There are three categories into which electrical load forecasting can be divided, they are Long-term (1–20 years), Medium Term (a week or month), Short term (one week to one hour). [10] The short term forecasting (STF) is used to provide the essential data for scheduling generator start-up and shutdown, and performing a complete examination of gearbox constraints. System security is also assessed using the STF. In order to economically prepare for future additions to the generation system and additional transmission planning, long-term forecasting (LTF) is essential. The medium-term forecasting (MTF) is useful for other elements, primarily for determining tariffs, planning maintenance and repairs, scheduling fuel supply, and financial management. [5][6].

### Table 1 Prediction Category and Duration

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Prediction Category</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Long Term</td>
<td>months to year</td>
</tr>
<tr>
<td>2</td>
<td>Mid Term</td>
<td>weeks to a month</td>
</tr>
<tr>
<td>3</td>
<td>Short Term</td>
<td>Hours to week</td>
</tr>
</tbody>
</table>
2. Factors Affecting on Load Forecasting
To get a reliable forecast from the system, researchers must carefully choose the dependent components. Each component has a varied impact on load forecasting in varying degrees, and each model may take a variety of factors into account, some of which are listed below. [11-15] though other major and small aspects influence every phase of forecasting, time, weather, and economic considerations may be thought of as the main influences to take into account. Some of the factors are discussed as follows.

Time Factor:
In order to anticipate loads, time is the most crucial aspect. The load curve contains "time of the day" property together with "day of week", "week of month", and "month of season" properties, according to [16] observations of the load curves of many distinct grid stations.

Weather Factors:
According to seasonal patterns, winter months (October–February) accounted for 20% of overall consumption, while the summer months (May–August) contributed 50% to the total. [17] According to a study, demand rose 1.5% in tropical Indian towns for every 1°C increase in summertime temperatures. [18]

Economic Factor:
System average load and maximum demand are significantly impacted by economic factors including power pricing, load management, and industrialization level. [19-23] the accuracy of forecasting is also greatly influenced by factors such as customer behavior, tariff changes, appliance descriptions, population density, and equipment age and employment levels.

2.1. Performance Analysis of Different Models used in Prediction
Each machine learning pipeline includes performance metrics. As per previous study, various distinct metrics are employed to evaluate the models’ performance, shown in the equations below.

Mean Absolute Error:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |x_i - x|$$

Where:
- $n$ = the number of errors,
- $\Sigma$ = summation symbol (which means “add them all up”),
- $|x_i - x|$ = the absolute errors.

Mean Absolute Percentage Error:

$$M = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|$$

Where:
- $N$ is the number of fitted points,
- $At$ is the actual value,
- $Ft$ is the forecast value.
- $\Sigma$ is summation notation (the absolute value is summed for every forecasted point in time).

Root Mean Squared Error:

$$RMS E = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{N - P}}$$

Where:
- $Y_i$ is the actual value for the $i$th observation.
- $\hat{Y}_i$ is the predicted value for the $i$th observation.
- $N$ is the number of observations.
- $P$ is the number of parameter estimates, including the constant.

