ISTANBUL UNIVERSITY –
JOURNAL OF ELECTRICAL & ELECTRONICS ENGINEERING

YEAR : 2004 (1051- 1056)
VOLUME : 4
NUMBER : 1

# MISSED DATA FORECASTING USING FAST NEURAL NETWORK ARCHITECTURE

**Abdullah BAL**

Department of Electrical and Computer Engineering, University of South Alabama,
Mobile, AL 36688

E-mail : abal@usouthal.edu

## *ABSTRACT*

*Wavelet function based feed forward neural network architecture is proposed for forecasting of missed data. A set of wavelet functions offers a multi-resolution approximation in signal analysis and provides localization in spatial domain. Wavelet neural network (WNN) is developed using these properties of wavelet functions in the traditional feed forward network. This network has multi-layer architecture and its each neuron includes wavelet functions in the first layer. The second layer entries are wavelet coefficients corresponding to first layer outputs. Wavelet functions provide high convergence capability and faster learning process in the multilayer networks. WNN performance is tested for function approximation and missed data forecasting problems.*

**Keywords:** *Wavelet theory, Feed Forward Neural Network, function approximation, forecasting*

## 1. INTRODUCTION

The approximation of general relation between input and output by nonlinear networks is very useful for system modeling and identification. Such approximation methods can be used to estimate of missed data in the real system. Neural networks are used extensively for this aim. Estimation of missed data is one type of function approximation problem. Neural network creates map through input data sets to desired data. This map can be call function between data groups. In the test phase, this map produces missed data estimation using adjusted weight coefficients and transfer functions [1].

In the past decade, researchers improved some solution for neural network problems such as local minimum catching, oscillation, optimum size adjustment, slow learning phase. For these aim, neural network architectures and learning algorithms are developed using fuzzy, genetic algorithm or wavelet transform techniques. Every proposed approach has alleviated different aforementioned problems [1-5].

In this study, a new approach proposed for the slow learning speed and some convergence problems using wavelet theory and multi layer neural network (MLNN). Firstly, the mexican hot wavelet function is selected from wavelet family to use as a transfer function in MLNN. Thus, weight coefficients of each neuron are obtained by adjustable wavelet parameters. As a result, the multi-resolution properties of wavelet offer high speed convergence capability in the MLNN architecture. To optimize these coefficients, gradient descent algorithm is employed in wavelet neural network.

*Received Date : 1.1.2003*
*Accepted Date: 10.12.2003*

This paper has been organized as follows: The wavelet transform is introduced in Section. 2. The inherent relation of feed forward neural network and wavelet transform is described in Section 3. Experimental results for function approximation and missed data forecasting are included in Section 4. Finally, concluding comments are included in Section 5.

## 2. SPATIAL LOCALIZATION OF WAVELET

Wavelet Transform (WT) is a popular spatial domain time-frequency theory especially during in the last ten years. Wavelet theory is based on wavelet functions family to express or approximate any given function. The WT is derived from one single function of wavelet by the operation of dilations and translations, and possesses localization performances in both time and frequency fields. So it can produce multi-resolution analysis of given one or two dimensional signals.

Wavelet family function can be formulated by dilating and translating of mother wavelet $\psi(x)$ as following [7],

$$\psi_{mn}(x) = 2^{n/2}\psi(2^n x - m) \qquad (m,n) \in Z^2 \quad (1)$$

The translation parameter *m* gives the position of the wavelet while the dilation parameter *n* governs its frequency. If *n* is large, the large features (i.e. low-frequency characteristics) of the signal are analyzed; conversely, if *n* is small, the small features (i.e. high-frequency characteristics) of the signal are analyzed. Thus, the wavelet decomposition allows us to view the signal at various frequency bands.

Wavelet transform is defined for any $f(x) \in L^2(R)$ by Equation 2.

$$W_{mn}(f) = < f(x), \psi_{mn}(x) > \quad (2)$$

The values of $(m,n)$ determines the local resolutions in time and frequency spaces.

A wide variety of functions can be chosen as the mother wavelet $\psi$, provided that $\psi(t) \in L^2$,

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \quad (3)$$

This flexibility in the choice of mother wavelet allows $\psi$ to be tailored to the specific application at hand. For neural network application, we would like $\psi$ to have compact support so that we can discriminate features of input data set in both time and scale.

