American Journal of Applied Mathematics and Statistics, 2019, Vol. 7, No. 1, 9-17
Available online at http://pubs.sciepub.com/ajams/7/1/2
Published by Science and Education Publishing
DOI:10.12691/ajams-7-1-2

Artificial Neural Network for Dynamic Iterative Forecasting: Forecasting Hourly Electricity Demand

K.A.D. Deshani[1,*], Liwan Liyanage-Hansen[2], Dilhari Attygalle[1]

[1]Department of Statistics, University of Colombo, Colombo 03, Sri Lanka
[2]School of Computing, Engineering and Mathematics, University of Western Sydney, Campbelltown, Australia
*Corresponding author: deshani@stat.cmb.ac.lk

Received November 07, 2018; Revised December 08, 2018; Accepted December 25, 2018

Abstract This paper presents the procedure of building a dynamic predictive model using an artificial neural network to perform an iterative forecast. An algorithm is proposed and named as “Artificial Neural Network Approach for Dynamic Iterative Forecasting”. The development of this algorithm focused on feature selection, identification of best network architecture for the model, moving window selection and finally the iterative prediction. This proposed algorithm was deployed to forecast next day’s hourly total demand in Sri Lanka as an illustration. Inclusion of a clustering effect that were based on the specialty of the day, as an input was investigated through this application, from which improved accuracies were shown.

Keywords: *dynamic forecast, neural networks*

Cite This Article: K.A.D. Deshani, Liwan Liyanage-Hansen, and Dilhari Attygalle, “Artificial Neural Network for Dynamic Iterative Forecasting: Forecasting Hourly Electricity Demand.” *American Journal of Applied Mathematics and Statistics*, vol. 7, no. 1 (2019): 9-17. doi: 10.12691/ajams-7-1-2.

1. Introduction

In the era of data science, prominence has been given to data driven research where static models become obsolete as new records are added to the database continuously over time. In such scenarios, most of the traditional statistical techniques become less capable to handle complex real world situations in a dynamic way. With the advancement of artificial intelligence approaches, these barriers can be overcome effectively.

In recent years, neural networks have gained much attention in the fields of pattern identification, classification and forecasting due to its capabilities of handling both linear and non-linear relationships efficiently. Thus, the use of neural networks in data driven research have gained much attention in the recent past.

2. Literature Review

When considering the literature regarding forecasting time series data, statistical models are static in nature so that a fitted model can only be used to forecast during a limited time frame [1,2,3,4]. When the forecasts have to be performed for a longer time span, the model parameters should be updated to represent the more recent situation and hence static models become obsolete over time. In this regard, artificial intelligence approaches plays a key role in implementing dynamic models using the concept of moving windows to update the models with respect to most recent data [5,6,7]. In almost all of the published work regarding dynamic forecasting, neural networks play a main role solely or in combination with another technique as a hybrid model.

Guan and Luh (2005) have combined fuzzy logic, neural networks, and autoregressive models-for very short-term power system load forecasting due to the high accuracy levels [5]. To forecast sunspot activates Xie et. al (2006) have used a hybrid neural network model, which integrated characteristics decomposition units, and a dynamic spline interpolation unit into the multiple adaptive time-delay neural networks [6]. An accurate and timely forecasting of traffic flow model has been implemented by Jiang and Adeli (2005) using a novel nonparametric dynamic time-delay recurrent wavelet neural network [7].

When the forecast has to be performed for many steps ahead, the researches have recommended iterative predictions than multi-step ahead predictions with experimental evidence [8,9]. Thus, this algorithm also proposes an iterative neural network model to increase the accuracy levels of forecasts which are several steps ahead.

3. Proposed Algorithm: Artificial Neural Network Approach for Dynamic Iterative Forecasting

When considering data driven research, many applications can be for forecasting future values. This paper proposes an algorithm that can be used to forecast time series data with many associative variables. The iterative forecast mechanism used in this algorithm enables the use of the recently forecasted values when forecasting the next immediate forecast, and is said to be superior to forecasting the series multi steps ahead.

PSEUDOCODE: ARTIFICIAL NEURAL NETWORK APPROACH FOR DYNAMIC ITERATIVE FORECASTING

This pseudocode can be used to forecast data iteratively using a neural network

INPUT:

Series = Series to be forecasted

OtherInputs = A set of other input variables for the NN

HiddenNeurons_K= Number of neurons in the hidden layer (K=1,….,k)

Replicate_K_N= Nth replicate of the network with K neurons (K=1,..,k and N=1,…,n)

OTHER VARIABLES:

LaggedInputs = Lagged terms to be added as inputs

K = Number of neurons in the hidden layer

MeanValidation_K = Mean error measure of the validation set having K neurons in the hidden layer (K=1,…, k)

MedianValidation_K = Median error measure of the validation set having K neurons in the hidden layer (K=1,…, k)

K = Next item to be forecasted

SelectedInputs = The set of selected input variables for the NN that resulted minimum average error measure (eg: Mean Squared Error). This set includes ***LaggedInputs*** also

MWsize = Size of the moving window

BestNetworks = A selected set of best neural networks to do the final forecast

t = Time of forecast

OUTPUT:

Forecast(t) = Forecasted value at time t

START

Feature selection

- Calculate mutual information estimates for lagged terms of the Series considering a set of suitable candidate lagged variables and sort them in descending order.
- Plot the estimates and select the set of lagged terms (***LaggedInputs***) prior the decrement rate become very low.
- Identify other possible input variables that can be incorporated to the model

Identification of best network architecture for the model and feature selection

For simplicity and as the most common case the pseudocode has been written for a neural network having only one hidden layer. This can be increases as desired. But with the increased number of hidden layers, over fitting problems arise.

