Analysis of Non-Stationary Time Series using Wavelet Decomposition

Lineesh M C*1, C Jessy John2

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 $Z_t$ into 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 nonlinear 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 Autoregressive (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)

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
\[
\{Z_t : t = 0, 1, 2, ..., N - 1\}
\]
can be decomposed as
\[
Z_t = X_t + Y_t, \ t = 0, 1, 2, ..., N - 1
\]
where \( X_t \) is the trend series and \( Y_t \) is the detail series given by,
\[
Y_t = \sum_{j=1}^{M} d_{j,t}
\]
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 Autoregressive Model Using Wavelet Techniques

The \( l \) threshold autoregressive model of order, \( k \) i.e. TAR (k) model (Priestley, 1988) is defined as,
\[
Y_t = a_0^{(j)} + \sum_{i=1}^{k} a_i^{(j)} Y_{t-i} + e_t^{(j)}
\]
where \( Y_{t-j} \in R^{(j)} \) for \( j = 1, 2, ..., l, \ R^{(j)} \)
being a given subset of the real line \( R^l \). In (2), \( e_t^{(j)}, \ j = 1, 2, ..., l \) and \( t \in Z \) (the set of integers) is a sequence of independent and identically distributed random variables with 0 mean, constant variance \( \sigma^2 \) and \( a_i^{(j)}, 1 \leq i \leq k \) 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;
\[
\phi_{j,k}(t) = 2^{-j/2} \phi(2^{-j} t - k), \quad j = 1, 2, ..., J; \ k = 0, 1, ..., 2^j - 1.
\]
\[
\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j} t - k), \quad j = 1, 2, ..., J; \ k = 0, 1, ..., 2^j - 1.
\]
where
\[
\phi(t) = \begin{cases} 
-2^{j} & \text{if } 2^j k \leq t < 2^j (k + 1/2) \\
2^{j} & \text{if } 2^j (k + 1/2) \leq t \leq 2^j (k + 1)
\end{cases}
\]
(4)

and
\[
\psi(t) = \begin{cases} 
2^{j} & \text{if } 2^j k \leq t < 2^j (k + 1/2) \\
-2^{j} & \text{if } 2^j (k + 1/2) \leq t \leq 2^j (k + 1)
\end{cases}
\]
(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)
\[
Y_t = \sum_{j=1}^{J} \sum_{k=0}^{j-1} \beta_{j,k} \psi_{j,k}(t)
\]
(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
\[
\lambda_j = \sqrt{2 \log \#d_{j,1}}, \quad \text{where} \quad \#d_{j,1}
\]
denotes the cardinality of \( \{d_{j,1}\} \). Also define
\[
\lambda = \sqrt{2 \log \#C_{M,1}}
\]

Here \( \lambda_j \) denotes the threshold for the \( j^{\text{th}} \) level detail series and \( \lambda \) denotes the threshold of the TAR model.

3.2.2 Estimation of TAR model

The Threshold Autoregressive model representing the detail series \( \{Y_t\} \) is given by,
\[
Y_t = \begin{cases} 
\sum_{j=1}^{J} \sum_{k=0}^{j-1} d_{j,k} \psi_{j,k}(t) & \text{if } Y_{t-d} < \lambda \\
T(t) + b_1 Y_{t-1} + b_2 Y_{t-2} + \ldots + b_k Y_{t-k} + e_t(1) & \text{if } Y_{t-d} \geq \lambda
\end{cases}
\]
(8)

where the coefficients \( \{b_{j,1}\} \) are defined by,
\[
b_{j,1} = \left( \sum_{j=1}^{J} \sum_{k=0}^{j-1} d_{j,k} \psi_{j,k}(t) \right) \psi_{j,1}(t)
\]
(9)

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,
\[
Z_t = \begin{cases} 
T(t) + b_1 Y_{t-1} + b_2 Y_{t-2} + \ldots + b_k Y_{t-k} + e_t & \text{if } Y_{t-d} < \lambda \\
T(t) + b_1 Y_{t-1} + b_2 Y_{t-2} + \ldots + b_k Y_{t-k} + e_t(2) & \text{if } Y_{t-d} \geq \lambda
\end{cases}
\]
(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.

