Introduction

For optimal power operation, electrical generation must follow electrical load demand. Load varies very frequently on a daily basis. If the generation is not able to cope up with the demand, there is a change in frequency. If the frequency exceeds the limits, there is a loss of synchronism which affects the power system on a great scale. Practically, the change in frequency should not exceed 0.5 Hz. In order to cope up with the demand, generation needs to predict values of the load, and this requires the need of load forecasting. The generation, transmission, and distribution utilities require some means to forecast the electrical load so they can utilize their electrical infrastructure efficiently, securely, and economically. Generation utilities use electrical, load forecasting techniques to optimize the power flow on the transmission network to reduce congestion and overloads. Distribution utilities would not have much interest in short-term electric load forecasts as their distribution systems are predominantly radial with predictable maximum load demands. Thus, the distribution systems are sized conservatively and short-term load changes have little effect on the distribution system.

Load Forecasting is a technique by which the future load demand can be predicted based on the physical factors like temperature, pressure, loading on lines, losses, weather conditions etc. With the help of load forecasting, it is desirable to meet the load demands in the future so that the electrical can be obviated to happen. Electrical failures like blackout, faults due to unnecessary loading can be avoided if proper take care can be taken to the load demands. Moreover, load demands are kept on increasing day-by-day. It varies a lot on hourly basis. The load forecasting techniques are categorised into three groups.

Three categories of Load Forecasting Techniques are:

1. **Short Term Load Forecasting (STLF):** The short-term load forecast (STLF) represents the electric load forecast for a time interval of a few hours to a few days [1].
2. **Medium Term Load Forecasting (MTLF):** The medium-term load forecast (MTLF) represents the electric load forecast for a time interval of a few months.
3. **Long Term Load Forecasting (LTLF):** The long-term load forecast (STLF) represents the electric load forecast over a time for years.

Overview of Load Forecasting

2.1 Load Forecasting Techniques

The techniques used and implemented to create STLF are:

1. Linear Regression
2. Stochastic time series
3. General Exponential Smoothing
4. State Space Method
5. Knowledge – based expert approach
6. Artificial Neural Network (ANN)

In this project, we use Artificial Neural Networks (ANN) to forecast load. However, the techniques listed above are also used and are very helpful in predicting load demands. Techniques 1 to 4 are statistical techniques while Technique 5 based on a set of rules and reshapes the day's load curve using other set of rules. ANN techniques is highly precise and useful technique. It uses an algorithm which combines previous system load data as well as weather data and based on these data it predicts the future load pattern. ANN mimics the neuron system present inside the brain of living beings [2].

STLF has been extensively researched and modelled at the generation, distribution as well as transmission level. But the problem arises at the distribution level perhaps because the distribution utilities are concerned with generation scheduling and system peaking and these utilities are not concerned with an individual customer's short term load demand.

2.2 Factors affecting Electrical Load

Typically, the usage of a single electrical device in a large power system is random and usage patterns of other devices may differ from each other. There is often a large diversity in individual loads, yet when these loads are summed into larger facility load, pattern emerge which can be statistically predicted [3].

There are four main factors that influence electrical load:

1. Economic Factors
2. Time Factors
3. Weather
4. Random Effects

1. Economic Factors

Economic factors consist of investment in the facility's infrastructure through construction of new buildings, labs, and experiments which add load to the electric system. Funding profiles for the site dictate how and when equipment, processes, and experiments can be operated. Utility programs such as demand charges and demand management plans affect the customer's electrical usage patterns during times of system peaking [3]. Economic factors will not influence the STLF as these factors typically change usage patterns over a longer time range than 24 hours [3]; however, economic factors can be the inspiration for studying a system's load pattern and implementing load reduction initiatives. As discussed earlier, the STLF is a useful tool for implementing demand management activities.

2. Time Factors

The three time factors that have the most influence on electric loads are:

1. Seasonal Effects
2. Weekly-daily cycle

3. Holidays

Seasonal effects account for long term changes in weather patterns [4]. These are not only weather patterns, but can include popular vacation dates and changes between Daylight Savings Time (DST) [3]. Weekly – Daily cycle are electric load patterns that are periodic over the course of week and each day. The daily electric load profile for a holiday is similar to that of a weekend profile. The magnitudes of the electric loads are lower. If the holiday is adjacent to a weekend, the load profile can be described as a "long weekend [3]."

3. Weather

Among the factors, weather influences the most. Weather factors have a significant effect on the short-term electric load profile of a power system [1], [3], [5]. Weather-sensitive loads, such as heating, ventilating, and air-conditioning (HVAC) equipment, will have a greater impact on smaller industrial/institutional power systems as these tend to be the larger loads on the system. HVAC equipment cycling on and off can produce electrical load profiles that appear to have random power swings. As the power load increases, there will be more load diversity, the effect of load cycling will be dampened, and the electric load will be smoother.

Weather factors that can affect the electric hourly load profile are humidity, solar irradiance, wind speed, barometric pressure, and precipitation. High humidity days will make cooling equipment operate for longer duty cycles to remove excess moisture out of the conditioned air. Long durations of high solar irradiance will radiantly heat the interior of buildings forcing the cooling equipment to operate longer and with less diversity. Precipitation has the tendency to reduce the air temperature and thus reduce the cooling load [3].

Wind speed and barometric pressure can also affect the hourly load profile, and often occur in tandem with other factors such as precipitation.

