2014 ASIA-PACIFIC
SOLAR RESEARCH CONFERENCE

Aziz Ahmad

Hourly Power Consumption Prediction for Residential Houses Using Artificial Neural Network Models

Aziz Ahmad[1], Timothy Anderson[1]

[1]*School of Engineering, Auckland University of Technology, Auckland, 1010, New Zealand*

E-mail: timothy.anderson@aut.ac.nz

Abstract

In this study several Artificial Neural Network (ANN) models were experimented to predict electricity consumption for a residential house in New Zealand. The effect of number of users in the house, day of the week and weather variables on electricity consumption was analyzed. Each model has been constructed using different structures, learning algorithms and transfer functions in order to come up with the best model which has better generalizing ability. Further each model has been experimented with different number of neurons in the hidden layers and different number of delays in the tapped layers, and their effect on prediction accuracy was analyzed. Subsequently the most accurate ANN model was used to study the effects of weather predictor variables on the electricity consumption. Actual input and output data were used in the training, validation and testing process. A comparison among the developed neural network models was performed to find the most suitable model. Finally the selected ANN model has been used to predict 24 hours in advance electricity consumption for a residential house in New Zealand.

Keywords-Power Consumption; Levenberg-Marquardt; Neural Network; Load Prediction

1. Introduction

Recently the need for precise modeling and prediction of power consumption has increased due to increased interest in renewable energy systems and their implementation worldwide. One reason for this is that predicting electric power consumption 24 hours in advance would help to efficiently optimize energy distribution between loads, particularly in buildings, and the local grid. Power consumption prediction is essential for generators, wholesalers and retailers of electric energy, who buy and sell, switch loads, plan maintenance and unit commitment and much more. However, with increasing costs being passed to consumers, there is also a need for consumers to be able to predict their requirements with a view to better utilizing on-site generators such as photovoltaic panels and grid-tied storage systems, thus delivering intelligent buildings.

In (Flax, 1991) an intelligent building was defined as the one which maximizes the efficiency of the service and minimizing the use of grid power. The author lists intelligent building components, with the energy management system (EMS) ranked most highly. Such a system controls and monitors energy consumption of the building. However, for effective operation of an EMS, an accurate prediction of power consumption is needed. Such a system would ideally be able to plan well in advance and take actions to avoid power shortages, as well as shift loads to off peak time when electricity prices are lower.

In this regard, short term load prediction has attracted significant attention from scientists and engineers. Various mathematical techniques have been widely used for load prediction including regressive analysis, wavelet analysis, fuzzy system modeling, neural network modeling and evolutionary algorithms (Xia et al., 2010). There are also a large variety of models presented in the literature, both simple and hybrid, created by combining two or more approaches. (Maia and Goncalves, 2009), (Niu et al., 2009).

In recent years, much research has been carried out on the application of ANN techniques to the load forecasting problem. As such, expert systems have been tried out (Ho et al., 1990), (Rahman and Hazim, 1993), and compared to traditional methods (Moghram and Rahman, 1989). The advantage of using ANN as compared to the other models is the ability to extract the implicit non-linear relationships among the variables by means of ''learning'' with training data. Many interesting ANN applications have been reported in power system areas, due to their computational speed, their ability to handle complex non-linear functions, robustness and great efficiency, even in cases where full information for the studied problem is absent Thus the works of (Khotanzad et al., 1997), (Khotanzad et al., 1998) are good examples. Also it appears that the use of ANN for load forecasting has been well accepted in practice, and is used by many utilities (Khotanzad et al., 1998).

In this work, an ANN model has been used to predict electric load in a typical NZ residential house with four occupants, 24 hours into the future. It is suggested that a similar model could be adopted for other locations with varying number of occupants.

2. Methodology

In this study nonlinear autoregressive network with exogenous inputs (NARX) recurrent neural network based predictive models were developed to forecast future values of electricity consumption, based on the previous values of electricity consumption and eight input variables. The predictive model can be expressed mathematically by predicting future values of the electricity consumption time series y(t) from past values of that time series and past values of input variables time series x(t). This NARX model is based on the linear ARX model, which is commonly used in time-series modelling.

