AN EXPLORATORY ANALYSIS ON HALF-HOURLY ELECTRICITY LOAD PATTERNS LEADING TO HIGHER PERFORMANCES IN NEURAL NETWORK PREDICTIONS

K.A.D. Deshani[1], M.D.T. Attygalle[2], L.L. Hansen[3], A. Karunaratne[4]

[1, 2, 4]Department of Statistics, University of Colombo, Colombo 03, Sri Lanka

[3] School of Computing, Engineering and Mathematics, University of Western Sydney, Australia

ABSTRACT

Accurate prediction of electricity demand can bring extensive benefits to any country as the forecasted values help the relevant authorities to take decisions regarding electricity generation, transmission and distribution appropriately. The literature reveals that, when compared to conventional time series techniques, the improved artificial intelligent approaches provide better prediction accuracies. However, the accuracy of predictions using intelligent approaches like neural networks are strongly influenced by the correct selection of inputs and the number of neuro-forecasters used for prediction. Deshani, Hansen, Attygalle, & Karunarathne (2014) suggested that a cluster analysis could be performed to group similar day types, which contribute towards selecting a better set of neuro-forecasters in neural networks. The cluster analysis was based on the daily total electricity demands as their target was to predict the daily total demands using neural networks. However, predicting half-hourly demand seems more appropriate due to the considerable changes of electricity demand observed during a particular day. As such clusters are identified considering half-hourly data within the daily load distribution curves. Thus, this paper is an improvement to Deshani et. al. (2014), which illustrates how the half hourly demand distribution within a day, is incorporated when selecting the inputs for the neuro-forecasters.

KEYWORDS

Clustering, Silhouette plots, Improve performance, Load curve prediction

1. INTRODUCTION

Predicting the future electricity demand is an essential task for a country, as a huge amount of money could be saved by utilizing the available electricity generation options. In this scenario, increasing the accuracy of short-term predictions is very crucial, as decisions regarding the required load, has to be taken within a short period of time. Literature regarding short-term load forecasting techniques consist of both conventional time series models and artificial intelligent approaches from many fields mostly in the field of engineering. To develop a dynamic forecasting system, intelligent approaches yield better results than conventional time series techniques as they could be adapted to suit novel conditions and handle more complex patterns in data. However, the accuracy of predictions using intelligent approaches like neural networks are strongly influenced by the correct selection of inputs and the number of neuro-forecasters used for prediction. (Farahat & Talaat, 2012; Barzamini, Hajati, Gheisari, & Motamadinejad, 2012; Nagi, Yap, Tiong & Ahmed, 2008). Deshani, Hansen, Attygalle, & Karunarathne (2014)

DOI : 10.5121/ijaia.2014.5303

suggested how a cluster analysis could be performed to group similar day types, which contribute towards selection of a better set of neuro-forecasters in neural networks. Their proposed cluster analysis was based on the total daily electricity demand and each day was assigned to one of the three clusters suggested from the analysis. However, electricity demand varies in accordance with consumers' activities with respect to time of the day and the day of the week. As a result of these variations, the hourly load requirement is never a constant throughout a particular day.

This paper presents a cluster analysis, performed to identify intra-day clusters and to group similar day types within those clusters respect to half-hourly electricity demand. Even though many external causes like metrological conditions such as temperature, rainfall, humidity, wind speed and cloud cover, economic and demographic factors influence the electricity demand, this paper has considered only a single input, which is the day type. The main focus has been given to illustrate how data mining techniques can be complimented by cluster analysis in giving efficient predictions.

A dataset consisting of half-hourly electricity demands in Sri Lanka was considered for the period of 01st January 2008 to 31st December 2012.

2. Literature Review

Literature related to load forecasting reveal that higher prediction accuracies could be obtained when using intelligent techniques when compared to using conventional statistical techniques (Farahat & Talaat, 2012; Barzamini, Hajati, Gheisari, & Motamadinejad, 2012; Nagi, Yap, Tiong & Ahmed, 2008). Many researchers point out the importance of using intelligent techniques in situations where quick weather changes lead to fail accurate predictions. (Seetha & Saravanan, 2007; Senjyu, Takara, Uezato, & Funabashi, 2002; Barzamini et al., 2012). Some of those popular intelligent techniques used in the literature are neural networks, fuzzy inference systems, genetic algorithms and expert systems.

Many researches had used the effect of different day types to enhance the load predictions considering their own country's situations. The literature reveals that, Soared and Medeiros (2008) had incorporated the maximum number of day types to their model, as Sunday - Saturday, holiday, working day after holiday, working day before holiday, working day between a holiday and weekend, Saturday after a holiday, working only during the mornings, working only during the afternoons and Special holidays. Another research considers seven days of the week and bank holidays as day types, and a principal component analysis had been performed accordingly and a segmentation scheme based on the first principal direction had been used to cluster similar months (Cho, Goude, Brossat, & Yao, 2013). Unlike these approaches, (Barzamini et al., 2012) had divided the weekly days into four categories based on unique load lags and had incorporated to the model. Considering the above, this research considers thirteen day types, which can be considered as different in Sri Lankan context.

