

Short Term Load Forecasting based on An Optimized Architecture of Hybrid Neural Network Model

Firas Shihab Ahmed

*Turkish Aeronautical Association University
Department of Information Technology
Ankara, Turkey*

Ministry of Electricity of Iraq, The Stat Company Electricity Production GEEP

Shadi Al Shehabi

*Turkish Aeronautical Association University
Department of Computer Engineering
Ankara, Turkey*

Abstract—Nowadays the predication of electricity demand is considered as a main criteria for estimating the electricity generation costs. Companies always sought to create balancing between the supply and demand of power. We are interested in this paper is short-term load forecasting (1 hour – 1 week). The important factors that are considered, in this paper, for the electricity supply are the temperatures and time. Then, different neural network models (NN) are used and developed for increasing the accuracy of the short term load forecasting such as backpropagation, Adaptive Neuro-Fuzzy Inference System (ANFIS), whereas Wavelets Neural Network (WNN) is used as a comparative model. An optimized neural network is found by defining the suitable number of hidden layers and their nodes and an optimized membership functions number associated to each input unit is proposed in ANFIS model. A comparison between these models shows that the optimised ANFIS model outperforms the other methods.

Keywords: Neural network (NN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Wavelets neural network (WNN)

I. INTRODUCTION

Power demand production is one the of most important factor for the management of electricity production, therefore, power demand is needed for capacity planning and maintenance arrangement [1], Long-term forecasts are required for electricity planning and maintenance scheduling in one year. Medium-term demand forecasts are required for energy system operation and planning, [2]. Short-term load forecasts are required for the control and configuration of energy systems. Indeed, accurate electricity demand prediction methods are important for reducing the cost of power production; also the distribution of electricity depends on the accurate forecasting of the demand modules. The time of load forecasting is divided into three types, short-term forecasts from one hour to one week, medium-term forecasts from a week to a year and long-term forecasts more than a year [3]. We are interested in this paper in short-term load forecasting which can help to estimate load flows and to make decisions that can prevent overloading. Short-term load forecasts depend on a table of historical information on hourly loads and temperature observations from ISO New England. One selected week hourly load is taken for training our models and hourly electricity consumption forecast for the next week is given by these models. We selected short-term load forecasting because a lot of companies of power production deepened on Short-term load forecasting to determine the capacity and the level of electricity supplying to meet the foreseeable load, also to determine the prices, five factors are considered: dry bulb temperature, wet bulb temperature, dew point temperature, daily hours and week days (work days and weekends). We propose to improve the short-term load forecasting performance of two neural network models such as back propagation neural network, Adaptive Neuro-Fuzzy Inference System (ANFIS), where three hidden layers with different number of neurons are used for the first model. The number of membership functions is determined by finding the relationship between the input variables. The powerful of the latest model is that it enables to make a right decision through the use of the fuzzy logic system. The third model is a hybrid model of wavelet transportation with the neural network, it is a two-stage prediction system which involves wavelet decomposition of input data at the first stage and the decomposed data with other input is trained using separate neural network to forecast the load. The forecasted load is obtained by reconstruction of the decomposed data. [4]