The equation for the NARX model is given by Equation 1.

$$y(t) = f\left(y(t-1), y(t-2), \dots, y\left(t-n_y\right), u(t-1), u(t-2), \dots, u(t-n_u)\right) \qquad (1)$$

where the next value of the dependent output signal y(t) is regressed on previous values of the output signal and previous values of an independent input signal. The NARX model is implemented using a feed-forward neural network to approximate the function f. A diagram of the resulting network is shown in Figure 1, where y(t) output series is predicted given past values of y(t) and another input series x(t).

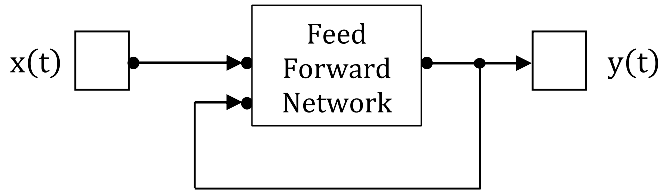


Figure 1. NARX block diagram.

There are different connection styles and learning algorithms in neural networks, the most common being the Back Propagation algorithm. The Back Propagation algorithm consists of two phases: a training phase and recall phase (Fatih et al, 2008). The training phase is important as it determines the success of the resulting network. In back propagation, there are two methods of updating the weights. In the first method, weights are updated for each of the input patterns using an iteration method. In the second method, used in this study, the mean value of input and output patterns of the training sets is calculated (Yousefizadeh and Zilouchian, 2001). As soon as the weight update values are obtained, the new weights and biases can be calculated using Equation 2.

$$W_{ij,n} = U_n + \alpha W_{ij,n} - 1 \qquad (2)$$

where $W_{ij,n}$ is a vector of current weights and biases, α is the momentum factor rate which determines how the past weights will reflect to the current value, and U_n is the update function which can be chosen according to the problem and data type.

According to Fatih et al, (2008) and Yousefizadeh and Zilouchian, (2001) the most commonly used equation solving algorithm is the Levenberg-Marquardt (LM) algorithm. It can be considered as an alternative to the conjugate methods for second derivative optimization. In the LM algorithm, the update function U_n can be calculated using Equation 3.

$$U_n = -[J^T \times J + \mu I]^{-1} \times J^T \times e \qquad (3)$$

Where J is the Jacobian matrix that contains the first derivatives of the network errors with respect to the weights and biases, and e is a vector of network errors. The parameter μ is a scalar number and I is the identity matrix. Depending on when the μ parameter is large, the update function U_n becomes identical to the basic back propagation (with a small step size). During processing the μ value decreases after each successful step and should be increased only when a tentative step increases the error term or performance function. Consequently the performance function is guaranteed to reduce or to become bounded at each iteration (Martin and Mohammad, 1994).

Now, the prediction accuracy of ANN models is dependent on the combination of input variables, training algorithm and ANN architecture configuration (Yadav and Chandel, 2014). One of the key tasks in time series prediction is the selection of the input variables. For the proposed non-linear ANN model there is no systematic approach that can be followed (Chernichow et al., 1996), so there is a risk of omitting important variables.

In this study two years of hourly data for eight input variables: Temperature ($_{Tair}$), Relative Humidity (RH), Air Pressure (P), Wind speed (W_s), Wind direction (W_d), Rain amount (R_a), Hour of Day (HD) and Day of Week (DW) along with two years hourly electricity consumption data from a residential house in west Auckland were used as inputs to the ANN. Input weather data were downloaded from the National Institute of Water and Atmospheres CliFlo database(2014) to train the ANN with electricity consumption as the target variable.

The data was presented in an unprocessed format, to study the effect of real input variables on target and predict output. Input and target data from 10 November 2011 to 10 November 2013 were used to train, validate and finally, test the network ability to predict 24 hours in advance electricity consumption for a residential house with four occupants.

As there were 256 possible combinations of eight weather predictor variables, testing the network with all combinations was not possible. Therefore for this study Moody et al., (1995), two-step sensitivity analysis technique was utilized to determine the most significant training variables. Once the most significant variables were determined, the network was trained with every selected variable, until the training error was minimized and the influence of each variable was removed by replacing it with its mean value or zero. To simplify the process, the twelve most significant combinations of the eight weather predictor variables were tested in order to investigate their effect on the electricity consumption prediction accuracy, as shown in Table 1.

