

# Short Term Load Forecasting for Erbil Distribution System Using ANN and Wavelet Transform

A. A. Rasool, A. A. Fttah, I.B.Sadik

**Abstract-** In this paper an approach is proposed for Short Term Load Forecasting (STLF) which combines Wavelet Transform (WT) and Artificial Neural Network (ANN). It is well known that the electrical load at any time can be considered as a linear combination of different frequencies. The daily load curves are decomposed into approximation part associated with low frequency and some detail parts associated with high frequency by means of (WT). Feed Forward Neural Networks are trained by low frequencies and corresponding average temperature or maximum and minimum temperature to predict the approximation part for the next seven days.

The short term load is forecasted by summing the predicted approximation part and the mean of the detail parts of the last three same type days.

This approach had been tested by load of Erbil, the capital of Iraq Kurdistan Region. The result have showed an encouraging accuracy, given that the available data is inaccurate.

**Key words**—Artificial Neural Networks, Short Term Load forecasting, Wavelet Transform.

## I. INTRODUCTION

Short Term Load Forecasting (STLF) plays an essential role in power system industry. Economical reliability of a power system depends significantly on the accuracy of forecasted load. With an accurate load forecasting system, operators in the control center will be able to make a design of operation of generators, energy transfer schedule of fuel and water supply, Energy Management System, Optimization of Power System Operation [1]. Moreover, load forecasting may also be helpful to the maintenance of generators [2].

To minimize the operation cost electric utilities use forecasted load to control the number of running generator units [3]. Forecast error in load prediction results in failure in providing the necessary reserves which leads to higher cost due to the use of expensive peaking unit. Over prediction of load, on the other hand, results in an unnecessary increase of reserves and hence increasing operation cost [4].

Two basic load models have been used for short term load forecasting: the peak load model and the curve model. The peak load models are not widely used since they do not give any information about the shape of the load curve. Various techniques for power system load forecasting have been employed in the past few decades.

Generally, the tow main approaches used for load forecasting previously are the time series and

regression approaches. The time series approach treats the load curve as a time series signal with known daily, weekly and seasonal periodicities and predict the future load by using various time series techniques such as Kalman filtering, the box-Jeenkins method, the Auto-Regressive moving-average (ARMA) model and the spectral expansion. These methods have an inaccurate result when there is an abrupt change in the environmental variables. The problem with the time series approach is that it uses a large number of equations which need a large computation time and possible numerical instabilities. The regression approach recognizes that the load curve is dependent on weather variables such as temperature, humidity, wind speed, cloud or rain and finds a functional relationship between the weather variables and the system demand. However, the functional relationship between the weather variables and load is not linear as this conventional regression approach has assumed. Hence, the load forecasting results produced by this approach is just an averaged result [5].

From the signal analysis point of view, the load can be considered as a linear combination of different frequencies. Every component of load corresponds to a range of frequency. It is apparent that some of the features of load are transient or temporary in nature. Although the Fourier transform is the most essential tool for stationary signal analyses and has been widely used in load forecasting as well, it is not so effective in capturing short duration transients.

In contrast to the Fourier Transform, the wavelet transform is especially suitable for transient analysis because of its time-frequency characteristics with adjusted window lengths. This enables the wavelet transform to provide better time resolution for high frequency components as well as better frequency resolution for low frequency components of a signal. Then the wavelet transform can be used as an effective tool for capturing important features and characteristics of the Load Curve. On the other hand, Artificial Neural Networks have been proved to be very powerful in finding internal representations of interdependencies within raw data that were not explicitly given.

## II. WAVELET TRANSFORM AND ANN

### A. Wavelet Transform

Wavelet Transform is a relatively new mathematical tool that has been the target of rich exploitation in many fields

such as image compression, signal processing and mechanical vibrations. Its rich temporal content and frequency isolation features has contributed to the successful results obtained by the application of Wavelet Transforms in the area of power system analysis [6], [7].