II. RELATED WORK

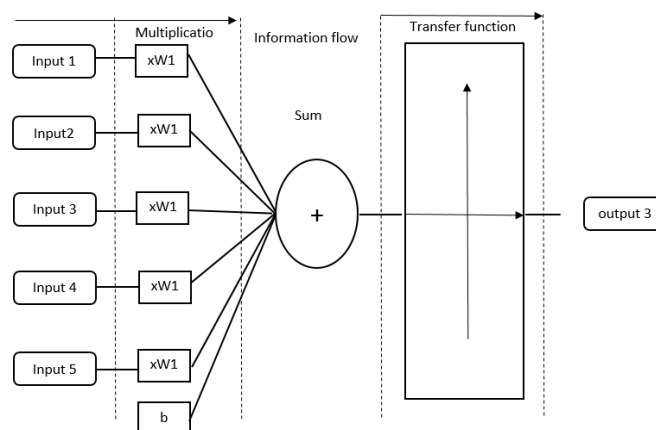
Actually, many researchers have published in the short-term load forecasting to predict the demand and the electricity price. Load forecasting is done by using some effective tools to determine the future demand that helps the companies to estimate the electricity production. In [5], different variants of Generalized Neural Network (GNN) trained with back-propagation and adaptive GAF were used to overcome the drawbacks of artificial neural network (ANN) and back propagation (BKP) training algorithm for improving the accuracy of short term load forecasting. A combined GNN model with wavelet transform and GNN-W models have been developed. In [6], the authors used multilayer perceptron (MLP) for one day ahead short term load forecasting. This study showed that MLP network has a good performance and reasonable forecasting accuracy, where they obtained a Mean Absolute Percentage Error (MAPE) between actual and predicted value equal to 1.066%. In [7], the authors added several factors that affect the behavior of the consumer load and also impact the total losses in transmission lines. The factors are: weather, economy and random disturbances. Different types of consumers are considered such as industrial, agricultural and domestic. It showed that domestic consumers give good statistical rules and are periodic in nature, but on the other hand industrial and agricultural loads are highly inductive and start up and shut down of such type of load induce huge spikes to the load curve. The spikes are disturbance of the startup and shot down because this load is quite random. It is impossible to predict the occurrence of these spikes so it will affect the short-term load forecasting. In [8], the author used Support Vector Machine (SVM) kernels for Short Term Load Forecasting on data. The obtained results of SVM are compared SVM with LDA and QDA. SVM kernel gave high performance of 99%. SVM can have the adaptability on different load situations such as working days and non-working days, also on some special events and different daily hours.

III. FORECASTING ALGORITHMS

Three methods are used in this paper, in order to find the short load forecasting for NEW England ISO, which are artificial neural network, Adaptive Neuro-Fuzzy Inference system (ANFIS) and wavelet transportation with a neural network (WNN) all those methods are used and evaluated.

3.1. Neural Network

Neural network (NN) is a mathematical model which is inspired from a biological neural network structure, the artificial neurons are basic building blocks of neural network in the mathematical model of neural network three steps (multiplication, summation and activation), where inputs of the neural network are weighted this means multiplication every input with weight, the second step sum function which sums all weights inputs, in third step summation and passing of previously weighted inputs by activation function (transfer function) as show in the figure (1) [9,10], .



Figur.1. Working principle of an artificial neuron [8]

The mathematical model of (*Neuron model*) is expressed as:

$$O_j = f_j \sum_k (w_{jk} x_k) \quad (1)$$

Where:

O_j : represents the output of a neuron

f_j : represents a transfer function

w_{jk} : represents an adjustable weight

x_k : represents the input of a neuron.

The activation function type (sigmoidal) which makes the output value in the range (0,1) [11]

3.1.1 Network Architecture:

there are three fully connected layers as a show in figure (2), those layers are an input layer, a hidden layer and an output layer, the signal comes from the input layer to the hidden layer after that multiplied weight with input value and sum those values than pass it to output layer by activation function

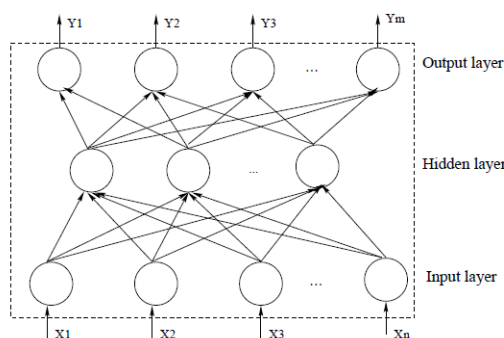


Figure.2. Schematic of the three-layer feed-forward ANN [11]

3.1.2 Training:

The aim of the training process is to get the desired output, through adjusting the network weights. The network is trained by using the Back-propagation algorithm, which consists of three steps [12]:

First step: is the forward step, where the output values are calculated for given inputs

Second step: is the propagation of the error of the outputs toward the input layer, with the fractional derivatives of the performance and Weights calculated in each layer.