Table 1. Models based on different combinations of input variables.

Model	Input variables	Model	Input variables
1	T_{air}, RH, P, W_s, W_d, R_a, HD, DW	7	T_{air}, RH, W_s, W_d, R_a, HD, DW
2	T_{air} RH, P, W_s, W_d, R_a, HD	8	T_{air}, P, W_s, W_d, R_a, HD, DW
3	T_{air}, RH, P, W_s, W_d, R_a, DW	9	RH, P, W_s, W_d, R_a, HD, DW
4	T_{air}, RH, P, W_s, W_d, HD, DW	10	T_{air}, RH, HD, DW
5	T_{air}, RH, P, W_s, R_a, HD, DW	11	T_{air}, RH, P, R_a
6	T_{air}, RH, P, W_d, R_a, HD, DW	12	T_{air}, RH, P, HD, DW

In order to determine the performance of developed ANN models quantitatively, and verify whether there was any underlying trend in performance of ANN models, the regression (R) (Pearsons correlation coefficient) and mean squared error (MSE) values were analyzed. The mean squared error (Equation 4) provides information on the short term performance and is a measure of the variation of predicated values around the measured data, where the lower the MSE, the more accurate the estimation.

$$\mathrm{MSE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}(\mathrm{I}_{\mathrm{p,i}} - \mathrm{I}_{\mathrm{i}})^2 \qquad (4)$$

Where $I_{p,i}$ is the predicted electricity consumption in kwh, I_i is the measured electricity consumption in kwh, and N denotes the number of observations.

3. Results and discussion

3.1. *Number of Hidden Neurons and Delays*

If the number of neurons or number of delays is increased, the network has a tendency to over-fit the data and also allows the network to solve more complicated problems, but on the other hand requires more computation. During experiments both the number of neurons in hidden layer and the number of delays in the tapped delay lines were varied until the network performed well in terms of the mean square error values. Therefore the effect of changing the number of neurons in the hidden layer, increasing and decreasing the number of delays was also investigated. Using tapped delay lines in the network is essential as it stores previous values of x(t) and y(t) sequences. The number of hidden neurons, network delays and time steps for training, validation and test were varied to determine which network exhibited the best performance. The number of neurons was changed between 10 and 250 and delays between 2 and 5 were tested in order to come up with the most suitable ANN prediction model. Taking Model 1 as an example, Table 2 shows the MSE and Regression values for various numbers of neurons in the hidden layer.

Table 2: MSE and Regression values for different number of neurons and delays.

Neurons	Delays	MSE	R	Time
10	2	0.0653	0.524	00:04
20	2	0.0672	0.509	00:06
30	2	0.0680	0.528	00:12
40	2	0.0692	0.476	00:13
50	2	0.0645	0.514	00:18
60	2	0.0695	0.472	00:23
60	3	0.0673	0.529	00:42
60	4	0.0676	0.489	01:05
60	5	0.0682	0.508	01:32
70	5	0.0634	0.517	01:55
70	2	0.0712	0.513	00:26
90	3	0.0731	0.479	01:04
90	2	0.0726	0.496	00:38
100	2	0.06787	0.515	00:50
120	2	0.0687	0.486	00:59
150	2	0.078	0.484	01:49
150	4	0.0754	0.455	08:58
200	2	0.0741	0.469	03:31
250	2	0.0758	0.458	04:36

Processing time was also observed and it was noted that time increased exponentially with increasing numbers of neurons or delays. After several trials, it was decided that the most suitable network, considering accuracy and processing time, had 50 hidden neurons and 2 delays in the tapped delay lines. Processing time was closely monitored because if the model were to be implemented on a hardware platform, processing power and memory would be limited compared to desktop resources.

3.2. Mean Squared Error Analysis

The MSE is the mean squared normalized error performance function which is the difference between the output and target values.

Network training can be stopped early by the validation vectors if the network performance on the validation vectors fails to improve or remains the same, as indicated by an increase in the mean square error of the validation samples. Test vectors are used as a further check that the network is generalizing well, but do not have any effect on training. The best validation performance for Model 4 is 0.0622 at epoch 4 with seven input variables as shown in Figure 2. It is shown that training, validation and testing errors decreased and merges with the dotted line on epoch 4 thus demonstrating the best validation performance.

