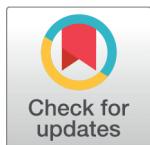




---

## RESEARCH ARTICLE



---

### OPEN ACCESS

Received: 20.10.2020

Accepted: 16.12.2020

Published: 29.12.2020

## Electricity requirement prediction using time series and Facebook's PROPHET

Rachel Joy Chadalavada<sup>1</sup>, Srinivasaiah Raghavendra<sup>2\*</sup>, V Rekha<sup>3</sup>

**1** PG Scholar, Department of CSE, School of Engineering and Technology, CHRIST Deemed to be University, Mysore Road, Bengaluru, 560074, India

**2** Associate Professor, Department of CSE, School of Engineering and Technology, CHRIST Deemed to be University, Mysore Road, Bengaluru, 560074, India. Tel.: +0-974-085-7501

**3** Assistant Professor, Department of CSE, School of Engineering and Technology, CHRIST Deemed to be University, Mysore Road, Bengaluru, 560074, India

**Citation:** Chadalavada RJ, Raghavendra S, Rekha V (2020) Electricity requirement prediction using time series and Facebook's PROPHET. Indian Journal of Science and Technology 13(47): 4631-4645. <https://doi.org/> 10.17485/IJST/v13i47.1847

\* Corresponding author.

Tel: +0-974-085-7501  
[raghav.trg@gmail.com](mailto:raghav.trg@gmail.com)

Funding: None

Competing Interests: None

Copyright: © 2020 Chadalavada et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment ([iSee](#))

ISSN  
Print: 0974-6846  
Electronic: 0974-5645

## Abstract

**Objectives:** The main objective of this research work is to forecast the electricity requirement of a particular household or an office or any building.

**Methods:** Forecasting is done using the PROPHET model which gives better results compared to other models like ARIMA and so on. Dataset considered here is a publicly available dataset called 'Appliance' dataset with what are all the appliances that are there in the particular household and number of appliances that are running on a day at every 10 minutes interval and so on. From the entire dataset, only two attributes are selected, and Log transformation is applied to the selected attributes. Finally, the PROPHET model is applied and the forecasting is done. **Findings:** The findings of the proposed models are: (i) Forecasting is done for the next 30 days based on different components like daily component, weekly component and trend component (ii) Wednesday is the lowest power utilization day and, power utilization increases till Saturday and Saturday is the highest power utilization day (iii) PROPHET model makes predictions very accurate based on the future data and is easy to make predictions compared to ARIMA. **Novelty:** Models were trained on the dataset from January 11, 2016, to May 27, 2016 interval which is at 10 min for about 4.5 months. The models anticipated qualities that could be viewed as effective in the 30-day forecast. The accomplishment of the model in the time series expectation was investigated and analyzed.

**Keywords:** Electricity; forecasting; PROPHET

---

## 1 Introduction

Forecasting is pretty much required everywhere nowadays. Forecasting helps in better future design. Energy consumption is increasing day after day, which in the future may lead to a lack of energy also. So, as and how product sale is being forecasted, the weather is being forecasted and so on, so should energy consumption be forecasted. This energy consumption forecasting will help in knowing the consumption of electricity that helps in securing the future. Electricity requirement forecasting helps in predicting

and analyzing the electricity consumption and taking proper actions accordingly. There have been many models used to forecast electricity consumption, where the results are not accurate.

The author proposes that the PROPHET performs better than ARIMA. Data processing steps such as timestamp conversion and feature selection are carried and, for selecting the model, a threefold splitting technique that gave good results for training, validation, and testing was employed. A 90-day forecast is done using both the models using the dataset. In the end, both models are compared in terms of performance metrics. So, by considering the result the PROPHET model beats ARIMA in  $R^2$  values and is chosen for forecasting<sup>(1)</sup>.

Two ARIMA models were used to forecast hourly prices of energy consumption which are the Spanish model required five hours in predicting future prices, and the California model took two hours. The error rate for the Spanish model was around 10% and the California model was around 5%. Nature and results reported earlier are some of the reasons for the error rate of price time series<sup>(2)</sup>.

A baseline model is used for analyzing the energy savings using the bills of electricity consumption in different seasons like cold winter and hot summer. Baseline models help in knowing the amount of energy used based on the energy baseline period. Eleven office building energy consumption bills are used for the research. From the results obtained it was noted that the monthly mean outside dry-bulb heat was the key variable and was enough to trace and use the baseline energy in winter and summer conditions<sup>(3)</sup>.

Generally, the functionalities of the power grid systems will be increasing which is directly proportional to access failures also. The model-fitting analysis, to reduce the frequency of access failures was employed. 404 failure code is used as train data and SARIMA, FbPROPHET, holt-winter, and GM algorithms are used to build a time series prediction model. Based on RMSE, SARIMA is recommended as a predictive model for power grid systems<sup>(4)</sup>.

A model based on the data-driven modeling for predicting the energy consumption of a building was proposed. The main model used is time series analysis to come up with a data-driven method. This model is used to forecast the energy consumption of a building. After launching, authenticating, examining, and assessing, it was found that the time series analysis is the best and accurate model for predicting building electricity consumption<sup>(5)</sup>.

A multi-model partitioning theory method was used for electricity demand load forecasting. Later the performance of the method is compared with other methods like AICC, AIC, and BIC. The proposed model is applied to the Hellenic power system which further proved to be reliable and effective and also mentions its usefulness in the analysis related to the electricity utilization and its price predictions<sup>(6)</sup>.

A new forecasting approach for time series data centered on the resemblance of pattern classifications is proposed. Grouping is done using clustering methods, and then labeling is done. The model was applied for predicting the electricity charges and need time series of Australian, Spanish, and American markets and provides good competitive results<sup>(7)</sup>.

Two forecasting models namely transfer function and dynamic regression based on time series analysis are proposed and applied to California and Spanish electricity markets. Based on the outcome of the implementation the average error for the Spanish market is 5% and the California market is 3%. The results are accurate which can be further exercised by both the producer and clients to create their bidding plans<sup>(8)</sup>.

Precise forecasting of electricity demand has many advantages: i). It saves the functioning and the maintenance price, ii). Improves the consistency of power allocation and distribution system and, iii). Can make the appropriate decisions for the future. In Malaysia, UTHM is a developing University and growing day by day and it is important for UTHM to predict the consumption of electricity. For this, the UTHM used different models such as SMA, WMA, SES, HL, HW, and CMA. Based on the outcome it was found that HW might forecast the consumption in a better way, centered on the smallest MAPE<sup>(9)</sup>.

The authors implemented electricity forecasting for Texas using R. The electricity data was collected from the ERCOT data bank. The data was arranged in the supply Vs demand format. The proposed model is implemented in R programming language and the results are compared with the existing information on energy consumption. Time series forecasting can provide information about the future electricity requirement for the utility operators. These predictions can assist as knowledge to utility operators to plan and determine when the highest shaving is needed<sup>(10)</sup>.