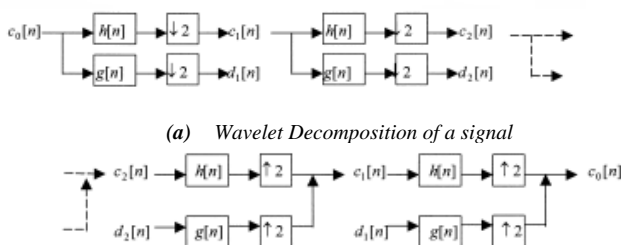
Several works have been developed in many areas with the aim of this tool, specially, in the last decade have been met the potential benefits of applying Wavelet Transform in the power systems.

The basic concept in Wavelet analysis is to select a proper wavelet (Mother Wavelet), then perform analysis using its translated and dilated versions, there are many kinds of wavelets which can be used as mother wavelet.

Similar to the Fourier transform, there are different wavelet transforms, called continuous wavelet transform (CWT) and discrete Wavelet Transforms (DWT) and fast discrete wavelet transform (FWT). The fast wavelet transform is known as multi-resolution analysis (MRA) or Mallat pyramidal algorithm. For a given square integrable (or finite energy) function (or signal)  $f(t)$ , its continuous wavelet transform is defined by equation (1).

$$(Wf)(a,b) = a^{-1/2} \int_{-\infty}^{\infty} f(t) \Psi\left(\frac{t-b}{a}\right) dt \quad \text{----- (1)}$$

Where  $\Psi(t)$  is the mother wavelet. The value of the wavelet transform  $(Wf)(a,b)$  is called the wavelet coefficient, which stands for the similar degree between the signal and the wavelet at the translation  $b$  and the dilation  $a$ . Translation means the time shift, and dilation means the time scale. In other words, it means how many components of the wavelet at dilation  $a$  are included in the original signal at translation  $b$ . For computer implementation, the discrete wavelet transform (DWT) is the most commonly used. There is a fast algorithm for the DWT, known as the fast DWT or Mallat and Daubechies' pyramidal algorithm. First, an original discrete signal  $C_0[n]$  is decomposed into two components,  $C_1[n]$  and  $d_1[n]$  by a low-pass filter  $h[n]$  and a high-pass filter  $g[n]$ , respectively. The transform is an orthogonal decomposition of the signal. The  $C_1[n]$ , named the approximation of the signal, contains the low frequency components of the signal  $C_0[n]$ , and the  $d_1[n]$ , named the detail of the signal, is associated with the high frequency components of the  $C_0[n]$ . Then, the approximation  $C_1[n]$  is again decomposed into a new approximation,  $C_2[n]$  and a detail  $d_2[n]$  by a bigger scale and continuing to a third scale, fourth scale and so on, according to the application. The original signal,  $C_0[n]$ , can also be reconstructed by these approximations and details. Fig.1 (a) and (b) illustrates the decomposing and reconstructing processes, respectively [8].



(b) Wavelet Decomposition of a signal

Fig.1 Wavelet Decomposition and Reconstruction

In the Fourier Transforms there is no way of judging whether the value of  $f(w)$  at a particular  $w$  is derived from a frequency present throughout the life of  $f(t)$  or during just one or a few selected periods.

The disadvantage can be overcome by the wavelet transform. For the windowed Fourier transform, the window width is always fixed, and the window is square shape. An important advantage of the wavelet transform is that the window widths of the wavelet transform can be adjusted. At low frequency, the window widths are longer, while at high frequency, the window widths are shorter. The shorter the window width, the better is the resolution. This means that the wavelet transform can provide better time resolution for high frequency components of a signal and better frequency resolution for low frequency components of the signal. Meanwhile, because of the shorter time window, the phase shift is hardly observed. [8], [9]. This is why Wavelet Transform is selected in this work.

B. Feed Forward Neural Network

Neural Networks have been used in a broad range of applications including: Pattern Recognition, Optimization, Prediction and automatic control. The applications of Neural Network in different power system operation and control strategies has lead to acceptable results. These applications are such as Load Forecasting, Fault Diagnosis/ Fault Location, Economic Dispatch, Security Assessment and Transient Stability [10].

