INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 4, Issue 3)
Available online at: www.ijariit.com

Long term load forecasting using K-mean clustering and ANN approach

Shelly Bhanot
shelly.sharma27@gmail.com
DAV Institute of Engineering & Technology, Jalandhar, Punjab

Shivani Mehta
shivanimehta7@gmail.com
DAV Institute of Engineering & Technology, Jalandhar, Punjab

ABSTRACT

This study implements the K-mean+ANN based Load forecasting technique to predict the load of Amritsar and Pathankot station. Artificial Neural Network (ANN) is one of the emerging methods used for forecasting the load. This method shown good results in different power station problems included planning, protection, designing, control & security analysis and fault diagnosing. In this paper, it is illustrated that from last few years the load forecasting has been widely adopted and this is due to increase in the demand for electric power and this had also resulted in an increase of generating sources expenditure. In most of the cases, it had been implemented in utilizes to determine the number of resources required to fulfill the demands of the project. The major focus area of the research was load prediction of two different stations for 9 upcoming years. In this study work past 18 years load information was collected and different years were select as base years and after that, the author had implemented the k-mean clustering and ANN method to predict load for upcoming nine years.

Keywords: *Load forecasting, Load data set of Amritsar-Pathankot station, ANN classification, K-means clustering*

1. INTRODUCTION

Forecasting is one of the needs for all industries that play the vital role in the growth. Electric utilities are taken as one of the most complex systems designed on earth in order to run the power grid that delivers the electricity to billions of people around the globe [1]. After turning on the switch we expect that light must turn on. But the system from generating the power to deliver that to the electrical appliances is not that simple. As in other industries, the product and services are stored in some kind of inventory, but this is not possible for the electrical power industry by using the currently present technology [2,3,4]. This is the reason electrical energy is delivered and consumed as soon as it is generated. Or it can be said that the supply and demand of electrical utility are needed to be balanced at every moment [5,6].

Load forecasting mainly concentrates on forecasting the demand for electricity that is used in different electrical power industry segments, that includes the transmission, generation, distribution, and retail [7]. Load forecasting application includes the planning of power supply, transmission, and distribution, management at demand side [8], operation and maintenance. The role of Load forecasting is important in order to plan [9], operate and control the power system. Forecasting means active load estimation for different load busses that is connected to the actual load occurrence. In a good forecasting model, some of the important features such as climate, weather, economy, human activities, and interaction are captured [10]. Certain lead time was required for the planning and operational application of the load forecasting and is known as the forecasting interval. This forecasting can become handy for analyzing support of strengthening and also for the expansion of the existing infrastructure, maintenance scheduling, controlling voltage and development of the infrastructure [11].

2. TYPES OF LOAD FORECASTING

On the basis of the time period, load forecasting technique can be classified into three categories:-

1) Short-Term Load Forecasting (STLF)
2) Medium Term Load Forecasting (MTLF)
3) Long-Term Load Forecasting (LTLF)

2.1 Short Term Load Forecasting (STLF)

In the STLF, daily or hourly values were covered for giving the necessary information of the system management in order to achieve the daily operations. This information was used for allocating the spinning reserve, planning the operation, committing the units and for scheduling the maintenance.

2.2 Medium Term Load Forecasting (MTLF)

In this type for forecasting, data from few days or few weeks was covered. That forecasting data was generally used for the fuel supply and unit maintenance scheduling. It is less time consuming than the long-term forecasting and consumes more time than the short term load forecasting. The basic application of the MTLF is to plan the seasonal peak load in the winter and summer seasons [12].

2.3 Long-Term Load Forecasting (LTLF)

In the LTLF, long time span such as in years are considered for planning. Prediction of future need such as expansion, staff recruitment, and equipment purchase is generally included in the long-term forecasting. The most common example of long-term forecasting is the General growth planning.