Third step: Weight adjustment, where the weights are adjusted in order to reduce the error

- sum-squared error SSE is computed by using the following function:

$$E = \frac{1}{2} \sum_p \sum_j (t_{pj} - O_{pj})^2 \quad (2)$$

Where:

O_{pj} : represents an output unit

t_{pj} : represents a desired output

P : represents an actual output

j : represents an input pattern

- In the training process uses the following learning function for adjusting the weights:

$$dW = m_c * dW_{prev} + (1 - m_c) * l_r * gW \quad (3)$$

Where:

dW_{prev} : represents the previous weight

gW : represents a weight gradient

l_r : represents the learning rate,

m_c : represents the momentum

3.2.3 Optimization of Neural Network

The first aim of this study is to find the optimal neural network architecture in order to use it to load forecasting with less mean absolute percentage value (MAPE). For this reason, we use an experimental way to find the optimal network architecture. So, the design of our network consists of three layers: an input layer, three hidden layers and an output layer. The input layer includes five input units, which are: the temperature dry, the temperature dew, the hours, the week work and the weekend.

3.3 Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS is a hybrid system that combines the human reasoning style of fuzzy logic and the connectionist learning style of neural network [13]. The hybrid model uses the fuzzy logic through multi layer neural network because ANFIS enables to combine training data through neural network and to take the decision through fuzzy inference system like the human brain. ANFIS consists of five layers as shown in figure (3).

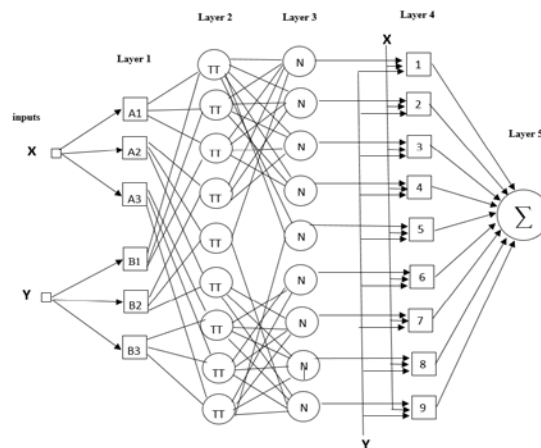


Fig.3. Structure of the ANFIS

First layer is the executing of the fuzzification process in order to convert the data into truth values making all the values between 0 and 1, where all the nodes in this layer are adaptive nodes. There is also a membership function which plays an important role in this paper. In our dataset, we have one week data in order to forecast the next week,

The node function:

$$Q_i^1 = \mu A_i(x) \quad , \quad i=1,2 \quad (4)$$

or

$$Q_i^1 = \mu B_{i-2}(y) \quad , \quad i=3,4 \quad (5)$$

Where:

Q_i^1 : represents an output
 x, y : represent the input values
 A_i, B_{i-2} : represent linguistic label associated with this node

In this layer, there are many different membership functions such as triangular-shaped- generalized bell functions A and B

generalized bell function:

$$\mu A(x) = \frac{1}{1 + \left| \frac{x - ci}{ai} \right|^{2b}} \quad (6)$$

Where ai, bi and ci are parameters

Second layer uses fuzzy rules (if/ and). In this step, we use the rules with all input values through (if/and) function. We combine the inputs which are ambiguous by fuzzy rules in order to find firing strength of a rule as output nodes. Every node in this layer is considered as a fixed node labeled W_i , the output is the multiplication of the incoming signal, here we modify these rules in order to reduce the error [14,15]

$$Q_i^2 = W_i = \mu A_i(x) \mu B_i(y), \quad i=1,2,3 \quad (7)$$

Third layer in the previous layer, the nodes are fixed and labelled, but here each node N is labelled in order to calculate the average of firing strength and sum of all these rules. The normalized firing strengths is the output of this layer:

$$Q_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2, 3 \quad (8)$$

Fourth layer in this layer we used the firing strength from the third layer with a set of parameters $\{ a_i, b_i, c_i \}$, where all the nodes are adaptive nodes with a node function, also in this layer the parameters represent consequent parameters \bar{w}_i represent the output of third layer .

