Analysis and Forecasting of Power Factor for Optimum Electric Consumption in a Household

Manu Mishra
x18127347
Group M
M.Sc. Data Analytics
Advance Data Mining

Niraj Shah
x18139108
Group M
M.Sc. Data Analytics
Advance Data Mining

Punit Lohani
x18127339
Group M
M.Sc. Data Analytics
Advance Data Mining

Shrikant Samarth
x18129137
Group M
M.Sc. Data Analytics
Advance Data Mining

***Abstract*— Household appliances operate on alternating current circuits and consist of reactive load, which is a combination of both capacitance and inductance. Due to presence of these kinds of loads, there is an existence of phase difference between the current and voltage. The cosine of phase angle between the current and voltage, in such situations, is called Power Factor, which is the ratio of the active and the apparent power. Theoretically, its value must lie in the range of (0, 1). Lower Power Factor affects the performance of electrical appliances. Therefore, it is important to perform a check on its value. This project focuses on the best short-term range of forecast for Power Factor so as to achieve optimum energy consumption for a household that can help in understanding the time series data in terms of seasonal components. Techniques like ETS method, Holt-Winters method and ARIMA model have been employed in this research. The result shows that, for forecasting the time-series data that contains seasonality Holt-Winters presented a significant result. In addition to this, best short-term range of forecast of Power Factor values have been achieved through ARIMA model with 6 and 12 months forecast periods.**

***Keywords*— Reactive load, capacitance, inductance, alternating current, phase difference, power factor, ETS method, Holt-Winters, ARIMA**

I. Introduction

A. Background and Motivation

The consumption of electricity in a household is the total combination of electrical energy consumed by the household appliances. Household electrical appliances are low rating electrical systems which operate on nominal levels of voltage and current. The significance of power factor can be understood in the case of AC circuits as they consist of resistance, capacitance and inductance [1]. Household appliances such as kitchen equipment, washing machine, light, air-conditioner etc. have a reactive load mixture of both capacitance and inductance. Due to the combined reactance offered by them, a phase difference exists between the source voltage and current. A better understanding of power factor can be given by the power factor triangle given below:

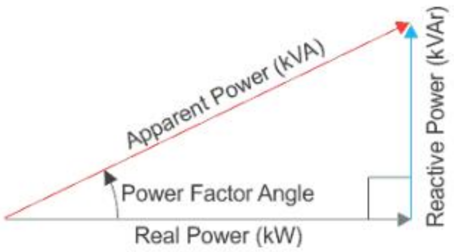

Figure 1: Power Factor Triangle

This can also be represented as:

$$\text{Cosine}(\emptyset) = \frac{\textit{Active Power}}{\textit{Apparant Power}}$$

where active power is the power consumed in an AC (Alternating Current) circuit and the apparent power is the total power across the element, given as the product of voltage and current across it.

Theoretically, the value for power factor lies in the range of (0, 1) and at all occasions, must be ideally close to 0.9 or more. This condition is the requirement for optimum performance of electrical equipment and transmission systems. Lower values of power factor must be avoided because it may lead to poor performance of equipment and sometime, to damaging it as well. Therefore, it is a good practice to keep a regular check on the value.

In practical situations, there can be variations in the value of the power factor due to imbalance in reactive loads. With the approach of machine learning methods, an attempt has been made to analyze the power factor trend for a household and forecast its power factor pattern so that best operating conditions can be maintained.

The project aims to provide an application of machine learning approach for power factor forecasting and can be useful for audience intended to get insight into the usefulness of power factor prediction and its implications.

B. Research Question

What can be the best short-term range of forecast for power factor patterns so that optimum energy consumption can be achieved for a household?

The objectives for this research are:

- To analyse the values of power factor for a household using a variety of electrical appliances over a period of time. This analysis will help in understanding the time-series data in terms of its seasonal components and other related parameters.
- To forecast the values of power factor using machine learning approach and select the best short-term forecasting range that can be used for referencing.

C. Structure of Paper

The paper has been divided into the following sub-sections:

- Related Work: Discussion and review of literature and work done in context to the project topic and research question.
- Methodology: Implementation of CRISP-DM model in implementation of data mining project.
- Result: Discussion of the results obtained in the methodology.
- Conclusion and Future Work: Consideration of the final outcome of the project and points for future work which can be performed with the scope of the project.

