Analysis of Non-Stationary Time Series using Wavelet Decomposition

Lineesh M C[1*], C Jessy John[2]

Department of Mathematics, National Institute of Technology Calicut,

NIT Campus P O – 673 601, Calicut, Kerala, India.

lineesh@nitc.ac.in , jessy@nitc.ac.in

Abstract: The increased computational speed and developments in the area of algorithms have created the possibility for efficiently identifying a well-fitting time series model for the given nonstationary-nonlinear time series and use it for prediction. In this paper a new method is used for analyzing a given nonstationary-nonlinear time series. Based on the Multiresolution Analysis (MRA) and nonlinear characteristics of the given time series a method for analyzing the given time series using wavelet decomposition is discussed in this paper. After decomposing a given nonstationary-nonlinear time series Z_t in to a trend series X_t and a detail series Y_t the trend series and the detail series are separately modeled. Model T(t) representing the trend series X_t and the Threshold Autoregressive Model of order k (TAR(k)) representing detail series Y_t are combined to obtain the Trend and Threshold Autoregressive(T-TAR) model representing the given nonstationary-nonlinear time series. The scale dependent thresholds for the T-TAR model are obtained using the detail series and using the trend series. Also simulation studies are done and the results revealed that the developed method could increase the forecasting accuracy. [Nature and Science 2010;8(1):53-59] (ISSN: 1545-0740).

Keywords: Non Stationary-nonlinear Time Series; Wavelet Decomposition; Trend Models; Threshold Autoregressive Models; Scaling Coefficients; Wavelet Coefficients.

1. Introduction

Many stochastic systems are observed to be nonlinear which governs to nonstationary non-linear time series or signals. The Annual sunspot time series, the Canadian lynx series (Priestley, 1988) Financial time series (Hyndman, 2008) are examples of nonstationary nonlinear time series. So modeling nonstationary-nonlinear time series/signals for prediction is need of the day. Curvilinear regression models, Threshold Auto-regressive (TAR) models, State Dependent Models, etc are used for modeling nonstationary-nonlinear time series (Makridakis, 1990; Priestley,1988). But accuracy in prediction of nonstationary-nonlinear time series/signals was one of the main issues associated with the existing models. The increased computational efficiency leads to the application of wavelet decomposition method as a tool for modeling nonstationary-nonlinear time series (Kants, 2003; Kuo, 1994; Minu, 2008; Nason, 1999; Papoulis, 1991). This method leads to high accuracy in prediction. This paper discusses the decomposition of a given nonstationary-nonlinear time series in to a trend series and detail series. The Wold decomposition theorem (Hayes, 2004) states that a given time series can be splitted in to trend series and detail series. It is established (Lineesh, 2008) that the resultant time series obtained by wavelet decomposition are the same as the trend and detail series due to Wold (Hayes, 2004). Here instead of using the conventional reconstruction of the time series using wavelet, the trend series and detail series are modeled separately and the model representing the given time series is obtained as a combination of both the models (Lineesh, 2008) which takes care of the time dependencies of the series and this combined Trend and Threshold Autoregressive model (T-TAR) is used for prediction.

2. Review of Literature

The fitting of models for nonstationary-nonlinear time series raises some complex issues like the determination of the best fitted model to the given time series. Strang (1998) discussed how to decompose a signal in to its wavelet coefficients and reconstruct the signal from the coefficients. Brockwell and Davis (1995) discussed a method for identifying the order of the time series model which is based on the patterns present in higher order cumulants. Sesay and Subba Rao (1988) derived Yule-Walker type difference equations for higher order moments and cumulants for certain class of Bilinear time series models. Subba Rao, M. Eduarda, et al. (1992) extended the idea for the tentative identification of the order of the Bilinear model. Yuan Li and Zhongjie Xie (1982) engaged in the study of the identification of the thresholds and time delay of TAR models by checking different empirical wavelets of the given data. Oyet (2001) introduced a new approach for modeling nonlinear time series based on wavelet smoothing. Skander Soltani (2002), Hayes (2004), Ko and Vannucci (2006) are also contributed to the wavelet analysis techniques for the analysis of nonstationary-nonlinear time series. Nason and Von Sachs (1999) give an overview of the work on wavelet applications of time series. Initially the applications of wavelet transform for time series analysis were focused on periodogram analysis and cycles evaluation.