Even though thirteen day types are considered, including all these day types into the model will complicate the prediction process. As such, the 'day type' will be clustered into similar day types in order to avoid complexities in the computation operations and to reduce forecasting error when training the neural networks (Barzamini et al., 2012; Seetha & Saravanan, 2007). They have discussed how accurate predictions are made when the inputs are wisely chosen to be fed into the neural network having different neuro-load forecasters to train similar featured loads. Literature also shows that in some research, similar days had been clustered based on experience of the experts of electricity supplying companies rather than performing any statistical analysis (Cho et al., 2013). Moreover, to understand energy consumption patterns in industrial parks, a cascade application of a Self-Organizing Map and a clustering k-means algorithm had been performed by

Hernandez, Baladron, Aguiar, Carro & Esguevillas (2012). Even though no study has considered performing a cluster analysis, this study focuses on a statistical analysis based on k-means clustering to complement the neural network approach.

3. METHODOLOGY

3.1. K-Means clustering

K-means clustering is a partitioning method. It partitions data into k mutually exclusive clusters. Unlike hierarchical clustering, k-means clustering operates on actual observations (rather than the larger set of dissimilarity measures), and creates a single level of clusters. The distinctions mean that k-means clustering is often more suitable than hierarchical clustering for large amounts of data.

Each cluster in the partition is defined by its member objects and by its centroid, or center. The centroid for each cluster is the point to which the sum of distances from all objects in that cluster is minimized. kmeans computes cluster centroids differently for each distance measure, to minimize the sum with respect to the measure that you specify.

Distance measure: In this paper, ‘kmeans’ function in Matlab software has been used with appropriate distance measures based on the clustering data.

Distance Measure	Description
sqEuclidean	Squared Euclidean distance. Each centroid is the mean of the points in that cluster.
cityblock	Sum of absolute differences. Each centroid is the component-wise median of the points in that cluster
correlation	One minus the sample correlation between points (treated as sequences of values). Each centroid is the component-wise mean of the points in that cluster, after centering and normalizing those points to zero mean and unit standard deviation.

Determining the number of clusters: To get an idea of how well-separated the resulting clusters are silhouette plot can be used. The silhouette plot displays a measure of how close each point in one cluster is to points in the neighboring clusters. This measure ranges from +1, indicating points that are very distant from neighboring clusters, through 0, indicating points that are not distinctly in one cluster or another, to -1, indicating points that are probably assigned to the wrong cluster.

Avoiding Local Minima: Like many other types of numerical minimizations, the solution that kmeans reaches often depends on the starting points. It is possible for kmeans to reach a local minimum, where reassigning any one point to a new cluster would increase the total sum of point-to-centroid distances, but where a better solution does exist. However using 'replicates' one can overcome that problem by taking the one with the lowest total sum of distances, over all replicates as the final answer. (The MathWorks)

4. ANALYSIS AND INTERPRETATION

4.1. Identifying Intra-day Clusters

A daily electricity load curve represents the electricity load as a function of time. Figure 1 displays fluctuations of half-hourly electricity demand of Sri Lanka from January 2008 to December 2012.

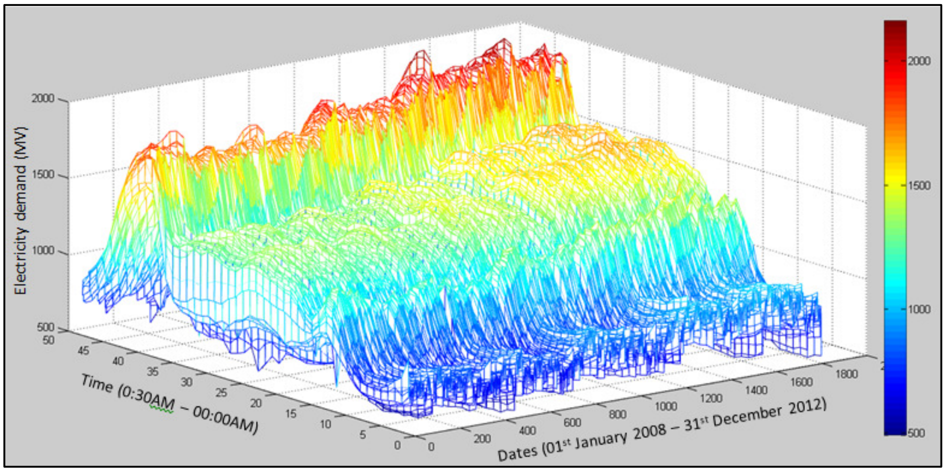


Figure 1. Fluctuations of electricity load curves over the years

From Figure 1, it is clearly shown that the daily demand curves appear to be having a similar pattern over the years, with a gradual increase of load from year to year. It is also noticeable that there are two sudden increments, one in the morning and the other at night, in each daily laod curve. The peak demand, starting around 6.30p.m and ending around 9.30p.m is identified as the most crucial aspect that needs to be addressed as the generation cost is very expensive during that time when compared to non-peak hours. A colour map scaling plot (Figure 2) shows that three main clusters can be seen.

