Applied Mathematical Sciences, Vol. 4, 2010, no. 50, 2485 - 2495

Wavelet Neural Networks for Nonlinear Time Series Analysis

K. K. Minu, M. C. Lineesh and C. Jessy John

Department of Mathematics
National Institute of Technology Calicut
NIT Campus P O - 673 601, India
lineesh@nitc.ac.in

Abstract

A wavelet network is an important tool for analyzing time series especially when it is nonlinear and non-stationary. It takes advantage of high resolution of wavelets and learning and feed forward nature of Neural Networks. Wavelets are a class of functions such that multiple resolution nature of wavelets provides a natural frame work for the analysis of time series. The power of this network to approximate functions from given input-output data is proved and it has utilized the localization property of a wavelet to focus on local properties. Guaranteed upper bounds on the accuracy of approximation is established. Here we are analyzing the time series of number of terrorist attacks in the world measured on monthly basis during the period February 1968 to January 2007 for establishing the superiority of this method over other existing methods. The simulation results show that the model is capable of producing a reasonable accuracy within several steps.

Mathematics Subject Classification: 37M10, 65T60, 92B20

Keywords: Non Stationary-nonlinear Time Series; Wavelet Networks; Scaling Coefficients; Wavelet Coefficients; Terrorist Attack Time Series; GARCH Model

1 Introduction

One of the greatest threats modern nations encounter is the terrorist attacks of various groups. It is usually unpredictable and unexpected. It may be national or trans national. The cause of it may vary from local to international but every nation faces it in one form or other. Prediction in time series is to model an existing data series in order to predict unknown future values accurately[14]. Linear models do not adequately represent nonlinear series and there are no single powerful tool for the analysis of nonlinear time series[12]. Wavelets can be considered as functions generated from a basic function by translations and dilations. The basic function is called the mother wavelet. Wavelet transforms involve representing a general function in terms of simple fixed building blocks at different resolutions[9]. They are generated from a single fixed function by changing translation and scale. The continuous wavelet transform considers a family $\{\Psi_t(a, b)\}$ and satisfies the admissibility condition expressed in terms of its Fourier transform. For Discrete Wavelet Transform(DWT), scale and translation parameters are chosen in such a way that at level m the wavelets are given by $\Psi_{m,n}(t) = 2^{-\frac{m}{2}}\Psi(2^{-m}t - n)$ where m and n are integers. Orthonormal basis and multiresolution analysis represents a function at various levels of resolution by projecting the function into an increasing sequence of subspaces.

Authors like Wei, W. W. W [16] concentrated on the analysis and forecasting of stationary time series. Many of these studies are based on the statistical concepts like correlation and regression analysis. On the other hand several researchers have been looking for better ways to design neural networks[14]. Hence it is of great importance to analyze the relationship between neural networks, approximation theory and functional analysis. In functional analysis any continuous function can be represented as a weighted sum of orthogonal basis functions. Such expansions can be easily represented as neural networks which can be designed for the desired error rate using the properties of orthonormal expansions. In order to take full advantage of orthonormality of basis functions, and localized learning, we need a set of basis functions which are local and orthogonal[15]. Wavelets are functions with these features. They have generated a tremendous interest in both theoretical and applied areas over the past few years. Wavelet networks are a class of neural networks that employ wavelets as activation functions[15]. These have been recently researched as an alternative approach to the neural networks with sigmoidal activation. In recent years, wavelet transformation is proposed for the analysis of time series. Researchers like Priestley, 1996; Morettin, 1997; Gao, 1997; Percival and Walden, 1999 focused on periodogram analysis of a time series. Bjorn, 1995; Soltani(2000); Renaud, 2003 are some groups studying time series prediction using wavelets.

2 Wavelets

A wavelet is a real or complex valued function $\Psi(.)$ satisfying the following conditions;
$\int_{-\infty}^{\infty} \Psi(u)du = 0$ and $\int_{-\infty}^{\infty} |\Psi^2(u)|du = 1$.