The simplest form of NN that is used to classify a particular type is known as the perceptron. The patterns classified in this form of NN are linearly separable [7]. MLP is a collection of simple processing units highly interconnected which process the information fed to the input units. These units are organized by layers. Typically, the neurons in the input layer serve only for transforming the input pattern to the rest of the network, without any processing. The information is processed by the hidden and output neurons. Fig. 2 shows the internal structure of a neuron. Each neuron has a certain number of inputs, but only one output. For a certain neuron the inputs could originate from external stimuli or could present the output of the other neurons. Signals travel from one neuron to other s through special links or connections. The importance of a connection between neurons  $i$  and  $j$  described by its weight  $W_{ij}$  [11].

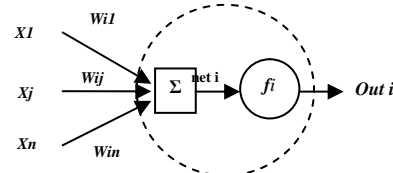


Fig. 2 The internal Structure of a Neuron

III. THE PROPOSED APPROACH

Firstly the hourly load data for the last three weeks were decomposed using wavelet transform into approximation

part associated with low frequencies and several detail parts associated with high frequencies. The three week period has been chosen empirically by using different periods and finding that this period is the optimal training period comparing with other periods.

Then the Neural Networks were trained by the load data for the last three weeks to predict the approximation part of the next seven days. Then the short term load is forecasted by summing the approximation part with the mean of the details of the last three same type days. Fig.3 illustrates the process.

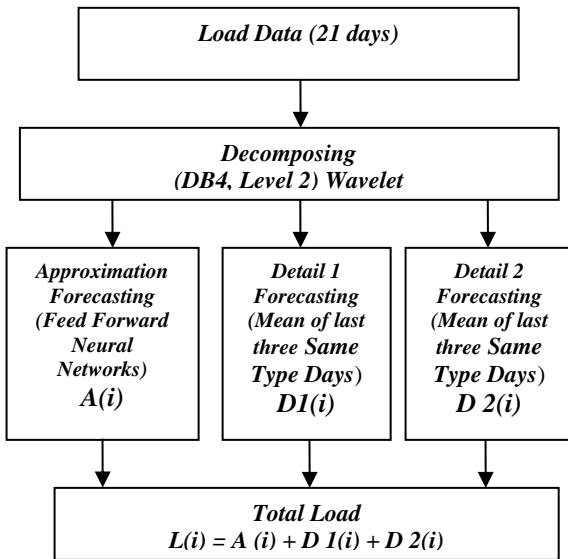


Fig. 3 Flow Chart of Load Forecast

A. *Decomposing the Load data*

The load data for the last three weeks have been decomposed into approximation and detail parts, the decomposing level is selected by decomposing the load curve into different levels such as level 1, level 2, level 3. Fig. 4 shows different level decomposition.

The smoother approximation and the decomposition its approximation part is close to the original curve should be selected [8], [12]. It is clear that level 2 is the closest one to the original curve and smoother than level 1 decomposition. Even though level 3 is smoother than level 2, but it can not be selected as its shape is different from the original curve

which means losing a lot of information in the curve, so level 2 is selected. Only the details of level 2 are showed in Fig. 5, because there are used for predicting the detail part of the load curve.

Because of the fact that the main information is in the approximation part associated with low frequencies which represents the trend of the normal load. So the approximation part is used for training the neural network which is used to predict the approximation part for the next seven days.