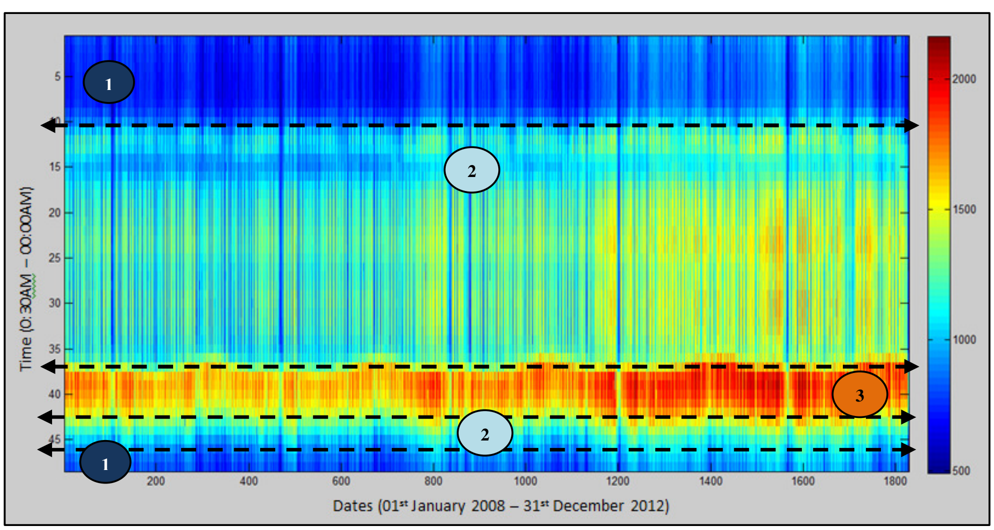


Figure 2. Colour map scaling plot of the load curves

In order to statistically validate the number of clusters within a day, 48 half-hours were clustered using the k-means algorithm. Even though there seems a trend in the data, it is assumed that the trend may not affect the clustering process as it is same for all the half-hour periods. The maximum average Silhouette Value of 0.8053 was resulted from the 3 clusters instance, and therefore three clusters were selected as the most appropriate number of clusters. In order to avoid the iterations to end up at local minimas, each clustering procedure was replicated 5 times and considered the one with the lowest total sum of distances.

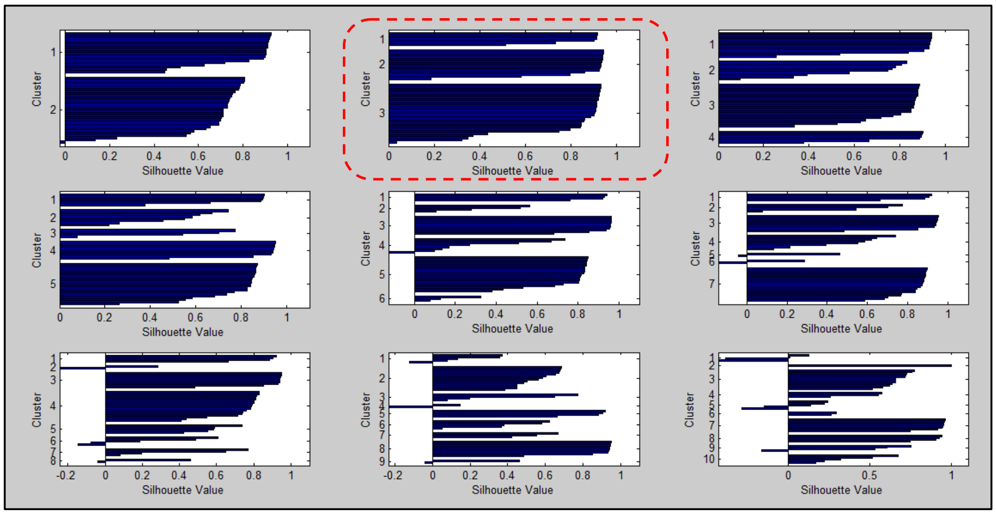


Figure 3. Silhouette Plots to determine the correct number of clusters

Based on the cluster analysis results, Table 1 displays how the 48 half-hours were grouped to the three clusters.