$$Q_i^4 = \bar{w}_i z_i = w_i (a_i x + b_i y + c_i) \quad , \quad i = 1, 2 \quad (9)$$

Fifth layer In this layer we computed the final output as the summation of all incoming signals Σ .

3.3.1 Optimization of ANFIS

In this paper, we propose a way to determine an optimized model of ANFIS by determining first the suitable number of membership functions of the fuzzy logic system. This can be done by finding the relationship between the input units. Such that, we can find a relationship between the number of hours in a day and the temperature (Dry and Dew).

Secondly, determining the type of the membership function is important for improving the performance of ANFIS, therefore it could be done experimentally by testing each type of the membership function in the suitable number of it.

3.4 Hybrid model of Wavelet Transportation with Neural Network

Wavelet transforms (WT) is a useful technique which decomposes the time series signal in terms of both time and frequency. Wavelet transform of a function is the improved edition of Fourier transform. WT as the name suggests, uses some small wave like function to analyze a signal and hence called wavelets (mother wavelet

function). Mathematically speaking, wavelet transform is the convolution of wavelet function and the signal. The translated and scaling of mother wavelet $\Psi(\mathbf{t})$ can be represented as follows in (10) [16]:

$$\Psi(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad (10)$$

- a : dilation parameter
- b : Translation parameter
- Ψ : wavelet

The signal is decomposed into approximation coefficient and detail coefficient. Wavelet decomposed components can be assembled back into original signal without loss of information. Reconstruction can be done by combining all the decomposed wavelet [16].

IV. EXPERIMENTAL RESULTS

Our test dataset is a table of historical hourly loads and temperature observations from the New England ISO. This dataset includes recorded values of two weeks with five variables, such that, the data of the first week is used as training data of the above-mentioned models (NN, ANFIS and WNN). The input variables are:

1. Temperature Dry
2. Temperature Dew
3. Hours: represent hours of each day (1-24) in the first week
4. Week work: represents the day number in the week (1-7)
5. Weekend: represents the two-type day: (1) represents the workday, (0) represents the weekend

The output value: represents the electricity load during the first week

The data of the second week is used for testing the short-term load forecasting. The forecast accuracy is evaluated using the mean absolute percentage error (MAPE). It is defined as follows:

$$MAPE = \frac{100}{n} \sum_{i=1}^n \frac{|Actual\ load_i - Forecasted\ load_i|}{Actual\ load_i}$$

4.1 Results of an optimized NN

As we above-mentioned, a neural network with one input layer, three hidden layers and one output layer is first determined then four models of different number of neurons in the hidden layers are tested (see Table 1), the training is done by using back propagation algorithm.

After finishing the training of the neural network by providing the desired load and some elements which affect the electricity load such as temperature, 24 hours of day and weekend, the network is ready for forecasting. It is done by applying a new input signal without knowing its desired load value; the neural network can predict the load value. We applied neural network with different model in order to find optimal neural network for our data through using different number of hidden layers' nodes of neural network as shown in the table 1. After the tests, we found that the optimal network is found when the number of nodes of the first layers of the hidden layers is more than the second and the third layers. Such that we get the less MAPE value with less time (MAPE = 0.13961) and (time = 21 min 14 sec) as compared to the other models. The optimal model contains 10 nodes in first hidden layer, 8 nodes in second hidden layer and 8 nodes in third hidden layer, so we proved experimentally that it is important to increase the number of nodes in the first hidden layer for reducing MAPE of the neural network. Finally, we used a neural network of three hidden layers because this technique makes network training faster, and with less MAPE value.