II. Related Work

A good approach for assessing the related work is by reviewing the literature related to the field of study. Literature review provides an exhaustive discussion on the past and present attempts on the related topic, which helps in acquiring a basis and overview of the concerned topic.

Considerable amount of work has been done with regard to the forecasting and analysis of electrical load consumption. In a research work by [2] an analysis was done on utilizing and reducing the electricity consumption by analysing NILM at appliance level. This paper aims at improving NILM using load Division and Calibration (INDiC) that simplifies NILM by separating appliances over numerous instrument points (meters/stages) and measured power was calibrated. The proposed method was used with the Combinatorial Optimization framework which shows significant improvement in disaggregation accuracy. [3] has created a model which could accurately predict the future electricity demand. It has assessed smart meter energy utilization data from multiple houses and used autocorrelation to set up decisive parameters for Gated Recurrent Unit (GRU) with the help of recursive Deep Learning.

In a study by [4], an explanation has been provided as to how time series approach, based on ARIMA (Auto Regressive Moving average) model is useful for repairable system failure analysis and forecasting. It has been suggested that for the analysis of system failure, statistical methods can bridge the gap in characterizing the failure process and found that the ARIMA model shows promising results for such analysis. Another work by [5] done in demand forecasting using time-series approach shows how information of historical demand can be used to create ARIMA models using time series. Based on Akaike criterion, Schwarz Bayesian criterion, maximum likelihood, and standard error performance criteria, a model was selected which corresponds to ARIMA (1,0,1) and then it was validated by another historical demand data. This research proves that the model could be used in forecasting future demands and can be a guideline for making decisions. Another approach by [6], identifies of the issue of power demand forecasting in residential micro-grids has been performed. In the research work, comparison of performances of a few methodologies utilizing ARMA models, support vector machines, and neural systems that perform one-step ahead predictions were extended to perform multi-step ahead prediction error and error variance. As a result, they found out that all machine learning approaches outperforms ARMA; but the researchers suggest the adoption of hybrid approach by using NAR (Nonlinear Auto-Regressive neural networks) which displays the best forecasting accuracy for short term forecasting, whereas SVM can be a potential choice for long term intervals. The suggestion for SVM and PCA, as a method, was suggested by [7] which analyses the electricity usage by a single home. For this, K-Means clustering algorithm was used for getting the ideal home utilization electricity data points. The results of the study indicate that Davis Boulden Index and Silhouette score obtains optimal, yet detailed number of clusters in the K-means algorithm. It was also been suggested that the house-hold power usage can be analysed using other machine learning techniques such as PCA and SVM algorithm. The literature on study involving the application of machine learning algorithms in energy load predictions can be further extended by considering the work of [8] which found that different machine learning approaches can be used to ensure the highest accuracy with the smallest computational cost for different time scales in forecasting electricity load. The performance-based metrics were compared using Artificial Neutral Network (ANN), Support Vector Regression (SVR), Random Forest Regression (RFR), Extreme Gradient Boosting (XGB) and Flexible Neural Tree (FNT) algorithms. A key point from the study shows how the prediction accuracy decreases with the increase in the time scale because of limitation of using all variables. Most of the comparisons produced XGB model as the best performer.

It is difficult to forecast short term load prediction accurately. Hence a model of hybrid ARIMA and SVM was proposed by [9] to forecast daily load while correcting the deviation of the previously forecasted values. The results show that ARIMA-SVMs hybrid model further improves ARIMA in simulation accuracy by declining 0.65% MAPE and 7.77MW decline in RMSE which is far better than the SVM model prediction. ARIMA and SARIMA (seasonal auto regressive integrated moving average) models were used in another paper by [10] for short term (i.e. at an hourly interval) load forecasting. It was found that ARIMA and SARIMA model outperforms other machine learning algorithms in predicting short term load forecasting. Forecasting error was compared and calculated by using multiplicative decomposition model and SARIMA model as well by [11] and the results showed that both of the time series model gives accurate prediction for short term demand forecasting. [12] used numerous algorithms including ARIMA and neural network as a bench marking method for short term forecasting of household's electricity demand on different levels and timescale. The forecasting horizon was considered from 15 minutes to 24 hours across two datasets which had power usage statistics of households. Overall results showed that MAPE (Mean Absolute Percentage Error) varied between 5 to less than 100% which can be further improved. Similar study on energy demand forecasting by [13], in his research work on New Zealand, compared six forecasting models, i.e. Logistic, ARIMA, combined, Harvey Logistic, Harvey and Val models. The comparison of accuracy was done for goodness-of-fit and showed that ARIMA is the best performing model for short term forecasting demand. Studies on India, Pakistan and Serbia by [14] and [15] [16] used metabolic grey model (MGM), ARIMA, MGM-ARIMA and back-propagation neural network and concluded that all models gave high accuracy i.e. above 95%.