There are two functions in wavelet transform namely the scale function(father wavelet) and the mother wavelet. These two functions give a function family that can be used for reconstructing a signal. This is the basic idea of Multiresolution Analysis(MRA). Some commonly used wavelet families are Haar Wavelet, Meyer Wavelet, Daubechies Wavelet, Mexican Hat Wavelet, Coiflet Wavelet and Last Assymetric[13].

3 Neural Networks and Time Series Analysis

Feedforward neural networks are composed of layers of neurons in which the input layer of neurons is connected to the output layer of neurons through one or more layers of intermediate neurons. The training process of the neural network involves adjusting the weights till a desired input/output relationship is obtained. The majority of adaptation learning algorithms are based on the Widrow-Hoff back-propagation algorithm. Feed forward neural networks have been proposed for analyzing a given time series.
The standard neural network method of performing time series prediction is to induce the function using any feedforward function approximating neural network architecture, such as, a standard MLP, an RBF architecture, or a Cascade correlation model, using a set of N-tuples as inputs and a single output as the target value of the network. This method is often called the sliding window technique as the N-tuple input slides over the full training set.
The neural network forecaster can be described as follows

$$z_{k+1} = NN(z_k, z_{k-1}, ..., z_{k-d}, e_k, e_{k-1}, ..., e_{k-d}); \tag{1}$$

where z is either original observations or processed data, and $\{e_k, e_{k-1}, ..., e_{k-d}\}$ are residuals.

4 Discrete wavelet transform - Decomposition and Reconstruction

The discrete wavelet transform of a given time series $\{X_{t+1} : t = 0, 1, 2, ..., N-1\}$ is defined by

$$\Psi(m, n) = 2^{-\frac{m}{2}} \sum_{k=0}^{N-1} X_k \Psi(2^{-m}k - n), \tag{2}$$

where $\Psi(k)$ need not be a sampled version of $\Psi(t)$.
It is computationally impossible to analyze a signal using all wavelet coefficients, so one may wonder if it is sufficient to pick a discrete subset of the upper half plane to be able to reconstruct a signal from the corresponding wavelet

coefficients. One such system is the affine system for some real parameters $a > 1, b > 0$. The corresponding discrete subset of the half plane consists of all the points $\{a^m, na^m b\}$ with integers $m, n \in \mathbb{Z}$. The corresponding baby wavelets are now given as

$$\Psi_{m,n}(t) = a^{-m/2}\Psi(a^{-m}t - nb). \tag{3}$$

A sufficient condition for the reconstruction of any signal X of finite energy by the formula

$$X(t) = \sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} < X, \Psi_{m,n} > .\Psi_{m,n}(t) \tag{4}$$

is that the functions $\{\Psi_{m,n} : m, n \in \mathbb{Z}\}$ form a tight frame of $L^2(R)$.

5 Wavelet Networks

Representing a continuous function by a weighted sum of basis functions can be made unique if the basis functions are orthonormal. It was proved that [3] neural networks can be designed to represent such expansions with desired degree of accuracy. Wavelets have many desired properties combined together like compact support, orthogonality, localization in time and frequency and fast algorithms. Neural Networks are used in function approximation, pattern classification and in data mining but they could not characterize local features like jumps in values well. The local features may be existing in time or frequency. The improvement in their characterization will result in data compression and subsequent modification of classification tools. Wavelet networks are a class of neural networks that employ wavelets as activation functions[15].

5.1 Wavelet Neural Network (WNN) Model for Prediction

The concept of time series forecasting by using wavelet is nothing but forecasting by using the data which is preprocessed through the wavelet transform, especially through DWT. By the presence of multiscale decomposition like wavelet, the advantage is automatically separating the data components such as trend component and irregular component in the data. There by forecasting of stationary or nonstationary data.

