



DEMAND FORECASTING OF ELECTRICITY CONSUMPTION USING DYNAMIC CLUSTERING OF TIME SERIES DATA

S.SIBI, PG Scholar

Dept. of Computer Science and Engineering
PSNA College of Engineering and Technology
Dindigul, TamilNadu
sibiselvam.94@gmail.com

Dr.S.PUSHPALATHA, Associate Professor

Dept. of Computer Science and Engineering
PSNA College of Engineering and Technology
Dindigul, TamilNadu
pushpa_pushpa@yahoo.com

Abstract—Utilization of power distribution networks enables the measurement of energy consumption of consumers over time. An advanced meter can measure energy consumption for every hour of every day to form a time series based load profile. In existing load demand forecasting techniques, demand is calculated by considering consumer system as a general class. By identifying various classes of consumers, demand forecasting can be made separately for different consumer classes. This project proposes a time series based dynamic clustering algorithm for identifying different classes of consumers using their load profile. The dynamic clustering algorithm uses the load profile of consumers to calculate similarities between them for clustering. Dynamic k-means clustering (DKMC) and Dynamic Fuzzy c-means (DFCM) clustering algorithms are used for clustering the load profiles of the different users. To calculate similarities between load profiles during clustering, Euclidean distance and correlation analysis are used. Dynamic clustering is implemented by considering different factors like seasonal consumption, type of consumer and area of consumption. Finally after clustering is executed for each cluster. Results from each cluster are modeled using Deep Belief Network (DBN) that uses Restricted Boltzmann Machines (RBMs). Predictions from each cluster are combined together to obtain final load prediction. The forecast of each cluster are combined together to obtain final load prediction. The forecast of each cluster shows the forecast of that particular factor representing the cluster. Also, electricity consumption for the next day is forecasted for each consumer within all clusters.

The forecasting cycle executes continuously using multi-forecast DBN that makes multiple forecasts to many future days using previous day forecasted value. To improve the accuracy in long-term forecasting, the Multi-Forecast DBN uses fine-tuning and regular pre-training process. The accuracy of the proposed Multi-Forecast DBN is measured using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The performance evaluation shows that the proposed multi forecasting model is more efficient compared to the existing load demand forecasting model.

Keywords—Time series data, Electricity consumption, dynamic clustering, Load demand forecast, Long-term forecast.

I. INTRODUCTION

There is an extensive measure of information accessible in the information industry. The information is of no utilization until the point when it is changed over into helpful data. It is important to examine this massive measure of information and concentrate helpful data from it. The data can be utilized in numerous applications, for example, Fraud Detection, Market Analysis, Production Control, Science Exploration and so forth. Forecasting of electricity load demand and electricity price has been a research focus since long before. Many

researches focus on improving the efficiency and performance of forecasting load demand to suit their applications. One major concern here is the exponentially increasing consumption and demand of electricity over time. As the field of technology has been improved tremendously the use of electrical appliances and technology has improved drastically over the past 10 years. It proposes a huge demand of electricity for the future and this demand keeps increasing more than ever. Many industries and organization uses automation and machinery that consumes large amounts of electricity and even at home, many appliances such as TV, A/C, etc. have been used.

Data mining is the process of extracting hidden predictive information from large database. Data mining is an intense new innovation with extraordinary potential to enable organization to consider on the most essential data in the information distribution centers. Data mining tool predict future trends and behaviors that involve knowledge driven decisions. Data mining technology can generate new business opportunities by providing automate prediction of trends, behaviors and automated discovery of previously unknown patterns. The architecture of a data mining system may have the following major components such as database, data warehouse, data warehouse server, data mining engine, pattern evaluation module and user interface.

Data analysis is the process to identify patterns within the data by extracting and modeling information. Since simplifying the amount of information to describe a huge set of data, features are extracted from the data. Feature selection plays an important role in several tasks, such as classification, cluster analysis, data mining, pattern recognition, etc. It is an important preprocessing technique for effective data analysis, in which the noise, irrelevant or redundant features are eliminated by choosing a subset from the original data features. By this process, the computational cost is reduced and accuracy of data analysis process is improved.