2.2. Comparative Analysis of Error Metrics of Different Forecasting Models
To demonstrate how accurate the load forecasting approaches are at predicting the actual load values, a number of criteria are employed as evaluation indices. [24-27] to measure the precision of their models, different researchers have relied on various statistical metrics. The most common static measurements used by academics around the world are shown in Table 2. According to the study, each of these measurements has benefits and drawbacks of its own. [28-32] mostly used metrics are MAPE, RMSE. The second-degree loss function provided by RMSE prioritizes larger errors rather than minor ones. It can measure average error naturally and is unambiguous. MAPE is simply applied to both large volume and low volume items and is not scale dependent. Figure 2 shows Prediction Matrices Studies.
### Table 2 Performance Matrices and Comparative Study of Various Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Reference</th>
<th>MAPE</th>
<th>RMSE</th>
<th>MAE</th>
<th>MSE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>[20]</td>
<td>3.38</td>
<td>4.51</td>
<td>2.19</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[21]</td>
<td>3.50</td>
<td>-</td>
<td>2.10</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[22]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.12</td>
<td>-</td>
</tr>
<tr>
<td>MLR</td>
<td>[23]</td>
<td>1.82</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>[24]</td>
<td>5.05</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.89</td>
</tr>
<tr>
<td>LSTM</td>
<td>[25]</td>
<td>51.45</td>
<td>0.8649</td>
<td>0.6278</td>
<td>0.7480</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[26]</td>
<td>-</td>
<td>3.337392</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[27]</td>
<td>6.27</td>
<td>0.0154</td>
<td>0.0107</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RF</td>
<td>[28]</td>
<td>4.60</td>
<td>6.10</td>
<td>-</td>
<td>-</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>[29]</td>
<td>5.45</td>
<td>3.219</td>
<td>-</td>
<td>-</td>
<td>0.576</td>
</tr>
<tr>
<td>SVM</td>
<td>[29]</td>
<td>1.9705</td>
<td>4.913</td>
<td>-</td>
<td>-</td>
<td>0.277</td>
</tr>
<tr>
<td></td>
<td>[30]</td>
<td>1.682</td>
<td>12.364</td>
<td>-</td>
<td>1.5438</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[31]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>KNN</td>
<td>[30]</td>
<td>3.6082</td>
<td>-</td>
<td>1.4826</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[32]</td>
<td>3.95</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FL</td>
<td>[33]</td>
<td>0.99</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[34]</td>
<td>1.94</td>
<td>202.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>XGBoost</td>
<td>[29]</td>
<td>-</td>
<td>2.914</td>
<td>-</td>
<td>-</td>
<td>0.711</td>
</tr>
<tr>
<td></td>
<td>[34]</td>
<td>2.95</td>
<td>368.4(kWh)</td>
<td>-</td>
<td>0.625(KW)</td>
<td>-</td>
</tr>
<tr>
<td>GBDT</td>
<td>[35]</td>
<td>-</td>
<td>4.748(%)</td>
<td>0.625(KW)</td>
<td>-</td>
<td>0.982</td>
</tr>
<tr>
<td>CNN</td>
<td>[36]</td>
<td>0.54</td>
<td>59.45</td>
<td>14745.41</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

#### 2.3. Models used for Predicting the Electricity Load Consumption

Table 3 shows comparative study of machine learning models for Prediction the review's objective is identifying the most popular algorithms used in electricity forecasting models. The sixty-one (61) article contains eleven (11) of the most frequently used algorithms. [33-37] the results of the study demonstrate how applying AI to different economic sectors can increase productivity and efficiency. According to a study, ANN models are frequently employed for prediction, and fuzzy logic is also growing in favor these days. Figure 3 shows the Comparative Studies.
Table 3 Comparative Study of Machine Learning Models for Prediction