Suppose we want to predict X_{t+1} where the data $\{X_j : j = 1, 2, 3, ..., t\}$ are given. The basic idea of WNN model is using preprocessed data that are obtained through the wavelet decomposition of X_t. Renaud et. al (2003) introduce Multilayer Perceptron (MLP) NN architecture(Feed Foreward Neural Network-FFNN) to process the wavelet coefficients. The FFNN architecture that is used for time series prediction consists of one hidden layer with P neurons defined as;

$$\hat{X}_{N+1} = \sum_{p=1}^{P} \hat{b}_p g[\sum_{j=1}^{J} \sum_{k=1}^{A_j} \hat{a}_{j,k,p} W_{j,N-2^j(k-1)} + \sum_{k=1}^{A_{J+1}} a_{J+1,k,p} v_{j,N-2^j(k-1)}], \qquad (5)$$

where j is the number of levels $\{j = 1, 2, 3, ..., J\}$, A_j orders of MAR model $(k = 1, 2, 3, ..., A_j)$; $w_{j,t}$ is the wavelet coefficient value, $v_{j,t}$ is the scale coefficient value and $a_{j,k}$ is the MAR coefficient value. Here g is an activation function in hidden layer of WNN.

6 Analysis of a Real World Nonlinear-nonstationary Time Series

In this paper we have considered the number of terrorist attacks in the world which is measured on monthly basis starting from February 1968 to January 2007. The plot of the data is given in figure 1.

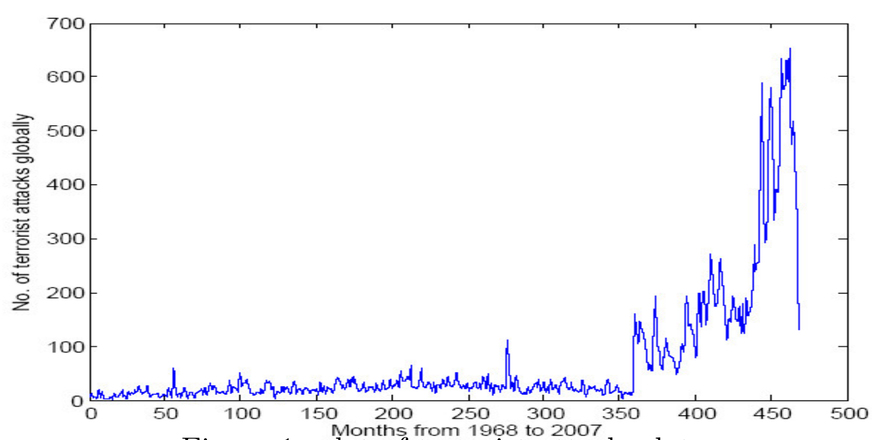

Figure 1: plot of terrorist attacks data

The plot of the autocorrelations of the terrorist attack data is given in figure 2. The plot shows that this is a nonstationary-nonlinear time series.

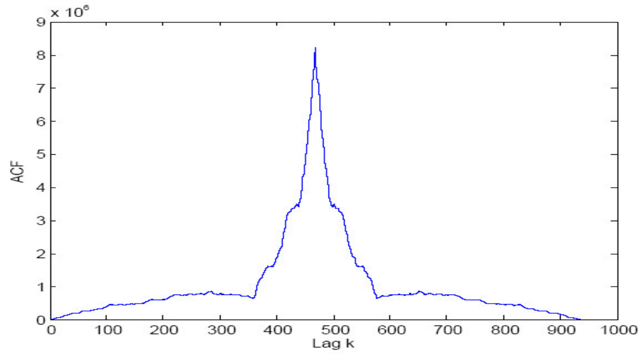

Figure 2: plot of acf of terrorist attack data

7 Analysis of Terrorist Attacks Data Using GARCH Model

The GARCH model is a prominent model in time series analysis that can represent a given nonstationary nonlinear time series up to a desired degree of accuracy. If an autoregressive moving average model (ARMA model) is assumed for the error variance, the model is a Generalized Autoregressive Conditional Heteroskedasticity (GARCH,Bollerslev(1986))model. A GARCH(p, q) model (where p is the order of the GARCH terms σ^2 and q is the order of the ARCH terms ε^2) is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + ... + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + ... + \beta_p \sigma_{t-p}^2. \tag{6}$$