Table 1. Time periods identified for the three clusters

Cluster	Time Period/s		
Cluster 1	11:00PM - 05:00AM	07:30AM	
Cluster 2	06:30PM - 09:00PM		
Cluster 3	05:30AM - 07:00AM	08:00AM - 06:00PM	09:30PM-10:30PM

AM0_30	AM1_00	AM1_30	AM2_00	AM2_30	AM3_00	AM3_30	AM4_00	AM4_30	AM5_00	AM5_30	AM6_00	AM6_30	AM7_00	AM7_30	AM8_00	AM8_30	AM9_00	AM9_30	AM10_00	AM10_30	AM11_00	AM11_30	00
1	1	1	1	1	1	1	1	1	1	3	3	3	3	1	3	3	3	3	3	3	3	3	3

PM12_30	PM1_00	PAM1_30	PM2_00	PM2_30	PM3_00	PM3_30	PM4_00	PM4_30	PM5_00	PM5_30	PM6_00	PM6_30	PM7_00	PM7_30	PM8_00	PM8_30	PM9_00	PM9_30	PM10_00	PM10_30	PM11_00	PM11_30	AM00_00
3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	3	3	3	1	1	1

The three clusters can be interpreted in the following way. Cluster 1 represents the time period where the demand is minimum, and is the time during which most people are a sleep. This is the time period 11:00PM – 05:00AM. In the same cluster, the time 7.30-8.00 am could possibly be the time where most people travel, and are in vehicles. Cluster 2 with the maximum demand is the peak demand period of 06:30PM - 09:00PM. This is the time where most families live in their homes engaged in various activities, and hence the usage of electricity is a maximum. Finally, Cluster 3 includes the rest of the hours of the day.

As the three clusters seems to justify the electricity usage in the Sri Lankan context, predicting half-hourly demands seems more appropriate. As Deshani et al. (2014) have found that there is a day effect for the daily electricity demand, the electricity usage patterns for different days within each of these chosen clusters will also be analysed.

4.2. Clustering Similar Day Types within Intra-day Clusters

In the dataset, each day had been assigned into one of the thirteen categories; Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Poyaday, PBM Holiday (Public, Bank, Mercantile), PB Holiday (Public, Bank), Working day before holiday, Working day after holiday, Working day between a holiday and a weekend, Saturday after holiday (Deshani et al., 2014). We use the same day classification in order to analyse the day effects within each of the identified clusters.

<u>Intra-day Cluster 1</u>:

- 11:00PM - 05:00AM

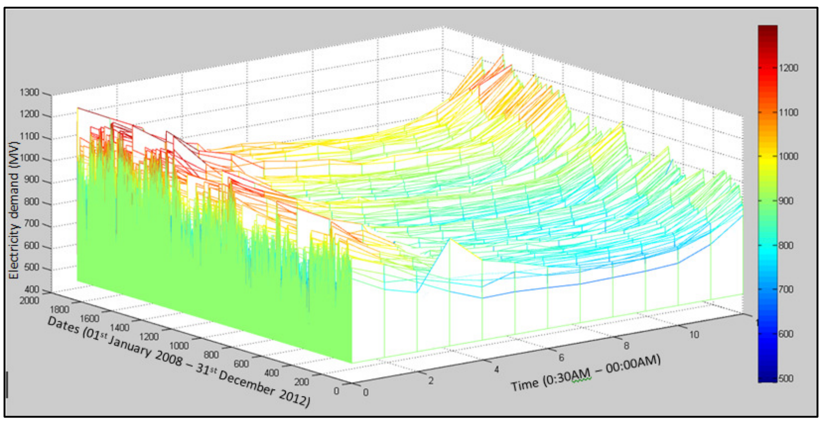
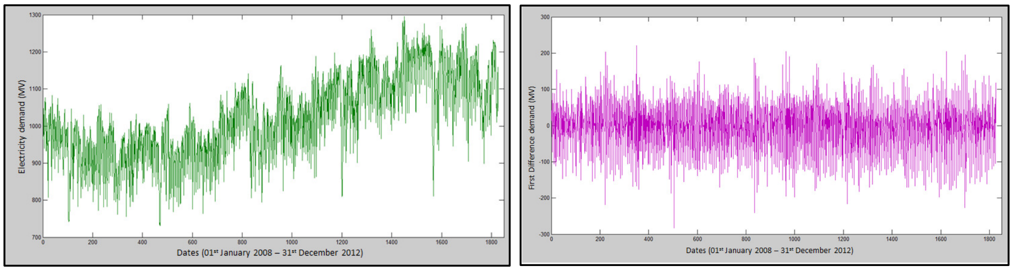


Figure 4. Fluctuation of electricity demand during the 11:00PM-05:00AM time period

Recall that the cluster 1 is times between 11:00PM - 05:00AM and the half-hour 7.30-8.00. Figure 4 displays how electricity demand varies within 11:00PM - 05:00AM time period. The load patterns during this period do not cluster based on the speciality of the day or month. As stated before, this seems reasonable, as during this time people do not engage in much activities.

- 07:30AM



(a) (b)

Figure 4. Fluctuation of electricity demand in the 07:30AM time period
(a) Original series (b) First differenced series

Figure 4 is drawn considering the half-hour 7:30AM - 8:00 AM. From the unit root test it was identified that there is a trend in the series (Figure 4 (a)) and was de-trended by taking the first difference (Figure 4 (b)). The de-trended data was clustered and two sub-clusters (maximum average Silhouette Value of 0.7012) could be identified.