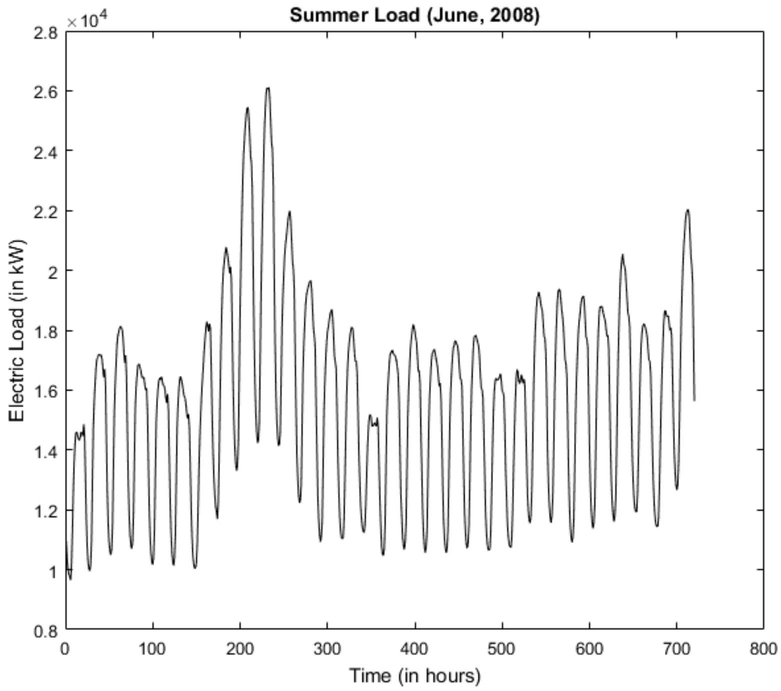


Fig. 1 Summer Load

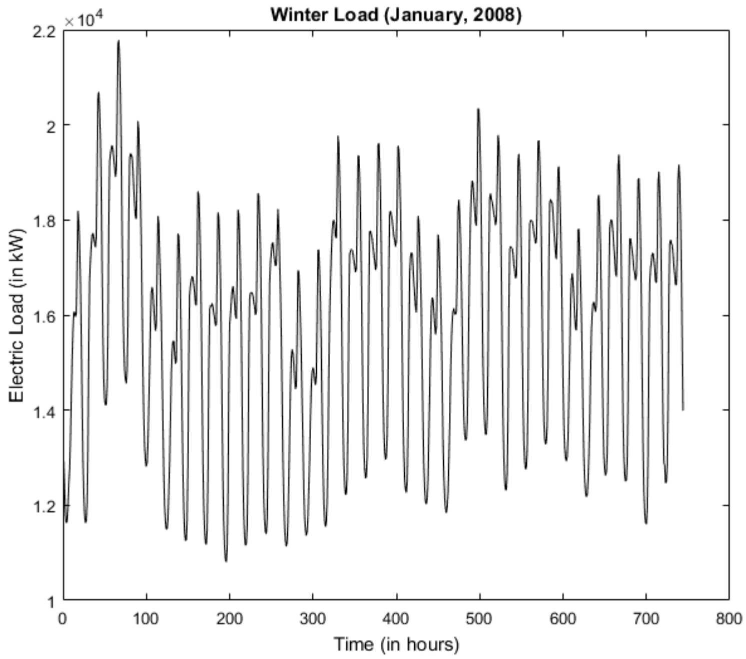


Fig. 2 Winter Load

4. **Random Effects**

   Random factors that influence the electrical load profile consist of all the other random disturbances in the load pattern that cannot be explained by the previous three factors [3]. These disturbances can consist of significant loads that do not have a set operating schedule which makes prediction difficult. Other disturbances such as widespread employee absences (due to sickness, inclement weather, etc.) and planned or unplanned utility system outages can have significant effects on the facility's load profile.

2.3 Artificial Neural Network (ANN)

An Artificial Neural Network can be described as a mathematical tool that mimics the thought processes of the human brain. ANNs were first applied to load forecasting in the late 1980's [6] - [7]. They are described as a multivariate, nonlinear, and nonparametric method which makes the good candidates for modelling complex nonlinear systems [6], [8] - [9]. ANNs have good performance in data classification and function fitting [10].

Neurons are the basic processing components of ANNs. The neurons are programmed to behave similarly to the neurons in the brain by receiving inputs, processing those inputs, and producing an output. The neuron is shown schematically in the Fig. 3 (shown below).

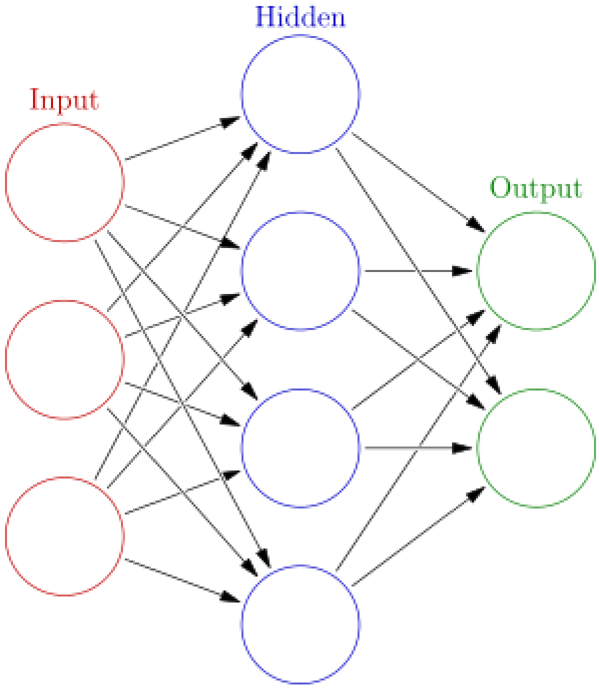


Fig. 3 Artificial Neural Network [11]

An ANN is based on a collection of connected units or nodes called artificial neurons (a simplified version of biological neurons in an animal brain). Each connection (a simplified version of a synapse) between artificial neurons can transmit a signal from one to another. The artificial neuron that receives the signal can process it and then signal artificial neurons connected to it [11].

In common ANN implementations, the signal at a connection between artificial neurons is a real number, and the output of each artificial neuron is calculated by a non-linear function of the sum of its inputs. The connections between neurons are called Edges. Artificial neurons and edges typically have a weight that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Artificial neurons may have a threshold such that only if the aggregate signal crosses that threshold is the signal sent. Typically, artificial neurons are organized in layers. Different layers may perform different kinds of transformations on their inputs. Signals travel from the first (input), to the last (output) layer, possibly after traversing the layer multiple times [11]

An ANN for load forecasting can be trained on a training set of data that consists of time-lagged load data and other non-load parameters such as weather data, time of day, day of week, month, and actual load data [5]. Some ANNs are only trained against days with data similar to the forecast day [12] - [13]. Once the network has been trained, it is tested by presenting it with predictor data inputs. The predictor data can be time-lagged load data and forecasted weather data (for the next 24 hours) [5]. The forecasted load output from the ANN is compared to the actual load to determine the forecast error. Forecast error is sometimes presented in terms of the root mean square error (RMSE) [5], [6], [8], [10] but more commonly in terms of the mean absolute percent error (MAPE) [1], [4]-[5], [6]-[10], [14], [15]-[16], [12]-[13]. It should be noted that an ANN trained on a specific power system's load and weather data will be system dependent. The ANN generated for that system will most likely not perform satisfactorily on another power system with differing characteristics. It is possible the same ANN architecture may be reused on the new system, but retraining will be required.
It should be noted that an ANN trained on a specific power system's load and weather data will be system dependent. The ANN generated for that system will most likely not perform satisfactorily on another power system with differing characteristics It is possible the same ANN architecture may be reused on the new system, but retraining will be required.