FOR (all the possible combinations of ***OtherInputs*** with the ***LaggedInputs***)

{

FOR (HiddenNeurons_K =1:k)

{

FOR (***Replicate_K_N*** =1:n)

{

- Train neural network with k neurons in the hidden layer
- Save training set error, testing set error, validation set error and execution time

}

Calculate the average and median of the above measures for network having K neurons in the hidden layer as ***MeanValidation_K***

}

IF (K resulting minimum(***MeanValidation_K***) = =K resulting minimum(***MedianValidation_K***))

Identify ***K*** as the number of neurons in the hidden layer

ELSE

Study variations in error measures of the replicates of both the cases for outlier observations and select the most suitable value for ***K*** with low variations.

}

Select the combination having minimum mean error measure as the best set of inputs to the final neural network as ***SelectedInputs***.

Selection of moving window

- Select the best size for the moving window varying the size of the moving windows for the neural network with ***SelectedInputs*** having minimum validation set error measure as ***MWsize***. The number of neurons in the hidden layer should be updated accordingly as K.
- Save 10 neural networks with ***SelectedInputs***, K neurons in the hidden layer to be used for the final iterative forecast as ***BestNetworks***.

Final dynamic forecast

- Do the final iterative forecasting up to ***h*** steps ahead taking the average output of the ***BestNetworks*** using the selected moving window size ***MWsize***, slide on with time with the updated model to forecast the next step ***Forecast(t)***.

END

4. An illustration of Artificial Neural Network Approach for Dynamic Iterative Forecasting

When considering power and energy industry, Short Term Load Forecasting (STLF) plays a critical role as the decision regarding the forecast should be taken within a very limited time span. Even though a vast range of research have been carried out on STLF, the challenging aspect is to increase the accuracy of the forecasts. Literature on STLF reveals that, using artificial intelligence based approaches for predictive models claimed to generate high accuracies compared to other conventional statistical techniques [10-14].

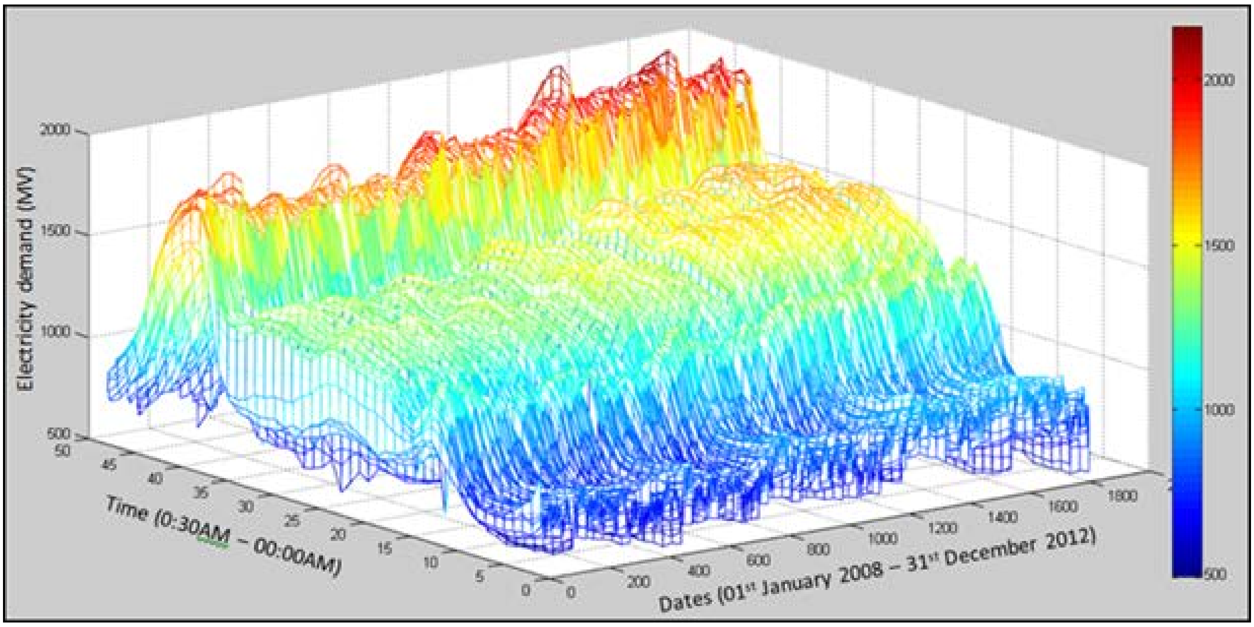

Figure 1. Fluctuations of electricity load curves over the years

Thus, to illustrate this algorithm, a dynamic model was implemented to forecast next day's hourly electricity demand in Sri Lanka. The dataset consisted of hourly electricity demands form 01st January 2008 to 31st December 2012. Data in the first four years were used for training purposes and the data of the final year was used for testing. Figure 1 displays how the hourly electricity demands vary over the five years.