Motivation of the work

In recent times, due to the drastic improvement in the field of technology and automation the consumption of electricity has increased exponentially compared to the past 10 years. The major reason behind this is the use of technology in various field and areas to complete work in a faster and efficient manner. Where there is automation, there is always a large consumption of power. Due to this there is a large demand for electricity and this demand will keep rising in the future. There is need to understand and forecast this demand to make preparation for meeting these demands. Load demand forecasting is an important research focus that is a need in many fields. The existing load demand forecast techniques can accurately predict the demand for short-term and mid-term periods. But in case of long-term load forecasting, the accuracy in predicting the load decreases as the period increases. The

demand and improvements in technology also keeps changing over time and so there is also a need to improve the accuracy in long-term forecasting that will keep up with these changes.

II. RELATED WORK

Yongbao Chen et al (2017) proposed a new SVR method based on forecasting models for calculating DR baselines of office buildings. SVR forecasting model with the other seven traditional forecasting models are used to calculate the DR baselines. The SVR model proposed is suitable for real-time calculation, to expend the application and improve its efficiency.

Hongye Guo et al (2017) discussed about an External Economic Factors (EEFs) may have greater effects on electricity consumption. A monthly electricity consumption forecasting framework called SAS-SVECM is proposed to improve the accuracy and applicability of midterm. It is based on the self-adaptive screening method (SAS) and the vector error correction method (VECM). The correlation and time lag between the EEF and electricity consumption is addressed by VECM. SAS method is used to identify the most influential EEF set based on the three steps that is applied for forecasting. In which two different type of EEF are identified to improve the performance, forecasting and short-run equilibrium relationship of VECM respectively.

The author Xueheng Qiu et al (2017) uses the Empirical Mode Decomposition (EMD) algorithm and deep learning approach. Load demand series were first decomposed into several intrinsic mode functions. To model each of the extracted IMFs, the Deep Belief Network (DBN) including two Restricted Boltzmann Machines (RBMs) is used, so that the tendencies of these IMFs can be accurately predicted.

Saima Hassan et al (2016) presented a novel design of interval type-2 fuzzy logic systems (IT2FLS) by utilizing the theory of extreme learning machine (ELM) for electricity load demand forecasting. ELM has become a popular learning algorithm for single hidden layer feed-forward neural networks (SLFN). Competitive prediction performance of the IT2FELM over IT2FKF with noisy data sets shows that IT2FELM can be utilized to model unpredictable consumer response in the smart grid environment.

E.E. Vogel et al (2017) proposed a time-series analysis to detect the areas where seismic energy can be occurred. It also gives an information theory approach to the study of seismicity. Three places were chosen, where there is a large occurrence of earthquake for the past 7years. These seismic data are analyzed by a tunable information recognizer known as wzip (word length zipper). At specific position in numeric information, only the exact matches of the sequence is recognized by the wzip. The time interval between the consecutive events is found in the variable by using the information theory. Thus the seismic data from geographical regions around the epicenter of important earthquake are considered on the extended time basis, before and after the earthquake. Each sequence contains two parameters such as

mutability and interval dilation that are related to the information in it. Mutability increases before the earthquake and it decreases after each events. The interval dilation increases maximum before the earthquake of magnitude 7.0 and above. It is decreased to the lowest value after the earthquake. In this period no false overshoot is found, since at different time period, it has the same behavior for three different geographical areas. By using this method it is possible to find the different way in the seismic activity.

Weihui Deng et al (2017) stated that nowadays water quality problem is a major issue in environment management department. The author describes a novel method and an analysis framework for large amount of water quality time-series data by using the time-series data mining method (TSDM). It is an important research for identifying the patterns hiding and knowledge in time-series data. TSDM is mainly used as an analysis framework. Dissolved oxygen (DO) is used in the analysis framework, which is for assessing the water quality of rivers. This DO is analyzed on weekly basis. The DO with low concentration leads to several problem and it is an unbalanced ecosystem. The cloud model is used to find the deviation of random phenomenon.

V. Prema et al (2015) proposed that the use of renewable energy source is increasing day-by-day. The main objective of the work is the short-term prediction of solar power from the solar irradiance by using the time series models. Basic time series models such as moving average and exponential smoothing methods were used. Moving average models is a simple method, that which smooth the past history data. The different moving averages are simple moving average, double and weighted moving average. Exponential smoothing is the extension to the moving average method, which is predicted by weighted moving average. The accuracy of forecasting was evaluated by estimating the mean absolute error and mean absolute percentage error (MAPE). The variation between the predicted solar irradiance and actual solar irradiance is called the error. The decomposition model of MAPE is as less as 9.28%. On cloudy days the error rate is increased because of the sudden change in solar irradiance.