Table (1) Forecasting evaluation of 4 different neural network models

Network	model 1	model2	model3	model 4
Input	5	5	5	5
Output	1	1	1	1
layers hidden	3	3	3	3
nodes in first layers	10	5	2	3
nodes second layers	8	10	15	6
nodes in third layers	8	10	15	10
learning rate	0.01	0.01	0.01	0.01
momentum confiscation	0.1	0.1	0.1	0.1
Transfer function	Hyperbolic tangent	Hyperbolic tangent	Hyperbolic tangent	Hyperbolic tangent
Training	200000	200000	200000	200000
target error	0.05	0.05	0.05	0.05
initialization threshold	random	random	random	random
Initialization weight	random	random	random	random
update interval cycles	500	500	500	500
MAPE	0.13961	0.1400	0.14012	0.1401
Time	21min 14 sec	20 min 41 sec	23 min 13 sec	15 min 51 sec

4.2 Results of an optimized ANFIS

As we previously mentioned, our aim is to determine an optimal number of membership functions of the fuzzy logic system through a neural network model. It is achieved into steps:

1) First, we suggest to find correlation between the daily hours and the temperature, such that the correlation value between Dew point and daily hours is $R = 0.9$, and the correlation value between Dry bulb and the daily hours is $R = 0.96$, where figure (3,4) shows an efficient correlation between them by taking the daily hours of one week.

2) Secondly, according to the figure (3,4) we note that the curve is a periodic curve, so we divide the daily hours into three periods, (Each period is 8 hours) in order to determine low, middle and high temperatures. Therefore three membership functions can be associated to three input variables which are temperatures Dew, temperature Dry and daily hours.

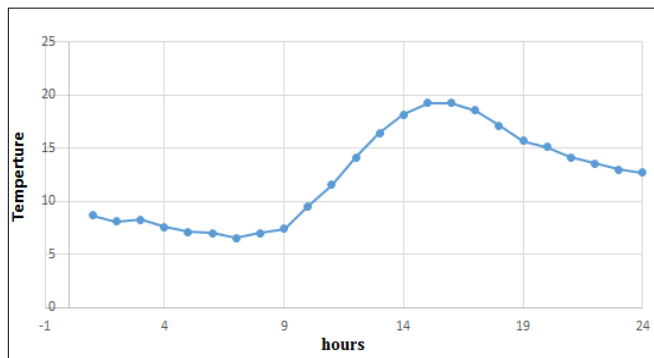


Figure .3. Correlation between the Dew point and the daily hours

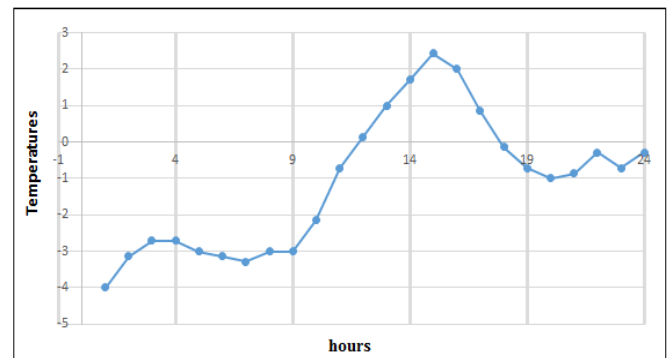


Figure.4. Correlation between the Dry bulb and the daily hours

Therefore, we found the optimal numbers of membership function associated to all inputs of the network are as follows:

- Three membership functions associated the first input (temperatures Dry)
- Three membership functions associated to the second input (temperatures Dew)
- Three membership functions associated to the third input (daily hours)
- Seven membership functions associated to the fourth input (week day) because it accepts only seven integer values
- Two membership functions associated to the fifth input (work day) because it accepts two integer values.