4.1. Feature Selection

Prior to designing the network, a feature selection procedure was conducted to identify the most relevant inputs for the model, as inclusion of irrelevant inputs can distort the forecasting performances, increase the training time and increase network complexity [15].

A filter based approach, namely, mutual information based feature selection procedure was used to choose the lagged demand terms as the inputs to the neural network. Because of the strong daily and weekly patterns in the data series, a candidate feature set consisting of 24*7=168 variables were used to select the most appropriate features for the model. It is important to note that Rana et. al (2014) have justified a sliding window of one week is sufficient for this choice of variables, as it captures all seasonal patterns across a week [14]. Figure 2 displays mutual information estimates sorted in descending order and the first 52 lagged terms were selected as the inputs to the neural network, due to no substantial changes with the mutual information estimates thereafter. The selected lagged variables were X_{t-1} to X_{t-14}; X_{t-16} to X_{t-26}; X_{t-28} to X_{t-31}; X_{t-35} to X_{t-37}; X_{t-47} to X_{t-49}; X_{t-71} to X_{t-73}; X_{t-95} to X_{t-97}; Xt_{-119} to X_{t-121}; X_{t-143} to X_{t-145}; X_{t-164} to X_{t-168}.

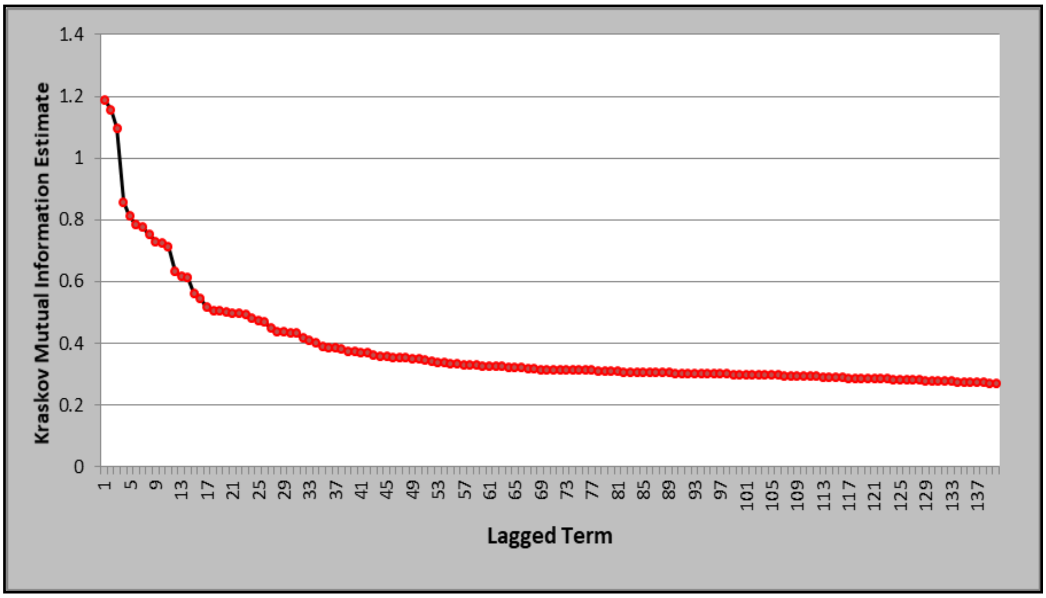

Figure 2. Graph of mutual information estimates of hourly electricity demands

4.2. Other Inputs

The exploratory analysis revealed that day of the week, month, specialty of the day and temperature plays a main role in STLF [16,17] and were used along with the lag variables as inputs in the model. In addition, results of the hourly data clustering [16,17] and curve clustering, were also considered as inputs to improve neural network performances.

4.3. Iterative Prediction Using Neural Networks

Designing the neural network

- To implement a dynamic forecasting model, the neural network should be efficient, adjustable and less complex. Among the relevant literature, back propagation neural networks are widely used for STLF [15,18,19,20,21] Thus an error back propagated neural network was selected to implement the forecasting model.
- As the requirement was to forecast next day's hourly demand dynamically, an iterative forecast was chosen as opposed to a multi-step ahead forecast. This choice was supported through literature in general, and has discussed extensively by [22].
- Given consistent data and enough neurons in its hidden layer, a two-layered feed-forward network with sigmoid hidden neurons and linear output neurons, fit multi-dimensional mapping problems arbitrarily well. The inputs as well as outputs were scaled in the range of [-1,1] [23].
- Due to the less training time, the Levenberg-Marquardt back propagation algorithm was used to train the network. As the size of the required memory capacity is manageable the choice of this training algorithm can be justified.
- This neural network consisted of three layers as input layer, hidden layer and an output layer. Even though researchers argue that one hidden layer is capable of modelling well avoiding overfitting, for the completeness of the experiment, all the test cases were done considering two hidden layers with same number of neurons in each layer.
  - The number of inputs in the input layer was varied with respect to different combinations of the input variables. The experiments are displayed in Table 1.
  - Only one neuron was in the output layer as one step ahead prediction will be done.
  - The number of neurons in the hidden layer was varied and the best combination was identified accordingly.