<table>
<thead>
<tr>
<th>Categories of model</th>
<th>Reference</th>
<th>Area</th>
<th>Time factor</th>
<th>DataSize</th>
<th>Reason of prediction</th>
<th>Categories of Features</th>
<th>Performance Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuzzy Logic ANFIS</td>
<td>[46]</td>
<td>Residential</td>
<td>Hourly</td>
<td>4year</td>
<td>Short Term Load Forecasting</td>
<td>Time, day, temperature, season, daylight</td>
<td>3.72%RMSE, 2.1%MAPE, 1.85%MAPE, 3.66%RMSE</td>
</tr>
<tr>
<td>ANOVA ANN</td>
<td>[21]</td>
<td>Non Residential</td>
<td>Hourly</td>
<td>24years</td>
<td>Energy consumption in industry sectors</td>
<td>heterogeneity of industries, products, equipment, technologies, processes</td>
<td>3.5MAPE, 2.1MAE</td>
</tr>
<tr>
<td>ANN</td>
<td>[38]</td>
<td>Residential</td>
<td>Hourly</td>
<td>5months</td>
<td>Building Energy Prediction</td>
<td>wind direction, sun radiation, humidity level, and temperature of the outside dry bulb</td>
<td>0.0323(CV)</td>
</tr>
<tr>
<td>ANN Time series</td>
<td>[41]</td>
<td>Non residential</td>
<td>Hourly</td>
<td>8years</td>
<td>Accurate peak forecasting</td>
<td>Temperature and humidity, windspeed, cloud cover, daylight Hours</td>
<td>N/A</td>
</tr>
<tr>
<td>ANN</td>
<td>[38]</td>
<td>Residential</td>
<td>Hourly</td>
<td>1year</td>
<td>Sunlight exposure and power usage of the load</td>
<td>sun radiation, surrounding temperature, wind direction, and sampling time stamp</td>
<td>25.50%WMAE</td>
</tr>
<tr>
<td>CNN SVM</td>
<td>[45]</td>
<td>Residential</td>
<td>Hourly</td>
<td>2years</td>
<td>Forecasting of residential buildings</td>
<td>The building's size, occupant level, electric appliances, humidity and temperature, and calendar element</td>
<td>0.29MAPE, 1.32MAE, 2.99MSE, 0.23RMSE</td>
</tr>
<tr>
<td>FF-DNNR-DNN</td>
<td>[39]</td>
<td>Residential</td>
<td>Hourly</td>
<td>4years</td>
<td>Most dominant factors for power consumption prediction</td>
<td>The distribution of the lagging electrical load across time, weather, time, and holidays, and the price of electricity</td>
<td>1.42%MAPE 306(MW/h)E RMS 210(MW/h)MAE</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>ANN</td>
<td>[40]</td>
<td>Non Residential</td>
<td>Hourly</td>
<td>4years</td>
<td>Management and Operation of residential power System</td>
<td>Temperature</td>
<td>44,111.39451MS 160.99752MAE</td>
</tr>
<tr>
<td>ANN</td>
<td>[22]</td>
<td>Residential</td>
<td>Hourly</td>
<td>1year</td>
<td>cloud-based application for electricity consumption analysis</td>
<td>Sensors, cloud storage, clustering</td>
<td>1.12MAE</td>
</tr>
<tr>
<td>ANN PCA</td>
<td>[42]</td>
<td>Residential</td>
<td>Hourly</td>
<td>6 Months</td>
<td>Electricity demand of a bio-climatic building and influence of temp and solar radiation</td>
<td>Temperature of the dry bulb, sun radiation, humidity ratio, and wind speed</td>
<td>N/A</td>
</tr>
<tr>
<td>ANN PSO</td>
<td>[20]</td>
<td>Non-Residential</td>
<td>Hourly</td>
<td>3 years</td>
<td>Short term Load prediction</td>
<td>N/A</td>
<td>0.02191MAE 1.26268e–003MSE 0.03388MAPE</td>
</tr>
<tr>
<td>ANN</td>
<td>[43]</td>
<td>Residential</td>
<td>Hourly</td>
<td>One Week</td>
<td>the exact capacity to produce the power and the demand from customers</td>
<td>Time factor, historical weather information, customer class, previous load that the region demanded, region growth, and amount of load increased</td>
<td>1.066%MAPE</td>
</tr>
<tr>
<td>ANN</td>
<td>[44]</td>
<td>Residential</td>
<td>Hourly</td>
<td>One day</td>
<td>Construction of component load model for household</td>
<td>Solar radiation, ambient temperature, humidity</td>
<td>N/A</td>
</tr>
</tbody>
</table>

3. Major Findings with Research Gaps

**Duration-based model performance analysis**

The majority of current models only anticipate loads for the short term, as opposed to long- or mid-term periods. [38-40] this could be due to the fact that a longer forecasting horizon greatly increases the likelihood that changes that are not yet known to us will have an impact on demand in the future.

**The influence of the temporal domain**

The prediction capabilities of many models, which rely on either residential or non-residential buildings, are inadequate. Achieving equilibrium between
residential and nonresidential models could be feasible if such a model could be created. [41]

**Model Flexibility Issues**

The models that have been studied during the research are intended for a certain use case, such as a particular sort of input parameter or geographic area. Throughout the investigation, no model has been discovered that can function with the best accuracy regardless of application changes. [42]

**Hybrid Models**

Offer greater accuracy than single-method approaches in predicting load consumption.