III. PROPOSED WORK

The major criteria to consider in case of long-term demand forecasting, is to increase the accuracy of forecasting through a long period of time. The proposed method in this paper uses the Dynamic Clustering algorithm and the Multi-Forecast Deep Belief Network (DBN) model for long-term forecasting of electricity load demand. Initially the input dataset collected is pre-processed to convert it into the required time series format. This time series is processed using the proposed Dynamic k-means Clustering algorithm and the Dynamic Fuzzy c-means Clustering algorithm separately. A comparison is made between these two clustering algorithms. The clusters are then processed using the proposed Multi-Forecast DBN to obtain long-term forecast values. The advantage of the proposed method is that fine-tuning and back propagation is employed to improve the forecasting accuracy. Another advantage is that,

the forecast are made based on three different classifications that make it useful to obtain deeper analysis for the future.

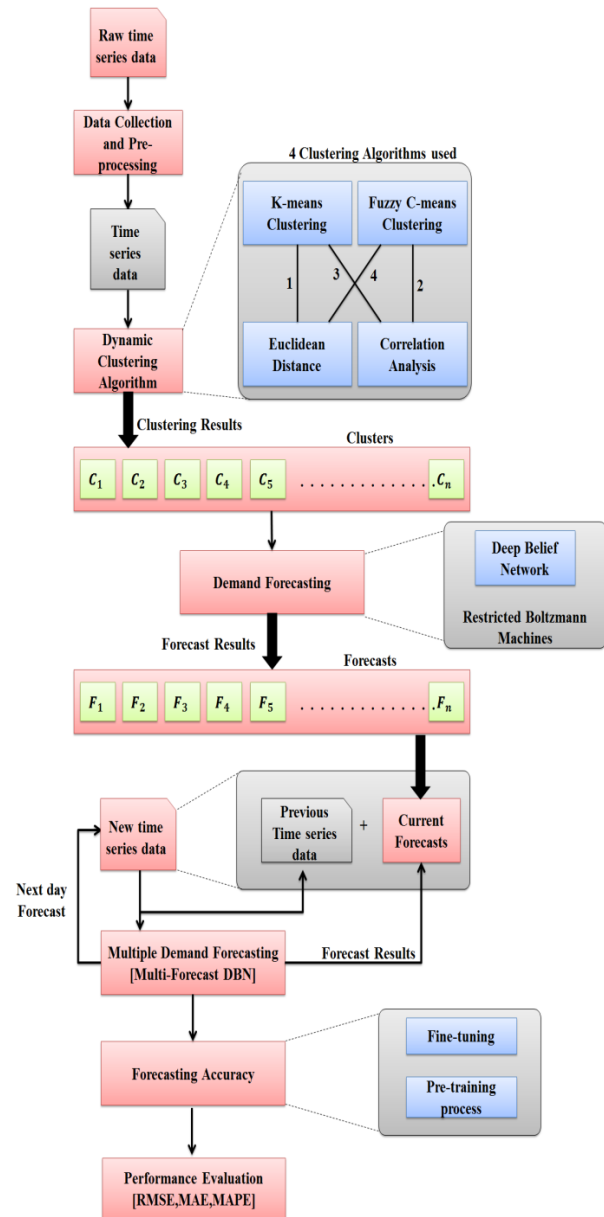


Figure 3.1 overall architecture diagrams

A. Data Collection and Pre-processing

Data collection is the process of collecting data from various sources for the purpose of analysis. Data can be collected from two sources, primary sources and secondary sources. After data collection the raw time series data is obtained that should be pre-processed to obtain the required time series data for our analysis. Pre-processing deals with

handling missing value, wrong values, removing unwanted values, alignment of data etc.,

B. Dynamic clustering algorithm

The dynamic clustering algorithm is the modified form of cluster algorithm for the purpose of handling the dynamic change in time series data over time. Due to this change, the cluster arrangement also keeps changing with time based on the changes in each of the instances that are being changed. In case of the electricity consumption data, the cluster can be formed based on various factors that affect the overall consumption of electricity. These factors can be classified into three as below:

i. Cluster based on seasonal changes – Change in season is a major cause that affects the consumption and production of electricity. Seasons are classified into four such as winter, summer, fall and spring.

ii. Cluster based on application area – The application or usage of electricity in various fields or area is different and this also affects its consumption. The area of application is classified into various fields such as Domestic, Commercial, Industry and Others.

iii. Cluster based on volume of consumption – Apart from these two categories, there is another simple classification factor that considers only the overall volume of electricity consumed. This can be classified into any number of categories such as Very Low, Low, Medium, High, Very High, etc.