In recent years, power factor improvement has been a point of attention for both customers and power generation organizations for efficient energy usage and consumption. Case studies on impact of power factor correction for the economies of Kuwait by [17] and Uganda by [18] pointed out that the value of power factor should be optimized in the range of 0.9 to 0.95 so that an improved performance of the electrical supply system can be achieved. Improvement in power factor may lead to reduction in overloading of transmission lines, high-voltage electrical equipment and other related household equipment which are served through the supply lines. A periodic measurement of power factor among the electricity consumers may lead to finding cases of low power factor and proper measures can be taken at the correct time to improve the situations. [19] discusses that If a case of low power factor is being encountered in electrical distribution system, the same can be remediated using shunt capacitor of appropriate value. A method of power factor correction, as provided by [20], can be implemented by using Integrated Circuits and microcontrollers. If the measurement of power factor falls below 0.85, the setup using microcontroller will cause triggering of relay which will then add the capacitors to the circuit, in order to correct the case of low power factor.

The main concerns related to low power factor values is of prime concern for discussion. According to [21] and [22], lower values of power factor increase the level of current in distribution system. This surge in current will directly lead to large line losses or copper losses, as per the formula:

Power Loss = $I^2 \times R$; where I is the current and R, resistance

In addition to this, conductor of large cross-section is required for flow of high current. High level of current flow through the conductor leads to large voltage drop. This is the cause of for poor voltage regulation. Machines operating on low power factor require higher kVA ratings and higher kVA ratings directly increase the cost of machines. Therefore, cost of machinery can be controlled by managing the level of power factor.

The above discussion in terms of literature review provides enough background for the problem being analysed in the project to find out the best short-term range of forecast for power factor patterns so that optimum energy consumption can be achieved for a household.

III.METHODOLOGY

Cross-Industry Standard Process for Data Mining, or in short, CRISP-DM [23] has been employed as the reference methodology for the data mining project. The selection for this methodology for the project has been made in accordance to the discussion that CRISP-DM channelizes the operation of data mining project into six phases which makes it easier to apply the concept of data mining in real-life scenarios and bring out the business outcomes from the same. Since the project approach is to scrutinize and forecast the power factor for households, therefore it stands out as the best model for this project.

There are six stages for CRISP-DM model, which are as follows:

- Business Understanding
- Data Understanding
- Data Preparation
- Modeling
- Evaluation
- Deployment

A. *Business Understanding*

The business objective for the project deals with the analysis and forecasting of power factor values for a household. This objective has been dealt in detail in the Research Question section. R has been used in this project for exploratory data analysis and model implementation. An attempt has been made in the project to produce accurate forecasting of power factor values, based on the dataset of 2 million electrical measurements. All the efforts have been made so that the results obtained from the forecasting and analysis can be useful in terms of business perspective.

B. *Data Understanding*

Data for the project has been sourced from UCI machine learning repository[1]. This dataset contains the electrical power consumption of an individual household located in Sceaux, which is seven kilometers from Paris, France. The time-series dataset contains measurements in the range of December 2006 to November 2010 and has approximately 2 million measurements, as a combination of following indicators:

VARIABLES	DESCRIPTION
global_active_power	minute-averaged active power, in kW
global_reactive_power	minute-averaged reactive power, in kW
voltage	minute-averaged voltage, in volts
global_intensity	minute-averaged current, in amperes
sub_metering_1	energy in sub_metering 1
sub_metering_2	energy in sub_metering 2
sub_metering_3	energy in sub_metering 3

Figure 2: Description of variables in dataset

[1] https://archive.ics.uci.edu/ml/datasets/individual+household+electric+power+consumption

Sub-metering 1refers to energy consumption in kitchen, sub-metering 2 is the measurement of energy consumed in laundry room while the measurement for electric water heater and air-conditioner is captured by sub-metering 3. The variables included in the dataset provide enough indicators to calculate and analyze the power factor and its prediction. Size of the dataset is pertinent enough for a data mining project to bring out useful insights related to the business requirements. The data quality in terms of missing values and other related parameters has been dealt in the stage for data preparation.