Selecting input variable combination and number of neurons in the hidden layer

- Sixteen separate neural networks were trained with different combinations of inputs.
- Each network was trained 50 times and the number of neurons in the hidden layer was chosen based on the minimum average MSE (Mean Squared Error).
  - The median MSE was also considered for this selection as the average is subjected to be influenced by extreme points.
  - For a particular network, when the minimum mean value and the minimum median value relates to the same number of neurons in the hidden layer, the choice was made on that number of neurons. When minimum average and minimum median acquired from two distinct sets of hidden neurons, both neuron sizes were observed and the one with the lower variations of the MSE was chosen.
- Each combination of inputs were considered for neural networks with 2 hidden layers too, but they generated worse errors. Thus it was decided to have only one hidden layer in the network.

A summary of results is displayed in Table 1. The categories of each input variable are listed below where the parenthesis gives the number of inputs under each category. Note that the categorical variables were used in the form of several binary variables.

- Lags (52) -> 52 variables with lagged observations
- Temp (1)-> Estimated hourly temperature measurement
- Day Type (7) -> Sunday, Monday, Tuesday, … , Saturday (7 indicator variables)
- Specialty (5) -> None, poya day, PBM (Public, Bank, Mercantile) holiday, one day after New Year, two days after New Year (5 indicator variables)
- Month (12) -> January, February, … , December (12 indicator variables)
- Cluster(6) -> Clusters identified based on hourly demand.
  - Cluster 1 – Demand during early morning
  - Cluster 2 – Demand during 5 a.m., 7 a.m., 8 a.m. and 11 a.m.
  - Cluster 3 – Demand during working hours on Sundays (in 2008, Saturdays were added in this cluster based on the cluster analysis results)
  - Cluster 4 – Demand during working hours on Monday – Saturday where specialty is none
  - Cluster 5 – Demand during working hours on Monday – Saturday where specialty is holiday
  - Cluster 6 – Demand during peak time
- CurveCluster(2) -> Clusters identified based the shape of the daily load curves (2 indicator variables)
  - Cluster 1 – No specialty (This includes PB, working day after holiday, etc.)
  - Cluster 2 – Holidays (PBM, poya day, Saturday after holiday)

Based on the experiments, two key findings were observed.

1. When exploring different input combinations as displayed in Table 1, with the inclusion of the cluster information as inputs, the Mean Squared Error of the validation set has reduced considerably.
2. With the inclusion of cluster information, the training times have also decreased.

It is therefore evident that, with the inclusion of clustering effect, the training process of the neural network becomes more efficient. Moreover, by using clusters as inputs, the network became less complex due to the reduction of dimensionality of a set of variables as inputs.

Table 1. Performance measures of neural networks with different input combinations

ID	Inputs (No of Variables)	No of Inputs	Neurons in the hidden layer resulting min MSE	Validation set MSE	Mean Execution Time
Only with available variables					
1	Lags (52)	52	25	1131.5	123.7
2	Lags(52) Temp(1)	53	13	1148.5	108.5
3	Lags(52) Day Type(7) Specialty(5)	64	23	1063.9	154.4
4	Lags(52) Day Type(7) Specialty(5) Month(12)	76	22	1058.2	137.5
5	Lags(52) DayType(7) Specialty(5) Month(12) Temp(1)	77	21	1086.0	204.9
With Cluster Results					
6	Lags(52) Day Type(7) Cluster(6)	65	18	1071.6	86.4
7	Lags(52) Day Type(7) Cluster(6) Month(12)	77	27	1075.2	80.37
8	Lags(52) DayType(7) Specialty(5) CurveCluster(2) Month(12) Temp(1)	79	26	1082.5	259.71
9	Lags only (52) CurveCluster(2) Temp(1)	55	28	1078.4	120.21
10	Lags only (52) CurveCluster(2) Cluster(6)	60	26	1055.6	120.8
11	***Lags only (52)*** ***CurveCluster(2)*** ***Cluster(6)*** ***Temp(1)***	***61***	***19***	***1048.5***	***100.02***
12	Lags only (52) CurveCluster(2)	54	18	1077.8	118.5
13	Lags(52) DayType(7) Specialty(5) Cluster (6) Month(12) Temp(1)	83	28	1081.0	147.8
14	Lags(52) Cluster (6) Month(12) Temp(1)	71	30	1091.6	180.6
15	Lags(52) Cluster (6) Temp(1)	59	26	1095.4	196.36
16	Lags(52) Cluster(6) Month(12)	70	20	1088.6	144.1

In conclusion, the minimum validation set error was obtained for the inputs, lags, cluster results based on hourly data, curve cluster results based on daily data, and temperature estimate. The results suggested that with the usage of clustering results, the performance of the neural network was increased.