## 3. INHERENT RELATION OF FNN AND WT

From the theory of functional analysis, functions can be represented as a weighted sum of basis functions. It is well known that the basic characteristic of FNN is that it needs a group of bases, whose linear combination is used to approach a demanded map to any desired degree. FNN architecture possesses an inherent dilation-translation and every node executes the same transfer (activation) function. On the other hand, WT offers powerful representation possessions by dilating and translating properties. Two methods are both powerful tools in studying nonlinear approximation problems. There is an inner relation between them. Their combination is considered a new method in handling the complexity and adaptability of problems. When the activation function of FNN selected the mother wavelet function in hidden layer, the weights and thresholds are obtained corresponding to the dilation and translation parameters of wavelet functions. Thus, the structure of the network and part of the weights are determined by wavelet coefficients. Furthermore, the local nature of activations makes it possible for the network to have better generalization and faster training speed.

The wavelet network can be considered as a particular case of the feed-forward basis function neural network model. The wavelet network is constructed by means of replacing the basis function with a multi-scale wavelet function $\psi$. The main architecture of WNN is illustrated in Fig. 1 [6]. This network exhibits a multi-input to multi-output nonlinear system realizing mapping $F : R^m \Rightarrow R^n$. In the simplest case, i.e. one input to one output, any one-dimensional nonlinear function can be approximated.

In the constructing of WNN architecture, the first step is the selection of wavelet function considering requirement of available derivation properties for gradient descent algorithm. Duration searching of activation function, it is decided on Mexican Hat function since it provides simplicity formulation and desired high performance. Mexican Hat function can be represent by the fallowing equation [8],

$$\psi = (1 - x^2).e^{-x^2} \qquad (4)$$

By using the wavelet for hidden layer and the linear function for output layer, WNN feed-forward structure may be formulated as fallow [9],

$$y = \omega^2(\psi(\omega^1.x + b^1)) + b^2 \qquad (5)$$

where $\omega^1$ represents connection weight matrices for first layer, $x$ is vector of input data set, $\omega^2$ denotes weight coefficients between wavelet neurons and output layer, $b^1$ and $b^2$ are biases of each layer and $\psi$ represents the wavelet function.

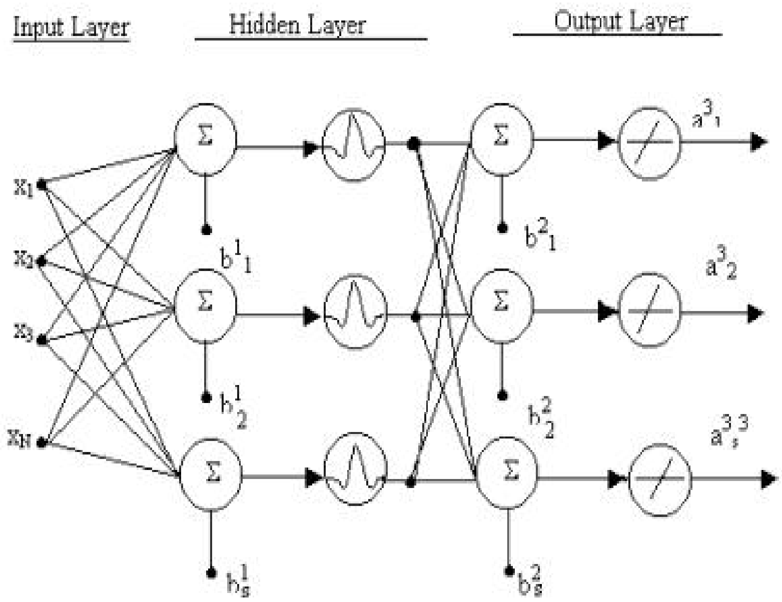


**Fig. 1.** Wavelet Neural Network architecture

# 4. EXPERIMENTAL RESULTS

In this section, the WNN performance is tested for function approximation and missed data forecasting problems. A detailed MATLAB software package has been developed to execute the proposed architecture. The results are compared with multilayer perceptron networks.

**A- Function Approximation**

In general, function approximation is used for unknown relation between input and output data. Similarly, a known function can be represented by another function. For this aim, the neural networks is used successfully. In this section, wavelet neural network is designed for function representation or approximation problem to evaluate performance of WNN. That's why, a known function that is created by mathematically is selected to be represented by weight coefficients and activation functions of WNN.