3. Estimation of T-TAR Models using Wavelet Decomposition Method

3.1 Wavelet Decomposition of Nonstationary-nonlinear Time series

To obtain a model for prediction of the given nonstationary-nonlinear time series Z_t it is required to decompose the given time series Z_t in to the trend series X_t and the detail series Y_t so that X_t and Y_t are orthogonal.

A given time series

$$\{Z_t : t = 0, 1, 2, \ldots, N-1\}$$

can be decomposed as

$$Z_t = X_t + Y_t,\ t = 0, 1, 2, \ldots, N-1$$

where X_t is the trend series and Y_t is the detail series given by,

$$Y_t = \sum_{j=1}^{M} d_{j,t} \tag{1}$$

where $d_{j,t}$ is the j^{th} level detail series, using wavelet decomposition technique. Lineesh (2008) proved that the components

$$X_t = C_{M,t} \quad \text{and} \quad Y_t = \sum_{j=1}^{M} d_{j,t}$$

of time series Z_t obtained by wavelet decomposition satisfy the requirements of Wold's decomposition of the time series.

3.2 Estimation of Threshold Auto-regressive Model Using Wavelet Techniques

The l threshold autoregressive model of order, k i.e. TAR (k) model (Priestley, 1988) is defined as,

$$Y_t = a_0^{(j)} + \sum_{i=1}^{k} a_i^{(j)} Y_{t-i} + e_t^{(j)} \tag{2}$$

where $Y_{t-j} \in R^{(j)}$ for $j = 1, 2, \ldots, l,\ \ R^{(j)}$

being a given subset of the real line R^1. In (2), $e_t^{(j)},\ j = 1, 2, \ldots, l$ and $t \in Z$ (the set of integers) is a sequence of independent and identically distributed random variables with 0 mean, constant variance σ^2 and $a_i^{(j)}, 1 \le i \le k$ are constant coefficients. The TAR (k) model is estimated for representing the detail series Y_t by applying wavelet decomposition method.

Determination of the coefficients and threshold are the main issues while analyzing a nonstationary-nonlinear time series using TAR model. In this paper the coefficients and thresholds are estimated as follows.

The scale and wavelet coefficients are defined as;

$$\phi_{j,k}(t) = 2^{-\left(\frac{j}{2}\right)} \phi\left(2^{-j} t - k\right),$$
$$j = 1, 2, \ldots, J;\ k = 0, 1, \ldots, 2^j - 1.$$
$$\psi_{j,k}(t) = 2^{-\left(\frac{j}{2}\right)} \psi\left(2^{-j} t - k\right),$$
$$j = 1, 2, \ldots, J;\ k = 0, 1, \ldots, 2^j - 1. \tag{3}$$

where

$$\phi(t)=\begin{cases} -2^{\left(\frac{-j}{2}\right)} & if\ 2^{j}k \le t \le 2^{j}(k+1/2) \\ 2^{\left(\frac{-j}{2}\right)} & if\ 2^{j}(k+1/2) \le t \le 2^{j}(k+1) \end{cases} \quad (4)$$

and

$$\psi(t)=\begin{cases} 2^{\left(\frac{-j}{2}\right)} & if\ 2^{j}k \le t \le 2^{j}(k+1/2) \\ -2^{\left(\frac{-j}{2}\right)} & if\ 2^{j}(k+1/2) \le t \le 2^{j}(k+1) \end{cases} \quad (5)$$

Define $\beta_{j,k} = \sum_{t=0}^{N-1} \psi_{j,k}(t).Y_t$ (6)

where Y_t is the detail series obtained by decomposing Z_t using wavelet decomposition.