Three membership functions are associated to the first, the second and the third input, because they depend on three periods of day. Each period represents 8 hours, seven membership functions are associated to the fourth input depending on seven values of week days, two membership functions are associated to the fifth input depending on two values, such as the working weekday is (1) and weekend is (0). We can experimentally show that our way of determining the optimal number of membership functions associated to each input unit is better than the original method, ANFIS, by implementing ANFIS with an arbitrary number of membership functions associated to each input unit and by using eight different types of membership functions

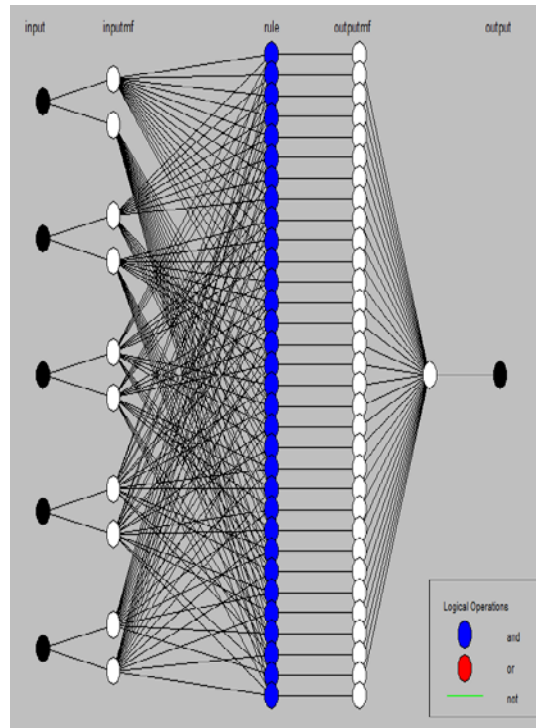


Fig. 5. The structure of the optimized ANFIS

The results in Table 2 show that the best membership function is (Pimf) with MAPE=0.0060, also we can see the training error of the Pimf membership function in the figure (6) and the Short load forecasting in the figure (7), for one week (168 hours)

Table (2) Comparison of Forecasting results when applying different numbers of different types of membership functions

Types of membership function	Number of membership functions			
	2,2,2,2,2	3,3,3,3,3	4,2,2,7,2	3,3,3,7,2
	MAPE	MAPE	MAPE	MAPE
trimf	0.1549	0.0216	0.1407	0.0167
trapmf	0.0114	0.0101	0.0152	0.0270
gbellmf	0.0137	0.0101	0.0383	0.0323
gaussmf	0.0936	0.0131	0.0129	0.0180
gauss2mf	0.0140	0.0233	0.0104	0.0133
pimf	0.0110	0.0072	0.0154	0.0060
dsigmf	0.0209	0.0247	0.0099	0.0072
psigmf	0.0209	0.0246	0.0098	0.0072
Types of membership function	Number of membership functions			
	2,3,2,4,2	3,2,2,4,2	2,3,4,5,2	2,2,3,2,2
	MAPE	MAPE	MAPE	MAPE
trimf	0.0100	0.0104	0.0095	0.0193
trapmf	0.0106	0.0163	0.0100	0.0112
gbellmf	0.0417	0.0372	0.0404	0.0112
gaussmf	0.0185	0.0190	0.0261	0.0128
gauss2mf	0.0113	0.0213	0.0108	0.0104
pimf	0.0105	0.0166	0.0096	0.0079
dsigmf	0.0127	0.0217	0.0107	0.0140
psigmf	0.0125	0.0210	0.0087	0.0140