Size of the moving window

- For the selected combination of inputs, the neural networks were trained by varying the moving window size. Figure 3 displays how the average MSE of the validation set varies with the size of the moving window. It can be clearly observed that, a moving window of 2 years resulted in the minimum average MSE when training the neural network. Thus, for the final prediction, the neural networks were trained using most recent data spanning 2 years. Considering the minimum MSE for the two year moving window period, 10 neurons were chosen for the hidden layer.

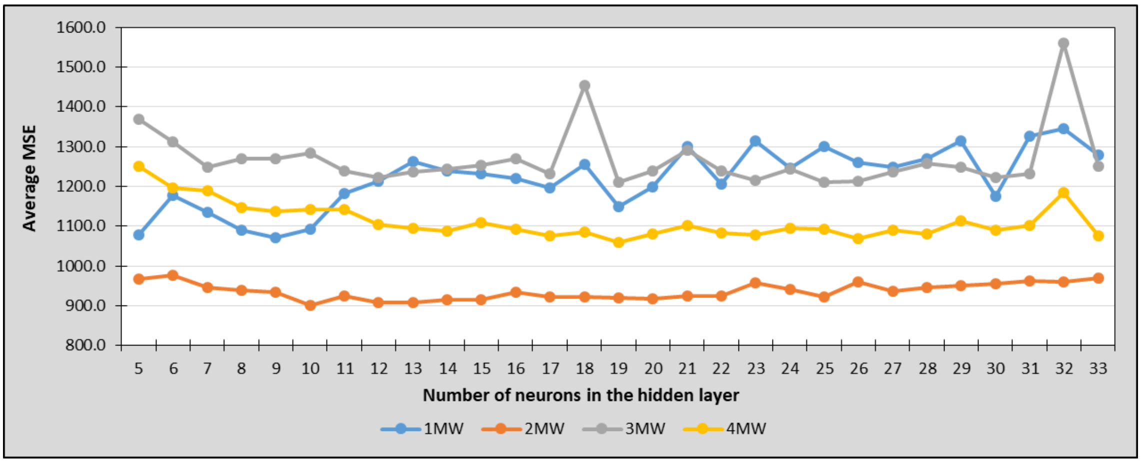

Figure 3. Average MSE of the validation set based on the moving window size for training

4.4. Accuracy of the Forecasts

The average MAPE for the 2012 forecasted values was 3.96%. Some selected descriptive statistics of the MAPE for the 2012 are shown in Table 2. The RMSE of the final dynamic prediction is shown in Figure 4. It can be clearly observed that for a majority of days the RMSE of the forecasted values were considerably low, except for a few days. Except the New Year's Day, all other days with very high errors were generated on days with no specialty.

It is learnt from Table 2, which 50% of the MAPE for the year 2012 were between 0.026 and 0.0368. Note that the prediction was made for an entire year to validate the suitability of this dynamic model for a longer time. The results have proven that the predictions are reasonably accurate even for this long period of one year. Thus, this model may be suitable even after a one year but, the applicability should be tested prior to using the implemented forecasting model.

Table 2. Descriptive statistics of the MAPE of the series

Minimum	Maximum	Percentiles		
		25	50	75
1.17%	14.9%	2.60%	3.68%	5.1%

In order to identify whether the observed errors can be grouped in some way, the RMSE values were clustered where the maximum Silhouette was obtained for 2 clusters. Figure 5 displays the distribution of the RMSE for the two clusters named as low error cluster (1) and high error cluster (2).

Table 3 shows the assignments with respect to the specialty of the day. It can be clearly seen that the majority of the PBM holidays and poya days were assigned to the group having low RMSE. Also it is to be noted that nearly 35% of the days having no specialty, have been assigned to the group with high errors.

Table 3. Two identified clusters based on RMSE of the dynamic predictions for 2012 using NN

Specialty	Cluster			
	1 -Low RMSE		2 - High RMSE	
	Count	Row %	Count	Row %
None	200	65.6	105	34.4
One day after New Year	0	0.0	1	100.0
PB Holiday	3	75.0	1	25.0
PBM Holiday	7	87.5	1	12.5
Poyaday	10	76.9	3	23.1
Saturday after holiday	1	50.0	1	50.0
Saturday after PB holiday	0	0.0	1	100.0
Two days after New Year	1	100.0	0	0.0
Working day after holiday	2	16.7	10	83.3
Working day after PB holiday	2	66.7	1	33.3
Working day before holiday	6	75.0	2	25.0
Working day before PB	1	100.0	0	0.0
Working day b/w PB & weekend	1	100.0	0	0.0
Working day b/w holiday & weekend	6	100.0	0	0.0
Total	240		126	