Electricity forecasting is done using data mining methods. The main aim of this is to make use of the past data and to determine the future electricity requirement. Data mining models like SVR, Random Forest, and M5P are used for forecasting. The paired T-test was used for identifying the best model based on the difference between the errors of each model. Based on the experimental outcome it was found that the SVR model performs better than the other models that are considered<sup>(11)</sup>.

PCA was used to find if any variables are correlated, which affects other variables like moisture, solar energy, and heat. A new climatic symbol Z was determined. For correlating the energy use and simulated daily cooling with the corresponding Z, regression models are used. On analyzing the errors, regression models gave close simulated values. These can be used in predicting climate change effects also<sup>(12)</sup>.

WEKA time series was used in forecasting energy demand for different seasons like rainy, summer, and winter. The data is from the TNEB of monthly energy consumption in the domestic category. Learning algorithms like Support Vector Machine, Multilayer Perceptron, Gaussian Process, Linear Regression are used. Later, Mean Absolute Error, and Direction Accuracy is calculated and compared with all the learning algorithms, and Support Vector Machine is chosen for electricity demand forecasting on seasonal based data<sup>(13)</sup>.

The authors are focusing on choosing the best forecasting periods among quarterly, daily, yearly, and monthly. Models like ARIMA and ARMA are used for data analysis. The best model with the best period to forecast is found by the smallest value of AIC and RMSE. From the outcome of the experiment, it was found that the ARIMA performed better with the monthly and quarterly data and ARMA performed better with the daily and the weekly data. The predicting time was appropriate for the short term such as 6 months, 28 days, 5 weeks, and 2 quarters<sup>(14)</sup>.

An approach using a model grounded on X12 and periodic and tendency decomposition using STL decomposition of scheduled electricity utilization predicting method is projected. Firstly, the STL model was used to generate the power utilization time series decomposition rendering to the attributes of electricity utilization per month. This divides the monthly electricity utilization into tendency, time of year, and arbitrary mechanisms. At last, the modification in the features of these mechanisms over time is selected for deliberation. To conclude the proper method is designated to forecast the mechanisms in the reshaping of the monthly electricity utilization estimate<sup>(15)</sup>.

The Household requirement of electricity is nonstop rising which directs to extreme greenhouse gas emission. A Thorough investigation of electricity utilization features for housing is desirable to advance efficacy, convenience, and to plot in advance for times of extreme electricity requests. A model based on ANN is projected to forecast the energy utilization of housing in Auckland 24h prior with further accurateness than the standard methods. The properties of five climate variables on energy utilization were studied. Additionally, the method was investigated with three dissimilar training methods, the scaled conjugate gradient, the Levenberg-Marquadt (LM), and Bayesian regularization and their consequence on forecast accurateness was studied<sup>(16)</sup>.

An exhaustive assessment of various statistical methods that can be employed to analyze and predict the electricity cost and loads is discussed with the help of 16 case studies. The study also covers detail regarding the seasonal decomposition, heavy distributions, spike processing, exponential smoothing, mean revision, and the different time-series models<sup>(17)</sup>.

The two main reasons for the increase in global energy usage are the rapid increase in population and urban development. The accurate prediction of the long-term peak load is most important for saving money and time corresponding to a country's power production. An investigation was made to compare the performances of the Holt-winters method with the PROPHET model for the long-time load predictions in Kuwait. The data generated from the Kuwait power plants between 2010 and 2020 was used for predicting the peak load between 2020 and 2030. Five metrics were used for comparing the performances of the said methods. From the experimental results, it was found that the PROPHET model achieves more accurate prediction when compared to the Holt-winters concerning the generalization test<sup>(18)</sup>.

Many factors such as time, electricity cost, season, economy, and weather are responsible for fostering electricity utilization. These factors have the features of many linear, periodic, and random changes. PROPHET model was proposed to do analyze the trend, and the results were compared with the ARIMA model. From the comparison, it was found that the PROPHET model achieved better accuracy. From the outcome, it was also found that the power utilization of shopping malls was most affected by high-temperature weather and holidays to some extent when compared with the utilization of the office buildings<sup>(19)</sup>.

To predict the long-range, and the short-range air pollution in Seoul, Korea, the PROPHET model was employed. The pollutants predicted in the work were PM10, SO2, CO, NO2, PM2.5, and O3. After long-time exposure, these pollutants leads to various health degradation. Present models based on the chemical methods for forecasting air pollution needs complex requirements and are difficult to use. ML model along with additional infrastructure was used to forecast the pollution depending on time. A three years data that measures air quality on an hourly basis was collected and the model was optimized using the PROPHET model. The forecasting for 2019 was done by using cross-validation on the 2017-18 data. Three statistical methods such as MSE, MAE, and RMSE were used to determine the performance of the model based on accuracy. From the output, it was found that the PROPHET model performed better compared to the other equivalent methods<sup>(20)</sup>.

For predicting energy utilization AI models are widely used in the past decade. The model based on the AI such as ANN, and SVR and the conventional models such as gray models, regression, and time series are compared based on the MAPE. Based on the experimental outcome, it was found that for the yearly analysis, conventional models are more favored<sup>(21)</sup>.

Different methods for predicting electricity requirement based on hourly demand was done for 10-50 years in advance. The challenges faced by these methodologies for future energy methods consisting of renewable energy bases and tight coupling between the power, building, and transport sectors<sup>(22)</sup>.

To find the model that is suitable for electricity utilization prediction about 113 dissimilar case studies are compared. For comparison, different criteria's such as inputs, time frame, outputs, error type, size of the data sample, scale, and value are considered. ML algorithms such as ANN, time-series, and SVM are considered. Based on different case studies considered, the meta-data analysis leads to the recognition of the finest traditions for predicting electricity utilization and power load forecasting. Based on the meta-data analysis, a nomenclature was defined to assist the researcher to make an informed decision and to select the best model for the problem<sup>(23)</sup>.

The quick transformation in the electricity divisions enhances both the chances and the requirement of Data Analytics (DA). As many methods and areas are emerging, it is necessary to combine and build scientific work. So, a systematic review of many fields mainly concerned with using the Data Analytics model in the setting of electricity was conducted. A qualitative review of about 200 case studies concerned with the analysis of the important application of DA in different areas of electricity divisions such as production, trading, distribution, and utilization was presented. For each area under consideration, the different DA methods and applications are reviewed, and the ideas for future research are also highlighted<sup>(24)</sup>.

Energy is the essence of today's world. In the past, the emission of CO<sub>2</sub> and the energy utilization was enhanced due to the population explosion and the luxury requirement of the individuals. Predicting energy utilization is the main requirement for planning, managing, and conserve energy. For the prediction of energy utilization, the data-driven methods provide practical models. ML algorithms are employed for the forecasting and evaluation is done using the performance procedures and the upcoming scope for energy prediction are emphasized<sup>(25)</sup>.

Load forecasting is considered as one of the vital parts in the scheduling and process of electricity services. Because of the technological enhancements, transformation in the economic circumstances, load prediction is most important. The load impact issues and the steps followed at various time limits may be affected by the prediction and vice versa. The most important challenge is to accurately predict future demand<sup>(26)</sup>.