The created function is formulated in Equation 6,

$$y = \frac{\sin(x) * \cos(x)}{2} \qquad (6)$$

The value of x are selected with 0.4 incremental values in the interval [-2,2] for learning phase input data set.. The desired output set y is obtained from Equatin 6 corresponding to input x values.Two different multilayer perceptron architectures (MLP) are amployed to compare the performance of WNN. First one is $MLP^1$ includes logaritmic sigmoid activation function and second one is $MLP^2$ includes hyperbolic tangant sigmoid activation function.

Learning rate is the other important factor which determines the convergence speed of networks. Although the high value of learning rate increases convergence speed, sometimes it can produce undesired problems. To demonstrate the aforementioned affect of learning rate, we selected $\alpha$=0.05 and $\alpha$=0.1 for three layers network architectures. Three networks results are given in Table 1 by evaluating iteration number, convergence time and learning rate affects.

**Table 1.** Learning results of $MLP^1$, $MLP^2$ and WNN networks

| *Learning Rate* | *$\alpha$ = 0.05* | | *$\alpha$ = 0.1* | |
|---|---|---|---|---|
| | *Iteration* | *Time(s)* | *Iteration* | *Time(s)* |
| **$MLP^1$** | 14687 | 1110,81 | 7687 | 502,95 |
| **$MLP^2$** | 3737 | 228,38 | 1621 | 84,47 |
| **WNN** | 1346 | 77,33 | 941 | 54,16 |

In the testing duration, input test data are selected with 0.1 step from range of [-2,2]. Consequently, WNN generates $y_1$ function that represents the original function $y$. The original function and approximated function are illustrated in Fig. 2.

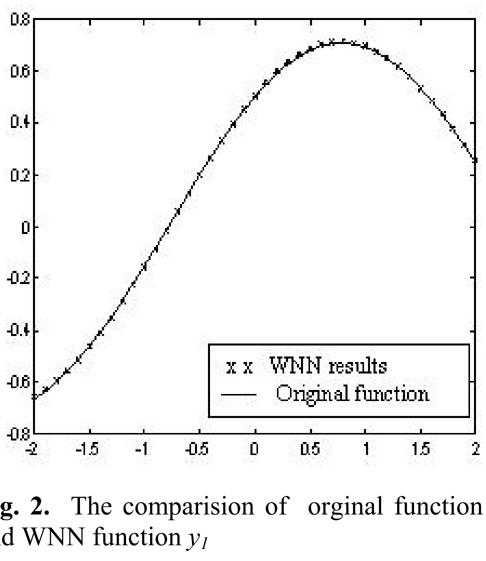


**Fig. 2.** The comparision of orginal function $y$ and WNN function $y_1$

**B-Missed Data Forecasting**

Data forecasting is used to estimate missed data using prior known information. In real word applications, some times all data can not be saved due to inevitable situation such as out of service of measurement tools, wrong process or personal mistakes. In these cases, the missed data forecasting is realizes using statistical or other mathematical techniques. The neural network approach provides wide spectrum representation for such a situation. The procedure is based on basic function approximation between known input and desired data set. Upon the learning phase, this function is used to forecast unknown data.

In this application, the real world meteorological measured data are used to estimate missed value. In the learning phase, weight and bias coefficients adjusted for moisture, pressure, evaporation and temperature data measured daily in Goztepe/Istanbul region in 1985. After learning phase, supposing some temperature data are missed, forecasting is executed using known moisture, pressure and evaporation data by WNN map function.

The WNN architecture has been designed as three layers which have 3:8:1 neurons distribution. (Fig.3).

For learning and testing applications, one month meteorological values are used , which are measured as a daily in meteorology station.. Among of this data, 24 days are selected for learning phase and rest data is used for testing phase. Fig 4 demonstrates the learning data set distribution in spatial domain.. The error rate of between measured and WNN output values are calculated as 8.11%, which is shown in Fig 5.