Table 2. Distribution of the days between the two clusters for 07:30AM

	Cluster	
	A	**B**
Monday	96.3%	3.7%
Tuesday	88.7%	11.3%
Wednesday	84.4%	15.6%
Thursday	85.9%	14.1%
Friday	75.5%	24.5%
Saturday	1.8%	98.2%
Sunday	98.0%	2.0%
Poyaday	79.0%	21.0%
PBM Holiday	79.5%	20.5%
PB Holiday	47.8%	52.2%
Working day before a holiday	32.3%	67.7%
Working day after a holiday	98.4%	1.6%
Working day between a holiday and weekend	59.3%	40.7%
Saturday after holiday	23.5%	76.5%

It could be seen that all the weekdays, Sundays and Holidays had been clustered as cluster A and Saturdays and most of the working days before holidays had been classified for to cluster B.

Intra-day Cluster 2: 06:30PM – 09:00PM

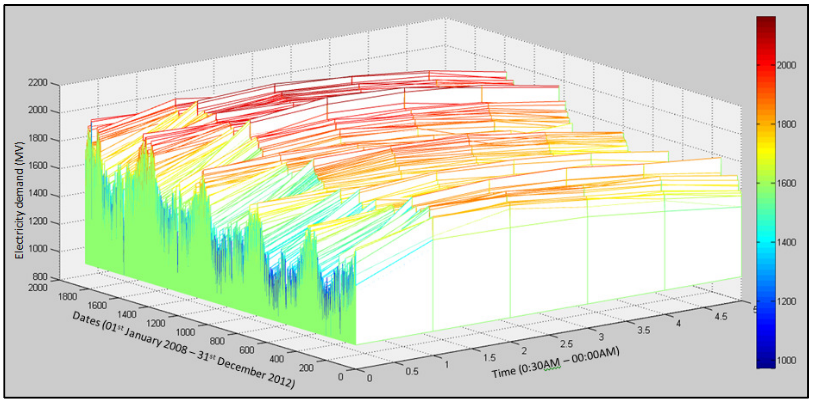


Figure 5. Fluctuation of electricity demand during the 06:30PM-09:30PM time period

According to Figure 5, load patterns within cluster 2's 06:30PM – 09:00PM time period seem to have a periodic pattern and a slight trend. The clustering was based on k-means clustering using the 'correlation' distance measure which resulted in two sub-clusters having the maximum average Silhouette Value of 0.8313. There was no need to remove the trend prior to clustering as the distance measure treated the point as a sequence of values and normalized values had been used for the calculations. The day types did not show any significant clustering result and hence considered months to identify for any clustering effect.

Table 3. Distribution of the days between the two clusters for 06:30PM – 09:00PM

		Cluster	
		G	H
Month	January	76.8%	23.2%
	February	96.5%	3.5%
	March	97.4%	2.6%
	April	94.7%	5.3%
	May	98.7%	1.3%
	June	100.0%	0.0%
	July	98.7%	1.3%
	August	98.7%	1.3%
	September	60.0%	40.0%
	October	1.3%	98.7%
	November	.7%	99.3%
	December	5.8%	94.2%

This peak time is shown to be clustered from January to August and October to December. September seems to be different than other months. The reason could be that part of September belonging to one cluster and the other part to the other cluster.

Intra-day Cluster 3:

- 05:30AM – 07:00AM

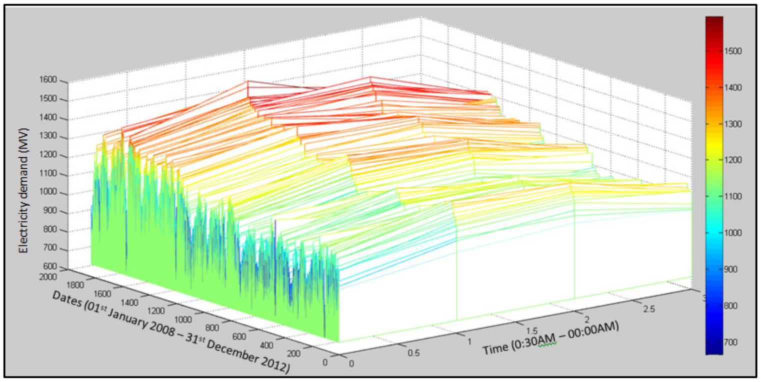


Figure 6. Fluctuation of electricity demand during the 05:30AM-07:00AM time period

Figure 6 displays how half-hourly electricity demands vary during the 05:30AM-07:00AM time period of the cluster 2. Two sub clusters (maximum average Silhouette Value of 0.0.7889) were prominent when clustering using the 'correlation' distance measure, for the cluster 3. Even though a clustering effect could be identified based on the type of the day, a better clustering affect could be seen considering both the type of the day and the month. It is to be noted that certain day types were combined in order to avoid complexities.