Planning, aim set, and irregularity recognition are the main tasks of data science in prediction. Even though the prediction is most important, it is accompanied by the challenge of generating more accurate predictions. To handle this challenge, a regression model with explainable constraints that can be adjusted by analysts was proposed along with the analysis to relate and evaluate the prediction procedures and automatically can indicate the modification<sup>(27)</sup>.

In recent days forecasting has become a major and most required technique which uses past and present data in making predictions for the future. Presently, many organizations are using information that is stored for forecasting. This data helps in predicting the future use of electricity which in turn makes us aware of the energy consumption and energy savings that are to be done in the future.

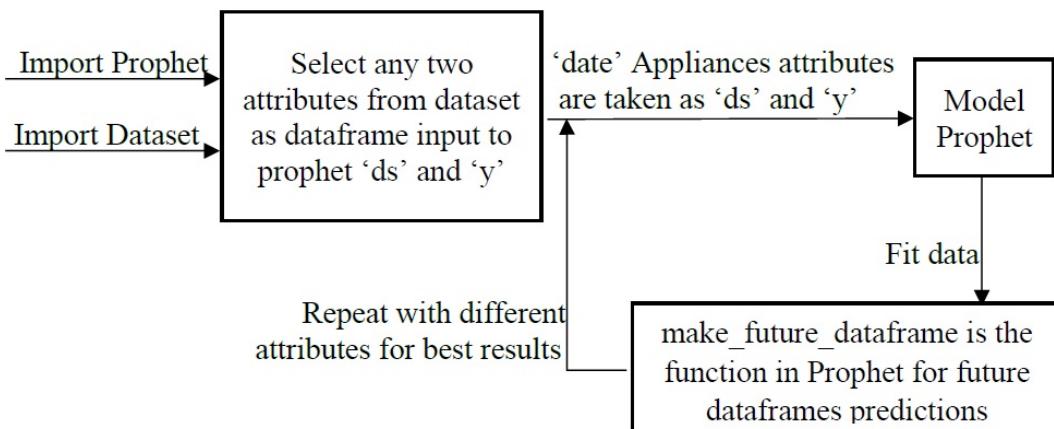
So, in our proposed work, the PROPHET model is used which gives better accuracy compared to other models like ARIMA. The paper is sorted out as, Section 2 explains about Facebook PROPHET, Section 3 describes the Framework used in the proposed work, Section 4 describes the Dataset considered, Section 5 details the Log transformation applied on the dataset, Section 6 explains the PROPHETs algorithm, in Section 7 we discuss the Implementation and Results and in Section 8 we conclude the proposed work.

## 2 Facebook PROPHET

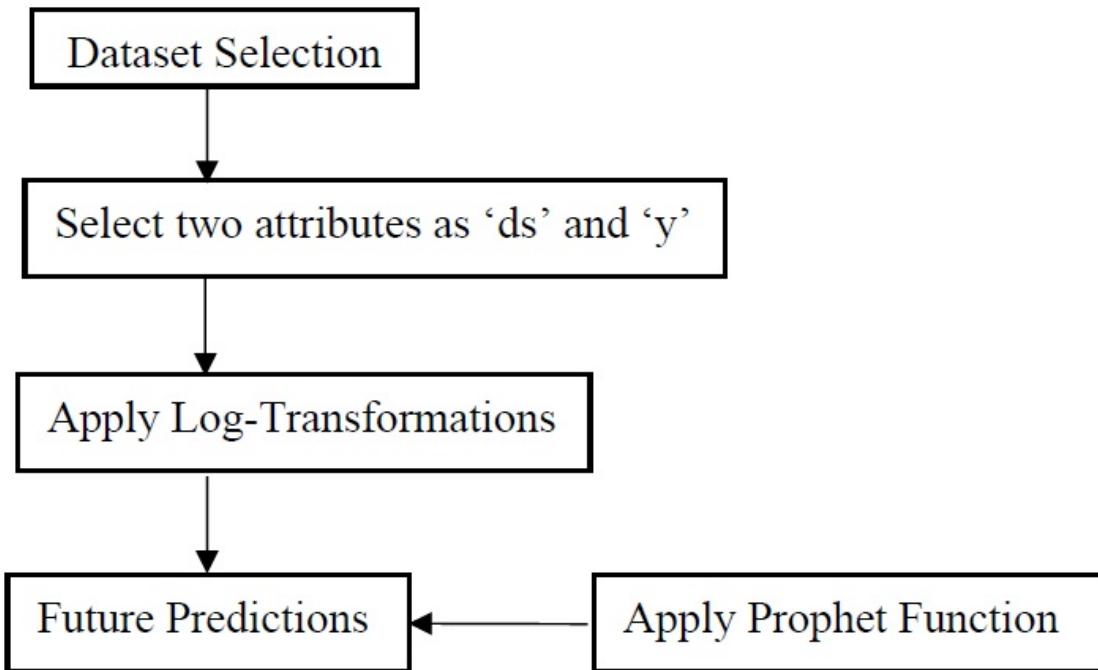
PROPHET is used in the forecasting process which is implemented in Python as well as R. PROPHET is quick and gives totally computerized estimates that can be tuned by hand by experts. It helps in business forecasting for Facebook with time, daily, weekly observations of history within a year, large outliers, etc. It helps in forecasting time series data using models in which non-linear patterns are fit with yearly, week by week, and everyday regularity in addition to holiday impacts. It gives the best results with time series which have strong seasonal impacts. It is strong to missing data and patterns in trends, changes in time series, and handles outliers well. It is an open-source library distributed by Facebook that depends on trend+seasonality+holiday models. It gives the capacity to make time series predictions with great exactness utilizing simple parameters and has support for including the effect of custom seasonality and occasions. It is used in various applications on Facebook in getting accurate forecasts for the future. It was found to perform better than any other models used for forecasting in many instances. The model implementation is the same in both R and Python with Stan code which is used for statistical modeling with high-performance computations for fitting and the forecasts to be done in seconds. It gives better forecasts on clumsy data with zero manual work. It adds on with various possibilities for the user in adjusting the forecasts. PROPHET uses the sklearn model which features many clustering algorithms, classification, and regression. PROPHET class can call the required fit and predict methods. Input datafram has two attributes 'ds' and 'y'. Attribute 'ds' is the datestamp column which has to be in Pandas library format, ideally YYYY-MM-DD for a date or YYYY-MM-DD HH:MM:SS for the timestamp. Column 'y' should be numeric and represents the measurement to be forecasted.

### 3 Framework

Dataset taken for this research is an ‘Appliance’ dataset with all the appliances that are there in the particular household and the number of appliances that are running on a day at every 10 minutes interval and so on. PROPHET which is the model being used for prediction is imported. From the entire dataset, only two attributes are to be taken as dataframe input. PROPHET has a dataframe input of only two attributes which are ‘ds’ and ‘y’ where the actual attributes are to be assigned or renamed as ‘ds’ and ‘y’. These attributes are assigned to the PROPHET model where the data is fit to the model. Later with make.future.dataframe function which is a function of the PROPHET model is used to make future data for which the forecast is to be done and predictions are made. For evaluating the results or for getting better results different attributes can be used to compare the results and go for the best predictions. [Figure 1](#) shows the model diagram of the proposed work.