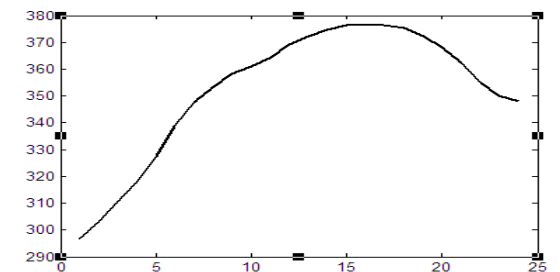
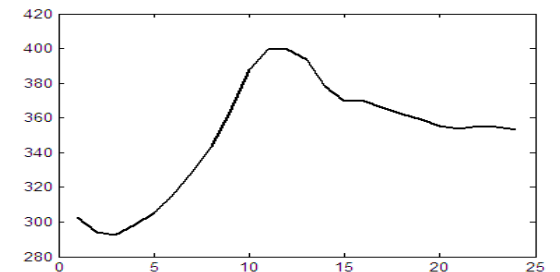
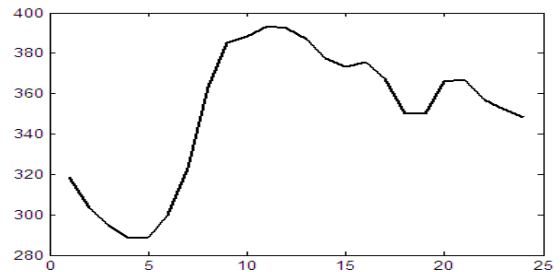
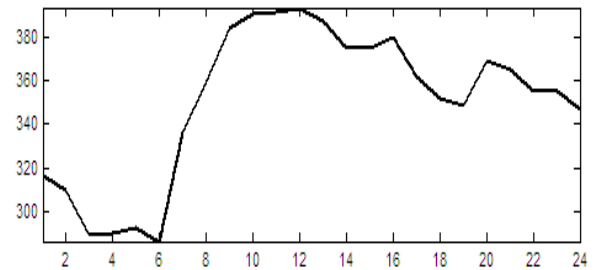


Fig. 4 DB 4 Approximation

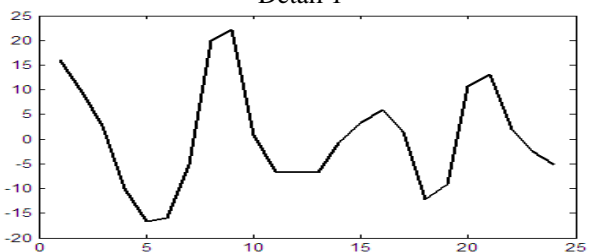
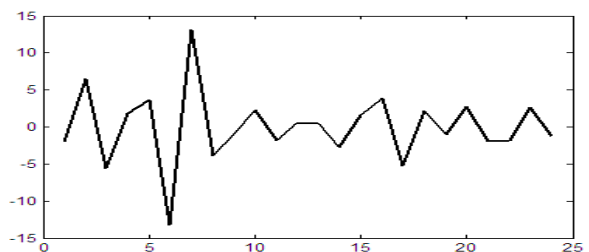


Fig. 5 Details of the level 2 Decomposition

**B. Predicting the Approximation**

The Feed Forward Neural Network is used to predict the approximation part, the steps followed are:

- 1) The approximation 2 (app (k,i)) for each hour of the last 21 days are normalized between (-1 & 1) using the equation 2 we denote the normalized values as App(k,i):

$$App(k, i) = \frac{2 \times [app(k, i) - \min app(k, i)]}{[\max app(k, i) - \min app(k, i)]} - 1 \quad \text{----- (2)}$$

- 2) The mean of the normalized approximation is computed for the last three same type days, for example, to predict the load curve for the next Friday, we compute the mean of the normalized approximation for the last three Fridays according the equation (3).

$$App(I) = \sum_{k=1}^3 \frac{App(k, i)}{3} \quad \text{----- (3)}$$

Where, k represents the day and i represents the hour.

- 3) The maximum and minimum of the predicted day were determined by two neural networks for minimum and maximum approximation respectively. Two models were tested, the first one uses average temperature and the other one uses the maximum and minimum temperatures Table 1 lists the input nodes and output nodes of the maximum and minimum model.

a. Maximum Model		b. Minimum model	
Input Nodes	Output nodes	Input Nodes	Output nodes
W	App <sub>max</sub>	W	App <sub>min</sub>
D		D	
T <sub>average</sub> or T <sub>max</sub>		T <sub>average</sub> or T <sub>min</sub>	
T(d-1) <sub>average</sub> or T(d-1) <sub>max</sub>		T(d-1) <sub>average</sub> or T(d-1) <sub>min</sub>	