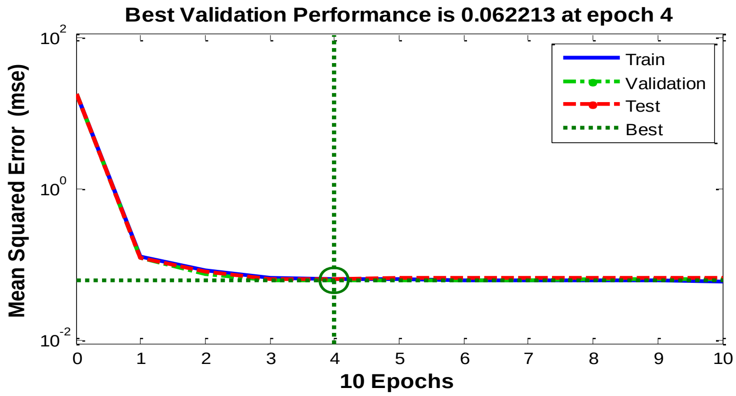


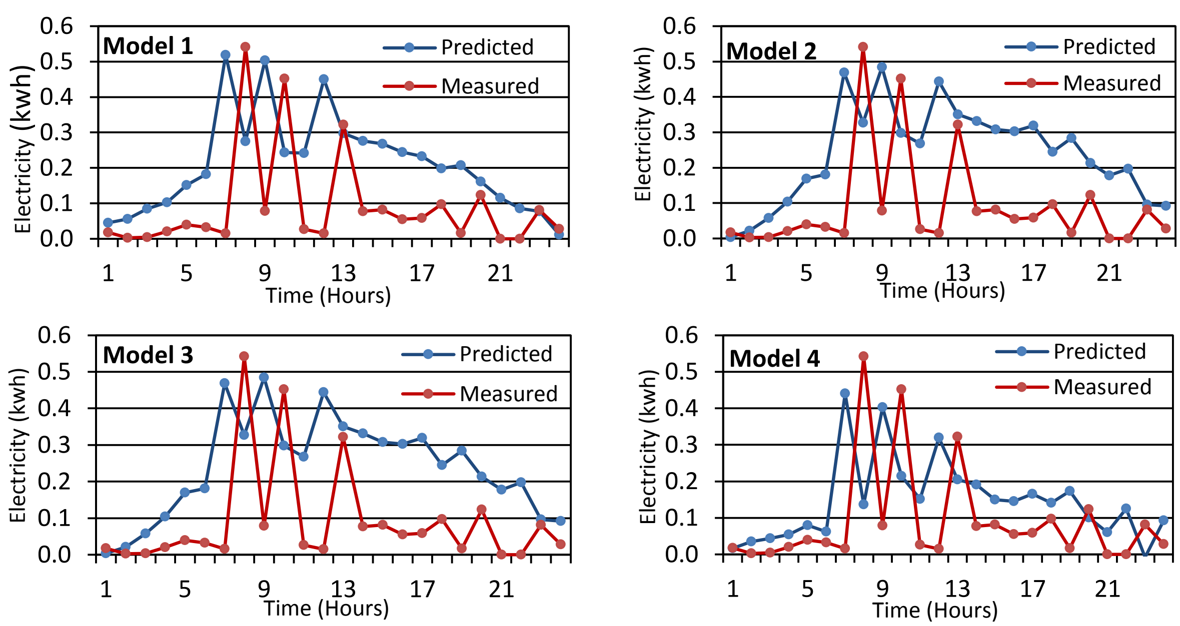
Figure 2. Mean Square Error (MSE) performance of the network.