A GARCH (1, 1) model is estimated for the terrorists attack time series with the parameters given by $\alpha_0 = 92.5749, \alpha_1 = 0.6375$ and $\beta_1 = 0.3314$. The plot of terrorist attack data and its prediction using the GARCH(1,1) model is given in figure 3.

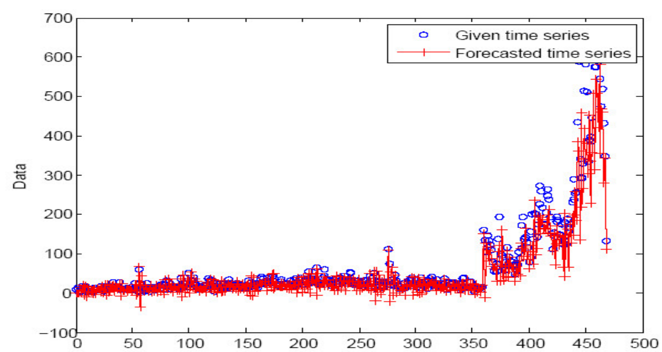

Figure 3: plot of forecasted data using GARCH Model

8 Analysis of Terrorist Attacks Data Using Neural Network(NN)

A feedforward neural network with back propagation is used to predict the last 100 month data values of the terrorist attacks time series. The plot of the training of the developed FFNN is given in figure 4.

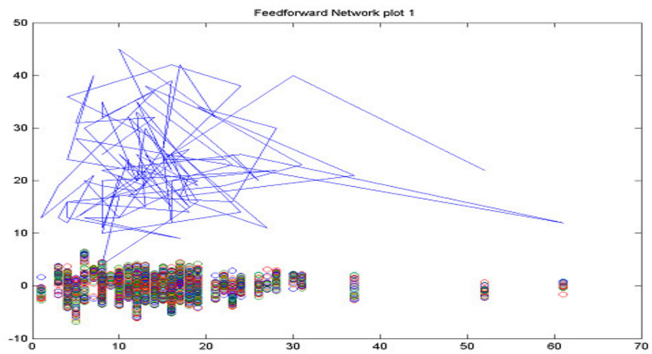

Figure 4: plot of learning FFNN

The terrorist attacks data and the predicted data using FFNN is given in figure 5.

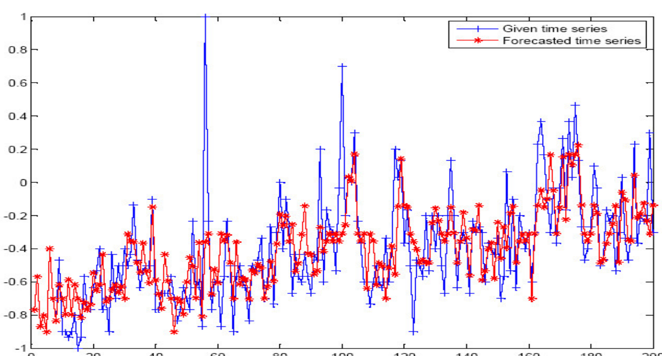

Figure 5: plot of forecasted data using FFNN

Also the plot of the error surface and contour are given in figure 6.