C. Data Preparation and Transformation

According to the analysis in previous section, the discussed data is suitable and has been taken up for the project. The description of the dataset, at the source, claims that the data contains missing values. Upon inspecting the same using R, it has been found that about 26,000 entries or 1.25% of the data has missing values. These values have been removed as their share is quite small as compared to the size of the dataset. Since the main motive of the analysis employs the use of power factor, in reference to the electrical measurements, therefore, a new column for power factor has been introduced. The formula for calculating the power factor is as follows:

Root Mean Square Voltage, V_{rms}=1.11× voltage ………. (1)

Apparent Power = V_{rms} × current ……………….………….. (2)

Power Factor = Active Power ÷ Apparent Power …..…... (3)

The value of voltage in equation (2) and current, as global_intensity in equation (3) has been provided in the parent dataset.

Another data pre-processing has been carried out in terms of the range of value of power factor. The value of power factor is theoretically always in the range [0,1] and practically, less than 1. The value of power factor as unity means that the referenced load is purely resistive, and 100% power is consumed. Since the data in reference has household appliances, which are reactive in nature i.e. inductive or capacitive, therefore there must be some power loss. Thus, the value of power factor can never be unity, and in this case, it must be less than one.

Therefore, a check has been performed to know the number of measurements having the value of power factor greater than 1. On analysis, 34,000 entries or 1.6% of the data have this anomaly. Owing to their share, they have been removed to produce the analysis-ready dataset.

The considered data for the analysis is a time-series data as the project deals with time-series forecasting of power factor. There is a possibility that the time series dataset may contain seasonal component. The seasonal data may iterate itself over time-range of day, week, month or year and this may interfere with the predictions given by the models.

Time-series data has been segregated in terms of four seasons: Autumn, Spring, Summer and Winter. From figure 3, it can be observed that the aggregated values of power factor, as per the seasons, are not same. This confirms the presence of seasonality component in the reference data.

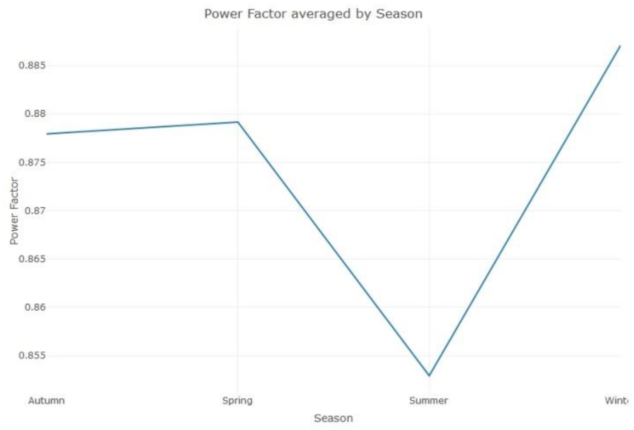

Figure 3: Power Factor Averaged by Seasons

The value of monthly averaged power factor is shown in figure 4, which shows that how aggregated value of power factor varies over months in a year, for the reference time period of 2007-2010. This variation in the values depicts that in addition to seasonal variation, the time series data has variation in terms of monthly intervals.

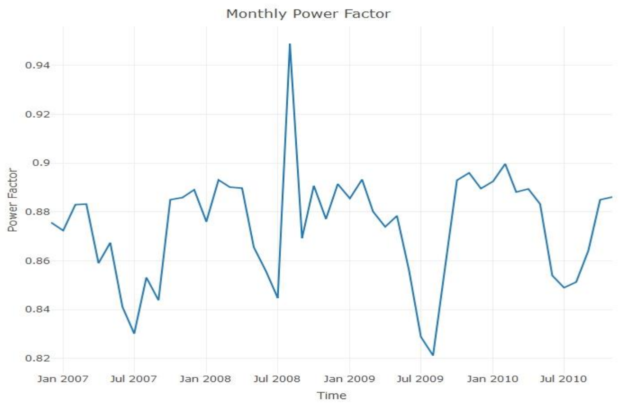

Figure 4: Power Factor Averaged by Months

Usage of energy for the household can be described in terms of the measurements from the three meters. Figure 5 shows that the sub-meter 1 and 2 have approximately same energy usage over the years 2007 to 2010 while the measurements of sub-meter 3 have the highest value.