Then using (6)

$$Y_t = \sum_{j=1}^{J} \sum_{k=0}^{2^j-1} \beta_{j,k}.\psi_{j,k}(t) \quad (7)$$

3.2.1 Estimation of the Threshold

The threshold of the TAR model is estimated as follows.
For $j = 1, 2, \ldots, J$ define

$\lambda_j = \sqrt{2\log(\# d_{j,t})}$, where $(\# d_{j,t})$ denotes the cardinality of $\{d_{j,t}\}$. Also define

$$\lambda = \sqrt{2\log(\# C_{M,t})}$$

Here λ_j denotes the threshold for the j^{th} level detail series and λ denotes the threshold of the TAR model.

3.2.2 E stimation of TAR model

The Threshold Autoregressive model representing the detail series $\{Y_t\}$ is given by,

$$Y_t = \begin{cases} b_1^{(1)}Y_{t-1} + b_2^{(1)}Y_{t-2} + \ldots + b_k^{(1)}Y_{t-k} \\ + e_t^{(1)}\ if\ Y_{t-d} < \lambda \\ b_1^{(2)}Y_{t-1} + b_2^{(2)}Y_{t-2} + \ldots + b_k^{(2)}Y_{t-k} \\ + e_t^{(2)}\ if\ Y_{t-d} \ge \lambda \end{cases} \quad (8)$$

where the coefficients $\{b_j^{(i)}\}$ are defined by,

$$b_j^{(i)} = \begin{cases} \sum_j \sum_t d_{j,t}^{(1)}.\psi^{(1)}_{j,t} \\ \sum_j \sum_t d_{j,t}^{(2)}.\psi^{(2)}_{j,t} \end{cases} \quad (9)$$

where

$$d_{j,t}^{(1)} = d_{j,t}\quad if\quad d_{j,t} < \lambda_j$$
$$and\ d_{j,t}^{(2)} = d_{j,t}\quad if\quad d_{j,t} \ge \lambda_j \quad (10)$$

$$\psi_{j,t}^{(1)} = \psi_{j,t}\quad if\quad d_{j,t} < \lambda_j$$
$$and\ \psi_{j,t}^{(2)} = \psi_{j,t}\quad if\quad d_{j,t} \ge \lambda_j. \quad (11)$$

3.3 Model for Trend Series

The best fitting ARMA (p, q) model, linear regression model and curvilinear regression model are considered for the analysis of trend series. The model thus obtained for trend series is denoted by T (t).

3.4 Trend and Threshold Autoregressive Model (T-TAR)

The T-TAR model representing the given nonstationary-nonlinear time series Z_t, using wavelet decomposition is obtained by combining the model representing the trend series and the detail series which is given by;

$$Z_t = \begin{cases} T(t) + b_1^{(1)}Y_{t-1} + b_2^{(1)}Y_{t-2} & \\ + \ldots + b_k^{(1)}Y_{t-k} + e_t^{(1)} & if\ Y_{t-d} < \lambda \\ T(t) + b_1^{(2)}Y_{t-1} + b_2^{(2)}Y_{t-2} & \\ + \ldots + b_k^{(2)}Y_{t-k} + e_t^{(2)} & if\ Y_{t-d} \ge \lambda \end{cases}$$

(12)

Here T(t) and TAR (k) preserves orthogonality.

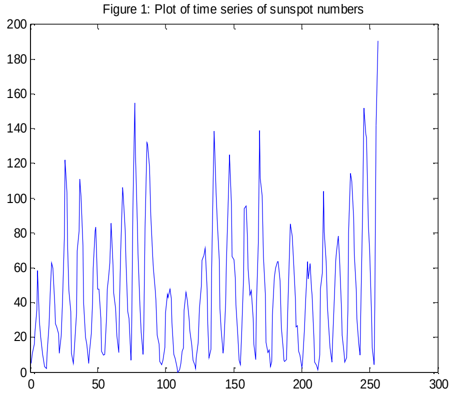
4. Application of T-TAR Models for Prediction

Prediction using time series originated from a stochastic system is the very aim of modeling a time series. The estimation of T-TAR model by applying wavelet theory is demonstrated with different real world time series and the results are presented here.

4.1 Analysis of the Time Series of Annual Sunspot Numbers

The time series of annual sunspot numbers during years 1700 – 1955 (Priestley, 1988) is taken for illustrating the estimation of T-TAR model explained in this paper. The plot of the time series is shown in figure 1.