**Fig 1.** Model diagram for the proposed work



**Fig 2.** Framework for the proposed work

**Figure 2** shows the framework of the way the research work is carried out. The steps followed in the proposed work are:

Step 1: The Appliance Energy Prediction Data Set with 19735 rows and 27 columns is imported from the UCI Machine Learning Repository.

Step 2: Required two columns are taken for data transformation. Only two columns are taken as the input framework of PROPHET has only two columns which are 'ds' and 'y'.

Step 3: Log-Transformation is applied to 'y' which is the 'Appliances' column from the dataset for data transformation.

Step 4: With 'ds' and 'y' attributes from the dataset, PROPHET is called and forecasting is done for the next 30 days.

Step 5: Future predictions are made using PROPHET with its basic functionalities like weekly component, daily component, and trend component.

## 4 Dataset Description

The dataset taken is from the UCI Machine Learning Repository. The 'Appliance' Energy Prediction Data Set consists of 19735 instances and 21 attributes. The dataset is at 10 minutes for about 4.5 months. The dataset is collected from an airport weather station of a household in Belgium and is shown in **Table 1**.

**Table 1.** Dataset description

Attribute	Meaning
Date	date and time, year-month-day hour:minute:second
Appliances	energy use, in Wh
Lights	energy use of light fixtures in the house, in Wh
T1	Temperature in the kitchen area, in Celsius
RH_1	Humidity in the kitchen area, in %
T2	Temperature in living room area, in Celsius
RH_2	Humidity in the living room, in %
T3	Temperature in laundry room area, in Celsius
RH_3	Humidity in the laundry area, in %
T4	Temperature in office room, in Celsius
RH_4	Humidity in the office room, in %
T5	Temperature in bathroom, in Celsius
RH_5	Humidity in the bathroom, in %
T6	Temperature outside the building, in Celsius
RH_6	Humidity outside the building (north side), in %
T7	Temperature in ironing room, in Celsius
RH_7	Humidity in the ironing room, in %
T8	Temperature in teenager room 2, in Celsius
RH_8	Humidity in the teenager room 2, in %
T9	Temperature in parents room, in Celsius
RH_9	Humidity in the parents room, in %

From the entire dataset, only two attributes are taken as the PROPHET model which is being used in the research has only two attributes as input for forecasting. The snapshot of the dataset with two attributes is shown in **Figure 3**.

1	date	Appliance
2	11-01-2016 17:00	60
3	11-01-2016 17:10	60
4	11-01-2016 17:20	50
5	11-01-2016 17:30	50
6	11-01-2016 17:40	60
7	11-01-2016 17:50	50
8	11-01-2016 18:00	60
9	11-01-2016 18:10	60
10	11-01-2016 18:20	60
11	11-01-2016 18:30	70
12	11-01-2016 18:40	230
13	11-01-2016 18:50	580
14	11-01-2016 19:00	430
15	11-01-2016 19:10	250
16	11-01-2016 19:20	100
17	11-01-2016 19:30	100
18	11-01-2016 19:40	90
19	11-01-2016 19:50	70
20	11-01-2016 20:00	80
21	11-01-2016 20:10	140
22	11-01-2016 20:20	120
23	11-01-2016 20:30	190

**Fig 3.** Two attributes taken as 'ds' and 'y'

## 5 Log transformation

Information change is the way toward taking a scientific capacity and applying it to the information. This segment talks about a typical change known as the log change. Every factor  $x$  is replaced with  $\log(x)$ , where the base of the log is surrendered over to the investigator. It is viewed as normal to utilize base 10, base 2, and the regular log.

This procedure is helpful for compacting the y-pivot while plotting histograms. For instance, on the off chance that has a huge scope of information, at that point, littler qualities can get overpowered by the bigger qualities. Taking the log of every factor empowers the representation to be more clear. A case of this is the number of ports on a framework. There are 65,535 ports accessible on a framework, in the event that are endeavoring to imagine traffic to every one of them, at that point this perception could conceal values on the lower scope of ports while endeavoring to show higher ports. Log change likewise underscores exceptions and permits us to possibly get a chime molded conveyance. The thought is that taking the log of the information can reestablish evenness to the information.

## 6 PROPHETs algorithm

PROPHET follows the sklearn model API and the working of the model is given below:

1. PROPHET just accepts the information as a dataframe with a ds (datestamp) and y. So first, how about converting the dataframe to the fitting organization.
2. Make an example of the PROPHET class and afterward fit the dataframe to it.

3. Make a dataframe with the dates for which need a prediction to be made with `makefuturedataframe()` function. At that point determine the number of days to gauge utilizing the `period's` parameter.
4. Call `predict` function to make a prediction and store it in the estimated dataframe. What's perfect here is that you can review the dataframe and see the predictions just as the lower and upper limits of the vulnerability interim.

PROPHET algorithm uses Eq. (1)

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad (1)$$

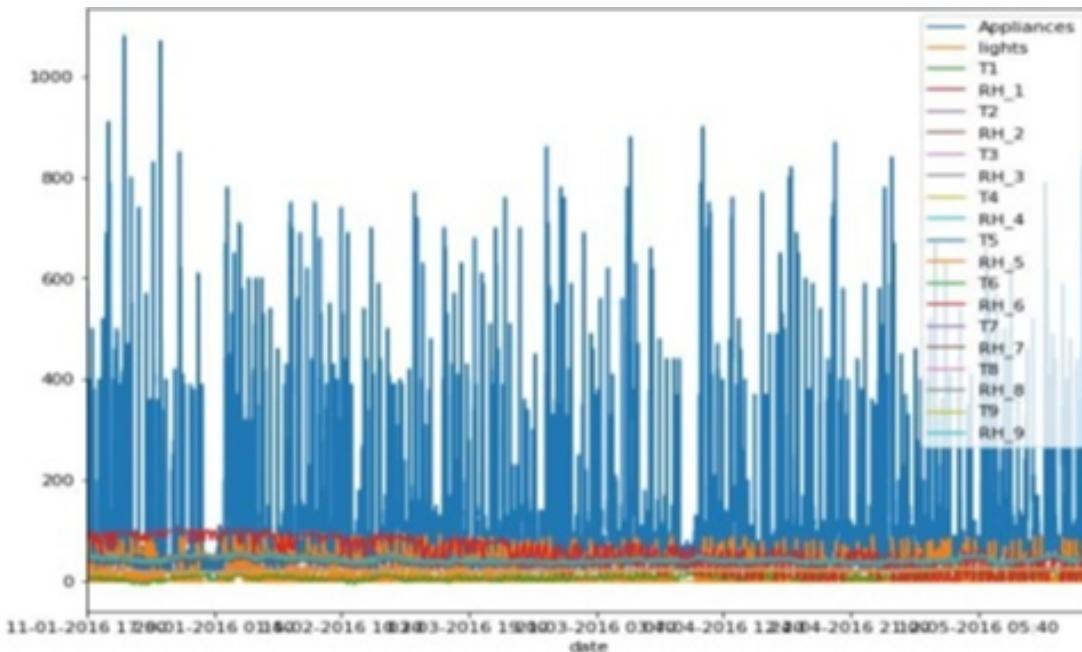
here,

1.  $g(t)$  models pattern, which portrays abatement in the information. PROPHET joins two pattern models: immersing development model, and a piece-wise direct model, contingent upon the sort of gauging issue..
2.  $s(t)$  models regularity with Fourier arrangement, which portrays how information is influenced via occasional factors.
3.  $h(t)$  models the impacts of occasions or huge occasions that sway business time series and

$\epsilon_t$  speaks to an unchangeable error term.