4. **Future Trends**

4.1. **Model Enhancement**

Research evaluation indicates that no model takes into account the customer class and concentrates on year-round peak hour prediction. So need to put further effort on clustering techniques.

4.2. **Cloud based Applications**

The greatest accuracy is reported for clustering approaches when real-time electricity consumption data is collected using a cloud-based application for electricity consumption analysis. However, additional work on real-time data is required in the future.

**Conclusion**

A thorough analysis of several load forecasting models covering a range of time periods and sectors has been provided in this paper. We went over the effectiveness and drawbacks of the conventional load forecasting techniques. Lastly, drawing attention to the constraints, patterns, and areas in need of further study will help the readers choose and assess the models in accordance with their needs and obtain guidelines for future work. The goal of future research will be to improve the drawbacks of the current load forecasting models.

**References**

[1]. IEA (2023), World Energy Outlook 2023, IEA, Paris ttps://www.iea.org/reports/world-energy-outlook-2023, License: CC BY 4.0 (report); CC BY NC SA 4.0 (Annex A)


[8]. Isaac Kofi Nti, Moses Teimeh, Adebayo Felix Adekoya and OwusuNyarko-Boateng, “Forecasting Electricity Consumption Of Residential Users Based On Lifestyle Data Using Artificial Neural Networks”, ICTACT JOURNAL ON SOFT COMPUTING, APRIL 2020, VOLUME: 10, ISSUE: 03


analysis, decision tree and neural networks” Energy Volume 32, Issue 9, September 2007, Pages 1761-1768


[15]. Chirag Deb a, Fan Zhang b, Junjing Yang a, Siew Eang Lee a, Kwok Wei Shah a, “A review on time series forecasting techniques for building energy consumption”, Renewable and Sustainable Energy Reviews Volume 74, July 2017, Pages 902-924


[20]. Zahra Shafiei Chafi and Hossein Afrakhte, Hindawi “Short-Term Load Forecasting Using Neural Network and Particle Swarm Optimization (PSO) Algorithm” Mathematical Problems in Engineering Volume 2021


[25]. Tae-Young Kim, Sung-Bae Ch, “Predicting residential energy consumption using CNN-LSTM neural networks”, Energy, 2019

[26]. Widyaning Chandramitasari, Bobby Kurniawan, Shigeru Fujimura, “Building Deep Neural Network Model for Short Term Electricity Consumption Forecasting”, International Symposium on Advanced Intelligent Informatics (SAIN), 2018

[27]. Oussama Laib, Mohamed Tarek Khadir , Lyudmila Mihaylov, “Toward efficient energy systems based on natural gas consumption prediction with LSTM Recurrent Neural Networks”, Energy 2019

[28]. Muhammad Waseem Ahmad *, Monjur Moursheed, Yacine Rezgui, “Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption”, Energy and buildings, 2017

[29]. Rita Banik · Priyanath Das, Srimanta Ray,
Ankur Biswas, “Prediction of electrical energy consumption based on machine learning technique”, Electrical Engineering 2020

[30]. Sathishkumar V E, Changsun Shin & Yongyun Cho, “Efficient energy consumption prediction model for a data analytic-enabled industry building in a smart city”, Building Research & Information 2021


[32]. Daniel Ramo, Mahsa Khorram, Pedro Faria and Zita Vale, “Load Forecasting in an Office Building with Different Data Structure and Learning Parameters”, Forecasting 2021


[34]. Rachna Jain, Nikita Jain, Yash Gupta, Tarun Chugh, Tushar Chugh, D. Jude Hemanth, “A Modified Fuzzy Logic Relation-Based Approach for Electricity Consumption Forecasting in India”, Springer 2019


[37]. Kangji Li, Xianming Xie, Wenping Xue, Xiaoli Dai, Xu Chen, Xinyun Yang, “A hybrid teaching-learning artificial neural network for building electrical energy consumption prediction”, Energy & Buildings, 2018