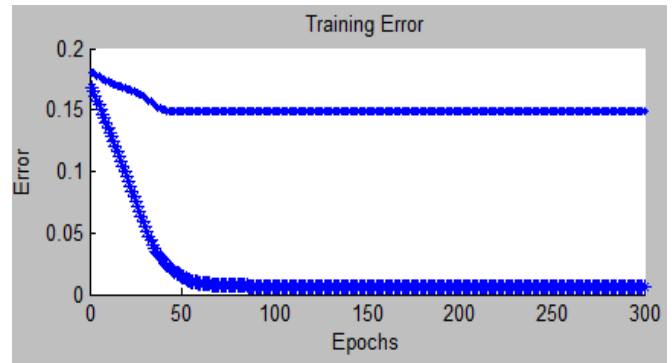
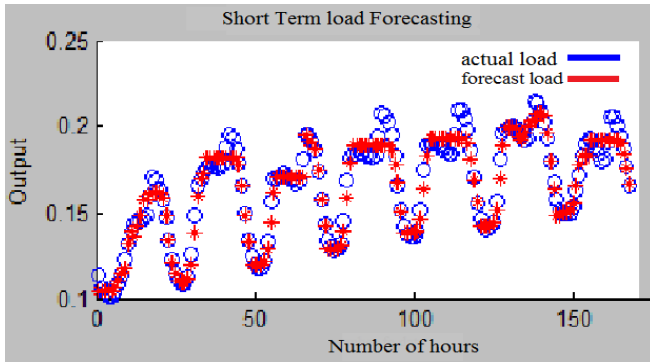


Fig. 6. Training error of the optimized model

Fig. 7. Short term load forecasting for the optimized model

4.3 Results of WNN

Three steps in WNN are used for short-term load forecasting:

First step: the five input variables decompose via wavelet into an approximation coefficient and a detail coefficient, then they are separated each input value to High frequency and low frequency as shown in the figure (8).

Second step: the output of each wavelet decomposition is divided into two levels in the neural network, such as high frequency (NNH) and low frequency (NNL).

Third step: in this step, the neural network is able to reconstruct high-frequency components and low-frequency components through combining those components together in order to get the load forecasting. The result shows that MAPE of WNN is 0.284

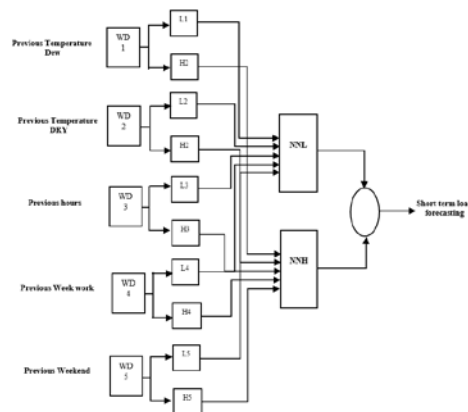


Figure.8. Structure of wavelet neural network

4.4 Comparing the performance of the forecasting methods

The mean absolute percentage error MAPE between the ‘forecast’ and ‘target’ loads presented in Table 1 for the three models, NN, optimized ANFIS and WNN is used to compare their performance. It shows that our optimized ANFIS model outperforms the other methods as well as it outperforms IT2 TSK FLS (A2-C1) model presented in [17], such that MAPE of the optimized ANFIS model is 0.006.

Table (3) Comparison of Forecasting results for several models

	Load forecasting methods		
	Optimized ANFIS	Neural Network	WNN
MAPE	0.0060	0.139	0.284

V. CONCLUSION

In this paper we have proposed first an optimal architecture of neural network model by finding a suitable number of hidden layers and a suitable number of their nodes, then an optimised model of ANFIS was proposed in order to get an accurate electricity load forecasting. The optimised ANFIS model is found by determining the suitable number of membership functions associated to each input node and by finding the suitable membership function type. The number of membership function is determined by finding the relationship between temperatures and the daily hours. The powerful of ANFIS comes from using fuzzy logic system that helps the neural network model to take the decisions. Wavelet transform with neural network model is used as a hybrid model, the powerful of this model comes from making the hidden layers of the neural network more active by using the mother function. A comparison between the three models showed that the error of short term load forecasting was 0.0060, while the error of short-term load forecasting of the neural network was 0.056, and the last model wavelet transportation with a neural network the error was 0.284. The results showed that our optimised ANFIS model outperforms the other models.