An ANN trained using the backpropagation technique can succumb to overfitting. Reference [10] defines overfitting of an ANN as "estimating a model that fits the data so well that it ends by including some of the error randomness in its structure, and the produces poor forecasts." The over fitted ANN can replicate in-sample (training) data with very low errors, but when presented with out-of-sample data, the results have high errors. This type of ANN is not capable of generalizing the input-output data relationships. Overfitting can occur due to overtraining and excessive network complexity [5], [6].

Overtraining can be avoided by using cross-validation. This process splits the original training data into a training set and validation set [6], [10]-[17]. The ANN parameters are trained on the training set and tested every few iterations on the validation set. When the performance of the validation set starts to decrease, training is ended [6].

Overly complex ANNs can over fit its training data. The ANN attempts to track down each data point in the training set [6]. Keeping the model simple will help produce smooth models that usually forecast better than complex models. ANNs that are designed with complex arrangements and classes on input data and high numbers of hidden neurons must estimate a very large number of weights and biases during training. An ANN trained under these conditions with a small input data set will have estimated many internal weights and biases on a comparatively small sample set of data. This will lead to a lack of generalization of the ANN's performance on out-of-sample data and result in high errors [6], [10]. A method, such as correlation analysis, should be applied to the input data in an effort to reduce the amount of data to only what is necessary [5], [17] - [14].

Implementation

We used MATLAB 2016a (version 9.0) by Mathworks Inc. to create and implement load forecast. The Neural Network Toolbox under Apps in MATLAB section provides the built-in functions and applications to assist in modelling non-linear systems. It supports ANN training, validation, testing, and simulation with hardcode and graphical user interface (GUI) application [14].

3.1 Input Data

We used one year (from January 1^{st}, 2008 to December 31^{st}, 2008) Australia's load data and weather data. The weather data includes:

1. Average air temperature (in ^{o}C)
2. Average global horizontal (in W/m^2) – the sum of direct normal irradiance, diffuse horizontal irradiance, and ground - reflected solar radiation
3. Average relative humidity (in %)
4. Average atmospheric pressure (in mBar)

The Load data as well as weather data have followed US' format. Therefore, the format is like: 20-11-2008 which means November 20^{th}, 2008. It should be noted that the year 2008 was a leap year so, the number of days was 366. The data is distributed in hours i.e., the data (electric load as well as weather data) is given in per hour of the day. So, the length of each column in the given data is 24 x 366 = 8784. And the total number of columns is 10. Thus, the total of cells in excel sheet is 87840.

All data including weather data and Load data was imported into the MATLAB to document the time of day for each input data. The data was imported as columns of length 8784 from the Excel Sheet. Each column has given a separate variable to specify its effect on load forecast. The Annual Data was plotted on hourly axis and the scatter plot is shown in Fig. 4 (shown below)