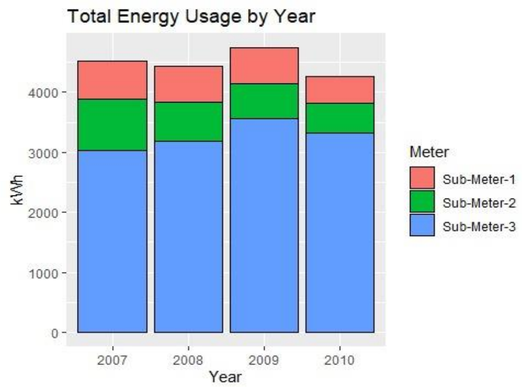

Figure 5: Total Energy Usage by Year, as per different sub-meters

The same trend can be seen in figure 6 where it can be interpreted that the sub-meter 3 has the highest proportion of energy consumption for the three years, as compared to sub-meter 1 and 2.

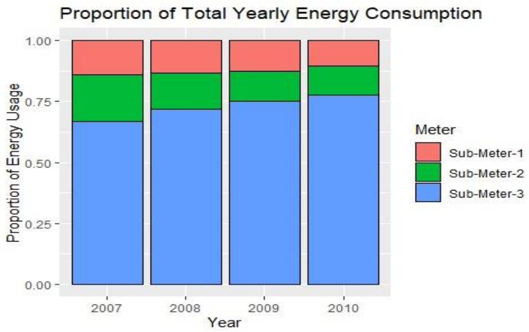

Figure 6: Proportion of Total Yearly Energy Consumption, as per different sub-meters

The appliances measured by sub-meter-1 are dishwasher, oven and microwave. Sub-meter-2 measures washing machine, refrigerator and light while consumption of electric water heater and air-conditioner is measured by sub-meter 3. According to the above description, the load mix for sub-meter-1and sub-meter-2 is somewhat resistive while it is a mixture of resistive and inductive load for sub-meter-3. This information is represented in order to explain that the overall household load is not completely resistive in nature and hence, the value of power factor at any given instance of time-series should not be equal to unity.

D. Modelling

A study by [24] shows the importance of data granularity. Breaking the data into the smallest possible and realistic levels or granules provides the opportunity to study and analyze the data at the micro level. Therefore, the reference time series has been broken down into daily, weekly and monthly levels so as to study the nature of data at each of the granular level.

Since the data has been divided in the possible levels in terms of day, week and month, the evaluation of randomness can be an added factor of study because this can provide the level at which the time series data can be fed to the forecasting models to get optimum results.

Figure 7 shows the randomness for the three levels of granularity i.e. day, month and week. The lowest value has been obtained for monthly granularity which means that time-series data, aggregated on monthly basis, possess lack of pattern and predictability in terms of prediction of power factor. Therefore, monthly aggregated data has been provided to the models, which has been discussed in detail.

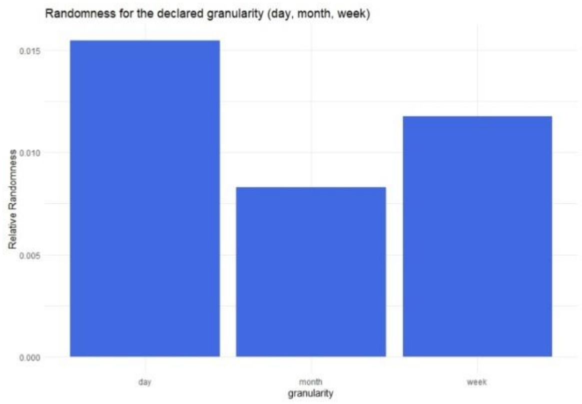

Figure 7: Randomness for the declared granularity (day, month and year)

For the purpose of forecasting and getting the result, below mentioned techniques have been used:

- *ETS Method*

Exponential smoothing method helps in the smart forecasting by taking the weighted average of the past values into consideration and assigning the higher weights to the recent values [25].