3. LITERATURE SURVEY

In [27], the author had proposed the data mining method for short-term load forecasting (STLF). First of all, the K-mean clustering algorithm had been implemented to categorize the load information into the four different types of the pattern on the basis of season. After this, the k-NN algorithm was implemented to disintegrate the categorized information into four patterns for Mondays, other weekdays, Saturdays, and Sundays. This categorized information was implemented to create a time series prediction model. After this, the prediction for load was done on the on weekdays and weekends, not including special holidays. Past load information can be implemented as inputs for load prediction.
In [28] author had presented a novice integrated technique for the short-term load forecasting of electric power systems by using the Fuzzy c-means (FCM) clustering technique, particle swarm optimization (PSO) and support vector regression (SVR) approaches. In this new technique proposed by the author, the training samples were made up of similar data type and these learning samples were used in the prediction process. These training samples were selected with the help of fuzzy clustering technique on the basis of the similarity index among the input samples and while selection the periodic features of load were also taken into consideration. In order to make the selected parameters optimum, the PSO technique was implemented. Nonlinear characteristics drawn among the parameters affecting the load as well as load prediction can be regressed using the SVR. In order to analyze the technique introduced by the author the actual load information from a city in Chongqing was collected and used for analysis purpose. The obtained results had shown that this method was highly precise.

In [28], the author had represented the K- mean clustering technique to forecast the load shedding and obtained the output for the pattern extraction process. Information was comprised of all loads in 40 electrical power stations in southern Libya. This data was collected for a two-year period. To determine the effectiveness of the k-means clustering algorithm a test was conducted in which the clustering of information was done so that the future forecasting for load shedding can be done in the SEGL. Five times all the clusters were developed in order to produce five different clusters with different size (1, 2, 5, 7 and 10) and these clusters were comprised of different seed values. Obtained pattern reveals the data on all the attributes. After testing the results had proved that produced clusters are optimum to be implemented in case of future load shedding schedule issue in case of the SEGL.

In [30], Author had represented the load characteristics on daily basis of a residential feeder and then this collected information was distributed in the different groups with the help of K-Means clustering algorithm to categorize the load characteristic curves. In this study, the relationship among the load profiles and the seasonal time periods were also developed in order to find out the types of season. In this study work, the truncated discrete Fourier transform coefficients were also derived which can be used in the load curves to minimize the number of problems in the clustering.

4. PROPOSED WORK

From last few years, the load forecasting has been widely adopted and this is due to increase in the demand for electric power and this had also resulted in an increase of generating sources expenditure. In most of the cases, it had been implemented in utilizes to determine the number of resources required to fulfill the demands of the project. The major focus area of the research was load prediction of two different stations for 9 upcoming years. These two stations were at Amritsar and Pathankot. In Amritsar station, the standard of living of the people increasing day by day, therefore, the load growth is predicted to increase at a fast pace. In case of Pathankot station, the load will be almost stagnant. In this study work past 18 years load information was collected and different years were select as base years and after that, the author had implemented the k means clustering and ANN method to predict load for upcoming nine years.

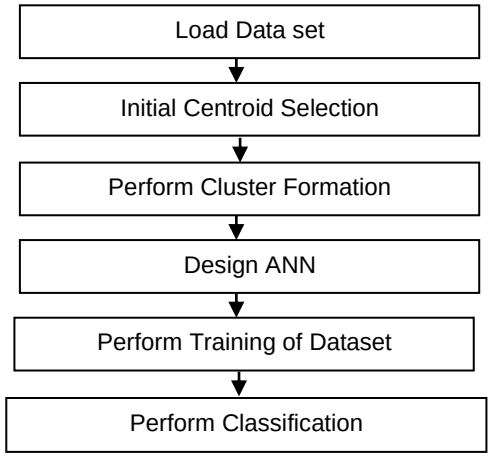

Fig. 1: Proposed framework

The above figure delineates the methodology of proposed work in a flowchart format.

1. **Load Dataset:** First step is to load the data set of the related work. In this work, we have loaded the data set of Amritsar and Pathankot stations of previous 14 years. On the basis of the gathered dataset, the further processing is done.

2. **Centroid Selection:** In this section, the centroid selection is done. The centroid is selected is done on the random basis. Then on the basis of the selected centroid, the cluster formation is initiated.

3. **Cluster Formation:** The Kmean clustering technique is used for creating the clusters of the created dataset. The cluster formation creates the clusters of the data set on the basis of the nature of the data. It categorizes the dataset in various clusters and each clusters comprised of the similar natured dataset.