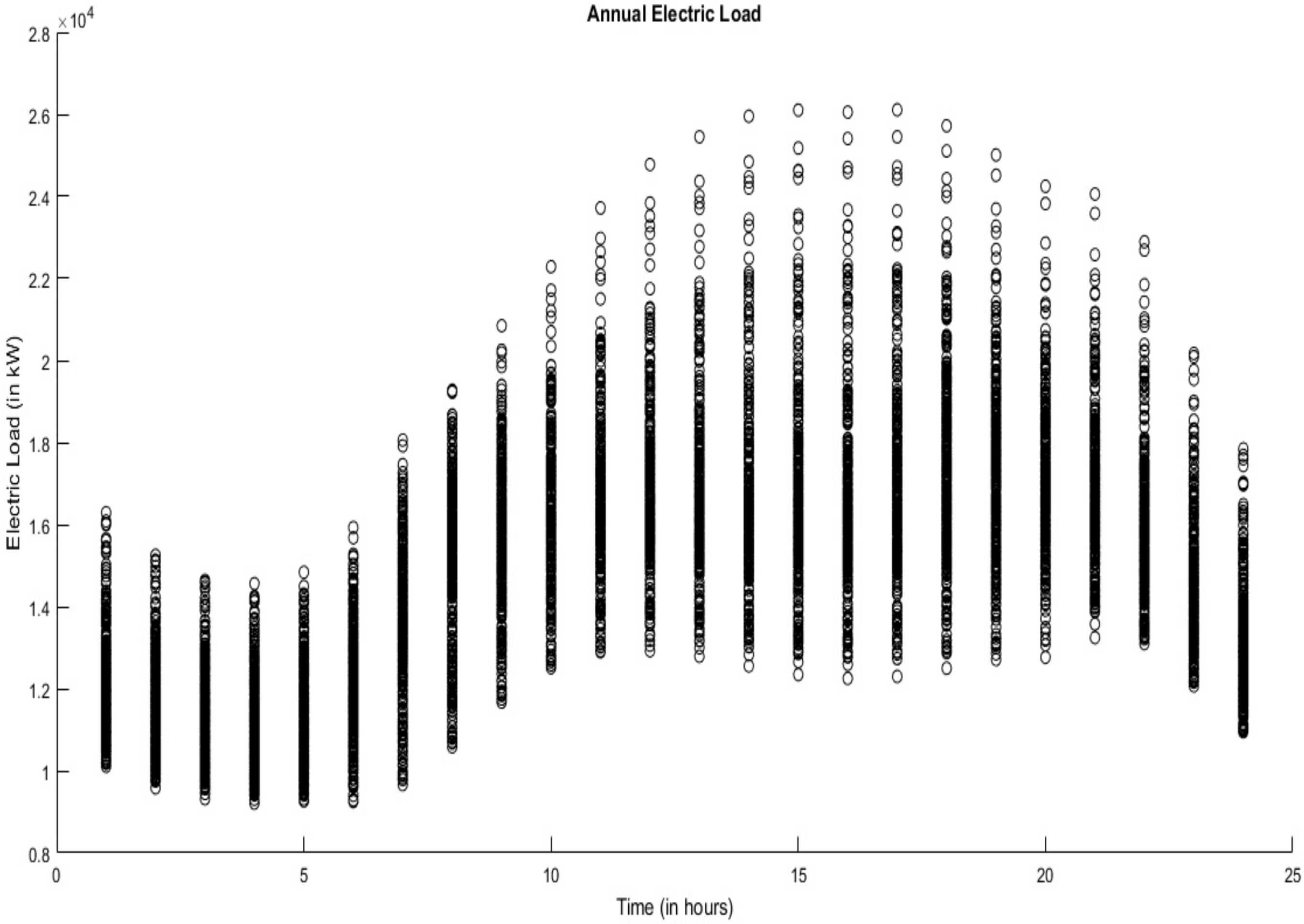


Fig 4. Annual Electric Load (hourly)

From the scattered plot, it is clear that the load was higher during the day than at the night. The variation of load is clearly visible. The load is clearly time dependent as stated above. So, we can conclude that the time of day is also important in determining the load forecast. However, the effect is low.

The simple plot was also plotted i.e., the plot of electric load (in kW) against the hours in a year (shown below). The plot was looking clumsy and the clear information was hard to extract.

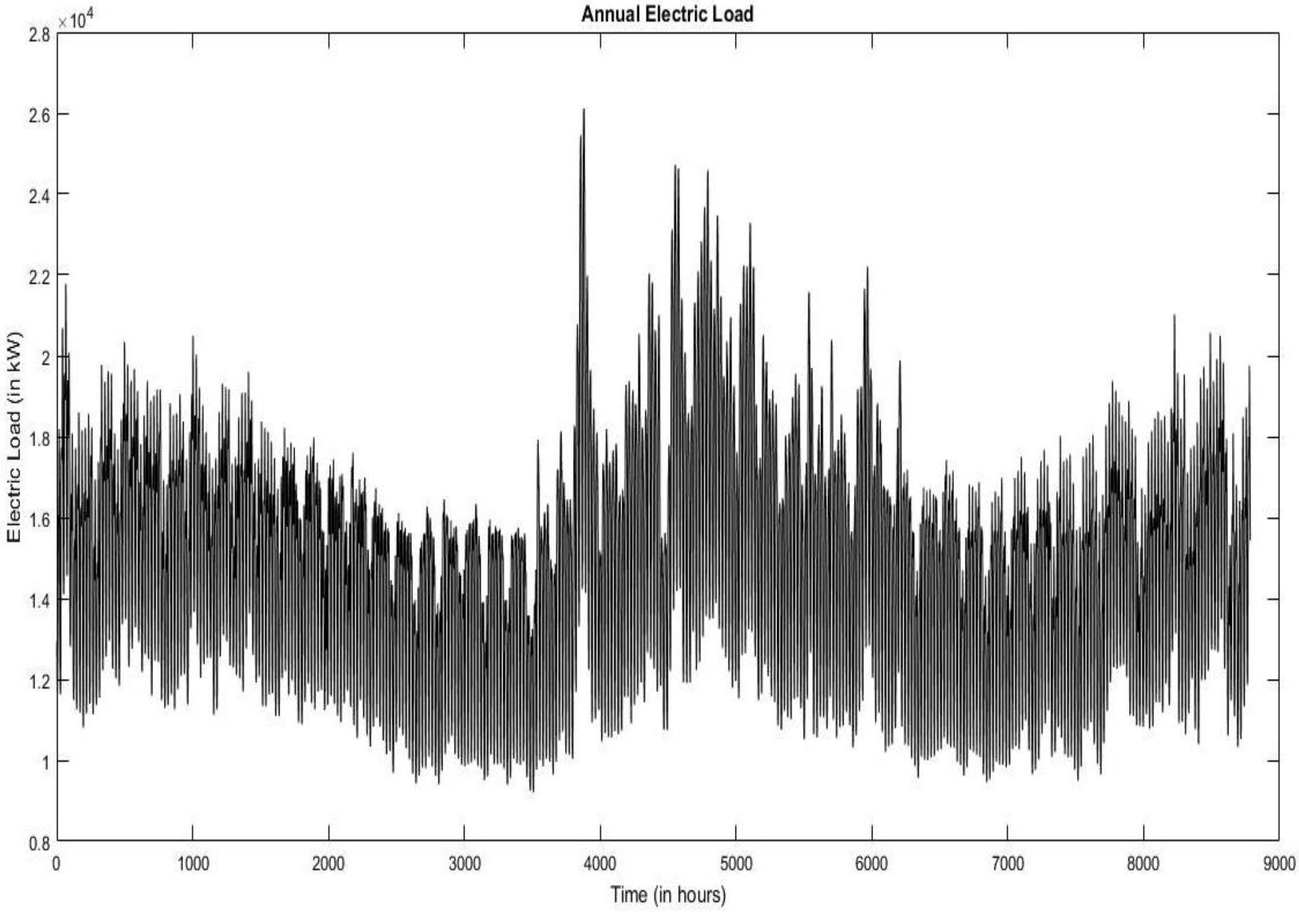


Fig. 5 Annual Hourly Electric Load

Fig.4 also reveals the significance a holiday has on the load profile. As an example, the dip in the load profile between approximately 6000 and 7000 hours is the load data for the week before, during, and after Christmas. The Monday following New Year's Day also has reduced load level, but the load profile returns to a more seasonal pattern the following Tuesday. Other holidays have a similar reduction in load on the day of the holiday, but none of them have as dramatic an effect as the Christmas holiday.