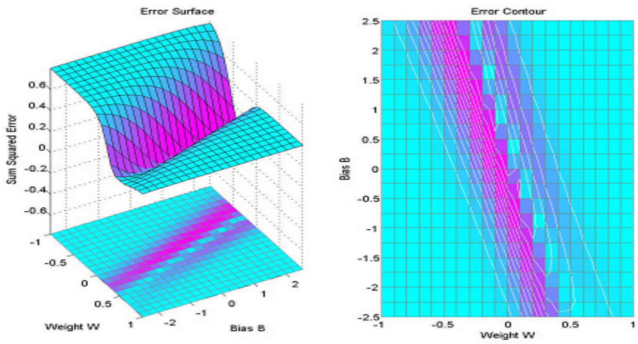

Figure 6: plot of the error surface and contour using FFNN

9 Analysis of Terrorist Attacks Data Using Wavelet Neural Networks(WNN)

Wavelet Neural Network is a feed forward neural network and it is trained using first 328 data points of the given terrorist attack time series and the last 100 points were predicted. The plot of the training of the developed WNN is given in figure 7.

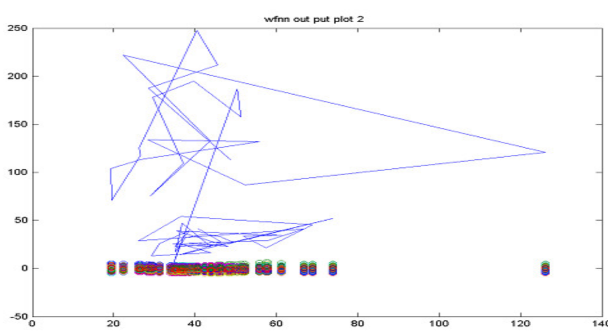

Figure 7: plot of learning WNN

The terrorist attack data and the predicted data using WNN is given in figure 8. Also the plot of the error surface and contour is given in figure 9.

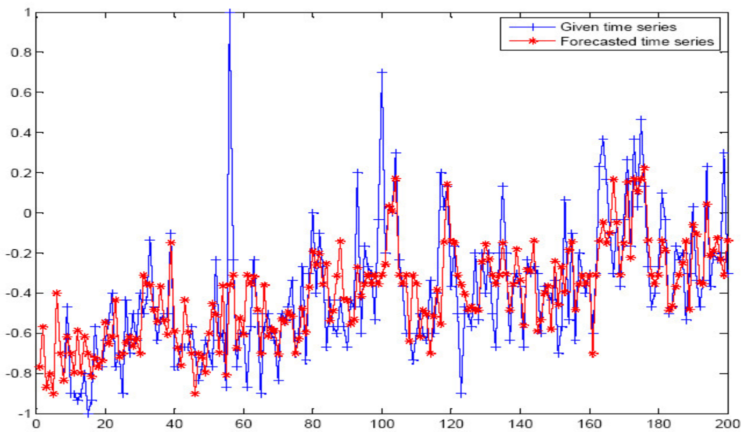

Figure 8: plot of forecasted data using WNN

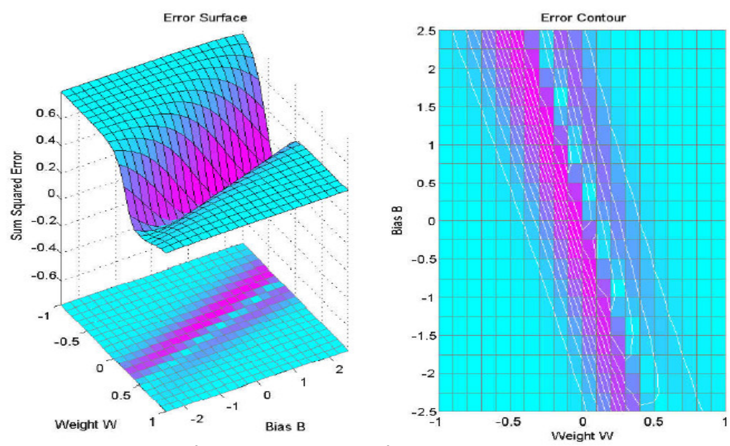

Figure 9: plot of the error surface and contour using WNN

10 Error Analysis

The details of the error analysis of the real world time series of terrorist attacks due to the methods discussed above are given in table 1. It is of great importance to observe that the distribution of the error follow a Gaussian distribution and there by verifying the efficiency of the developed method.