## 7 Implementation and Results

PROPHET follows the sklearn model API. The input dataframe for PROPHET is 'ds' and 'y' where 'ds' is datestamp - YYYY-MM-DD, timestamp - YYYY-MM-DD HH:MM:SS and 'y' is the forecast that is to be done. The 'y' section must be numeric and comprises the estimation to be determined. PROPHET helps in simple forecasting considering two fields, from which we wish to conjecture. [Figure 4](#) shows the diagram representing the informational index before the change. The diagram in [Figure 4](#) shows the date and rest of the other segments, which is incomprehensible in forecasting. This is one reason PROPHET is liked, where it considers two segments 'ds' and 'y' which would be the input dataframe for PROPHET. Along these lines, considering the dataframe, 'ds' which is the datestamp will take the 'date' segment and 'y' will take the 'Appliances' segment from the dataset.



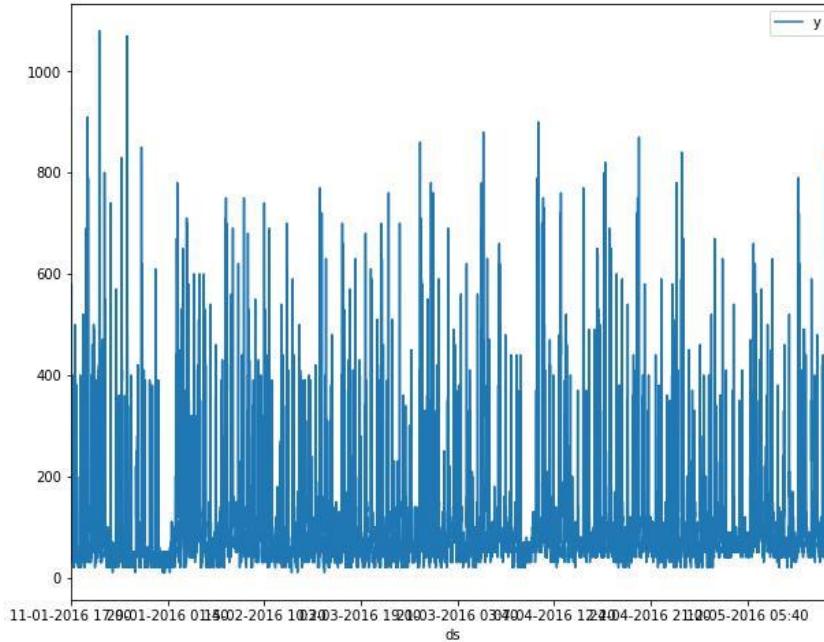
**Fig 4.** Graph representing the original dataset

After the selection of the two attributes 'ds' and 'y' the log-transformation is applied. The log-transformed values for the 'Appliances' column which is 'y', the second input for the PROPHET dataframe showed in [Figure 5](#). PROPHET helps in easy forecasting considering two fields, from which we wish to forecast.

0	4.094345
1	4.094345
2	3.912023
3	3.912023
4	4.094345
5	3.912023
6	4.094345
7	4.094345
8	4.094345
9	4.248495
10	5.438079
11	6.363028
12	6.063785
13	5.521461
14	4.605170
15	4.605170
16	4.499810
17	4.248495
18	4.382027
19	4.941642
20	4.787492
21	5.247024
22	4.700480
23	4.700480
24	4.700480
25	4.700480

**Fig 5.** Log-transformation values

The log change is, seemingly, the most acclaimed among the different kinds of changes used to change slanted information to generally acclimate to typicality. In case the original information follows log-ordinary dissemination, by then, the log-changed information follows a typical or near typical dispersion. Thus, a log change is applied to section 'Machines' and a chart is plotted on the changed information for 'date' and 'Appliances'. [Figure 6](#) shows the chart of 'ds' and 'y' which is log-changed information.



**Fig 6.** Graph representing 'ds' and 'y'

Figure 7 shows the parameters for which the model is built. Now with these parameters, the forecasting is done. The parameters include 'k', 'm', 'delta', 'sigma-obs', 'beta', 'trend', 'Y', 'beta-m', and 'beta-a'.

```

{'k': array([[-0.13604014]]),
 'm': array([[0.09801003]]),
 'delta': array([[ -7.79787879e-02, -1.20078350e-02,  1.51467569e-06,
                  8.71539786e-02,   1.11039576e-01,  1.07615717e-01,
                 -1.83975018e-06,   7.00099874e-07,  -2.76452265e-06,
                 -2.34488941e-07,  -5.69703859e-03,  -9.80645905e-03,
                 -3.60260208e-06,  -1.69209844e-02,  -1.32060472e-02,
                 -4.69592646e-02,   3.67271536e-07,   8.81426999e-07,
                 2.36516190e-08,   1.34404534e-06,  -1.47968545e-06,
                 3.42679201e-07,  -2.70485517e-06,  -6.24118932e-03,
                 7.37256970e-08]]),
 'sigma_obs': array([[0.0872592]]),
 'beta': array([[ 0.00589354, -0.0055046 , -0.00080662,  0.00039308, -0.00154423,
                  0.00057854, -0.03490154, -0.02737013, -0.01489706, -0.00934373,
                  0.01591459, -0.00960604,  0.00455795,  0.00555485]]),
 'trend': array([[0.09801003, 0.09800725, 0.09800446, ..., 0.08221363, 0.08221324,
                  0.08221285]]),
 'Y': array([[ 0.06577952, 0.06591064, 0.06611376, ..., 0.03614451, 0.03568925,
                  0.035449 ]]),
 'beta_m': array([[ 0., -0., -0.,  0., -0.,  0., -0., -0., -0.,  0., -0.,  0.,
                  0.]])},
 'beta_a': array([[ 0.00589354, -0.0055046 , -0.00080662,  0.00039308, -0.00154423,
                  0.00057854, -0.03490154, -0.02737013, -0.01489706, -0.00934373,
                  0.01591459, -0.00960604,  0.00455795,  0.00555485]])}

```

**Fig 7.** Parameters for the model

Expectations are then made on a dataframe with a segment 'ds' containing the dates for which a forecast is to be made. You can get an appropriate dataframe that stretches out into the future a predefined number of days using the aide strategy PROPHET.make-future-dataframe. Of course, it will likewise include the dates from the history, so will see the model fit too.