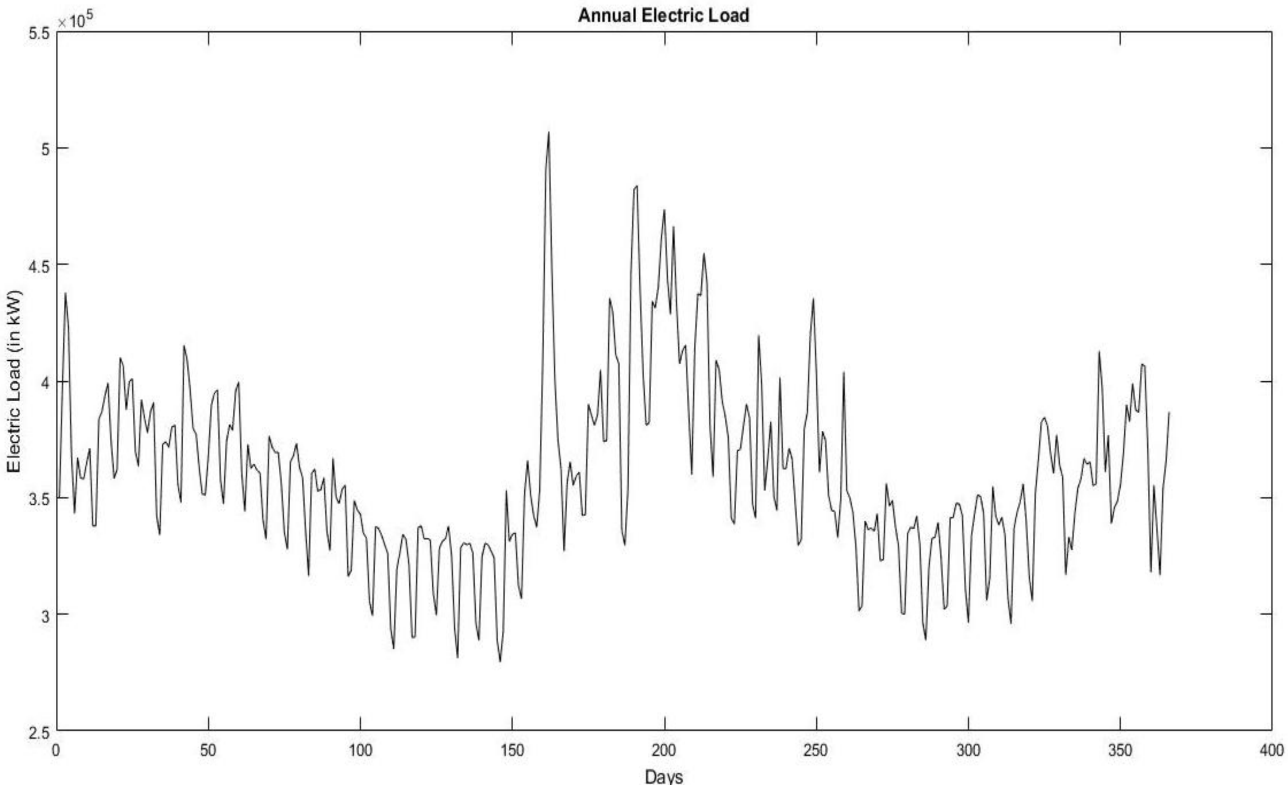


Fig. 5 Annual Daily Electric Load

The daily load data was extracted from hourly separated data by a formula. The daily annual load was also plotted against its corresponding days. As it is clear from the graph (shown above), the load reaches its peak during the 160th day of the year and that is approximately in the month of May. Here, the load varies according to the season as stated above. Hence, the seasons affect the load quite a lot. So, the seasonal affects have to take into consideration to predict the electric load.

To get the clear insight about the effects of seasons on the electric load. We have plotted the electric load against time (in hours) of the four seasons: Winter, Autumn, Summer, and Winter. The graphs are shown below