The proposed Dynamic clustering algorithm takes into consideration all three types of classification to perform clustering. This way the load demand can be forecasted based on these different factors. For this purpose the Dynamic k-means clustering (DKMC) and the Dynamic Fuzzy c-means clustering (DFCM) algorithms have been employed (Ignacio Benitez et al 2016). The pre-processed time series data is given as input to the Dynamic clustering algorithm. Here for the purpose of calculating distance, the Euclidean distance measure and the Correlation analysis measure are used.

k-means Clustering

The k-means clustering is a distance based clustering mechanism that initially chooses random cluster values and then clusters each instance or items within the data based on their distance to the cluster centers. The k-means clustering algorithm will calculate the mean of all items within a cluster to calculate its cluster center value. In most cases the k-means algorithm uses Euclidean distance for clustering but in case of time series data other measures such as correlation analysis or cosine similarity can also be used. The proposed Dynamic k-means clustering (DKMC) algorithm uses the k-means approach dynamically to cluster electricity time series data based on certain pre-defined factors as discussed previously.

Fuzzy c-means clustering

The Fuzzy c-means clustering algorithm calculates the membership values of each instance or data within the time series by making use of fuzzy sets. The fuzzy set value here corresponds to the cluster center or the centroid. The number of fuzzy sets can be set to any value based on the need and this number is the same as the number of clusters that will be

formed. The Fuzzy c-means clustering makes use of a fuzzification parameter called as 'm' within the range [1, N] and this will determine the degree of fuzziness of the clusters. If the value of m is set to be 1, then the effect will be the crisp clustering of points into the clusters. As the value of m increases the degree of fuzziness between the points or data also increases. Similar to the k-means clustering, the Fuzzy c-means clustering also selects random cluster centers initially and then uses the Euclidean distance or Correlation measure to calculate the distance of each data or instance to each of the clusters. After this step, the Fuzzy c-means clustering will calculate the membership matrix that includes membership values of each instance with that or all the clusters. The sum of membership of a single instance with all clusters will always be 1. Then using these membership values of the instances or data, the new cluster centers are calculated and the algorithm keeps iterating till the optimal cluster split has been generated.

C. Demand Forecasting

The proposed demand forecasting model here makes use of the Deep Belief Network (DBN) for forecasting the load demand (Ping Zhong et al 2016). The Deep Belief Network (DBN) is a graphical model that consists of many layers with each layer consisting of hidden units or latent units. All these layers are connected between each other but there is no connection between the hidden units of one layer and the other. The DBN makes use of the Restricted Boltzmann Network (RBM) stacked up on one another for the unsupervised learning process. An RBM is an energy based model that consists of one visible layer and one hidden layer that are connected together. By stacking up many RBMs over one another, the hidden layer of one RBM becomes the visible layer of the next RBM. This way the unsupervised training process takes place using a greedy algorithm and this is called as the Deep Belief Network (DBN).

D. Multiple Load Demand Forecasting

In the proposed approach, the Deep Belief Network (DBN) training model is enhanced by making use of fine-tuning and back-propagation algorithm, this process of enhancement is known as multiple load demand forecasting. All the parameters used within the deep learning architecture to improve the efficiency in the training process and to improve the forecasting accuracy. The fine-tuning here is done with respect to a supervised training criterion, which makes use of additional learning process to obtain accurate forecasting from the learned data. For this purpose the supervised gradient descent based fine-tuning is done on the output of the last hidden layer of the DBN that makes use of the logistic regression analysis and log-likelihood cost function. This will enhance the obtained forecasting result and provide better accuracy. Apart from the fine-tuning, the back-propagation algorithm is also used within the proposed DBN. According to the back-propagation algorithm, whenever the training results obtained from one layer is not up to the satisfactory level then the process goes back to the starting of that layer to implement the training process again. The back-propagation algorithm makes sure that the training result of one layer is more efficient

and provides better training results compared to the previous layer. This way the final output obtained at the last layer is already an optimized one and this is further fine-tuned to obtain the most optimal forecasting result. This way the proposed DBN based demand forecasting can be used for long-term demand forecasting with enhanced accuracy.