4. **Design ANN and perform Training:** In this step, the artificial neural network designs is modeled. After creating the ANN model, next step is to perform training. The training of dataset is done on the basis of the created clusters. On the basis of the trained dataset, the ANN would perform the testing and generate a result or decision.

5. **Perform Classification:** This step performs the classification of the dataset on the basis of the trained dataset. After performing the classification, next step is to perform the performance evaluation of the proposed technique. In order to evaluate the performance, various performance metrics have been used in this work such as an actual load of the plant, calculated plant and error rate.

5. RESULTS

This study implements the Kmean Clustering and ANN mechanism to predict the load for Amritsar-Pathankot stations. The load forecasting s done on the basis of the dataset of the previous 18 years and it is predicted for the next upcoming 9 years. For achieving the objective, the author gathers the data from previous years and is shown in the table below:

Table 1: Load dataset for Amritsar station

Years	Actual Load	Population
2000	45.73	9,10,012
2001	43.21	9,15,144
2002	62.02	9,20,111
2003	60.43	9,25,010
2004	63.58	9,34,010
2005	63.42	9,35,190
2006	73.01	9.48,190
2007	71.92	9,62,190
2008	105.09	9,76,453
2009	111.02	9,91,063
2010	114.49	11,04,069
2011	114.44	11,18,761
2012	129.04	11,40.017
2013	129.21	11,51,273
2014	140.36	11,90,940
2015	146.42	12,06,327
2016	156.49	12,20,456
2017	163.42	12,84,845

Table 2: Load dataset for Pathankot station

Years	Actual Load
2000	6.00
2001	6.08
2002	6.02
2003	6.11
2004	6.46
2005	6.6
2006	6.61
2007	6.73
2008	6.76
2009	6.86
2010	7.9
2011	10.5
2012	9.26
2013	10.21
2014	10.27
2015	10.85
2016	10.96
2017	11.26

Table 3: Load dataset for Amritsar station (Proposed work)

Years	Actual Load	Calculated Load	% error
2000	45.73	45.73	0.00
2001	43.21	43.21	0.00
2002	62.02	62.02	0.00
2003	60.43	60.43	0.00
2004	63.58	63.58	0.00
2005	63.42	63.42	0.00
2006	73.01	73.01	0.00
2007	71.92	71.92	0.00
2008	105.09	80.79	24.56
2009	111.02	111.02	0.00
2010	114.49	114.49	0.00
2011	114.44	114.44	0.00
2012	129.04	129.04	0.00
2013	129.21	129.21	0.00
2014	140.36	140.36	0.00
2015	146.42	142.89	-3.53
2016	156.49	156.49	0.00
2017	163.42	163.42	0.00