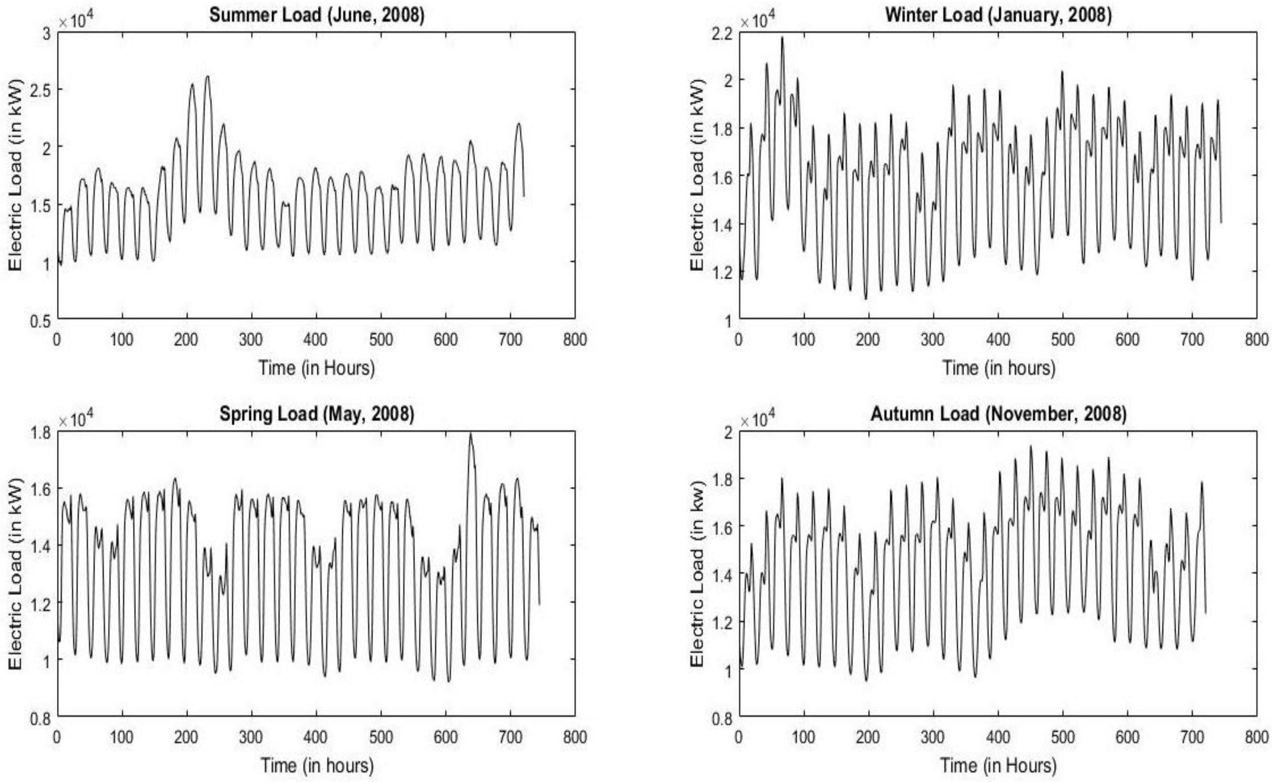


Fig. 6 All Seasons Electric load

An explanation for the profiles of summer and winter plots is the fact that cooling demand is high in the summer and heating demand is high in the winter. During the hot afternoons in the summer, cooling systems will be operating causing high electricity demand. Also, during the cold nights and early mornings during the winter, heating systems will be operating causing high electricity demand. The fall and spring season plots have similar profiles. They show a positive correlation between load and temperature, but the temperature profile has a slight offset, so the correlation for fall and spring seasons will often be less than that for the summer season.

3.2 Training and Testing with ANN

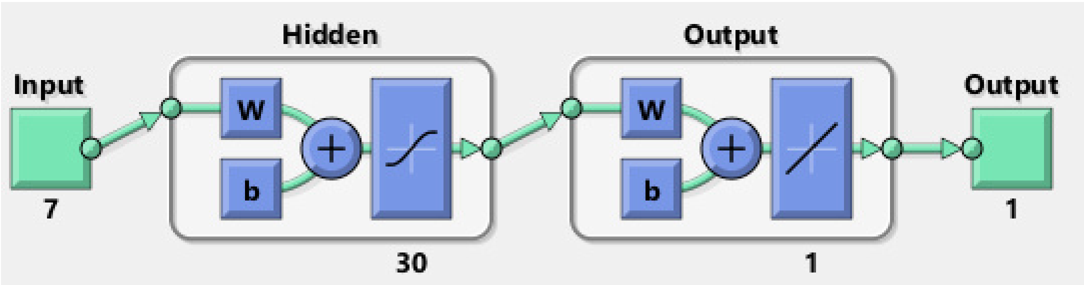
The whole data set was divided into two sets: Training set and Test Set. Training set consists of 80% of whole data and Test set contains the rest data. The training set was used to make a model which, therefore, predicts the load in the future. The model is made by a MATLAB app Neural Net Fitting. The training set has got inputs which are as follows:

1. Temperature (in ^{o}C)
2. Humidity (in %)
3. Pressure (in mBar)
4. Time (in hours)
5. Global Horizontal (in W/m^2)
6. Previous Day Same Hour Load (in kW)
7. Previous Week Same Day Same Hour Load (in kW)

The data was performed on the training and forecast predictor datasets, the number of hidden layers, or neurons, in the ANN was defined to be 30 neurons, and the ANN was created for the user-defined forecast day. The built-in MATLAB Levenberg - Marquardt optimization training function was used to perform the backpropagation training of the feed-forward ANN [19]. This process iteratively updated the internal weight and bias values of the ANN to obtain a low error output when utilizing the training predictor dataset and a target dataset. The target dataset consists of the actual load values for a given predictor dataset.

After training, the ANN was tested using only the training predictor dataset. This step allows the user to verify the trained ANN can produce low error forecasts on in sample data. After testing, the ANN forecasted plot was plotted against test set data plot and MAPE was calculated. The results of this forecast were stored, and the entire ANN training, testing, and forecasting process was repeated a set number of times with the intention of reducing the forecast error.

The Simulink Model was extracted from the net fitting toolbox is shown below.



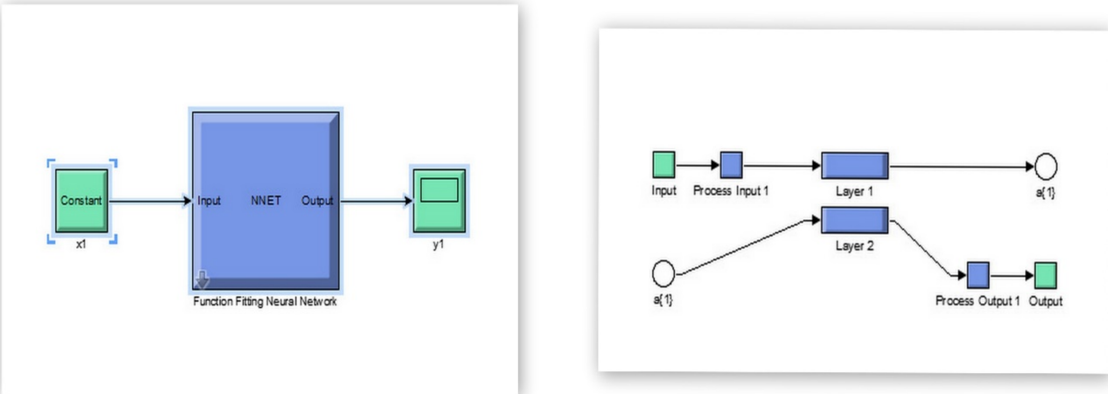


Fig. 7 Simulink Model

Result

A graph of forecasted load was plotted against the time (in hours) and comparison was made against the Actual Load (test data load). A part of this graph is shown in Fig 8 (shown below).