E. Performance Analysis

Performance analysis is the process of calculating the performance of the algorithm used by using certain metrics. These metrics are called as performance metrics and different metrics are user for different types of algorithms and applications. In case of forecasting, the performance of the forecasting algorithm can be shown by identifying its accuracy and error rate. The lesser the error rate, the higher is the performance and accuracy. The most commonly used performance analysis metrics for forecasting algorithms are:

i. Root Mean Squared Error (RMSE)

The RMSE shows the error that is prevailed between the actual forecast obtained f and the expected forecast e .

$$RMSE = \sqrt{\frac{1}{N} \sum_{(i)} (f_i - e_i)^2} \quad (3.1)$$

Here N is the total number of instances or users and $i = 1, 2, \dots, N$.

ii. Mean Absolute Error (MAE)

The MAE shows the average absolute difference between the actual forecast obtained f and the expected forecast e .

$$MAE = \frac{1}{N} \sum_{(i)} |f_i - e_i| \quad (3.2)$$

iii. Mean Absolute Percentage Error (MAPE)

The MAPE is the average absolute percentage difference between the actual forecast obtained f and the expected forecast e .

$$MAPE = \frac{1}{N} \sum_{(i)} \frac{|f_i - e_i|}{|e_i|} \quad (3.3)$$

IV. RESULT

For the purpose of implementation, the USA electrical dataset of 2004 has been taken that consists of many cities. For implementation only 3 cities has been taken and each of these cities has electricity consumption for different buildings as given in the list below.

FullService Restaurant	QuickService Restaurant
Hospital	Secondary School
Large Hotel	Small Hotel
Large Office	Small Office
Medium Office	Stand-alone Retail Shop
Mid-rise Apartment	Strip Mall
Out Patient Hospital	Super Market
Primary School	Warehouse

So, each city has a total of 16 files for each of the buildings as above and so for 3 cities the total number of files is 48. Each file contains the electricity and gas consumption data for every hour for the whole year 2004. Also the electricity consumption and gas consumption are split based on the different rooms.

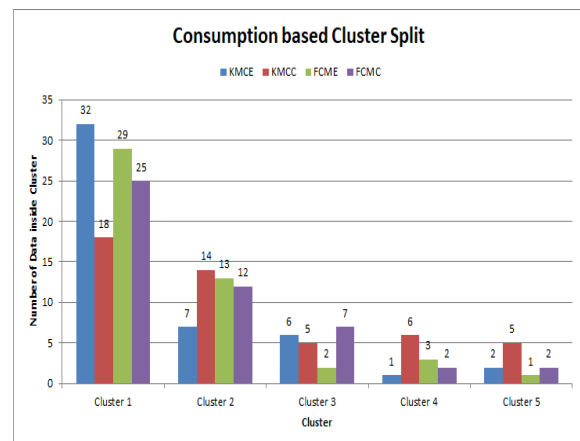


Fig.2. Cluster result based on electricity consumption

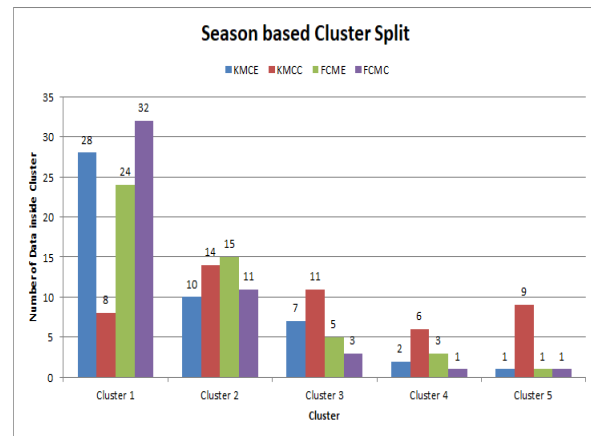


Fig.3. Cluster result based on seasons

During the pre-processing stage, the gas consumption values are removed and the electricity consumption values for different rooms are added. The final preprocessed file contains total electricity consumption for every hour. During the first stage of implementation the clustering algorithm is implemented to cluster all the 48 files and for this the k-means

clustering and Fuzzy c-means clustering is used with Euclidian distance and Correlation analysis separately.