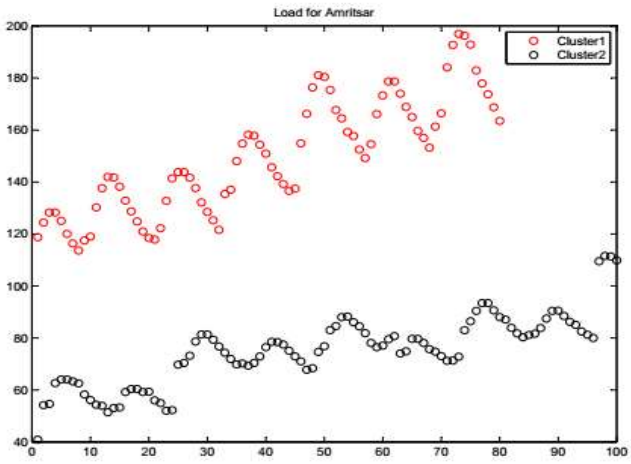

Fig. 2: K-mean clustering with Amritsar load

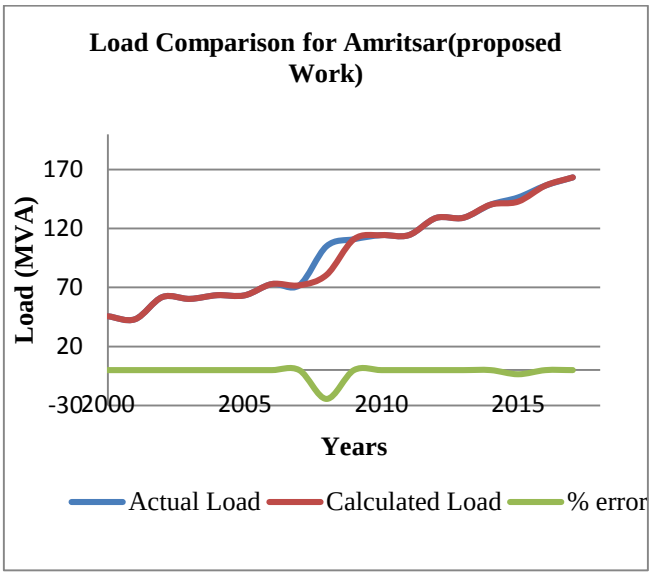

Fig. 3: Load comparison using Kmean-ANN technique for Amritsar station

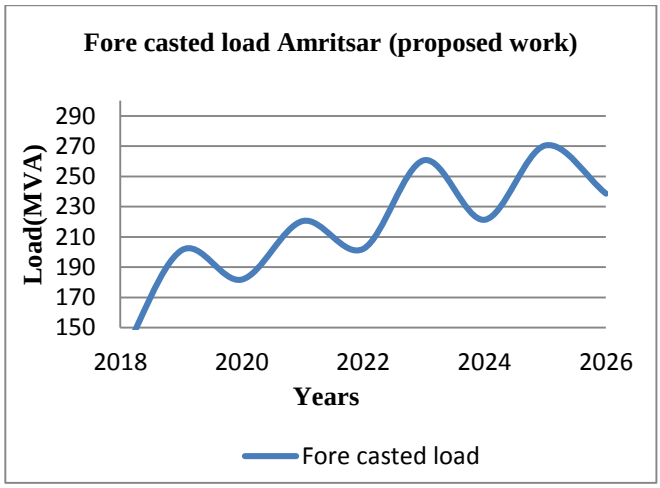

Fig. 4: Forecasted load for next nine years

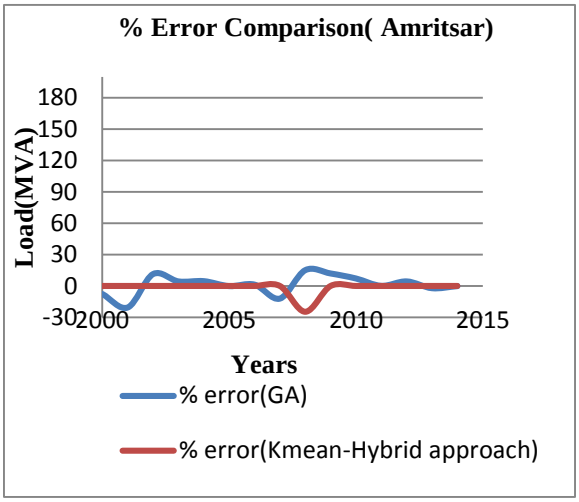

Fig. 5: Error comparison of Amritsar load (k-mean-NN&GA [11])

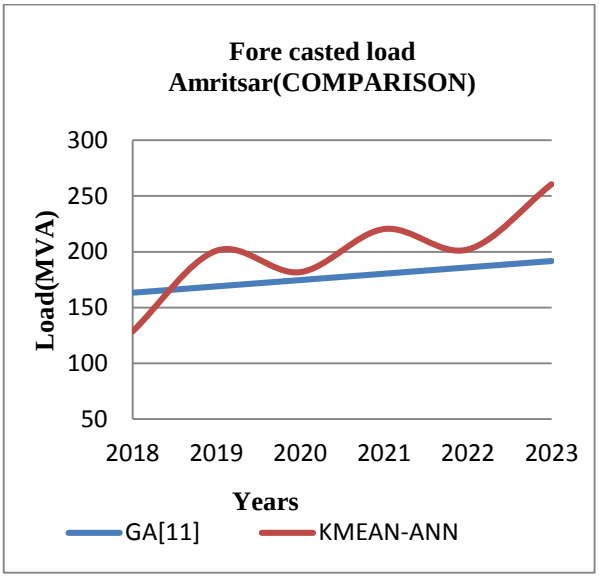

Fig. 6: Forecasted load comparison of Amritsar from 2018-2023 (k-mean-ANN & GA [11])