As a part of this project, ets() in R, has been used in the ETS method.

- *Holt-winters method*

Double exponential smoothing type of method has been used as part of the Holt-winters method. The Holt-Winters method is basically employed for forecasting the data which consist of correlation in terms of seasons, change in trends and seasonality [26].

- *ARIMA model*

ARIMA model has been used for forecasting the result and it consists of three parameters i.e. (p,d,q) which is represented as ARIMA(p,d,q). The AR part of the model is based on the concept that the current estimate of the series xt, which is explained as p, where p represents the total number of past steps needed to forecast the present value. Level of differencing required by the original time series to transform to stationary is represented by d. q represents the order of the moving average model.

Here, Auto ARIMA has been implemented using auto.arima() in R, that helps in obtaining the best ARIMA model which is performed in reference to single variable i.e. Power factor [27].

E. Evaluation

For the evaluation of the forecasted model, three metrics have been taken into consideration and they are:

- Mean absolute percentage error (MAPE)
- Root mean square error (RMSE)
- Mean absolute error (MAE)

If the value of MAPE falls in the bracket of 10%-20% then the model will be categorized as an acceptable model while the value of less than 10% shows that the model has performed well [28]. The mean absolute error is an average of errors in the measurement. Moreover, lower value of the RMSE signifies that the model is good enough.

Below are the results obtained:

1) ETS Method

The performance metrics for ETS method is represented below:

```
> accuracy(ETS_prediction_total, total_test) #check metrics comparing prediction in test
                       ME       RMSE        MAE        MPE     MAPE      MASE      ACF1 Theil's U
Training set 5.829531e-06 0.02364467 0.01809745 -0.0725242 2.079681 0.9817027 0.3306948        NA
Test set     3.228226e-03 0.01776414 0.01686827  0.3282878 1.927717 0.9150255 0.6938210  1.350313
```

Figure 8: Table for accuracy statistics for ETS method

The forecast of the power factor for the next 12 months from the ETS method is depicted below:

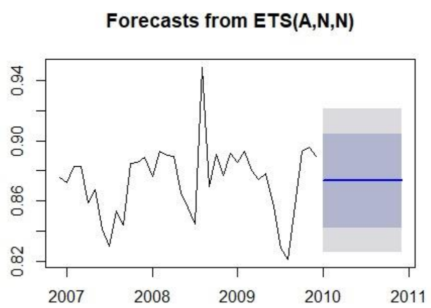

Figure 9: Forecast curve for ETS method

2) Holt-Winters Method

The performance metrics for Holt-Winters method is represented below:

```
> accuracy(Holtw_prediction, total_test)
                       ME       RMSE        MAE        MPE     MAPE      MASE       ACF1 Theil's U
Training set -0.003780850 0.02263265 0.01489047 -0.4760895 1.687889 0.8077387 0.07693017        NA
Test set     -0.006428431 0.01484467 0.01145581 -0.7437867 1.309794 0.6214243 0.25977313  1.176375
```

Figure 10: Table for accuracy statistics for Holt-Winters method

The forecast of the power factor for the next 12 months from the Holt-winters method is depicted below:

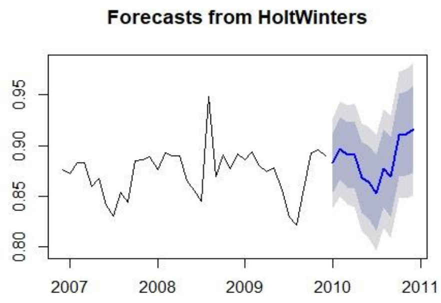

Figure 10: Forecast curve for Holt-Winters method

3) ARIMA Model

The performance metrics for ARIMA model is represented below:

```
> accuracy(ARIMA_prediction_total, total_test)  #check metrics comparing prediction in test(12)
                        ME       RMSE        MAE         MPE     MAPE      MASE        ACF1 Theil's U
Training set -2.146605e-05 0.02229564 0.01600477 -0.06720018 1.831786 0.8681846 -0.02040226        NA
Test set      2.285011e-03 0.01699610 0.01610484  0.22219290 1.842661 0.8736129  0.68259633  1.326773
```