Figure 1: Plot of time series of sunspot numbers



4. 1.1 T-TAR Model Estimation of the Time Series of Annual Sunspot Numbers

Using the method explained in this paper T-TAR model is estimated for the time series of sunspot numbers using the wavelet method and it is given in Table 1.

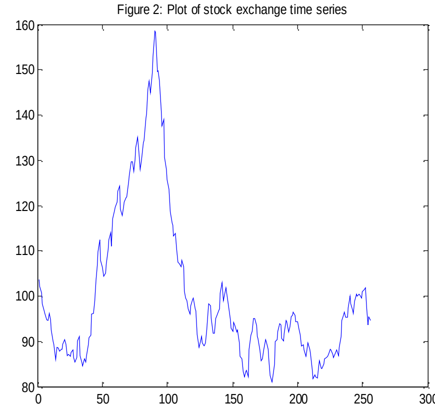
4.1.2 Model Estimation of the Time Series of Sunspot Numbers using the Existing Method

The commonly used method for analyzing nonstationary-nonlinear time series is due to Priestley. Using this method the model representing the time series of annual sunspot numbers is estimated. The analysis results using Priestley's method is included in Table 2.

4.2 Analysis of Stock Exchange Time Series

To see variety of applications the method is applied for the analysis of stock exchange time series. The time series representing monthly weighted-average exchange value of U. S. Dollar starting from September 1977 to December 1998 is taken for illustrating the method discussed in this paper. This is a secondary data (Hyndman, 2008). The plot of the data is given in figure 2.

Figure 2: Plot of stock exchange time series



4.2.1 T-TAR model Estimation of the Stock Exchange Time Series

Using the method explained in this paper T-TAR model is estimated for stock exchange time series using the wavelet method and it is given in table 3.

4.2.2 Model Estimation of Stock Exchange Time Series using the Existing Method

Using the existing method the model representing the time series is estimated. The analysis results of the stock exchange time series using the existing method due to Priestley is included in Table 4.

4.3 Analysis of IBM Stock Price Time Series

The time series of daily closing IBM stock prices (Hyndman, 2008) is taken for illustrating the estimation of T-TAR model explained in this paper. The plot of the data is shown in Figure 3.

Figure 3: Plot of IBM stock price time series

4.3.1 T-TAR Model Estimation of the IBM Stock Price Time Series

Using the method explained in this paper T-TAR model is estimated for IBM stock price time series using the wavelet method and the T-TAR model estimated is given in Table 5.

4.3.2 Model Estimation of IBM Stock Price Time Series using the Existing Method

Using the existing method due to Priestley the model representing the IBM stock price time series is estimated. The analysis results using Priestley's method is included in Table 6.

5. Conclusions

In this paper a new method for analyzing nonstationary-nonlinear time series using wavelet decomposition is introduced. Under this method the given nonstationary-nonlinear time series is decomposed into trend and detail series. After decomposition of the given time series the resultant series are modeled separately and then the T-TAR model for the given time series is obtained by combining the models representing the trend series and detail series. This method gives a comprehensive algorithm for analyzing nonstationary-nonlinear time series which is an advantage over the existing method.

The developed method is verified using different time series. The developed method is compared with the existing method and the error analysis in Table 7 shows the efficiency of the method in improving the accuracy in prediction.

Table 1: Estimated T-TAR model for the time series of annual sunspot numbers

Threshold	Estimated Model	MAPE	MSE
3.33	$Z_t = \begin{cases} 0.99X_{t-1} - 0.003X_{t-2} - 5.6Y_{t-1} - 8.89Y_{t-2} - 21.7Y_{t-3} + e_t^{(1)} & if \quad Y_{t-1} < 3.33 \\ 0.99X_{t-1} - 0.003X_{t-2} + 4.28Y_{t-1} + 7.09Y_{t-2} + 5.48Y_{t-3} + e_t^{(2)} & if \quad Y_{t-1} \geq 3.33 \end{cases}$	0.7901	4.4814