Table 4. Distribution of the days between the two clusters for 05:30AM-07:00AM

Month	Day Type	X	Y
January	Week day	58.3%	41.7%
	Week end	0.0%	100.0%
	Holidays	0.0%	100.0%
February	Week day	74.7%	25.3%
	Week end	0.0%	100.0%
	Holidays	6.7%	93.3%
March	Week day	95.2%	4.8%
	Week end	0.0%	100.0%
	Holidays	0.0%	100.0%
April	Week day	63.8%	36.2%
	Week end	7.9%	92.1%
	Holidays	22.2%	77.8%
May	Week day	99.0%	1.0%
	Week end	36.6%	63.4%
	Holidays	80.0%	20.0%
June	Week day	97.1%	2.9%
	Week end	9.8%	90.2%
	Holidays	100.0%	0.0%

July	Week day	99.1%	.9%
	Week end	2.3%	97.7%
	Holidays	20.0%	80.0%
August	Week day	17.1%	82.9%
	Week end	0.0%	100.0%
	Holidays	37.5%	62.5%
September	Week day	87.5%	12.5%
	Week end	10.3%	89.7%
	Holidays	42.9%	57.1%
October	Week day	98.0%	2.0%
	Week end	16.3%	83.7%
	Holidays	50.0%	50.0%
November	Week day	93.0%	7.0%
	Week end	2.5%	97.5%
	Holidays	0.0%	100.0%
December	Week day	16.5%	83.5%
	Week end	0.0%	100.0%
	Holidays	0.0%	100.0%

When considering cluster 3,the early morning time period, all weekends over the year had been clustered to cluster Y. Moreover, holidays in many months have also been clustered to Y. In addition, the weekdays of August and December, which are school holiday times in Sri Lanka, are clustered in Y. In contrast, in May and June, the holidays are clustered into cluster X as normal working days, as the electricity demand is comparatively higher in Wesak and Poson festival season.

- 08:00AM – 06:00PM

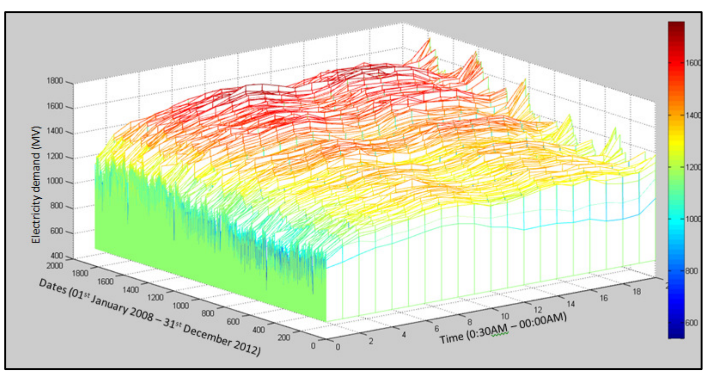


Figure 7. Fluctuation of electricity demand during the 08:00AM – 06:00PM time period

The electricity demand during 8.00am 6.00pm is shown in Figure 7. During this period, a slight trend can be seen and this period is the normal working time period of the day. As earlier time periods two sub-clusters could be identified for this period also.

Table 5. Distribution of the days between the two clusters for 08:00AM – 06:00PM

Day Type	Cluster	
	P	Q
Monday	1.9%	98.1%
Tuesday	3.8%	96.2%
Wednesday	3.3%	96.7%
Thursday	4.7%	95.3%
Friday	5.2%	94.8%
Saturday	52.4%	47.6%
Sunday	98.8%	1.2%
Poyaday	95.2%	4.8%
PBM Holiday	100.0%	0.0%
PB Holiday	34.8%	65.2%
Working day before a holiday	9.7%	90.3%
Working day after a holiday	9.7%	90.3%
Working day between a holiday and weekend	14.8%	85.2%
Saturday after holiday	82.4%	17.6%

It could be clearly seen that all working days cluster to Q whereas Sundays, Poyadays, PBM Holidays, Saturday after holidays clustered into P.

- 09:00PM – 10:30PM

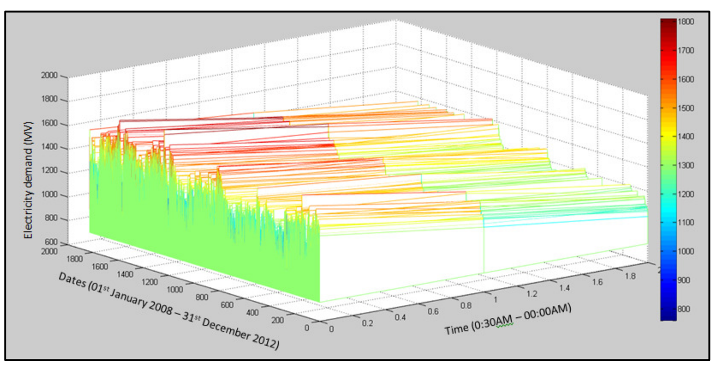


Figure 8. Fluctuation of electricity demand during the 09:00PM – 10:30PM time period

Figure 8 displays how electricity demand gradually increases within cluster 3’s 09:00PM – 10:30PM time period. During this period two sub-clusters were identified as the most appropriate number of clusters.