Table 4: Load dataset for Pathankot station (Proposed work)

Years	Actual Load	Calculated Load	% error
2000	6	6	-1.67
2001	6.08	6.08	-2.32
2002	6.02	6.02	-5.15
2003	6.11	6.12	-5.75
2004	6.46	6.5	-9.2
2005	6.6	6.66	0
2006	6.61	7.06	-4.69
2007	6.73	7.66	-8.69
2008	6.76	7.43	-14.2
2009	6.86	9.9	-19.44
2010	7.9	11.97	-7.84
2011	10.5	14.42	3.92
2012	9.26	9.26	0.00
2013	10.21	10.14	0.00
2014	10.27	10.27	0.00
2015	10.52	13.17	2.65
2016	10.75	12.54	1.79
2017	11.20	12.05	0.85

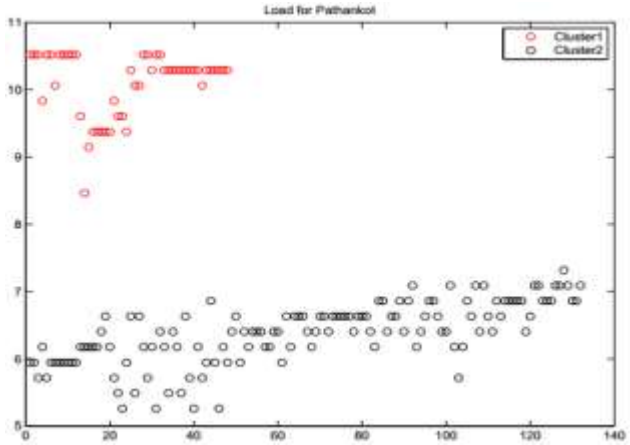

Fig. 7: K-mean clustering with Pathankot load

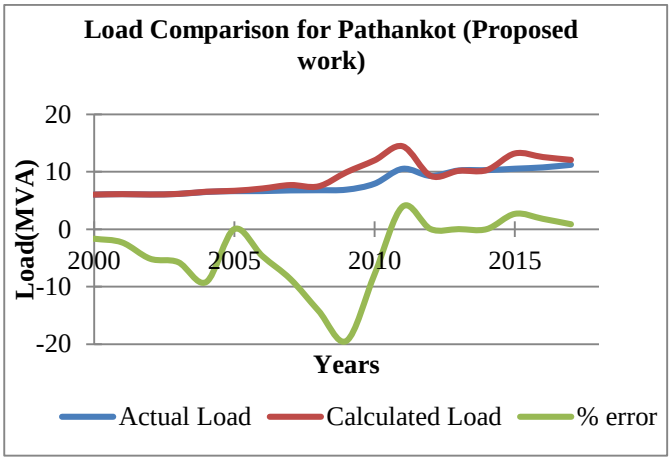

Fig. 8: Load comparison using kmean-ANN technique for Pathankot station

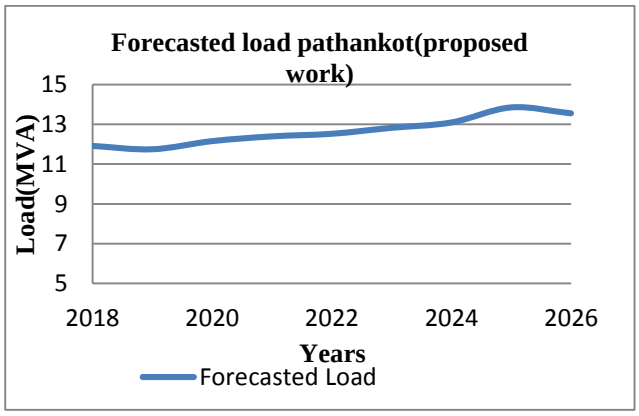

Fig. 9: Forecasted load for next nine years

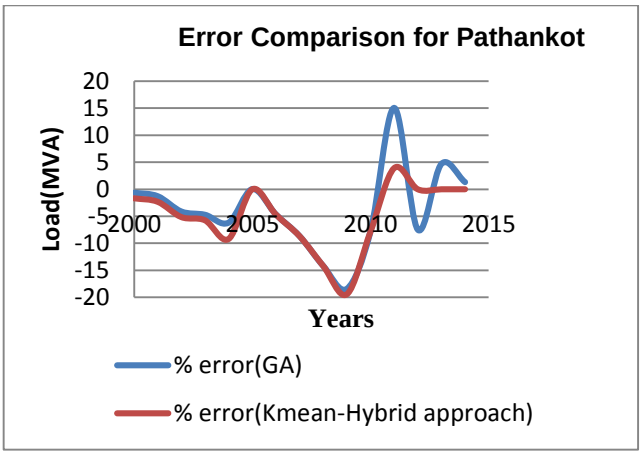

Fig. 10: Error comparison of Pathankot load (k-mean-ANN& GA [11])

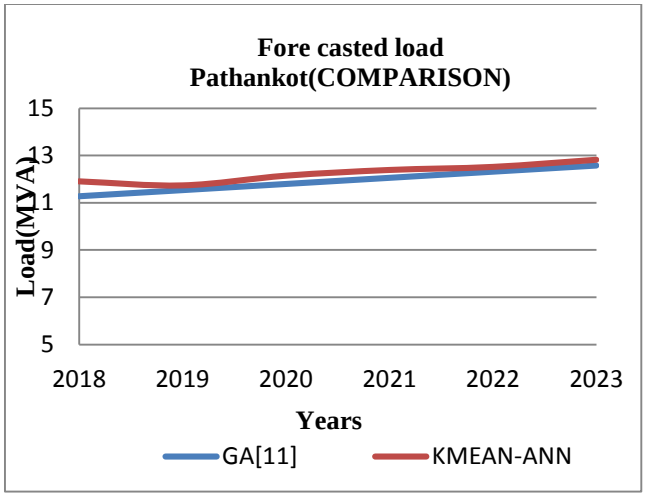

Fig. 11: Forecasted load comparison of Amritsar from 2018-2023 (k-mean-ANN& GA[11])

6. CONCLUSION

This study develops a k-Mean and ANN based load forecasting technique for Amritsar and Pathankot station. The analysis for Amritsar and Pathankot station has been done individually. The parameters like Actual Load of the station, Calculated Load of the station and Error is measured to prove the proficiency of proposed work.

To sum up, the dataset of previous year is gathered for future analysis and it has been proved beneficial for future prediction. In this study the dataset of previous 18 years have been used as an input to the system. Along with this the K-Mean Clustering and ANN mechanism is also applied to the dataset. On the basis of the results and tables that are shown in above section, it is concluded that