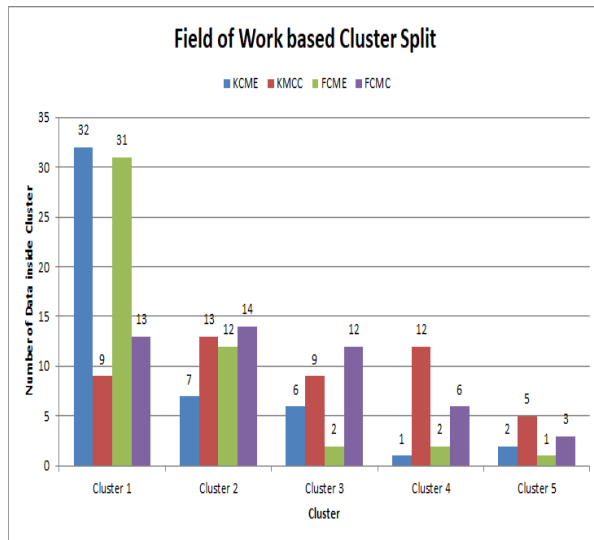


Fig.4. Cluster result based on field of work

The cluster results are shown below in Fig. 2, Fig. 3 and Fig. 4 respectively for Consumption based, Season based and Field based clustering.

After the clustering process the multiple forecasting is implemented for all three types of cluster results separately to forecast the electricity consumption for the next year 2005. Fig. 5, Fig. 6 and Fig.7 shows the overall forecast error for each month of forecast for the year 2005 respectively for Consumption based, Season based and Field based clustering.

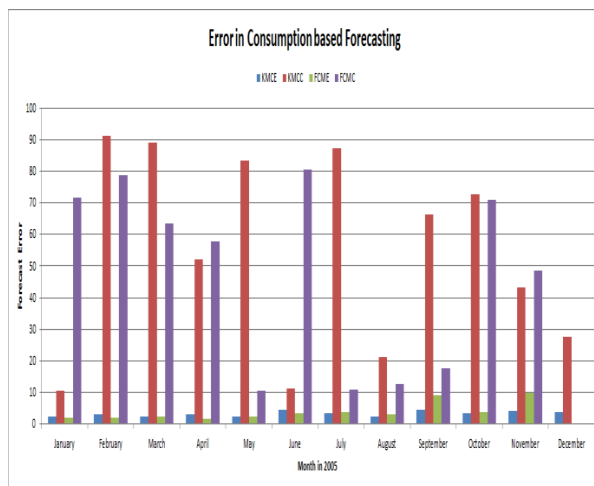


Fig.5 Forecast error in electricity consumption based forecasting

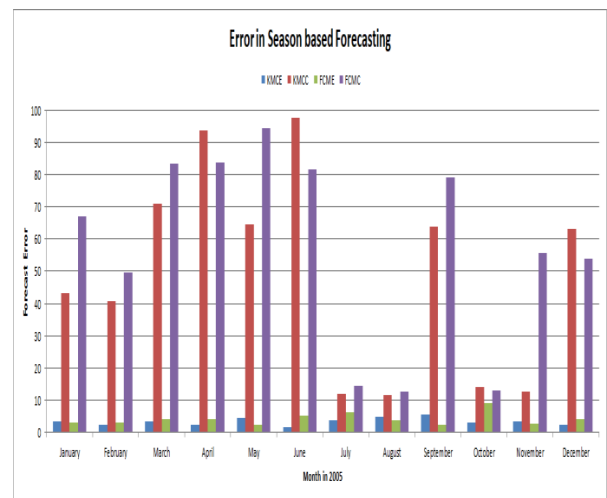


Fig.6 Forecast error in season based forecasting

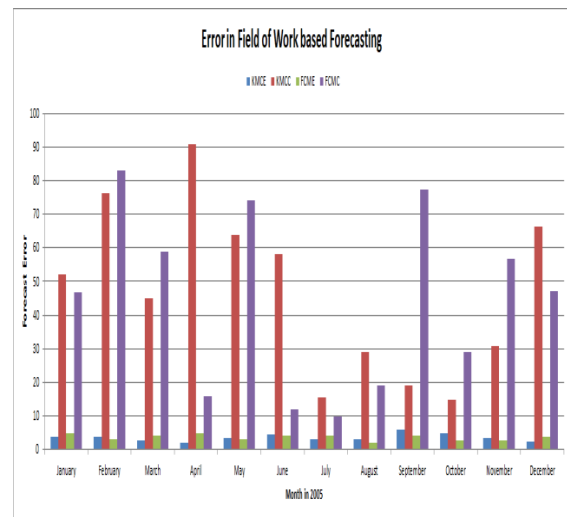


Fig.7 Forecast error in field of work based forecasting

The results show that the forecast value obtained by using the Euclidean distance has relatively lesser error compared to the forecast made using the Correlation analysis. Also it can be seen that different forecast result has been obtained for different type of clustering options. So based on this the detailed forecast can be obtained for analysis.