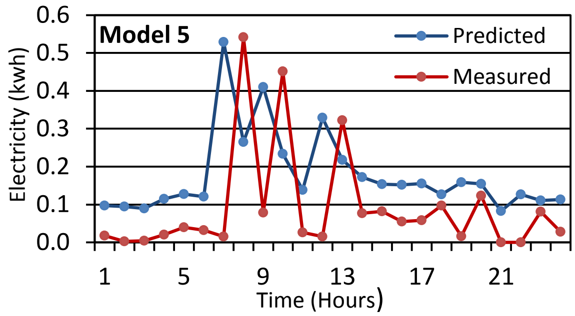
For the twelve models described in Table 1, the NARX network architecture with LM training algorithm was trained, validated and tested. Values of MSE and Regression were closely monitored to find the best model; Table 3 shows the MSE and Regression values for 50 neurons in the hidden layer with 2 delays in the tapped layers.

Table 3: MSE and Regression values for all 12 ANN models.

Model	MSE	Regression (R)	Model	MSE	Regression (R)
1	0.0642	0.514	7	0.0699	0.514
2	0.0795	0.504	8	0.0704	0.487
3	0.0693	0.470	9	0.0681	0.516
4	**0.0639**	**0.527**	10	0.0673	0.514
5	0.0666	0.508	11	0.0700	0.491
6	0.0683	0.519	12	0.0669	0.518

Figure 3 illustrates this point further, by showing for a single day, that for the first six models there is close correlation between the measured and ANN predicted values for electricity consumption in Auckland. However, in Table 3, it can be seen that Model 4 is the best among all 12 models with 0.0639 MSE and 0.527 Regression value.





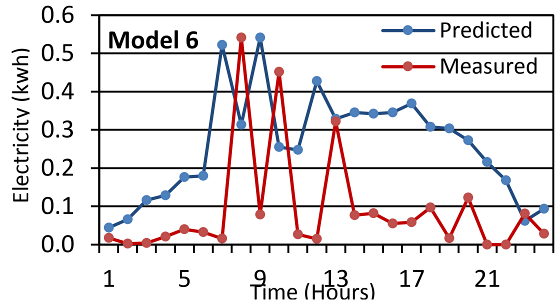


Figure 3. Measured and predicted electricity consumption values for the first 6 models.

3.3. Electricity consumption prediction for a Residential house in west Auckland

Having determined the most suitable configuration of ANN, Model 4 was used to predict electricity consumption of a house with four occupants (2 adults and 2 children) in west Auckland, New Zealand to test the model, as shown in Figure 4. In Figure 4 it can be seen that using real data to train the ANN gives predicted values of electricity consumption similar to those measured by the electricity meter. In this regard, it suggests that the ANN with the LM training algorithm offers a suitable predictive tool for electricity consumption in New Zealand. Moreover, it shows that training neural networks with real data can deliver satisfactory prediction of the output variable.

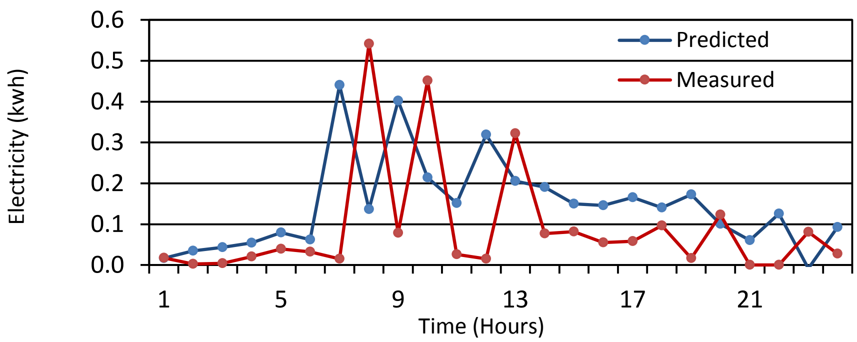


Figure 4. Measured and predicted values for a residential house in West Auckland.

4. Conclusion

This paper proposed a predictive model based on recurrent neural networks trained with the Levenberg-Marquardt back-propagation learning algorithm to forecast hourly electricity consumption using two years of historical electricity consumption and weather data. Twelve different combinations of eight weather predictor variables were used to train, validate and test twelve ANN models. Real-time input and target data were used without normalizing to study the real effects of input variables on outputs. Subsequently, one model, with the lowest Mean Square Error value, highest Regression value and lowest processing time was used to predict electricity consumption for a house with four occupants in west Auckland, New Zealand. Predicted values were compared with measured data from the house electricity meter which showed close correlation. Based on the experimental results including mean squared error analysis, error autocorrelation function analysis, regression analysis and time series response, the proposed ANN model illustrated the capability to predict electricity consumption values at a later time. These results further demonstrated the generalization capability of this approach and its ability to produce accurate estimates and forecasts for electricity consumption.