the load prediction rate for Amritsar Station is quite effective in comparison to the Pathankot Station. Since the error rate in Pathankot station is quite higher in comparison to the Amritsar Station As this study implements the K-mean Clustering and ANN classification technique for load forecasting and also leads to the outstanding results but still more amendments can be done in this work by applying the concept of deep learning. It will help to analyze the data more precisely and could perform the load forecasting in quite better manner.

7. REFERENCES

[1] Muhammad Usman Fahad(2014), "Factor Affecting Short-Term Load Forecasting", Journal of Clean Energy Technologies, Vol. 2, No. 4, pp 305-309

[2] M. A. Farahat (2010), "A New Approach for Short-Term Load Forecasting Using Curve Fitting Prediction Optimized by Genetic Algorithms, International Middle East Power Systems Conference, vol 14, pp 106-110

[3] Miloˇs B. Stojanovi´c,(2010), "Mid-Term Load Forecasting Using Recursive Time Series Prediction Strategy With Support Vector Machines", ELEC. ENERG. vol. 23, no. 3, pp 287-298

[4] Abbas Khosravi (2010), "Construction of Optimal Prediction Intervals for Load Forecasting Problem", IEEE Transactions on Power Systems, vol: 25, no. 3, pp1496-1503

[5] James W. Taylor(2002), "Neural Network Load Forecasting with Weather Ensemble Predictions" IEEE Transactions on Power Systems, vol: 17, no. 3, pp 626-632

[6] Eisa Almeshaiei (2011), "A methodology for Electric Power Load Forecasting", Alexandria Engineering Journal, Vol 50, No. 2, pp 137-144

[7] Saeid Creighton(2010), "Load forecasting and neural network a prediction interval-based perspective" Computational Intelligence in Power Engineering, vol 302, pp 131-150.