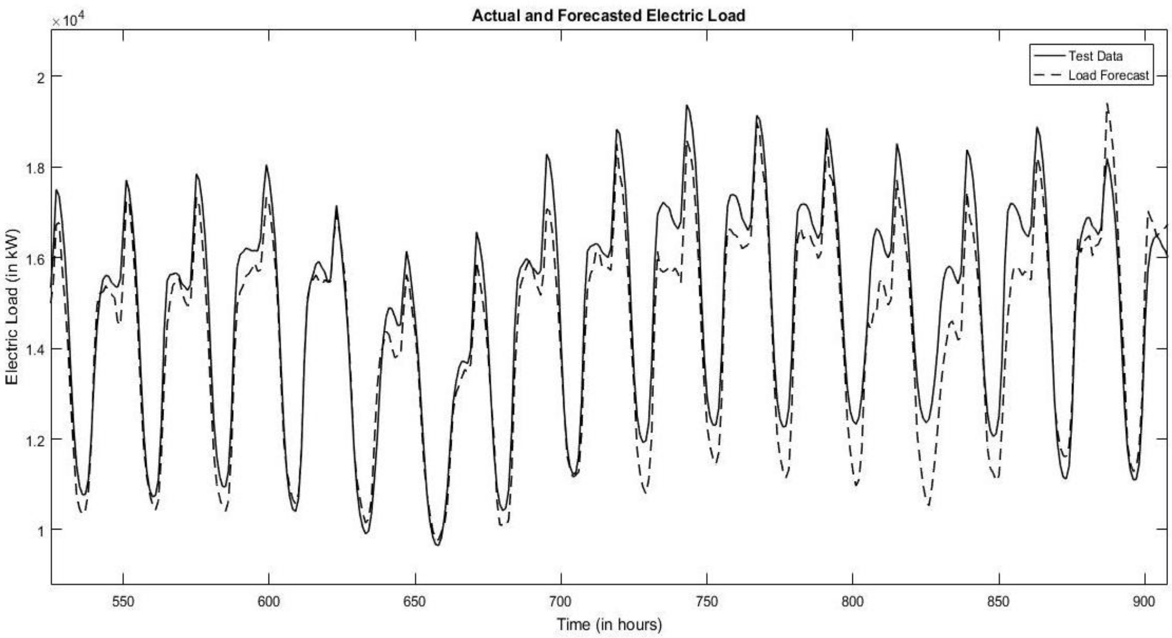


Fig. 8 Actual and Forecasted Load

The graph shows a little deviation of the forecasted plot from the Test data load. The MAPE (mean absolute percentage error) came out to be 5.1440 % which is bearable.

However, a simpler comparison was also made and plotted as scattered graph (shown below). The graph shows a little deviation of the fitted line (linear) from the ideal fitted line (T = F). The regression value of the test set was R = 0.95968 while the regression value of the training set was R = 0.95419

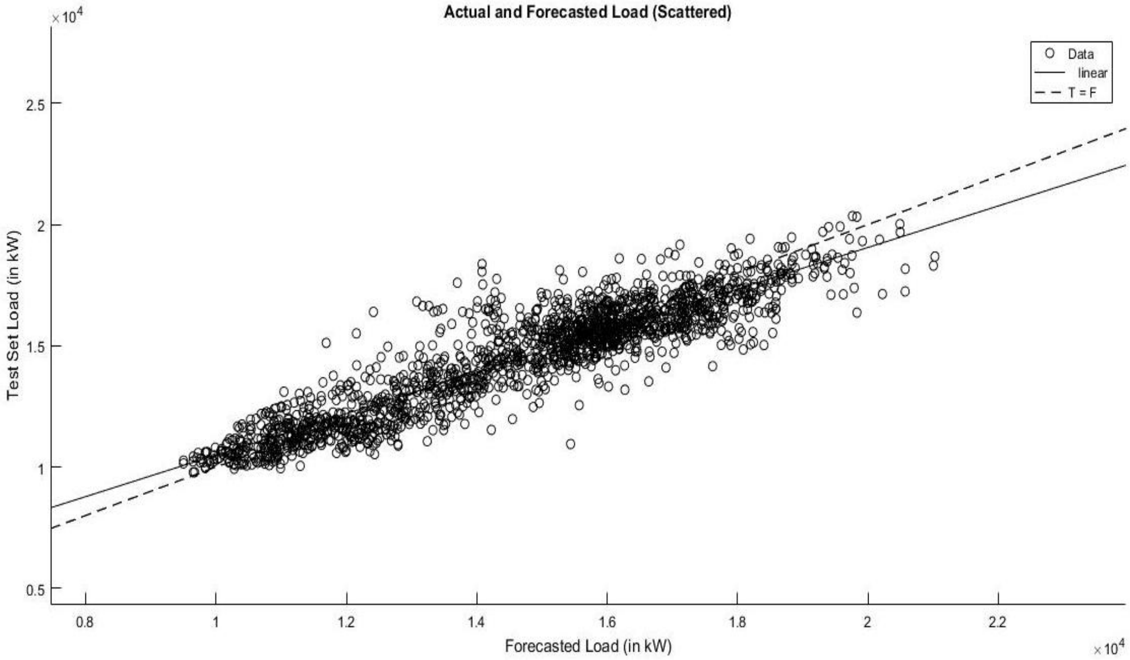


Fig. 9 Actual and Forecasted Load (Scattered Plot)

The error values represent days where the actual load profile experienced a planned or unplanned outages or other sudden load change. This is drawback to using an ANN to forecast the load. The ANN cannot predict these events, so a one or two hours' outage on a major distribution feeder will have a significant impact on the resulting error. If these events did not occur, then the error would have been much less. However, the high error values occurred during low load levels. This illustrates the chaotic load profiles experienced during the winter season. The winter load profiles are not as smooth as the profiles during the other seasons, so the ANN has a harder time forecasting the changes in the profile. The number of hidden layer neurons in each ANN was set at 30. This value was selected by trial and error to determine the minimum number of hidden-layer neurons that would produce the lowest forecast error.

The values selected for the minimum inputs, and number of hidden-layer neurons had a direct effect on the MATLAB program's ANN training runtime. This added complexity would increase the program runtime and would also present the ANN with so much input data that the resulting forecast error would increase. Increasing the training weeks also required the ANN to process a larger amount of data, which increased the program runtime. If the input data set was too large, the ANN could not generalize its output, and the resulting forecast error would increase. Increasing the number of hidden-layer neurons increased the program runtime because the weights and biases of each neuron have to be calculated and optimized during network training. An overly complex network could not generalize on the out of-sample data set and would "over fit" the in sample data.

Conclusion

It is a goal for every power system manager to have their power system operate efficiently, securely, and economically. To meet this goal, the behaviour of their power system must be understood. Analysis of the system’s normal operating bounds, response to customer demands, and reaction to weather events will provide insight on system loading. Short-term electric load forecasting can provide that insight for the following day to assist in making power system operational decisions. The demand charges and associated system loading can be reduced by energy demand management techniques such as conservation, on-site generation, or implementing demand-response agreements with the serving utility. Energy demand management is a process where the energy use for a facility is planned and coordinated with the system’s various load centres.