[Figure 8](#) shows the next future data instances for the 'date' attribute which is 'ds' from the original dataset using make-future-prediction function. In this case next 10 data instances are predicted using the function, which is originally 19734 instances, so now the next 10 i.e., 19744 are predicted. In PROPHET normally the future data can be made the maximum of 1/4 of the data, which gives the best results within the range, i.e., if the data is 100, then only next the 25 future data can be predicted. More than that can also be predicted but 1/4 of the data gives more accurate and best results from the future data.

```
19725 2016-12-05 22:20:00
19726 2016-12-05 22:30:00
19727 2016-12-05 22:40:00
19728 2016-12-05 22:50:00
19729 2016-12-05 23:00:00
19730 2016-12-05 23:10:00
19731 2016-12-05 23:20:00
19732 2016-12-05 23:30:00
19733 2016-12-05 23:40:00
19734 2016-12-05 23:50:00
19735 2016-12-06 23:50:00
19736 2016-12-07 23:50:00
19737 2016-12-08 23:50:00
19738 2016-12-09 23:50:00
19739 2016-12-10 23:50:00
19740 2016-12-11 23:50:00
19741 2016-12-12 23:50:00
19742 2016-12-13 23:50:00
19743 2016-12-14 23:50:00
19744 2016-12-15 23:50:00
```

[19745 rows x 1 columns]>

**Fig 8.** Future Data Instances for 'ds'

The foresee technique will relegate each line in future an anticipated worth which it names yhat. On the off chance that you go in verifiable dates, it will give an in-test fit. The conjecture object here is another dataframe that includes a section yhat with the figure, just as segments for parts and uncertainty intervals. [Figure 9](#) shows the future predicted values with yhat.

	ds	trend	yhat_lower	yhat_upper	trend_lower	trend_upper
0	2016-01-02 00:00:00	105.850835	-56.349522	194.567426	105.850835	105.850835
1	2016-01-02 00:10:00	105.847825	-49.234318	193.325046	105.847825	105.847825
2	2016-01-02 00:20:00	105.844815	-40.306570	190.024278	105.844815	105.844815
3	2016-01-02 00:30:00	105.841806	-46.654795	184.220173	105.841806	105.841806
4	2016-01-02 00:40:00	105.838796	-54.822014	196.958646	105.838796	105.838796

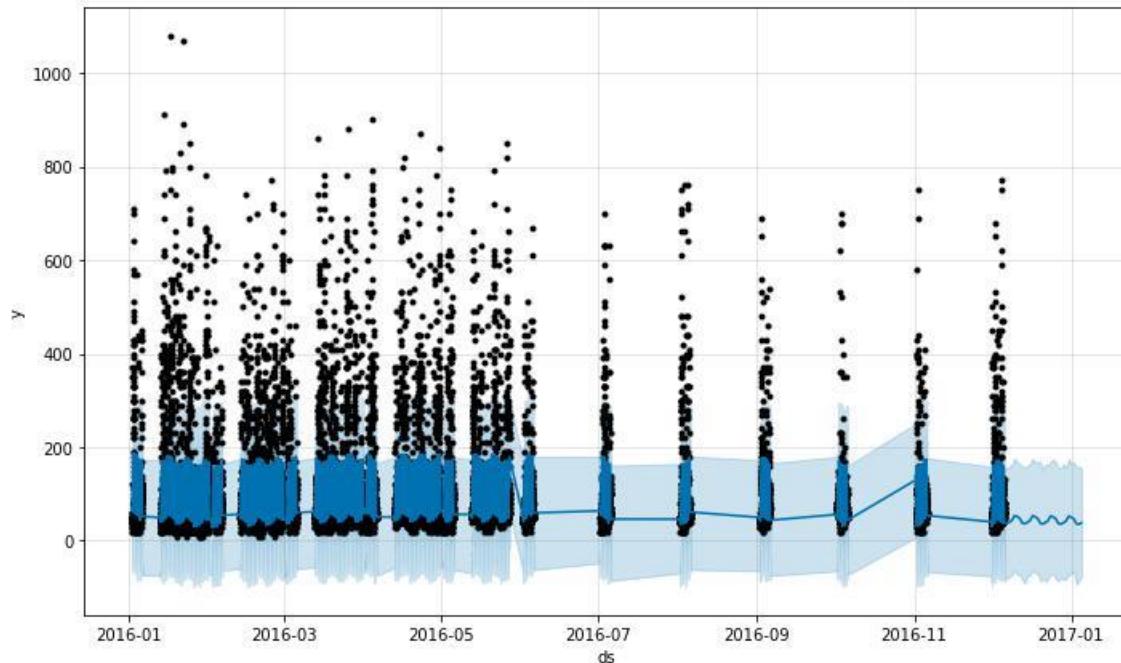
**Fig 9.** Future yhat attributes

**Figure 10** shows the forecast that is generated for 'date' and 'Appliances' attributes which are 'ds' and 'y'. Now the future generated instances are now plotted. **Figure 11** shows the forecasting for the next 30 instances or in this case, the next 30 days at the same time interval every day.

19750	2016-12-21 23:50:00	38.911294
19751	2016-12-22 23:50:00	42.709553
19752	2016-12-23 23:50:00	52.870935
19753	2016-12-24 23:50:00	50.947981
19754	2016-12-25 23:50:00	47.270927
19755	2016-12-26 23:50:00	37.010010
19756	2016-12-27 23:50:00	35.636394
19757	2016-12-28 23:50:00	38.486323
19758	2016-12-29 23:50:00	42.284582
19759	2016-12-30 23:50:00	52.445963
19760	2016-12-31 23:50:00	50.523010
19761	2017-01-01 23:50:00	46.845956
19762	2017-01-02 23:50:00	36.585039
19763	2017-01-03 23:50:00	35.211423
19764	2017-01-04 23:50:00	38.061352