Table 1 List of Input nodes and Output Node for The Feed Forward Neural Networks

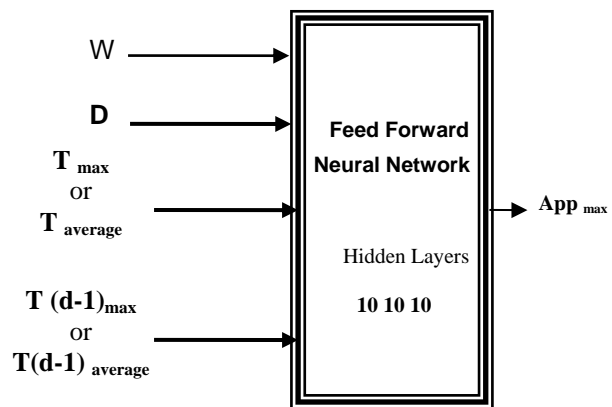
Where:

W Day Index in a week (1 – 7), for Sunday to

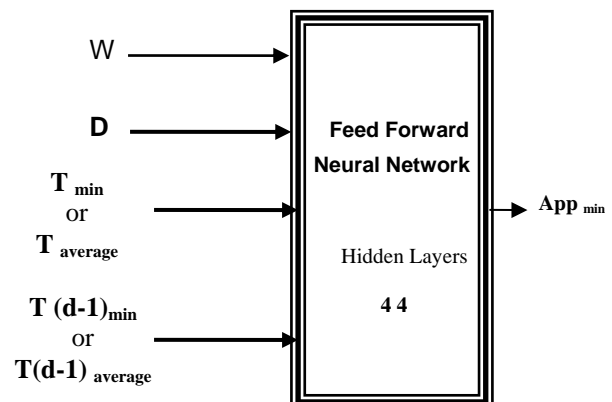
Saturday respectively.

- D Type of day Index (1 for weekend & 0 for weekdays).
- T<sub>average</sub> The average temperature of the predicted day.
- T(d-1)<sub>average</sub> The average temperature of one day before.
- T<sub>max</sub> Maximum Temperature of the Predicted day.
- T<sub>min</sub> Minimum Temperature of the Predicted day.
- T(d-1)<sub>max</sub> Maximum Temperature of one day before.
- T(d-1)<sub>min</sub> Minimum Temperature of one day before
- App<sub>max</sub> Maximum Approximation of the predicted Day.
- App<sub>min</sub> Minimum Approximation of the predicted Day.

Fig.6 shows the two Neural Network architectures used. The number of hidden layers and the neurons in each layer in the neural networks were given empirically so it was found that for max model the best one was three hidden layers, each layer contains 10 neurons and for minimum model, the best design was two hidden layers, each layer contains 4 neurons, Matlab 7.0 program is used for the simulation.



(a) Maximum Model



(b) Minimum Model

Fig. 6 Neural Networks Architectures

- 4) Now we have the normalized hourly approximation for the predicted day App(i), which are the mean of the normalized hourly approximation for the last three same type day and we have also the maximum and minimum approximation of the predicted day which are the outputs of the two Neural Networks, so we can predict the actual value of the hourly approximation of the predicted day according to equation (4).

$$A(i) = 0.5[App(i) + 1] \times [App_{max} - App_{min}] + App_{min} \quad \text{-----(4)}$$

A(i) is the approximation of ith time of the predicted day

C. Predicting the Detail

The detail of the predicted day are forecasted, the mean of the details of the last three same type days represents the details of the predicted day, as we have chosen level 2 for decomposing the load curve we have to predict two details

$$D(i) = \sum_{k=1}^3 \frac{D(k, i)}{3} \quad \text{-----(5)}$$

D (i) is the detail of the ith time of the predicted day

D. Summing the approximation and the details

The summation is the total load that should be predicted.

$$L(i) = App(i) + D(i) \quad \text{-----(6)}$$

L(i) is the load of the ith time of the predicted day.