Figure 11: Table for accuracy statistics for ARIMA Model

The forecast of the power factor for the next 12 months from the ARIMA (1, 0, 0) is depicted below:

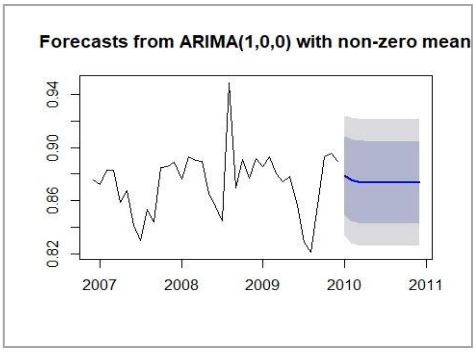

Figure 12: Forecast curve for ARIMA Model

On evaluating the results, for forecasting the data that contains seasonality, Holt-winters presented a good result with the RMSE value of 0.014. In addition, the ARIMA model with 6 and 12 months forecasts period proved to be an effective model for predicting the power factor usage for achieving the optimum energy consumption of households with acceptable values of MAPE, RMSE and MAE respectively. Further, the forecast months and the values of respective RMSE, MAPE and MAE are represented below in the tabular format:

Model	Forecast_months	RMSE	MAPE	MAE
ARIMA	3	0.018	1.94	0.017
	6	0.017	1.83	0.016
	9	0.017	1.95	0.017
	12	0.017	1.84	0.016

Figure 13: Evaluation Statistics for different forecasting months of ARIMA

F. Deployment

In the phase of deployment of the project, the results derived from the evaluation section can be employed to provide insights regarding the business objectives outlined. The report of the project is a complete discussion of the implementation of solution of research question, and thus, can be referred as a handy documentation by the intended audience.

IV. RESULT

The evaluation phase of the CRISP-DM methodology deals with assessment of the various methods applied to the data set, i.e. ETS method, Holt-Winters method and ARIMA model, in order to get useful results with reference to the research question. For forecasting the time-series which contains seasonality, Holt-Winters presented a good result with the RMSE value of 0.014.The best short-term range of forecast for power factor values and its patterns has been achieved through ARIMA model with 6-months and 12-months forecasting period.

V. CONCLUSION & FUTURE WORK

The best short-term range of forecasts for power factor patterns and values has been achieved for 6 months and 12 months duration, using ARIMA model. Satisfactory result has been achieved using Holts-Winters method as well. These results can help the organizations which deal with power factor measurement so that the overall power quality of the electrical distribution and transmission systems can be maintained. The scope of this project can be further extended by application of hybrid model for enhancing the performance and accuracy of the forecasts. The forecastHybrid package in R provides hybridModel function which fits multiple models from the auto.arima (), ets (), thetaf (), nnetar (), stlm () and tbats () functions using either equal weights or weights based on in-sample errors [28].

A different approach of using various combinations of resistive and reactive load can be used for checking the ability of ARIMA and hybrid model for forecasting the power factor patterns of households.

REFERENCES

[1] J. Grainger, W. Stevenson and W. Stevenson, Power system analysis, 2003.

[2] N. Batra, H. Dutta and A. Singh, "INDiC: Improved Non-intrusive Load Monitoring Using Load Division and Calibration," IEEE, vol. 1, pp. 79--84, 2013.

[3] C. Heghedus, A. Chakravorty and C. Rong, "Energy Load Forecasting Using Deep Learning," 2018 IEEE International Conference on Energy Internet (ICEI), no. IEEE, pp. 146-151, 2018.

[4] S. Ho and M. XIE, "The use of ARIMA models for reliability forecasting and analysis," Computers & industrial engineering, vol. 35, pp. 213--216, 1998.

[5] J. Fattah, L. Ezzine, Z. Aman, H. E. Moussami and a. A. Lachhab, "Forecasting of demand using ARIMA model," International Journal of Engineering Business Management, vol. 10, p. 1847979018808673, 2018.

[6] R. Bonetto and M. Rossi, "Machine Learning Approaches to Energy Consumption Forecasting in Households," arXiv preprint arXiv:1706.09648, 2017.

[7] H. W. Choi, N. M. F. Qureshi and D. R. Shin, "Analysis of Electricity Consumption at Home Using K-means Clustering Algorithm," in 2019 21st International Conference on Advanced Communication Technology (ICACT), 2019.