[8] Samuel Humeau(2013), "Electricity Load Forecasting for Residential Customers: Exploiting Aggregation and Correlation between Households", Sustainable Internet and ICT for Sustainability (SustainIT), vol 1, pp 1-10.

[9] Che Guan,(2013), "Wavelet Neural Network Based Very Short-term Load Forecasting and Prediction Interval Estimation", DigitalCommons, pp 1-104.

[10] Simone Paoletti, "Load forecasting for active distribution networks", research gate, pp 2-7.

[11] Dogra Sonika1, Sidhu Darshan, Kaur Daljeet3 (2015)"Long-Term Load Forecasting using soft computing Techniques", vol:6, Issue 6 ISSN 2229-5518.

[12] A.G. Bakirtzs et al. ," A Neural Network short-term load forecasting model for the greek power system", IEEE, Vol 11, Issue 2, Pp 258-263, August 2002.

[13] Dongxiao Niu et al, "Power load forecasting using support vector machine and ant colony optimization", ELSEVIER, Vol 37, Issue 3, pp 2531-2539, March 2010

[14] Hugo Morais et al, "Optimal Scheduling of a renewable micro-grid in an isolated load area using mixed integer linear programming", Vol 35, Issue 1, Pp 151-156, January 2010.

[15] Changhao Xia et al, "Short, Medium and long-term load forecasting model and virtual forecaster based on radial basis function neural network", ELSEVIER, Vol b32, Issue 74, Pp 743-750, September 2010

[16] M.R.Airashidi et al, "Long-term electric load forecasting based on particle swarm optimization", ELSEVIER, Vol 87, Issue 1, pp 320-326, January 2010.

[17] S.Sp.Pappas et al. "Electricity demand load forecasting of the Hellenic power system using ARMA model", ELSEVIER, Vol 80, Issue 3, Pp 258-284, March 2010.

[18] Hong-ze Li et al, " an A hybrid power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm", ELSEVIER, Vol 37, Pp 378-387, January 2013.

[19] Julio Usaola, "Probabilistic load flow with correlated wind power injections", ELSEVIER, Vol80, Issue 5, Pp 528-536, May 2010.

[20] G.J. Tsekouras et al, "Short Term Load Forecasting in Electric Power Systems with ANN", Springer, Vol 343, pp 19-58, 2010.

[21] G.A. Adepoju et al, "Application of Neural Network to Load Forecasting Nigerian Electric Power System", Pacific journal of science and technology, Vol 8, Issue 1, Pp 7-68-72, May 2007.

[22] Qi Wi, "Power load forecasts based on hybrid PSO with Gaussian and adaptive mutation and Wv-SVM", ELSEVIER, Vol 37, Issue 1, Pp 194-201, January 2010

[23] Jianzhou Wang et al, "Combined modeling for electric load forecasting with adaptive particle swarm optimization", ELSEVIER, Vol 35, Issue 4, Pp 1671-1678.

[24] Carolina Garcia Ascanio et al, "Electric power demand forecasting using interval time series: A comparison between VAR and iMLP", ELSEVIER, Vol 38, Issue 2, Pp 715-725, February 2010

[25] M. Foley et al, "A strategic review of electricity system models", ELSEVIER, Vol 35, Issue 12, Pp 4522-4530, December 2010

[26] Wei Chiang Hong, "Application of Chaotic ant swarm optimization in electric load forecasting", ELSEVIER, Vol 38, Issue 10, pp 5830-5839, October 2010.

[27] Cheol-hong Kim et al, "Short-term load forecasting using data mining technique", J Electr Eng Technol, Vol 7, Issue 6, Pp 807-8013, 2012.

[28] Pan Duan et al, "Short Term Load Forecasting for electric power System using PSO-SVR and FCM clustering techniques", Energies, Vol 4, Pp 173-184, 2011

[29] Ahmed alkaline et al, "Application of K-Mean clustering algorithm to predict load shedding of the southern electrical grid of Libya", IEEE, October 2014

[30] Salahuddin A Azad et al, "Identification of typical load profiles using K-means Clustering Algorithm", IEEE, March 2015.