REFERENCES

- [1] J. W. Taylor and P. E. McSharry, Senior Member, "Short-Term Load Forecasting Methods: An Evaluation Based on European Data ", IEEE Transactions on Power Systems, 22, 2213-2219, 2008.
- [2] S. M. Barakati, A. A. Gharaveisi and S. M. Reza Rafiei, "Short-term load forecasting using mixed lazy learning method", Turkish Journal of Electrical Engineering & Computer Sciences, Turk J Elec. Eng. & Comp Sci (2015) 23: 201-211.
- [3] O. T. Altinoz, E. Mengusoglu, "Cloud-based Long Term Electricity Demand Forecasting using Artificial Neuro-Fuzzy and Neural Networks", IEEE ELECO 2015 - November 2015.
- [4] S. Gupta, V. Singh, A. P. Mittal and A. Rani, "A Hybrid Model of Wavelet and Neural Network for Short Term Load Forecasting", International Journal of Electronic and Electrical Engineering. ISSN 0974-2174, Volume 7, Number 4 (2014), pp. 387-394.
- [5] D. K. Chaturvedi, A. P. Singh and O. P. Malik, "Short term load forecast using fuzzy logic and wavelet transform integrated generalized neural network", International Journal of Electrical Power & Energy Systems. May 2015.
- [6] S. K. Sheikh1 and M. G. Unde, "SHORT-TERM LOAD FORECASTING USING ANN TECHNIQUE", International Journal of Engineering Sciences & Emerging Technologies, Feb 2012. ISSN: 2231 – 6604 Vol 1, Issue 2, pp: 97-107.
- [7] M. U. Fahad and N. Arbab, "Factor Affecting Short Term Load Forecasting", Journal of Clean Energy Technologies, Vol. 2, No. 4, October 2014.
- [8] L. Hussain, M. S. Nadeem and S. A. Ali Shah, " SHORT TERM LOAD FORECASTING SYSTEM BASED, " International Journal of Computer Science & Information Technology (IJCSIT) Vol 6, No 3, June 2014.
- [9] A. Krenker, J. Bešter and A. Kos, "Introduction to the Artificial Neural Networks", İçinde: Suzuki, K. (Ed). Artificial Neural Networks-Methodological Advances and Biomedical Applications. InTech, Croatia, 2011.
- [10] Mrs. J. P. Rothe, Dr. A. K. Wadhvani and Dr. Mrs. S. Wadhvani, "Short Term Load Forecasting Using Multi Parameter Regression", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 6, No. 2, 2009.
- [11] G.J. Tsekouras, F.D. Kanellos, and N. Mastorakis, "Short Term Load Forecasting in Electric Power Systems with Artificial Neural Networks", Lecture Notes in Electrical Engineering pp 19-58, Volume 343 2015.
- [12] H. Chen, C. A. Cañizares and A. Singh, "ANN-based Short-Term Load Forecasting in Electricity Markets", Power Engineering Society Winter Meeting, 2001. IEEE, 28 Jan. – 1 Feb. 2001.
- [13] M. Mordjau, B. Boudjema, M. Bouabaz, R. Daira, "Short Term Electric Load Forecasting Using Neuro-fuzzy Modeling for Nonlinear System Identification", LRPCI Laboratory, University 20 August, Skikda, 2011.
- [14] J. P. Rothe, A.K. Wadhvani and S. Wadhvani, "Artificial Neural Network and ANFIS Based Short Term Load Forecasting in Real Time Electrical Load Environment", International Journal of Current Engineering and Technology, Accepted 27 May 2014, Available online 01 June 2014, Vol.4, No.3 (June 2014).
- [15] K. Yang, L. Zhao, "Load Forecasting Model Based on Amendment of Mamdani Fuzzy System", IEEE pp:24-26, 2009.
- [16] S. Gupta, V. Singh, A. P. Mittal, Asha Rani, "A Hybrid Model of Wavelet and Neural Network for Short Term Load Forecasting", International Journal of Electronic and Electrical Engineering. ISSN 0974-2174, Volume 7, Number 4 (2014).