An artificial neural network (ANN) is a mathematical model that mimics the decision-making processes of the human brain. The fundamental component of the ANN is the neuron. Neurons are programmed to behave similarly to the neurons in the brain by receiving inputs, processing those inputs, and producing an output. The neurons are connected together to form a network that can be used to solve nonlinear problems. It should not be assumed that an ANN trained with inputs and targets for one power system will produce low error forecasts when used on a different system. The structure and size of the ANN should only be as complex as required to produce acceptable forecast on out-of-sample data. Overly complex ANNs can over fit their training data producing very low error on in-sample input data, but high error on out-of-sample input data. Overfitting reduces the ANN’s ability to generalize on data it has not been trained on. In addition, a complex ANN will have a high training runtime as there are more weights and biases to be estimated during the learning process.

One of the difficulties in applying ANNs to load forecasting occurs when there is an outage on the system or a significant event that causes the load to change such as the start-up or shutdown of a high demand load. The ANN cannot be trained to unplanned outages because there is no way to predict when these will occur. There is also difficulty in training the ANN to planned outages because these to not occur often, therefore the ANN will not be exposed to many of them during its training phase.

Unplanned outages will still result in higher error forecasts, but planned outages could be used as input to the ANN. If an outage is planned, the estimated affected load could be supplied as a dedicated input to the ANN. It would be better for the outage value to be supplied to the ANN as an input than just subtracting the affected outage load from the ANN forecast. When a significant load change event occurs on a power system, this has the effect of changing the load pattern for the rest of the day. This effect is called load inertia. For example, if a large load is switched into a power system, the load profile will have a step change at the switching time. Throughout the rest of the day, the load profile will have higher values for longer time spans. In addition to the new large load, the system will experience additional losses, causing equipment to run hotter, and require additional cooling. The inertia of this load change will diminish as the system load decreases for the day. However, if the load change event was large enough, the load inertia might carry over to the following day which would affect that day’s load forecast.

The same would be true if a large load was switched off of the power system. The remaining load profile for the day would be less than normal, but the difference in load levels would not be exact value of the load that was removed. If a power system experiences these kinds of large load change events on a regular period, then the ANN training algorithm should be capable of predicting this inertia and load step changes. However, if these events do not occur regularly, then they should be supplied as part of the forecast predictor data set so the ANN can adjust its load forecast accordingly.

References

1. Y. Al-Rashid and L.D. Paarmann, "Short-term electric load forecasting using neural network models," *Circuits and Syst.*, Ames, IA, 1996, pp. 1436-1439.
2. I. Moghram and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," *IEEE Trans. Power Syst.*, vol. 4, pp. 1484 1491, Nov. 1989.
3. G. Gross and F.D. Galiana, "Short-Term Load Forecasting," *Proc. IEEE*, vol. 75, pp. 1558-1573, Dec. 1987.
4. T. Hong, M. Gui, M. Baran, and H.L. Willis, "Modeling and forecasting hourly electric load by multiple linear regression with interactions," *Power and Energy Soc. General Meeting*, Minneapolis, MN, 2010, pp. 1-8.
5. P.J. Santos, A.G. Martins, and A.J. Pires, "Short-term load forecasting based on ANN applied to electrical distribution substations," *Universities Power Engineering Conf.*, Bristol, UK, 2004, vol. 1, pp. 427-432.
6. H.S. Hippert, C.E. Perdreira, and R.C. Souza, "Neural networks for short term load forecasting: a review and evaluation," *IEEE Trans. Power Syst.*, vol. 16, pp. 44-55, Feb. 2001.
7. M. Ramezani, H. Falaghi, and M. Haghifam, "Short-term electric load forecasting using neural networks," *Int. Conf. on Computer as a Tool*, Belgrade, 2005, pp. 1525- 1528.
8. K.Y. Lee, Y.T. Cha, and J.H. Park, "Short-term load forecasting using an artificial neural network," *IEEE Trans. Power Syst.*, vol. 7, pp. 124-132, Feb. 1992.
9. T. Senjyu, P. Mandal, K. Uezato, and T. Funabashi, "Next day load curve forecasting using recurrent neural network structure," *Proc. Inst. Elect. Eng.*, vol. 151, pp. 388-394, May 2004.

10. Z.H. Osman, M.L. Awad, and T.K. Mahmoud, “Neural network based approach for short-term load forecasting,” *IEEE/PES Power Systems Conference and Exposition,* Seattle, WA, Mar. 15-18, 2009.
11."*Artificial Neural Networks as Models of Neural Information Processing | Frontiers Research Topic*". Retrieved 2018-02-20.
12.T. Senjyu, H. Takara, and K. Uezato, “Two-hour-ahead load forecasting using neural network,” *Int. Conf. on Power Syst. Technology*, Perth, WA, 2000, pp. 1119-1124.
13.T. Senjyu, H. Takara, K. Uezato, and T. Funabashi, “One-hour-ahead load forecasting using neural network,” *IEEE Trans. Power Syst.*, vol. 17, pp. 113-118, Feb. 2002.
14. W. Charytoniuk and M. Chen, “Very short-term load forecasting using artificial neural networks,” *IEEE Trans. Power Syst.*, vol. 15, pp. 263-268, Feb. 2000.
15.D.C. Park, M.A. El-Sharkawi, R.J. Marks II, L.E. Atlas, and M.J. Damborg, “Electric load forecasting using an artificial neural network,” *IEEE Trans. Power Syst.*, vol. 6, pp. 442-449, May 1991.
16. B.S. Kermanshahi et al, “Artificial neural network for forecasting daily loads of a canadian electric utility,” *Applications of Neural Networks to Power Systems*, Yokohama, Japan, 1993, pp. 302-307.
17. F. Liu, R.D. Findlay, and Q. Song, “A neural network based short term electric load forecasting in Ontario Canada,” *Int. Conf. on Computational Intelligence for Modeling,*
18.*Control, and Automation*, Sydney, NSW, 2006, pp. 119-126. H. Demuth, M. Beale, and M. Hagan. (2009, Mar.). *Neural network toolbox*™ *6 user’s guide*