[8] T. a. Vantucha, A. G. Vidalb, A. P. Ramallo-Gonz´alezb, A. F. Skarmetab and S. M. s´aka, "Machine learning based electric load forecasting for short and long-term period," IEEE, pp. 511--516, 2018.

[9] H. NIE, G. LIU, X. LIU1 and Y. WANG, "Hybrid of ARIMA and SVMs for short-term load forecasting," Energy Procedia, vol. 16, pp. 1455--1460, 2012.

[10] H. Nguyen and C. K. Hansen, "Short-term electricity load forecasting with Time Series Analysis," 2017 IEEE International Conference on Prognostics and Health Management (ICPHM), pp. 214--221, 2017.

[11] J. Deng and P. Jirutitijaroen, "Short-term load forecasting using time series analysis: A case study for Singapore," 2010 IEEE Conference on Cybernetics and Intelligent Systems, pp. 231--236, 2010.

[12] A. Veit, C. Goebel, R. Tidke, C. Doblander and H.-A. Jacobsen, "Household Electricity Demand Forecasting--Benchmarking State-of-the-Art Methods," arXiv preprint arXiv:1404.0200, 2014.

[13] Z. Mohamed and P. Bodger, "Forecasting electricity consumption: A Comparison of models for New Zealand," University of Canterbury. Electrical and Computer Engineering., no. University of Canterbury. Electrical and Computer Engineering., 2004.

[14] F. Jiang, X. Yang and S. Li, "Comparison of Forecasting India's energy demand using an MGM, ARIMA model, MGM-ARIMA model, and BP Neural Network model," Sustainability, vol. 10, p. 2225, 2018.

[15] S. A. U. Rehman, Y. Cai, R. Fazal, G. D. Walasai and N. H. Mirjat, "An integrated modeling approach for forecasting long-term energy demand in Pakistan," Energies, vol. 10, p. 1868, 2017.

[16] R. Janković, "Forecasting Energy Consumption in Serbia using ARIMA Model," Virtual international conference on science, no. Technology and Management in Energy, 2017.

[17] O. A. Al-Naseem and A. K. Adi, "Impact of Power Factor Correction on the Electrical Distribution Network of Kuwait – A Case Study," The Online Journal on Power and Energy Engineering (OJPEE), vol. 2, pp. 173-176, 2011.

[18] G. Okoboi and J. Mawejje, "The impact of adoption of power factor correction technology on electricity peak demand in Uganda," Journal of Economic Structures, vol. 5, p. 3, 2016.

[19] S. CS, H.Hasim and A. A. AS, "Impact study on Power Factor of Electrical Load in Power Distribution System," Malaysian Nuclear Agency, Bangi, Kajang, Selangor (Malaysia), Technical Support Div., 2015.

[20] N. G. N. Daud, F. R. Hashim and M. H. A. Tarmizi, "Power Factor Regulation for Household Usage," AIP Conference Proceedings, vol. 1930, p. 020047, 2018.

[21] D. Kothari and I. Nagrath, Power system engineering, Tata McGraw-Hill New Delhi, 2008.

[22] V. D. Toro, Electrical engineering fundamentals, Prentice-Hall, 1986.

[23] C. Shearer, "The CRISP-DM Model: The New Blueprint for Data Mining," Journal of Data Warehousing, vol. 5, pp. 13--22, 2000.

[24] G. Eibl and D. Engel, "Influence of data granularity on smart meter privacy," IEEE Transactions on Smart Grid, vol. 6, pp. 930--939, 2014.

[25] G. Jain and B. Mallick, "A Study of Time Series Models ARIMA and ETS," SSRN Electronic Journal, 2017.

[26] S. Gelper, R. Fried and C. Croux, "Robust Forecasting with Exponential and Holt–Winters Smoothing," Journal of Forecasting, vol. 29, pp. 285--300, 2010.

[27] M. Smith and R. Agrawal, "A Comparison of Time Series Model Forecasting Methods on Patent Groups," MAICS, pp. 167--173, 2015.

[28] D. Shaub and P. Ellis, "CRAN," 12 02 2019. [Online]. Available: https://cran.r-project.org/web/packages/forecastHybrid/forecastHybrid.pdf.