Table 6. Distribution of the days between the two clusters for 09:00PM – 10:30PM

	Cluster	
	M	N
Monday	37.0%	63.0%
Tuesday	43.9%	56.1%
Wednesday	42.9%	57.1%
Thursday	50.7%	49.3%
Friday	81.6%	18.4%
Saturday	93.3%	6.7%
Sunday	6.1%	93.9%
Poyaday	22.6%	77.4%
PBM Holiday	25.6%	74.4%
PB Holiday	47.8%	52.2%
Working day before a holiday	87.1%	12.9%
Working day after a holiday	45.2%	54.8%
Working day between a holiday and weekend	81.5%	18.5%
Saturday after holiday	76.5%	23.5%

Even though many day types did not exhibit a clear clustering approach based on the day type, Friday, Saturday, Working day before a holiday, Working day between a holiday and a weekend and Saturday after holidays clustered to cluster M. Sundays, Poyadays and PB holidays clustered to cluster N.

5. CONCLUSION

The electricity demand varies in accordance with consumers' activities with respect to time of the day and the day of the week. When predicting half-hourly electricity demand in a short term manner, these patterns influence the prediction process a lot. Therefore, a cluster analysis was carried out to identify intra-day clusters and to group similar day types within those clusters. The results of the cluster analysis were used to select a better set of neuro-forecasters for neural network predictions.

The K-means clustering algorithm was used to cluster data and three intra-day clusters were identified. Cluster 1 represents the time period where most people are a sleep (11:00PM – 05:00AM) and the time where most people may possible be traveling, and are in vehicles (07:30AM-08:00AM). Cluster 2 is the peak demand period of 06:30PM - 09:00PM where most families live in their homes engaged in various activities. Finally, cluster 3 includes the rest of the hours of the day. Based on the day type or the month, the three main intra-day clusters were further grouped and most appropriate number of clusters for each time period was considered as the number of neuro-forecasters used for prediction in that period. The number of neuro-forecasters and training set selection criteria has been shown in Table 7. Using this input selection process, the prediction performances of the neural networks can be improved.

Table 7. Section of neuro-forecasters based on the cluster analysis

Cluster	Time Period/s	Number of Neuro-forecasters	Training set selected based on
1	11:00PM - 05:00AM	1	All data can be used
	07:30AM	2	day type
2	06:30PM - 09:00PM	2	month
3	05:30AM - 07:00AM	2	month and day type
	08:00AM - 06:00PM	2	day type
	09:30PM-10:30PM	2	day type

If the considered time duration could be expanded, results can be improved and will be more accurate as there will be more observations for the sub categories.

5. ACKNOWLEDGEMENTS

I would like to thank my supervisors Dr. M.D.T. Attygalle, Dr. Liwan Liyanage Hansen and Ms. A. Karunaratne for the invaluable guidance, advice and the support given to do my research. Also I would like to thank the Vice Chancellor of University of Colombo, Dr. W.K. Hirimburegama, the Dean of the faculty, Prof. K.R.R. Mahanama for financially supporting this research through the University of Colombo Research Grant. I am grateful for the Ceylon Electricity Board of Sri Lanka for providing the data set and the background information about the electricity demand. At last but not least to my loving husband, my little son, my mother my father and my sister for their invaluable help and patience during the period I do my research.

REFERENCES

[1] Barzamini, R., Hajati, F., Gheisari, S., & Motamadinejad, M. B. (2012). Short Term Load Forecasting using Multi-layer Perception and Fuzzy Inference Syatems for Islamic Countries. Journal of Applied Sciences , pp40-47.

[2] Deshani, K.A.D, Hansen, L. L., Attygalle, M.D.T, & Karunarathne, A. (2014). Improved Neural Network Prediction Performances of Electicity Demand: Modifying Inputs through Clustering. Second International Conference on Computational Science and Engineering (pp. 137-147). India: AIRCC.

[3] Farahat, M. A., & Talaat, M. (2012). Short-Term Load Forecasting Using Curve Fitting Prediction Optimized by Genetic Algorithms. International Journal of Energy Engineering , pp23-28.

[4] Hernandez, L., Baladron, C., Aguiar, J. M., Carro, B., & Esguevillas, A. S. (2012). Classification and Clustering of Electricity Demand Patterns in Industrial Parks. Energies , pp5215-5227.

[5] Nagi, J., Yap, K. S., Tiong, S. K., & Ahmed, S. K. (2008). Electrical Power Load Forecasting using Hybrid Self-Organizing Maps and Support Vector Machines. International Power Engineering and Optimization Conference, (pp. 51-56). Selangor.

[6] Seetha, H., & Saravanan, R. (2007). Short Term Electricity Load Prediction Using Fuzzy BP . Journal of Computing and Information Technology , pp267-282.

[7] Soares, L. J., & Medeiros, M. C. (2008). Modeling and Forecasting short-term electricity load: A comparison of methods with an application to Brazilian data. International Journal of Forecasting , pp630-644.

[8] The MathWorks, I. Statistics Toolbox.