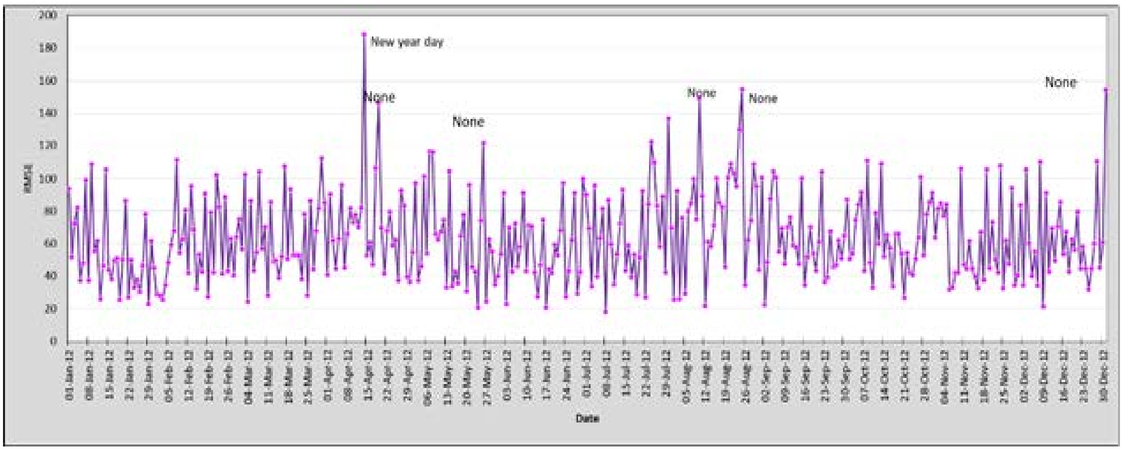

Figure 4. RMSE of the dynamic forecast for the year 2012 using iterative NN

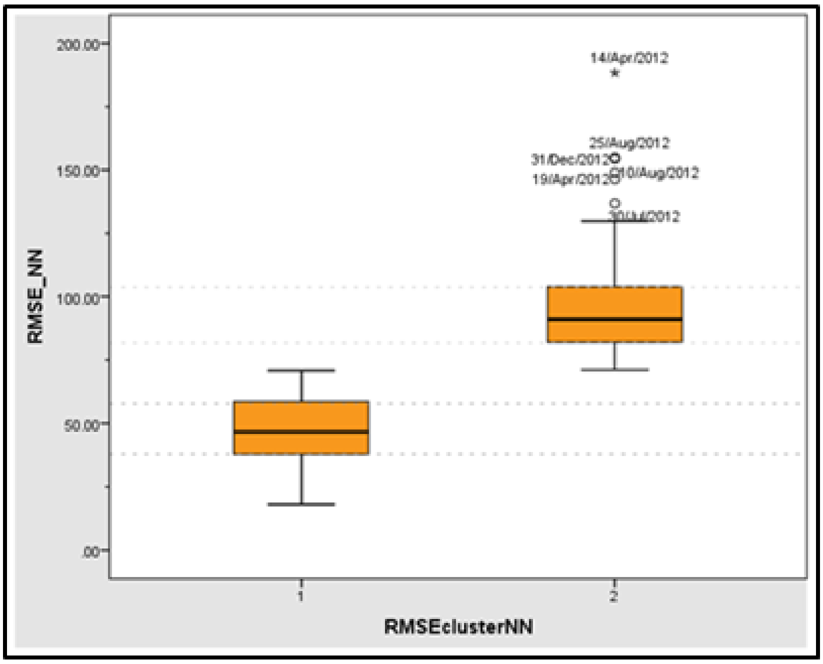

Figure 5. Distribution of the RMSE of the two identified clusters

- When exploring errors based on the day of the week, it was identified that Saturdays have high errors compared to other days of the week. It is anticipated that this may be as a result of the unclear assignment of Saturdays to a particular cluster when clustering hourly electricity demands [16,17].
- To address this issue, another neural network was trained with the same set of inputs, but replaced the six dummy variables with seven, to represent hourly demand clustering results. The seventh variable indicated whether a certain demand is during a working hour on a Saturday or not. As such, the working hours on weekdays were thus separated from the demand on Saturday working hours.
- When there is a holiday (PBM holiday or poya day) the demands during working hours were assigned to the cluster including working hours of holidays during weekdays.

After this adjustment for Saturdays, the process was repeated to dynamically forecast electricity demand for the year 2012. The model resulted in substantially low errors, for both Saturdays and weekdays when compared to the previous model. However, the new model, when compared to the previous neural network, was not able to improve accuracies for certain holidays. The new errors were compared with the previous forecasting errors as shown in Figure 6.

- It can be clearly seen that some of the high errors resulted from the previous model have been reduced with the new Saturday adjusted model.
- The new model is also incapable of accurately predicting the demands of one day after the New Year's day.
- Out of the 366 days of the year 2012, 66% of the days resulted with low MAPE for the new model when compared to the previously model.
- The average MAPE for the year 2012 decreased from 3.969% to 3.522% (decreased by 11.26%) after addressing the effect of Saturday working hour demands.
- Thus the recalibrated model was selected to be used in the final forecasting process.

Figure 7 displays examples for accurately, moderately and poorly forecasted days using the Saturday adjusted model. For accurate forecasts, MAPE values were less than 2% in general. Based on the overall deviation for the entire day, the examples were classified into moderately or poorly forecasted days.