IV. RESULTS

The proposed approach was tested on the 2006 Load data for Erbil Governorate, the capital of the Iraq Kurdistan Region, the data from (1 – 21) August 2006 were used for training the Artificial Neural Networks to predict the maximum and minimum approximation for the predicted days and the period from (25 – 31) August 2006 for testing, Fig. 7, the results have showed a reasonable Percentage error equation (7), it was found that using the average temperature gives better results. Table 2 shows the percentage error for the seven forecasted days hourly. Due to the limited scope of the paper only the detailed results of using average temperature have been showed The Mean Absolute Error MAPE% for this period was found according to equation (8) which was 2.99%, while it is 4.56% by using the maximum and minimum temperature Table 3 shows the MAPE% for each day using the average temperature and maximum and minimum temperature. The results are reasonable, given that the data in hand is bad and inaccurate due to the deficiency in the electrical generations as a result

of bad economic and security condition,

$$ERROR \% = \frac{|X_i - Y_i|}{X_i} \times 100 (\%) \quad \text{-----(7)}$$

$$MAE \% = \frac{1}{N} \sum_1^N \left( \frac{|X_i - Y_i|}{X_i} \right) \times 100 (\%) \quad \text{-----(8)}$$

Where N is number of hours (168 for one week), Xi Is the actual load and Yi is the forecasted load.

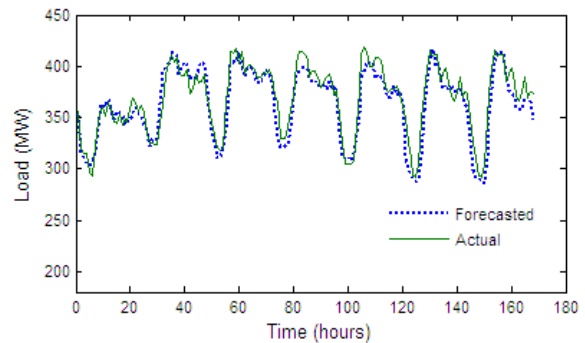


Fig. 7 Actual and Forecasted Load of 25 – 31 August 2006 using average temperature

Table 2 Absolute Percentage Error for the predicted period (25 – 31) August 2006 Using Average Temperature

Day Hour	25 Aug	26 Aug	27 Aug	28 Aug	29 Aug	30 Aug	31 Aug
1	1.22	0.89	0.37	0.65	7.91	0.01	4.84
2	2.06	5.46	0.90	4.21	5.57	4.98	3.57
3	2.52	4.15	0.67	2.84	0.04	8.46	2.07
4	0.63	5.47	0.73	2.76	0.97	3.99	4.75
5	6.68	4.38	0.66	3.24	0.22	6.45	3.80
6	8.72	3.00	2.31	0.80	2.28	7.88	2.82
7	1.44	5.43	6.50	0.61	2.84	1.58	7.37
8	0.50	6.88	2.78	5.42	5.24	1.40	9.64
9	3.31	0.79	4.14	2.41	4.21	0.17	5.79
10	1.14	5.00	5.54	0.28	3.19	3.01	0.33
11	1.95	5.46	7.34	1.36	0.36	1.57	1.13
12	0.93	3.74	1.74	2.20	3.48	5.12	1.39
13	0.74	5.37	7.22	2.50	2.86	3.45	0.47
14	1.44	6.28	4.54	4.99	4.27	2.37	7.10
15	4.51	2.50	3.58	1.43	3.38	1.93	2.51
16	1.42	0.89	4.48	2.44	3.19	2.85	0.06
17	4.05	3.09	6.91	3.14	0.33	0.56	1.22
18	0.45	0.57	4.63	6.02	0.42	3.77	1.96
19	1.69	0.55	6.42	5.03	0.08	3.84	2.17
20	1.23	5.07	4.12	1.26	0.93	1.97	0.03
21	2.52	0.24	4.04	2.55	1.47	0.59	4.12
22	0.33	0.41	2.43	1.39	2.13	3.54	5.71
23	0.57	2.20	5.45	8.26	0.34	0.75	5.01
24	2.80	0.07	5.93	4.35	1.50	1.05	1.45

Table 3 MAPE% for the two cases (using Average Tem & using Maximum and Minimum Tem)

Day	Using Average Tem.	Using Max. & Min Tem.
Fri	2.2	3.23
Sat	3.25	3.18
Sun	3.89	4.77
Mon	2.92	4.09
Tue	2.38	9.13
Wed	2.97	3.02
Thur	3.3	4.5
<b>Weekly MAPE%</b>	<b>2.99</b>	<b>4.56</b>

## V. CONCLUSION

Short Term Load Forecasting, with lead time from a hour to seven days, plays a key role for economic and secure operation of power system.