Authors

Ms. K.A.D. Deshani is a Lecturer (Probationary) in the Department of Statistics, University of Colombo since March 2011. Before she was absorbed to the permanent cadre, she has been working in the same department as an Assistant Lecture, obtaining a B.Sc Special Degree in Statistics with Computer Science in 2008. She has a keen interest in developing computer systems to incorporate the dynamic nature to statistical interpretations. Her research interests are in the areas of Operational Research and Data Mining. Currently, she is a member of the research team of the project titled “Developing an Economical Strategy for the Future Electricity Generation Procedure in Sri Lanka”; which received the University of Colombo research grants 2011 which is carried out in collaboration with University of Western Sydney, Australia and Ceylon Electricity Board, Sri Lanka. In June 2012 she registered for a M.Phil under the said grant. With her interest in research, she was a contributed speaker at the International Statistics Conference on the publication titled “Analysis of Efficiency of a Multi-Queue against a Single Queue with Many Servers: A Study on Advertisement Counter Queues at a Leading Newspaper Company” and was published in the Proceedings of the International Statistics Conference 2011, Colombo Sri Lanka. In 2013 a research paper titled “A Study of the Dynamic Behaviour of Daily Load Curve for Short Term Predictions” was published in the proceedings of the International Symposium for Next Generation Infrastructure (ISNGI) Australia.

Dr. Attygalle has been Head of the Department of Statistics from 2010 to present. As a Senior Lecturer attached to the Department of Statistics from 2006 to present, she has been routinely involved in Teaching, Research and many other administrative roles such as being the Coordinator of the MSc in Applied Statistics, BSc Special degree and Joint Special degree programmes conducted by the Department of Statistics. She holds professional memberships of the Sri Lanka Association for the Advancement of Science (SLASS) and the Institute of Applied Statistics -Sri Lanka.

Dr. Attygalle obtained her PhD in Statistics from the Lancaster University, UK, and a MSc in Statistics from the Warwick University, UK, in 2006 and 1996 respectively. Prior to this she completed a Diploma in Applied Statistics, from the University of Colombo in 1992. She obtained her first degree majoring in Statistics, Applied Mathematics and Pure Mathematics also from the University of Colombo, graduating with a first class in 1987. As professional qualifications she has obtained the Staff and Educational Development Association (SEDA)-UK accreditation as a teacher in higher education in 2005 and the Certificate in Teaching in Higher Education (CTHE) by the Staff Development Centre of the University of Colombo also in 2005. Her key research areas are Statistical Modelling, Model Diagnostics, Data Visualization and Sports Statistics. As a senior lecturer she has supervised many undergraduate and postgraduate research projects. Currently she is so-supervising two MPhil/PhD research students. She had been instrumental in developing industry links with many private and government organisations over the years and thus has carried out many consultancy projects and other training programs through the Department of Statistics. She had also won one of the University of Colombo research grants in 2011.

Dr Liwan Liyanage joined University of Western Sydney in the year 1989 with university level teaching experience at University of Colombo, University of Wollongong and King Saud University Riyadh, totalling 12 years. Qualifications: B. Sc (First Class), Graduate Diploma in Applied Statistics, Masters Degree in Theoretical Statistics and the Ph. D. in the area of Applied Probability gives her the breadth of coverage across the statistics disciplines. At UWS she has been instrumental in developing many degree programs in particular the integrated degree B. Maths and IT using data mining as the integrating tool. Senior lecturer (1995) and head of program of B. Maths & IT (1999). Her PhD was in random walk models, diffusion and related applications namely queuing theory and game theory. Thus her initial research was in applied probability, namely random walk models with difference equations, the master equation models with partial differential equations, and queuing models. This leads to differential equations representing diffusion and double diffusion. Her research, bridge the probabilistic models to the differential equation models of diffusion. Her passion to integrate disciplines and research methods have led her to the current research areas which include innovative work in “Operational Statistics” a new area developed in collaboration with UC Berkeley; Optimisation Techniques and Data Mining. Application areas include bio security, public health, climate change, electricity production and demand. From her 10 PhD/Masters students 8 had completed the research successfully. She has established ongoing national and international linkages and research collaborations and 30+ publications and a book chapter. Her paper on Operational Statistics was the 2nd most downloaded paper in April 2006 from Science Direct’s TOP 25 articles.

Mrs. A. Karunarathne is a former head of the Department of Statistics, and also had served as the former head of Department of Statistics and Computer Science. She had been in the service for more than 40 years and has been the key person to start the Special Degrees in Statistics and also initiate the Internship program in the Department. As a senior lecturer she is conducting lectures mainly in the field of Operational Research and Stochastic Processes.

She obtained a Diploma di. Sp.(Operational Research) from University of Rome and her first degree was B.Sc.(Mathematics) from the University of Colombo. She has contributed to the continuous development and transmittance of statistical knowledge through many diverse avenues, a key example of which is her involvement in the publication of a book on basics of statistics titled “Moolika Sankayanaya” written for University entrants and A/L Science Students in Sept 1997. As a senior lecturer she has supervised many undergraduate and postgraduate research projects.. Her key research areas are Stochastic Processes, Simulation, Queuing Models and Performance Modelling of Communication Networks.