References

Chernichow, T., Piras, A., Imhof, K., Caire, P., Jaccard, Y., Dorizzi, B., Germond, A, 1996, 'Short term electric load forecasting with artificial neural networks', *Engineering Intelligent Systems for Electrical Engineering and Communications*, Volume 4, p85-99.

Fatih, O. H., Omer, N. G., Mehmet, K, 2008, 'Hourly solar radiation forecasting using optimal coefficient 2-D linear filters and feed-forward neural network', *Solar Energy*, Volume 82, p714-726.

Flax, B, 1991, 'Intelligent Buildings', *IEEE Communications Magazine*, Volume 29, p24-27.

He, Y., Yang, W., Zhang, Y., Li, D., F. Li, 2009, 'Urban electric load forecasting using combined cellular automata', *Journal of Computers*, Volume 4, p207-215.

Ho, L. K., Hsu, Y. Y., Chen, F. C., Lee, E. T., Liang, C. C., Lai, S. T., Chen, K. K, 1990, 'Short term load forecasting of Taiwan power system using a knowledge-based expert system', *IEEE Transactions on Power Systems*, Volume 5, pp. 1214-1221.

Khotanzad, A., Afhkhami-Rohani, R., Lu, T., Abaye, A., Davis, M., Maratukulam, D, 1997, 'ANNSTLF a NN-based load forecasting system', *IEEE Transactions on Neural Networks*, Volume 8, p835-845.

Khotanzad, A., Afhkhami-Rohani, R., Maratukulam, D, 1998, 'ANNSTLF artificial neural network short-term load forecaster generation three', *IEEE Transactions on Neural Networks*, Volume 13, p1413-1422.

Maia, C., Goncalves, M, 2009, 'A methodology for short-term electric load forecasting based on specialized recursive digital filters', *Computers & Industrial Engineering*, Volume 57, p724-731.

Martin, T. H., Mohammad, B. M, 1994, 'Training feed-forward networks with the Marquardt Algorithm', *IEEE Transactions on Neural Networks*, Volume 5, p989-993.

Moghram, I., Rahman, S, 1989, 'Analysis and evaluation of five short-term load forecasting techniques', *IEEE Transactions on Power Systems*, Volume 4, pp. 1484-1491.

Moody, J., Utans, J., Rehfuss, S., Siegelmann, H, 1995, 'Input variable selection for Neural Networks: application to predicting the US business cycle', *Computational Intelligence for Financial Engineering, Proceedings of the IEEE/IAFE 1995*, p118-122.

National Institute of Water and Atmospheric Research (http://cliflo.niwa.co.nz/) [cited 05/09/2014]

Niu, D., Gu, Z., Zhang, Y, 2009, 'An AFSA-TSGM based wavelet neural network for power load forecasting', *Lecture Notes in Computer Science*, Volume 5553, pp 1034-1043.

Rahman, S., Hazim, O, 1993, 'A generalized knowledge-based short-term load-forecasting technique', *IEEE Transactions on Power Systems*, Volume 8, pp. 508-514.

Xia, C., Wang, J., McMenemy, K, 2010, 'Short, medium and long term load forecasting model and virtual load forecaster based on radial basis function neural networks', *International Journal of Electrical Power & Energy Systems*, Volume 32, p743-750.

Yadav, A. K., Chandel, S. S, 2014, 'Solar radiation prediction using artificial neural network techniques: A review', *Renewable and Sustainable Energy Reviews*, Volume 33, p772-781.

Yousefizadeh, H., Zilouchian, A, 2001, 'Chapter 3: Neural network architectures', , *Intelligent control systems using soft computing methodologies*, Edited by A. Zilouchian, M Jamshidi, CRC Press.

Acknowledgements

The authors would like to thank Ian Sumner and Phuong Le at Energy Conscious Design Ltd for helping with collecting power consumption data.

This work was funded in part by a Callaghan Innovation Education Fellowship.