The approach used in this paper for Short Term Load Forecasting is based on combining the Wavelet and Neural Network, Wavelet Transform is used to enhance the learning capability of the Artificial Neural Networks.

The Temperature has a close relationship with the load. Due to the lack of weather information only the average or maximum and minimum Temperatures are included for the Neural Network, hence they are used as inputs to the Neural Networks to predict the maximum and minimum approximation respectively.

The details of the predicted day are calculated by computing the mean of the details of the last three same type days.

The total load is calculated by summing the approximation and the details, the approach has been tested with the real data of Erbil in the Iraq Kurdistan Region, for the data of 2006, the MAE% for the last week of August was 2.99 % by using average temperature, while it is 4.56% by using maximum and minimum temperature which is encouraging.

## REFERENCES

- [1] Anata Oonsivilai and El-Hawary, M.E. " wavelet Neural Network Based Short Term Load Forecasting of Electric Power System Commercial Load" Proceedings of the 1999 IEEE Canadian Conference on Electrical and Computer Engineering, Alberta, Canada May 9-12 1999.
- [2] Hyan Xu, Wei Ji Chen, "Artificial Neural Network Short-term Electrical Load Forecasting Techniques", IEEE TENCON, 1999, pp. 1458-1461.
- [3] Jie Bao, "Short-term Load Forecasting based on Neural network and Moving Average", Artificial Intelligence lab, Iowa State University.
- [4] A.G. Baklrtzis, V. Petridis, S. J. Klartzls and M. C. Alexlads "A Neural Network Short Term Load Forecasting Model for the Greek Power System" IEEE Transactions on Power Systems, Vol. 11, No. 2, May 1996. pp. 858-863.
- [5] Kinh D. Pham, P.E. "Load Forecasting Using Artificial Neural Network" Conference Paper, No. 95 B4, 1995 IEEE.
- [6] E. G. Swee, Terence and S.Elangovan, " Wavelets Based Analysis of Non-Uniformly Sampled Data Power Forecasting", Department of Electrical engineering, National University of Singapore.
- [7] Ruey Hwa Loh, "Time Series Forecast With Neural Network and Wavelet Techniques", thesis of Bachelor Degree, department of Electrical and Computer Engineering, University of Queensland, October, 2003.

- [8] S.J. Yao, Y.H. Song and L.Z. Zhang, X.Y. Cheng "Wavelet transform and neural networks for short-term
- [9] Electrical load forecasting", Energy Conversion & Management, 7 February, pp. 1975-1988.
- [10] RobiPolikar, "The Wavelet Tutorial" Department of Electrical and Computer Engineering, Rowan University.
- [11] M. Tarafdar Haque, and A.M. Kashtiban, "Application of Neural Networks in Power Systems; A Review", Transaction on Engineering Computing and Technology, V6 JUNE 2005, ISSN 1305-5313.
- [12] Mihai Gavrilas, "Neural Network Based Forecasting for Electricity Markets", Technical University of Iasi, Romania.
- [13] Li Zhang, and Peter B. Luh, Fellow, "Neural Network-Based Market Clearing Price Prediction and Confidence Interval Estimation With an Improved Extended Kalman Filter Method" IEEE Transactions on Power System, VOL. 20, NO. 1, FEBRUARY 2005.

**A. A. RASOOL** was born in Baghdad 1972, he obtained B.Sc. In Electrical engineering from the University of Salahaddin in 1994. He is currently pursuing his Master degree in power Engineering at the Salahaddin University.

**A. A. FATTAH** was born in Iraq-Kurdistan 1953, he obtained B. Sc. And Master degrees from Baghdad University in 1976 and 1979 respectively, he is presently assistant professor in Electrical Engineering, University of Salahaddin.

**I. B. SADIK** was born in Halabja in Kurdistan Region in 1973, he obtained B. Sc. In Electrical Engineering from university of Salahaddin, and MSc.,and PH.D from UPV in 2002 and 2006 respectively.