Table 2: Analysis of time series of sunspot numbers using Priestley's method

Threshold	Estimated Model	MAPE	MSE
35	$X_t = \begin{cases} 0.539X_{t-1} - 0.196X_{t-2} + 0.483X_{t-3} + e_t^{(1)} & if \quad X_{t-2} < 35 \\ 0.542X_{t-1} - 0.127X_{t-2} + 0.017X_{t-3} + 0.051X_{t-4} + 0.029X_{t-5} + 0.45X_{t-6} + e_t^{(2)} & if \quad X_{t-2} \geq 35 \end{cases}$	3.1828	6.5317

Table 3: Estimated T-TAR model for the stock exchange time series

Threshold	Estimated Model	MAPE	MSE
3.52	$Z_t = \begin{cases} 0.99X_{t-1} - 0.000157X_{t-2} - 0.79Y_{t-1} - 1.258Y_{t-2} + e_t^{(1)} & if \quad Y_{t-1} < 3.52 \\ 0.99X_{t-1} - 0.000157X_{t-2} + 1.121Y_{t-1} + 2.499Y_{t-2} + e_t^{(2)} & if \quad Y_{t-1} \geq 3.52 \end{cases}$	1.12	0.48

Table 4: Analysis of stock exchange time series using Priestley's method

Threshold	Estimated Model	MAPE	MSE
90	$X_t = \begin{cases} 1.003X_{t-1} - 0.44X_{t-2} + 0.25X_{t-3} - 0.19X_{t-4} + 0.31X_{t-5} - 0.31X_{t-6} + 0.23X_{t-7} - 0.18X_{t-8} + 0.39X_{t-9} - 0.49X_{t-10} + 0.58X_{t-11} - 0.28X_{t-12} + 0.4X_{t-13} + e_t^{(1)} & if \quad X_{t-3} < 90 \\ 1.169X_{t-1} - 0.16X_{t-2} + e_t^{(2)} & if \quad X_{t-3} \geq 90 \end{cases}$	1.9164	0.5215

Table 5: Estimated T-TAR model for the IBM stock price time series

Threshold	Estimated Model	MAPE	MSE
8.744	$Z_t = \begin{cases} 0.99X_{t-1} + 0.0000274X_{t-2} - 3.02Y_{t-1} - 2.659Y_{t-2} - 8.021Y_{t-3} - 37.34Y_{t-4} + e_t^{(1)} & if \quad Y_{t-1} < 8.744 \\ 0.99X_{t-1} - 0.0000274X_{t-2} + 2.86Y_{t-1} + 2.726Y_{t-2} + 1.9Y_{t-3} + 2.36Y_{t-4} + e_t^{(2)} & if \quad Y_{t-1} \geq 8.744 \end{cases}$	0.0401	27.696

Table 6: Analysis of IBM stock price time series using Priestley's method

Threshold	Estimated Model	MAPE	MSE
560	$X_t = \begin{cases} 1.293X_{t-1} - 0.293X_{t-2} + e_t^{(1)} & if \quad X_{t-1} < 560 \\ 1.13X_{t-1} - 0.338X_{t-2} + 0.176X_{t-3} + 0.145X_{t-4} - 0.28X_{t-5} + 0.016X_{t-6} - 0.106X_{t-7} + 0.257X_{t-8} + e_t^{(2)} & if \quad X_{t-1} \geq 560 \end{cases}$	1.466	84.482

Table 7: Error Comparison of T-TAR model and Model due to Priestley

Sr. No.	Time Series	T-TAR Model		Model due to Priestley	
		MAPE	MSE	MAPE	MSE
1.	Sunspot	0.7901	4.4814	3.1828	6.5317
2.	Stock Exchange	1.12	0.48	1.9164	0.5215
3.	IBM Stock Price	0.0401	27.696	1.466	84.482

Correspondence to:
Lineesh M C, C Jessy John,
Department of Mathematics, NIT Calicut,
Kerala, India
Telephone: 91-495 2286507
Cellular Phone: 919495328642
Emails: lineesh@nitc.ac.in, jessy@nitc.ac.in

References:

[1] Brockwell, P.J., Davis, R.A. Time Series: Theory and Methods. Springer. 1995.
[2] Hayes, M.H. Statistical Digital Signal Processing and Modeling. John Wiley and Sons. 2004.
[3] Hyndman R. J. Time Series Data Library,www.robhyndman.com/TDSL/. 2008.
[4] Ip, W. C, Wong, H, Li, Y, Xie, Z. Threshold Variable Selection by Wavelets in Open loop-Threshold Autoregressive Models. Statistics and Probability Letters. 1999. Vol 42:4: pp.375-392(18).
[5] Kants, H., Schreiber, T. Nonlinear Time Series Analysis. 2nd Edition, Cambridge University Press. 2003.
[6] Ko, K., Vannucci, M. Bayesian wavelet analysis of autoregressive fractionally integrated moving-average processes. Elsevier, Science Direct. 2006: 136: 3415-3434.
[7] Kopsinis, Y., McLaughlin, S. Empirical Mode Decomposition Based Soft Thresholding. Proceedings of the 16th European Signal Processing Conference, EUSIPCO. 2008.
[8] Kuo, R. J. Automated Surface Property Inspection using Fuzzy Neural Networks and Time Series Analysis. World Congress on Neural Networks, San Diego. Vol 1. 1994.
[9] Lineesh, M. C., Jessy John, C. Identification of Threshold Autoregressive Models using Wavelet Basis. Proceedings of the National Conference on Recent Developments and Applications of Probability Theory, Random Process and Random Variables in Computer Science. 2008.
[10] Makridakis, S. Sliding Simulation: A New Approach to Time Series Forecasting. Management Science.1990: Vol 36: No. 4.
[11] Minu, K. K., Jessy John, C. A Mathematical Discussion on Outlier Elimination and Gibbs Error in Wavelet Neural Networks. Proceedings of the National Conference on Recent Developments and Applications of Probability Theory, Random Process and Random Variables in Computer Science. 2008.
[12] Nason, G. P., von Sachs, R. Wavelets in Time Series Analysis. Phil. Trans. R. Soc. Lond, A, 1999: 357, 2511-2526.
[13] Oyet, A. J. Nonlinear Time Series Modeling: Order Identification and Wavelet Filtering. Interstat. Journals. 2001.
[14] Papoulis, A. Probability, Random Variables, and Stochastic Processes, 3rd Edition. McGraw-Hill. 1991.
[15] Percival, D. B., Walden, A. T. Wavelet Methods for Time Series Analysis. Cambridge University Press. 2000.
[16] Priestley, M. B. Non-linear and Non-stationary Time Series Analysis. Academic Press.1988.
[17] Rao, R. M., Bopardikar, A.S. Wavelet Transforms – Introduction to Theory and Applications. Pearson Education. 1998.
[18] Rao, T. S., Eduarda, S. Identification of Bilinear Time Series Models. Statistica Sinica .1992:2: 465 - 478.
[19] Sesay, S. A. O., Rao, T.S. Yule-Walker Type Difference Equations for Higher Order Moments and Cumulants for Bilinear Time Series Models. Journal of Time Series Analysis.1988: 9: 385-401.
[20] Soltani, S. On the use of the wavelet decomposition for time series prediction. Elsevier, Neurocomputing.2002: 48: 267-277.
[21] Strang. Long-Term Prediction, Chaos and Artificial Neural Networks. Journal of Roayal Statistical Society. 1998.
[22] Sysel, P., Misurec, J. Estimation of Power Spectral Density using Wavelet Thresholding. Proceedings of the 7th Conference on Circuits. Systems, Electronics, Control and Signal Processing, 2008: p.207-211. Tenerife Canary Islands, Spain.
[23] Wei, W.W.S. Time Series Analysis Univariate and Multivariate Methods. Addison-Wesley Publishing Company. 1990.
[24] Mallat, S. A wavelet tour of signal processing. Academic Press. 1999.
[25] Christopouloul, E.B., Skodras, A. N., Georgakilas, A. A. Time Series Analysis of Sunspot Oscillations Using the Wavelet Transform. Digital Signal Processing. 2002:2: pp.893-896.
[26] www.FreeLunch.com-http://www.eonomy.com/freelunch.

Date of Submission: 27-11-2009