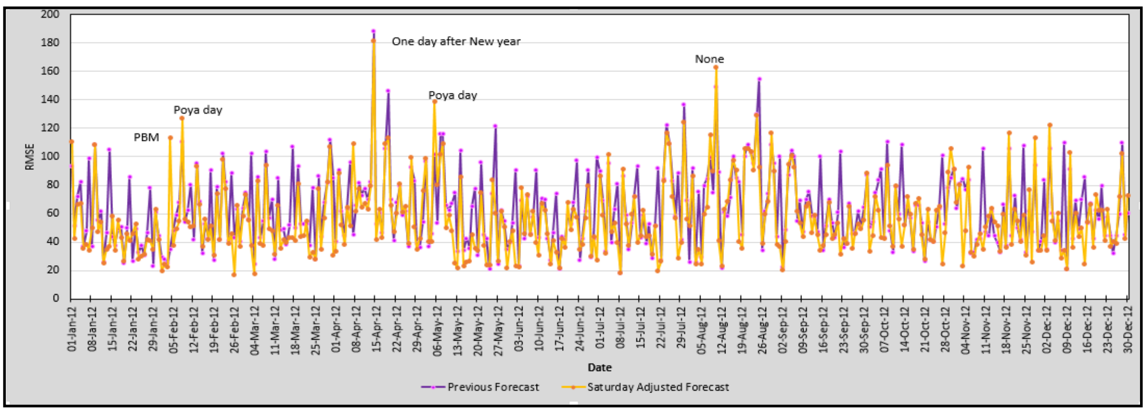

Figure 6. Comparison of the RMSE after adjusting Saturday working hour demand allocations

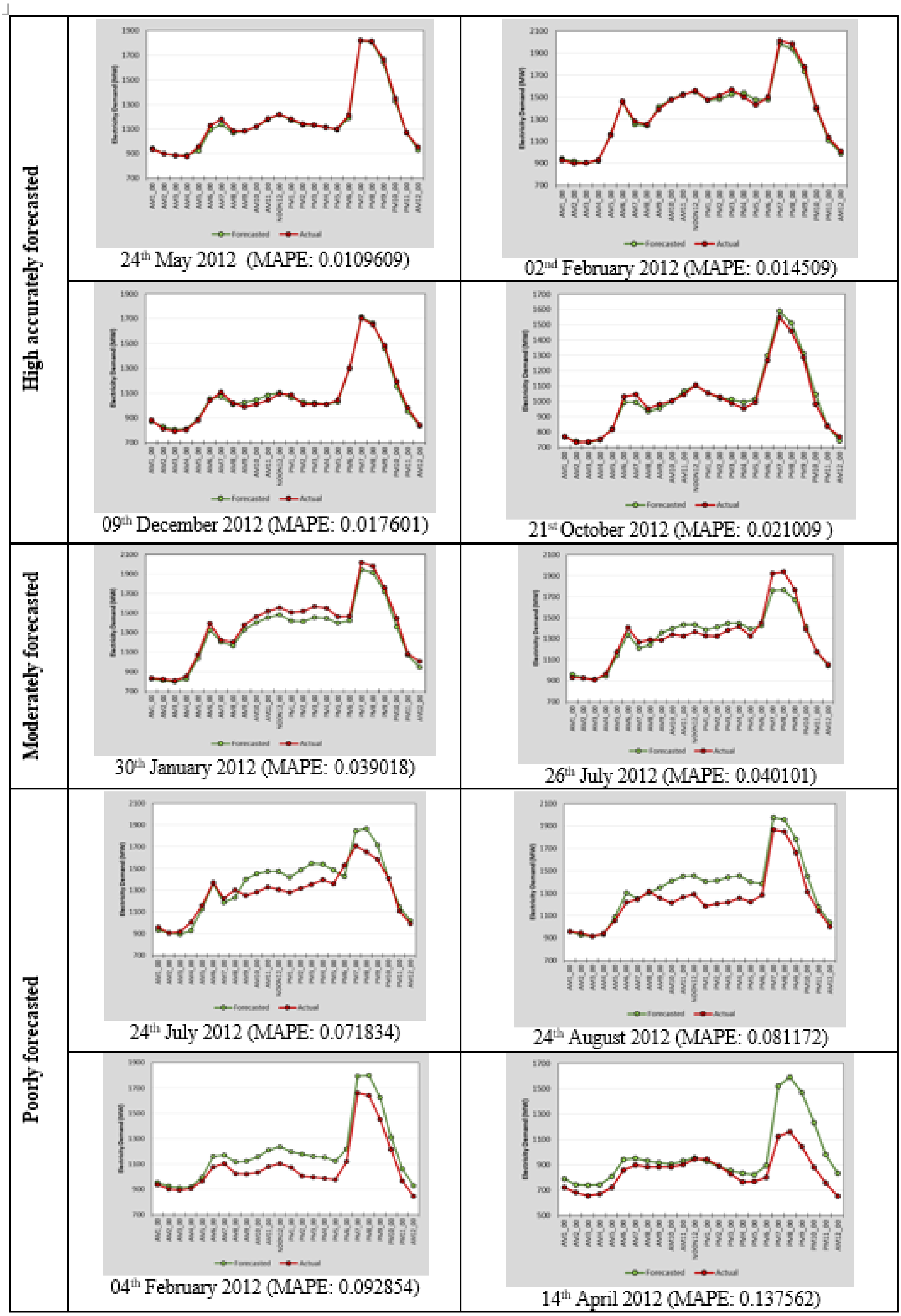

Figure 7. Some well, moderately and poorly forecasted days using the final iterative neural network

5. Synopsis

In this paper, an algorithm named "Artificial Neural Network Approach for Dynamic Iterative Forecasting" was formulated to iteratively forecast time series data using artificial neural networks in a dynamic way. As an illustration, hourly electricity demand in Sri Lanka were forecasted. It was identified that with the inclusion of clustering effects as inputs of the neural network, the training process became more efficient and accurate. It is important to emphasize that, a majority of holidays were well forecasted through this NN approach, except for a few specific holidays.