V. CONCLUSION

Load forecasting is an important field in recent times and there is a need to forecast the demand of electricity for the future concern. In this project the load forecasting is done based on different criteria such as the season, electricity consumption and the field where the electricity is consumed. By forecasting the load based on different aspects deeper analysis and findings can be obtained. The detailed study on electricity load forecasting is done by studying the various types of forecasting and the various algorithms used. The need

for load demand forecasting and its applications has also been studied by taking various research papers. A detailed survey has been made with the existing methods and their advantages and disadvantages have been noted. A suitable dataset has been downloaded for the proposed multiple load forecasting model that contains all the three required factors. The data is collected and then pre-processed to make it suitable for implementing the next phase for demand forecasting. The pre-processed data will be split based on the three proposed factors using the dynamic clustering algorithm. Both the k-means and the Fuzzy c-means clustering will be used separately. After clustering, the Deep Belief Network (DBN) based load demand forecasting will be implemented for each cluster and then the Multi-Forecast DBN will be implemented to forecast long-term load demand. To improve the accuracy in long-term demand forecasting, the fine-tuning and back propagation algorithms will be employed within the Multi-Forecast DBN model. In the future, the accuracy in forecasting will be further enhanced by using an intelligent and automated Fuzzy based Logical System that takes into consideration many aspects and conditions that apply to real time reasoning of electricity consumption that depends on factors such as people, climate, area, type of organization, type of work and the social behavior.