## 5. CONCLUSIONS

In this paper, fast neural network architecture is proposed using wavelet function as an activation function in the multi layer feed forward neural network. This architecture is called wavelet neural network. The convergence performance and error ratio of WNN are tested for two different applications. In the first application, WNN is run for function approximation to represent a known function. This function is formulated by mathematically and then its input and output values are used for learning phase data set. Learning duration, consumption time and convergence speed is compared with the other traditional multilayer networks. From Table 1, it is evident that the WNN speed is significantly more than the other MLP networks. For second application, real measured meteorological data set are selected. In this case, the goal is to forecast unknown or missed data using previous information. Upon learning phase, missed data of seven days are forecasted by the WNN. When evaluating of results with real values, the forecasting error is calculated as 8.11%. This ratio indicates that WNN has high learning capability in quite short response time.

**Acknowledgement**
The author wishes to thank Dr. O. Ucan and Dr. H. Pastaci for many rewarding discussions.

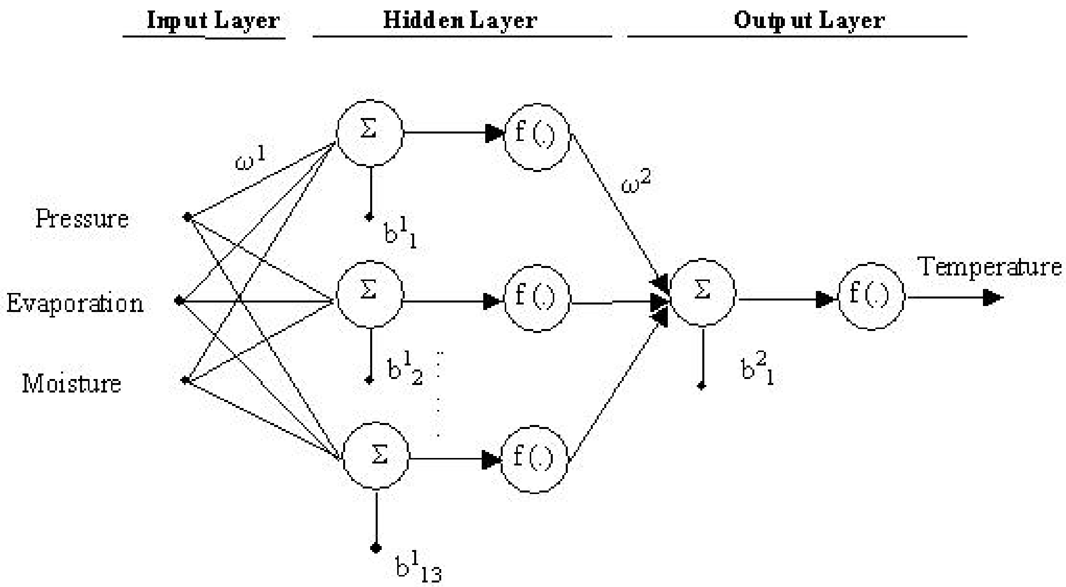


**Fig. 3.** Three layers WNN designed for missed data forecasting

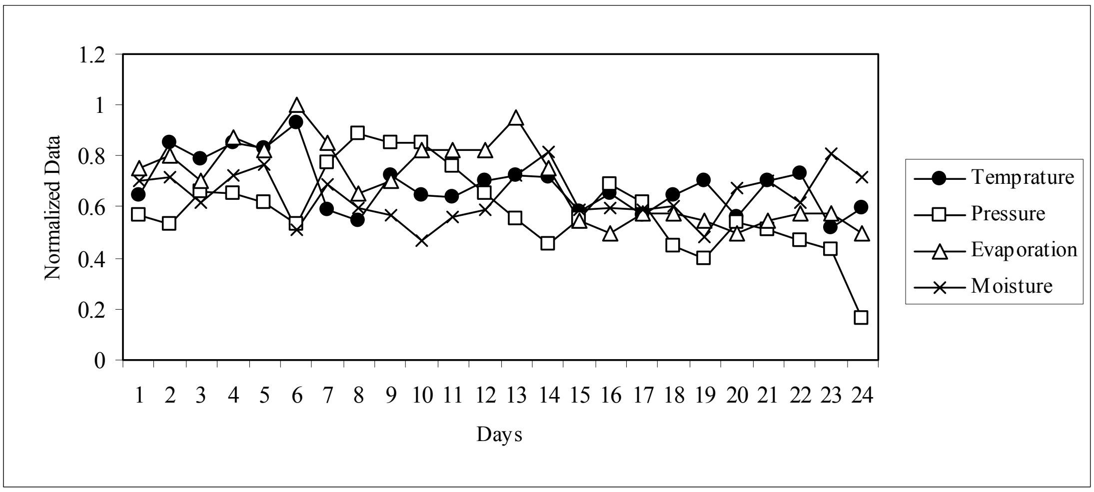


**Fig. 4.** WNN learning data set.

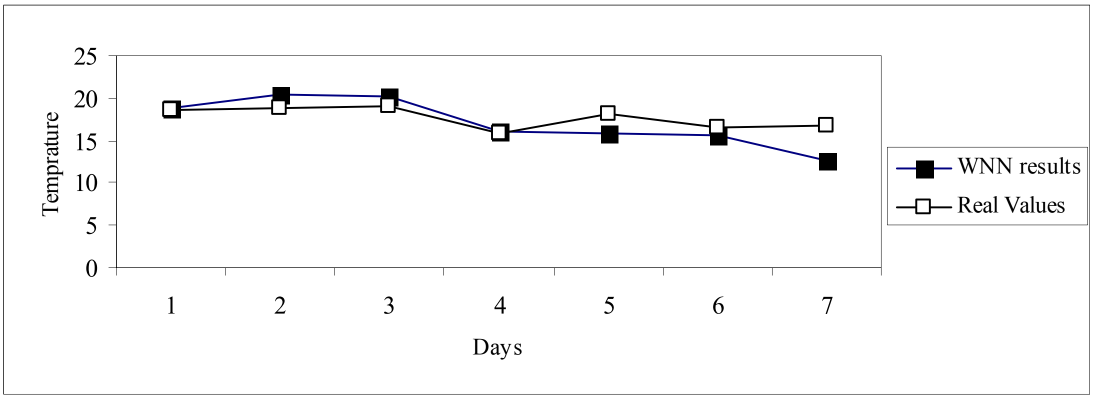


**Fig. 5** WNN test results and real measured values.

## REFERENCES

[1] Quing, W., Xueqin, S., Quingxin Y., and Weilli,Y., "Using Wavelet Neural Networks for the Optimal Design of Electro Magnetic Devices", IEEE Trans. on Magnetics, March 1997, 33(2):1928-19930