Acknowledgements

We acknowledge University of Colombo for the financial assistance provided for this research.

References

[1] Turner, L., & Witt, S. (2001). Forecasting Tourism Using Univariate and Multivariate Structural Time Series Models. Tourism Economics, 7(2), 135-147.

[2] Prasad, A., Chai, L., Singh, R., & Kafatos, M. (2006). Crop yield estimation model for Iowa using remote sensing and surface parameters. International Journal of Applied Earth Observation and Geoinformation, 8(1), 26-33.

[3] Snyder, R. (2002). Forecasting sales of slow and fast moving inventories. European Journal of Operational Research, 140(3), 684-699.

[4] Taylor, J. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. Journal of the Operational Research Society, 799-805.

[5] Guan, C., Luj, P., Michel, L., Wang, Y., & Friedland, P. (2013). Very Short-Term Load Forecasting: Wavelet Neural Networks with Data Pre-Filtering. Power Systems IEEE Transactions on, 28(1), 30-41.

[6] Xie, J., Cheng, C., Chau, K., & Pie, Y. (2006). A hybrid adaptive time-delay neural network model for multi-step-ahead prediction of sunspot activity. International Journal of Environment and Pollution, 28(3/4), 364-381.

[7] Jiang X., Adeli H. (2005) Dynamic Wavelet Neural Network Model for Traffic Flow Forecasting. Journal of Transportation Engineering, 131(10).

[8] An, N., & Anh, D. (2015). Comparison of Strategies for Multi-step-Ahead Prediction of Time Series Using Neural Network. International Conference on Advanced Computing and Applications. Ho Chi Minh City, Vietnam: IEEE.

[9] Boné, R., & Crucianu, M. (2009). Multi-step-ahead prediction with neural networks: A review.

[10] Zhang, G., Patuwo, B., & Hu, M. (1998). Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting, 35(62), 35-62.

[11] Park, D., El-Sharkawi, M., Marks II, R., Atlas, L., & Damborg, M. (1991). Electric Load Forecasting Using An Artificial Neural Network. IEEE Transadions on Power Systems, 6(2), 442-449.

[12] Othman, M. M., Harun, M. H., Salim, N. A., & Othman, M. L. (2015). Sequential Process of Feature Extraction Methods for Artificial Neural Network in Short Term Load Forecastng. ARPN Journal of Engineering and Applied Sciences, 10(19), 8830-8838.

[13] Mandal, P., Senjyu, T., Urasaki, N., & Funabashi, T. (2006). A neural network based several-hour-ahead electric load forecasting using similar days approach. Electrical Power and Energy Systems, 28, 367-373.

[14] Rana, M., Koprinska, I., & Troncoso, A. (2014). Forecasting Hourly Electricity Load Profile Using Neural Networks. Internationall Joint Conference on Neural Networks, (pp. 824-831). Beijing

[15] Mishra, D. K., Dwivedi, A., & Tripathi, S. (2012). Efficient Algorithm for Load Forecasting in Electric Power System using Artificial Neural Network. International Journal of Latest Research in Science and Technology, 1(3), 254-258.

[16] Deshani, K., Attygalle, M., Hansen, L., & Karunarathne, A. (2014, May). An Exploratory Analysis on Half-Hourly Electricity Load Patterns Leading to Higher Performances in Neural Network Predictions. Interanational Journal of Artificial Intelligence and Applications, 5(3), 37-51.

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

[18] Filik, U. B., Gerek, O. N., & Kurban, M. (2011). Hourly Forecasting of Long Term Electric Energy Demand using Novel Mathematical Models and Neural Networks. International Journal of Innovative Computing. Information and Control, 7(6), 3545-5357.

[19] Lee, K., Cha, Y., & Park, J. (1992). Short Term Load Forecasting using Artificial Neural Networks. Transactions on Power Systems, 124-132.

[20] Lu, C. N., & Wu, H. T. (1993). NEURAL NETWORK BASED SHORT TERM LOAD FORECASTING. IEEE Transactions on Power Systems, 336-342.

[21] Baliyana, A., Gauravb, K., & Mishra, S. K. (2015). A Review of Short Term Load Forecasting using Artificial Neural Network Models. Procedia Computer Science, (pp. 121-125). Odisha.

[22] McElroy, T. (2015). When are Direct Multi-step and Iterative Forecasts Identical? Journal of Forecasting, 34, 315-336

[23] Demuth, H., & Beale, M. (2002). Neural Network Toolbox User's Guide - Version 4. Natick: The MathWorks.