REFERENCES

- [1] Abderrezak Laouafi, Mourad Mordjaoui, Farida Laouafi, Taqiy Eddine Boukelaia (2015), 'Daily peak electricity demand forecasting based on an adaptive hybrid two-stage methodology', *Electrical Power and Energy Systems*, Vol.77, pp.136-144.
- [2] Abderrezak Laouafi, Mourad Mordjaoui, Salim Haddad, Taqiy Eddine Boukelaia, Abderahmane Ganouche (2017), 'Online electricity demand forecasting based on an effective forecast combination methodology', *Electric Power Systems Research*, Vol.148, pp.35-47.
- [3] Eric Burger M., Scott Moura J. (2015), 'Gated ensemble learning method for demand-side electricity load forecasting', *Energy and Buildings*, Vol.95, pp.427-439.
- [4] Fatemeh Chahkoutahi, Mehdi Khashei (2017), 'A seasonal direct optimal hybrid model of computational intelligence and soft computing techniques for electricity load forecasting', *Energy*, doi: 10.1016/j.energy.2017.09.009.
- [5] Fayssal Beloufa, Chikh M.A. (2013), 'Design of fuzzy classifier for diabetes disease using Modified Artificial Bee Colony algorithm', *Computer methods and programs in biomedicine*, Vol.112, pp.92-103.
- [6] Francisco Martı́nez-Alvarez, Alicia Troncoso, Jose Riquelme C., and Jesu's Aguilar-Ruiz S. (2011), 'Energy Time Series Forecasting Based on Pattern Sequence Similarity', *IEEE TRANSACTIONS ON knowledge and data engineering*, Vol. 23, No. 8, pp.1230-1243.
- [7] Ghasemi A., Shayeghi H., Moradzadeh M., Nooshyar M. (2016), 'A novel hybrid algorithm for electricity price and load forecasting in smart grids with demand-side management', *Applied Energy*, Vol.177, pp.40-59.
- [8] Hesam Izakian, Witold Pedrycz (2013), 'Agreement-based fuzzy C-means for clustering data with blocks of features', *Neurocomputing*, Vol.127, pp.266-280.
- [9] Hisashi Takeda, Yoshiyasu Tamura, Seisho Sato (2016), 'Using the ensemble Kalman filter for electricity load forecasting and analysis', *Energy*, Vol.104, pp.184-198.
- [10] Hongye Guo, Qixin Chen, Qing Xia, Chongqing Kang, Xian Zhang, (2017), 'A monthly electricity consumption forecasting method based on vector error correction model and self-adaptive screening method', *Electrical Power and Energy Systems*, Vol.95, pp.427-439.
- [11] Hyojoo Son, Changwan Kim (2016), 'Short-term forecasting of electricity demand for the residential sector using weather and social variables', *Resources, Conservation and Recycling*, Vol.123, pp.200-207.
- [12] Mashud Rana, Irena Koprinska (2015), 'Forecasting electricity load with advanced wavelet neural networks', *Neurocomputing*, Vol.182, pp.118-132.
- [13] Omer Faruk Ertugrul (2015), 'Forecasting electricity load by a novel recurrent extreme learning machines approach', *Electrical Power and Energy Systems*, Vol. 78, pp.429-435.
- [14] Prema V., Uma Rao K. (2015), 'Development of statistical time series models for solar power prediction', *Renewable Energy*, Vol.83, pp.100-109.
- [15] R.Karthikeyan, A.Prabha, "A Novel Approach For Load Balancing In The Cloud Computing", *International Journal of Innovations in Scientific and Engineering Research (IJISER)*, Vol.2, No.4, pp.83-88, 2015.
- [16] Saima Hassan, Abbas Khosravi, Jafreezal Jaafar, Mojtaba Ahmadi Khanesar (2016), 'A systematic design of interval type-2 fuzzy logic system using extreme learning machine for electricity load demand forecasting', *Electrical Power and Energy Systems*, Vol.82, pp.1-10.
- [17] Ujjwal Kumar, Jain V.K. (2010), 'Time series models (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India', *Energy*, Vol.35, pp.1709-1716.
- [18] Vogel E.E., Saravia G., Pasten D., Munoz V. (2017), 'Time-series analysis of earthquake sequences by means of information recognizer', *Tectonophysics*, doi: 10.1016/j.tecto.2017.06.031.
- [19] Vu D.H., Muttaqi K.M., Agalgaonkar A.P. (2014), 'A variance inflation factor and backward elimination based robust regression model for forecasting monthly electricity demand using climatic variables', *Applied Energy*, Vol.140, pp.385-394.
- [20] Warren Liao T. (2005), 'Clustering of time series data—a survey', *Pattern Recognition*, Vol.38, pp.1857-1874.
- [21] S. Manjappa and Dr.M.Govindaraj, Effects of the Varied Regimens of Training Program on Explosive Power and Flexibility of Industrial Training Students", *International Research Journal of Multidisciplinary Science & Technology (IRJMRS)*, Vol.2, No.1, pp.43-46, 2017.
- [22] Weihui Deng, Guoyin Wang (2017), 'A novel water quality data analysis framework based on time-series data mining', *Journal of Environmental Management*, Vol.196, pp.365-375.
- [23] Xueheng Qiu, Ye Ren, Ponnuthurai Nagarathnam Suganthan, Gehan Amaratunga A.J. (2017), 'Empirical Mode Decomposition based Ensemble Deep Learning for Load Demand Time Series Forecasting', *Applied Soft Computing Journal*, <http://dx.doi.org/doi:10.1016/j.asoc.2017.01.015>.
- [24] Xue-Hui Meng, Yi-Xiang Huang, Dong-Ping Rao, Qiu Zhang, Qing Liu (2013), 'Comparison of three data mining models for predicting diabetes or prediabetes by risk factors', *Journal of Medical Sciences*, Vol.29, pp.93-99.

- [25] Yan Zhang, Rui Wang, Tao Zhang, Yajie Liu, Bo Guo (2016), 'Model predictive control-based operation management for a residential microgrid with considering forecast uncertainties and demand response strategies', IET Generation, Transmission & Distribution, Vol. 10, Iss. 10, pp. 2367–2378.
- [26] Ye Rena, Suganthana P.N., Srikanth N., Gehan Amaratunga (2015), 'Random vector functional link network for short-term electricity load demand forecasting', Information Sciences, Vol.367–368, pp. 1078-1093.
- [27] Yongbao Chen, Peng Xu, Yiyi Chu, Weilin Li, Yuntao Wu, Lizhou Ni, Yi Bao, Kun Wang (2017), 'Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings', Applied Energy, Vol.195, pp.659-670.