[2] Bal A., Pastaci H., Ucan O.N., (2001)"A Fast Wavelet Network Algorithm for EMG Classification", NEU-CEE 2001, 23-25 May. 2001, pp. 293-295, Lefkosa, Cyprus.

[3] Zhang, Q. and Benveniste A., (1992), "Wavelet Networks", IEEE Trans. On Neural Networks, 3(6): 889-898, Nov. 1992,

[4] Albora M, Bal A. and Ucan O.N.,(2002), "Wavelet Based Design of Dynamical Neural Networks for Magnetic Map Processing", International Conference on Earth Science and Electronics, ICESE'2002, Istanbul-Turkey, 24-25 October 2002.

[5] S. Pittner , S. V. Kamarthi and Q. Gao, "Wavelet Networks for sensor signal classification in flank wear assessment", Journal of Intelligent Manufacturing , 9,315-322, 1998.

[6] Bal A. (2002) "Wavelet Neural Network and Learning Algorithm", PhD Dissertation, YTU, 2002

[7] Rao, R.M. and Bopardikar, A.S., Wavelet Transforms Introduction to Theory And Applications, Addison Wesley Longman,USA

[8] Vetterli Martin, Kovacevic Jelena, Wavelets and Sub-band Coding, Prentice Hall, New Jersey, USA.

[9] Hagan, T.M.,Demuth, H. B. and Beale, M. Neural Networks Design, PWS Publishing Company, Boston, USA.

**Dr. Bal** received his BS degree from Electronics and Communication Engineering Department at Istanbul Technical University in 1993. He received his MS and PhD degrees in Electrical Engineering Department at Yildiz Technical University, Istanbul, Turkey in 1997 and 2002, respectively. Currently, he is a researcher scholar at the University of South Alabama. His research interests includes image processing, pattern recognition, wavelet theory, artificial neural network, cellular neural network, machine learning, biomedical image and signal processing, geophysical image analysis and optical pattern recognition.

*Abdullah BAL*