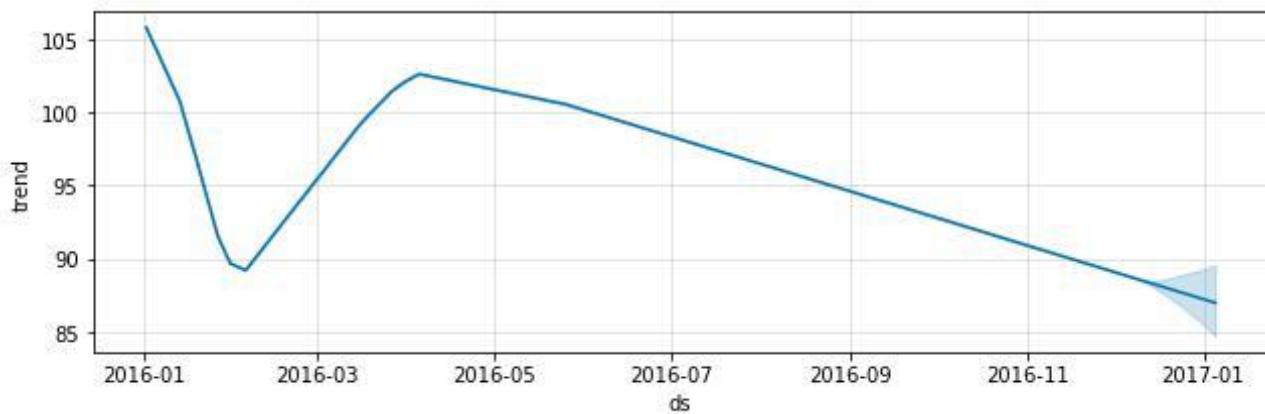
19765 rows × 2 columns

**Fig 10.** Future data instances for 'ds' and 'y'



**Fig 11.** The result of 30-day forecast

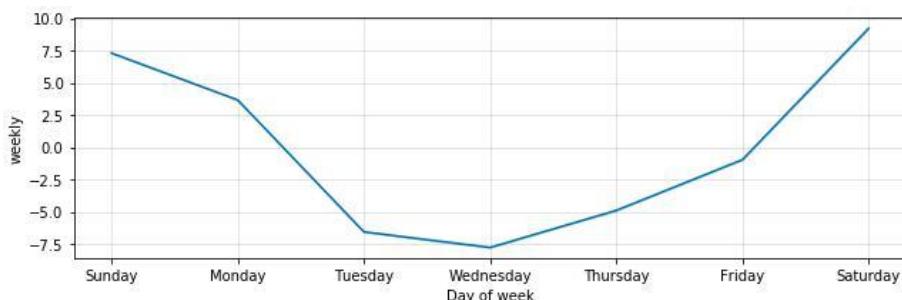
PROPHET gives forecast segments naturally like the pattern, yearly and week by week seasonality. Presently in the dataset, the date segment has hourly, minutes, and seconds which are not accessible in the PROPHET model as of presently. Along these lines, forecasting is done day by day for the following 30 days which are the following 30 instances at 23 Hours 50 Minutes, what's more, 0 Seconds each day, as PROPHET can't do a forecast for hourly, minutes and seconds information. [Figure 12](#) shows the forecasting for the following 30 instances in the pattern segment. The forecasted qualities on pattern segments appear that vitality utilization is progressively decreasing for quite a while for the 30 instances.



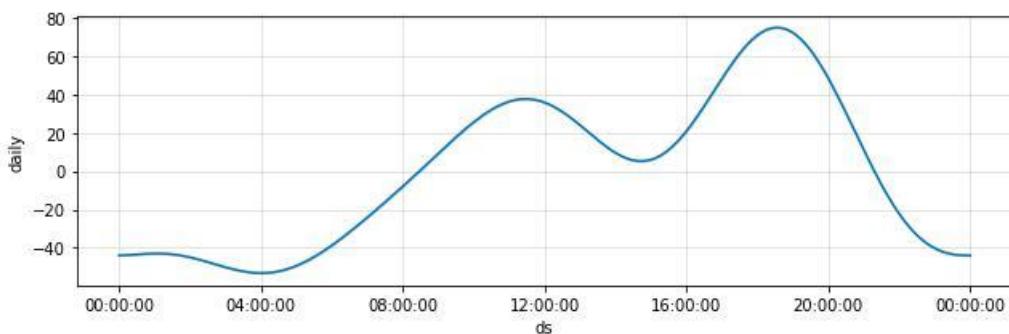
**Fig 12.** Trend component for the 30-day forecast

[Figure 13](#) gives the forecasted diagram on a week by week segment. The week after week part shows that the vitality utilization continued decreasing from Sunday till Wednesday and began increasing marginally from Wednesday till Saturday. On Saturday the vitality utilization rate is greatest contrasted with the remainder of the week.

[Figure 14](#) gives the day by day part forecasted diagram which begins at 0, 4, 8, 12, 16 and 20 hours and closes at 0 hours of the day. [Figure 14](#) shows that at any day for the forecasted 30 instances, on a normal, the vitality utilization is high between 16 hours, and 20 hours.



**Fig 13.** Weekly component for the 30-day forecast



**Fig 14.** Daily component for the 30-day forecast

## 8 Conclusion

The target of this work was to think about vitality request forecast esteems using the PROPHET and ARIMA techniques. To this end, the informational collection has been examined thoroughly. Models were trained on the dataset from January 11, 2016, to May 27, 2016 interval which is at 10 min for about 4.5 months. The models anticipated qualities that could be viewed as effective in the 30-day forecast. In this regard, the ARIMA and PROPHET bundles in Python were seen as very significant elements for the current task. The accomplishment of the model in the time series expectation was investigated and analyzed. While the PROPHET model makes predictions very accurate based on future data and was easy to make predictions compared to ARIMA.

The proposed work forecasts energy prediction for any buildings, office, household, and many more. So, from the literature survey, among many forecasting models like ARIMA, SARIMA, PROPHET, and many more, it was found that for the chosen dataset for research work, PROPHET works better than ARIMA. Based on the dataset, the accuracy of different models varies. PROPHET is very much beneficial to the ones that are not much experienced with forecasting and predicting. PROPHET helps when anyone is not having technical knowledge in predictions.

Forecasting is done for the next 30 days using a PROPHET. The 30-day forecast is done in different components like daily component, weekly component, and trend component which are the basic functionalities of PROPHET. Facebook PROPHET as of now is not available in hourly forecasting. So, PROPHET is the best model and the anticipated outcomes are referenced above where, the pattern part shows the utilization rate kept continuously decreasing, week by week segment shows the utilization rate was most extreme on Saturday and the everyday segment shows the high vitality utilization rate was between 16 hours and 20 hours.

From the forecast results, we can say that Wednesday is the lowest power utilization day, and power utilization increases till Saturday and Saturday is the highest power utilization day.

The future scope of this research is to forecast for a minute and hourly forecasts also using PROPHET, which is the components not available.

## Acknowledgement

We would like to thank CHRIST Deemed to be University for providing us the environment and the opportunity for completing this work.