Table 1: Error Comparison

M.S.E(GARCH)	M.S.E(FFNN)	M.S.E.(WNN)
3.2395	0.6536	0.0015

11 Conclusion

Trend and Threshold Autoregressive(T-TAR) model using wavelet decomposition method were developed by Lineesh M C and C Jessy John(2010). Wavelet Neural Networks(WNN) for forecasting purpose were introduced by Minu K K and C Jessy John(2010). In this paper WNN for forecasting is implemented by using the T-TAR model and it is applied to the terrorist attacks time series which is a nonstationary nonlinear time series. The analysis results were compared with that of existing methods. Simulation plots are given for GARCH model, FFNN and Wavelet Neural Network. The comparison study with the help of error table shows that Wavelet Neural Networks provide the best model for analyzing the terrorist attacks time series.

References

[1] E. B. Christopouloul, N. K. Athanassios and A. A. Georgakilas, Time Series Analysis of Sunspot Oscillations Using the Wavelet Transform, *Proceedings of the 14th International Conference on Digital Signal Processing,* **2**(2002), 893 - 896.

[2] M. H. Hayes, *Statistical Digital Signal Processing and Modeling,* John Wiley and Sons, 2004.

[3] S. Haykins, *Neural Networks-A comprehensive Foundation,* McGraw Hill, 1993.

[4] K. Ko and M. Vannucci, Bayesian wavelet analysis of autoregressive fractionally integrated moving-average processes, *Elsevier, Science Direct,* **136** (2006), 3415-3434.

[5] Y. Kopsinis and S. McLaughlin, Empirical Mode Decomposition Based Soft-Thresholding, Proceedings of the 16^{th} European Signal Processing Conference, *EUSIPCO* (2008).

[6] Y. Li and Z. Xie, The Wavelet Identification of Thresholds and Time Delay of Threshold Autoregressive Models, *Statistica Sinica,* **9** (1999), 153-166.

[7] M. C. Lineesh and C Jessy John, Analysis of Nonstationary Time Series using Wavelet Decompositon, *Nature and Science,* **1**(2010), 53-59.

[8] M. C. Lineesh, K. K. Minu and C. Jessy John, Analysis of Nonstationary Nonlinear Economic Time Series of Gold Price - A Comparative Study, *International Mathematical Forum,* Vol. 5, no. 34(2010),1673-1683.

[9] S. Mallat, *A wavelet tour of signal processing,*, Academic Press, 1999.

[10] A. Papoulis, *Probability, Random Variables, and Stochastic Processes,* (3rd Edition), McGraw-Hill, 1991.

[11] D.B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis,* Cambridge University Press.

[12] M.B. Priestley, *Non-linear and Non-stationary Time Series Analysis,* Academic Press, 1988.

[13] R.M. Rao and A. S. Bopardikar, *Wavelet Transforms - Introduction to Theory and Applications,* Pearson Education, 1998.

[14] S. Soltani, 2002, On the use of the wavelet decomposition for time series prediction, *Elsevier, Neurocomputing,* **48**(2002)267-277.

[15] Suhartono and Subanar, Development of Model Building Procedures in Wavelet Neural Networks for Forecasting Non-Stationary Time Series,*European Journal of Scientific Research,* **34** (2009),No. 3, 416-427.

[16] W.W.S. Wei, *Time Series Analysis-Univariate and Multivariate Methods,* Addison-Wesley Publishing Company.

[17] A. Zhang and Benveniste, *Wavelet Networks,* IEEE transactions on Neural Networks, Vol.3, No.6, 1992.

Received: November, 2009