## References

- 1) Yenidogân I, Çayir A, Kozañ O, Dağ T, Arslan C. Bitcoin forecasting using ARIMA and PROPHET. In: and others, editor. 3rd International Conference on Computer Science and Engineering (UBMK), Sarajevo. 2018;p. 621–624. Available from: <https://doi.org/10.1109/UBMK.2018.8566476>.
- 2) Contreras J, Espinola R, Nogales FJ, Conejo AJ. ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems*. 2003;18(3):1014–1020. Available from: <https://dx.doi.org/10.1109/tpwrs.2002.804943>.
- 3) Lei F, Hu P. A baseline model for office building energy consumption in hot summer and cold winter region. In: and others, editor. International Conference on Management and Service Science, Wuhan;vol. 1. 2009;p. 1–4. Available from: <https://doi.org/10.1109/ICMSS.2009.5301031>.
- 4) Yan Y, Li B, Xiao J, Liang Y, Shang Y, Zhou K. Comparative study on prediction algorithms for power grid system access failure times. In: and others, editor. IOP Conference Series: Earth and Environmental Science;vol. 252 of 3. 2019;p. 1–11. Available from: <https://doi.org/10.1088/1755-1315/252/3/032183>.
- 5) Zhou R, Pan Y, Huang Z, Wang Q. Building energy use prediction using time series analysis. In: IEEE 6th International Conference on Service-Oriented Computing and Applications. 2013;p. 309–313. Available from: <https://doi.org/10.1109/SOCA.2013.14>.
- 6) Pappas SS, Ekonomou L, Karampelas P, Karamousantas DC, Katsikas SK, Chatzarakis GE, et al. Electricity demand load forecasting of the Hellenic power system using an ARMA model. *Electric Power Systems Research*. 2010;80:256–264. Available from: <https://dx.doi.org/10.1016/j.epsr.2009.09.006>.
- 7) Alvarez FM, Troncoso A, Riquelme JC, Ruiz JSA. Energy time series forecasting based on pattern sequence similarity. *IEEE Transactions on Knowledge and Data Engineering*. 2011;23(8):1230–1243. Available from: <https://dx.doi.org/10.1109/tkde.2010.227>.
- 8) Nogales FJ, Contreras J, Conejo AJ, Espinola R. Forecasting Next-Day Electricity Prices by Time Series Models. *IEEE Power Engineering Review*. 2002;22(3):58. Available from: <https://dx.doi.org/10.1109/mper.2002.4312063>.
- 9) Lee YW, Gaik TK, Yee CY. Forecasting electricity consumption using time series model. *International Journal of Engineering and Technology*. 2018;7:218–223. Available from: [https://doi.org/10.14419/ijet.v7i4.30\\_22124](https://doi.org/10.14419/ijet.v7i4.30_22124).
- 10) Ngabesong R, McLauchlan L. Implementing “R” Programming for time series analysis and forecasting of electricity demand for Texas, USA. In: IEEE Green Technologies Conference(GreenTech);vol. 1. 2019;p. 1–4. Available from: <https://doi.org/10.1109/GreenTech.2019.8767131>.
- 11) Pandya R, Parikh A. Predicting domestic electricity consumption using data mining for rural area of Gujarat. In: 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS). 2018;p. 167–173. Available from: <https://doi.org/10.1109/CSITSS.2018.8768772>.
- 12) Lam JC, Wan KKW, Wong SL, Lam TNT. Principal component analysis and long-term building energy simulation correlation. *Energy Conversion and Management*. 2010;51:135–139. Available from: <https://dx.doi.org/10.1016/j.enconman.2009.09.004>.
- 13) Usha TM, Balamurugan SAA. Seasonal Based Electricity Demand Forecasting Using Time Series Analysis. *Circuits and Systems*. 2016;07(10):3320–3328. Available from: <https://dx.doi.org/10.4236/cs.2016.710283>.
- 14) Chujai P, Kerdprasop N, Kerdprasop K. Time series analysis of household electric consumption with ARIMA and ARMA Models. *Proceedings of the International MultiConference of Engineers and Computer Scientists*. 2013;1:295–300. Available from: [http://www.iaeng.org/publication/IMECS2013/IMECS2013\\_pp295-300.pdf](http://www.iaeng.org/publication/IMECS2013/IMECS2013_pp295-300.pdf).
- 15) Sun T, Zhang T, Teng Y, Chen Z, Fang J. Monthly Electricity Consumption Forecasting Method Based on X12 and STL Decomposition Model in an Integrated Energy System. *Mathematical Problems in Engineering*. 2019;2019:1–16. Available from: <https://dx.doi.org/10.1155/2019/9012543>.
- 16) Ahmad A, Anderson TN, Rehman SU. Prediction of electricity consumption for residential houses in New Zealand. Smart Grid and Innovative Frontiers in Telecommunications. *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*. 2018;245:165–172. Available from: [https://doi.org/10.1007/978-3-319-94965-9\\_17](https://doi.org/10.1007/978-3-319-94965-9_17).
- 17) Weron R. Modeling and forecasting electricity loads and prices: A statistical approach. and others, editor;John Wiley & Sons. 2007.
- 18) Almazrouee AI, Almeshal AM, Almutairi AS, Alenezi MR, Alhajeri SN. Long-Term Forecasting of Electrical Loads in Kuwait Using Prophet and Holt-Winters Models. *Applied Sciences*. 2020;10(16):5627. Available from: <https://dx.doi.org/10.3390/app10165627>.
- 19) Gong F, Han N, Li D, Tian S. Trend analysis of building power consumption based on PROPHET algorithm. *Asia Energy and Electrical Engineering Symposium (AEEES)*. 2020;p. 1002–1006. Available from: <https://doi.org/10.1109/AEEES48850.2020.9121548>.
- 20) Shen J, Valagolam D, McCalla S. Prophet forecasting model: a machine learning approach to predict the concentration of air pollutants (PM2.5, PM10, O<sub>3</sub>, NO<sub>2</sub>, SO<sub>2</sub>, CO) in Seoul, South Korea. *PeerJ*. 2020;8:e9961. Available from: <https://dx.doi.org/10.7717/peerj.9961>.
- 21) Wei N, Li C, Peng X, Zeng F, Lu X. Conventional models and artificial intelligence-based models for energy consumption forecasting: A review. *Journal of Petroleum Science and Engineering*. 2019;181. Available from: <https://doi.org/10.1016/j.petrol.2019.106187>.
- 22) Lindberg KB, Seljom P, Madsen H, Fischer D, Korpås M. Long-term electricity load forecasting: Current and future trends. *Utilities Policy*. 2019;58:102–119. Available from: <https://dx.doi.org/10.1016/j.jup.2019.04.001>.
- 23) Kuster C, Rezgui Y, Moushedi M. Electrical load forecasting models: A critical systematic review. *Sustainable Cities and Society*. 2017;35:257–270. Available from: <https://dx.doi.org/10.1016/j.scs.2017.08.009>.
- 24) Scheidt FV, Medinová H, Ludwig N, Richter B, Staudt P, Weinhardt C. Data analytics in the electricity sector - A quantitative and qualitative literature review. *Energy and AI*. 2020. Available from: <http://dx.doi.org/10.1016/j.egyai.2020.100009>.
- 25) Amasyali K, El-Gohary NM. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*. 2018;81:1192–1205. Available from: <https://dx.doi.org/10.1016/j.rser.2017.04.095>.
- 26) Khuntia SR, Rueda JL, van der Meijden MAMM. Forecasting the load of electrical power systems in mid- and long-term horizons: a review. *IET Generation, Transmission & Distribution*. 2016;10(16):3971–3977. Available from: <https://dx.doi.org/10.1049/iet-gtd.2016.0340>.
- 27) Taylor SJ, Letham B. Forecasting at Scale. *The American Statistician*. 2018;72:37–45. Available from: <https://dx.doi.org/10.1080/00031305.2017.1380080>.