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

Using the existing method the model representing the time series 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.

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.
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>
<thead>
<tr>
<th>Threshold</th>
<th>Estimated Model</th>
<th>MAPE</th>
<th>MSE</th>
</tr>
</thead>
</table>
| 3.33      | \( Z_t = \begin{cases} 
0.99X_{t-1} - 0.003X_{t-2} - 5.6Y_{t-1} \\ 
- 8.89Y_{t-2} - 21.7Y_{t-3} + e_t^{(1)} \\ 
0.99X_{t-1} - 0.003X_{t-2} + 4.28Y_{t-1} \\ 
+ 7.09Y_{t-2} + 5.48Y_{t-3} + e_t^{(2)} \end{cases} \) if \( Y_{t-1} < 3.33 \) | 0.7901 | 4.4814 |

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Estimated Model</th>
<th>MAPE</th>
<th>MSE</th>
</tr>
</thead>
</table>
| 35        | \( X_t = \begin{cases} 
0.539X_{t-1} - 0.196X_{t-2} + 0.483X_{t-3} + e_t^{(1)} \\ 
+ 0.542X_{t-1} - 0.127X_{t-2} \\ 
+ 0.017X_{t-3} + 0.051X_{t-4} \\ 
+ 0.029X_{t-5} + 0.45X_{t-6} + e_t^{(2)} \end{cases} \) if \( X_{t-2} < 35 \) | 3.1828 | 6.5317 |
Table 3: Estimated T-TAR model for the stock exchange time series

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Estimated Model</th>
<th>MAPE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.52</td>
<td>[ Z_t = \begin{cases} 0.99X_{t-1} - 0.000157X_{t-2} &amp; \text{if } Y_{t-1} &lt; 3.52 \ -0.79Y_{t-1} - 1.258X_{t-2} + e^{(1)}<em>t \ 0.99X</em>{t-1} - 0.000157X_{t-2} \ +1.12Y_{t-1} + 2.499Y_{t-2} + e^{(2)}<em>t &amp; \text{if } Y</em>{t-1} \geq 3.52 \end{cases} ]</td>
<td>1.12</td>
<td>0.48</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Estimated Model</th>
<th>MAPE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>[ 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} &amp; \text{if } X_{t-3} &lt; 90 \ -0.49X_{t-10} + 0.58X_{t-11} - 0.28X_{t-12} \ +0.4X_{t-13} + e^{(1)}<em>t \ 1.169X</em>{t-1} - 0.16X_{t-2} + e^{(2)}<em>t &amp; \text{if } X</em>{t-3} \geq 90 \end{cases} ]</td>
<td>1.9164</td>
<td>0.5215</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Estimated Model</th>
<th>MAPE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.744</td>
<td>[ Z_t = \begin{cases} 0.99X_{t-1} + 0.0000274X_{t-2} \ -3.02Y_{t-1} - 2.659Y_{t-2} &amp; \text{if } Y_{t-1} &lt; 8.744 \ -8.021Y_{t-3} - 37.34Y_{t-4} + e^{(1)}<em>t \ 0.99X</em>{t-1} - 0.0000274X_{t-2} \ +2.86Y_{t-1} + 2.726Y_{t-2} + 1.9Y_{t-3} &amp; \text{if } Y_{t-1} \geq 8.744 \ +2.36Y_{t-4} + e^{(2)}_t \end{cases} ]</td>
<td>0.0401</td>
<td>27.696</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Estimated Model</th>
<th>MAPE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>560</td>
<td>[ X_t = \begin{cases} 1.293X_{t-1} - 0.293X_{t-2} + e^{(1)}<em>t &amp; \text{if } X</em>{t-1} &lt; 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} &amp; \text{if } X_{t-1} \geq 560 \ -0.106X_{t-7} + 0.257X_{t-8} + e^{(2)}_t \end{cases} ]</td>
<td>1.466</td>
<td>84.482</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Time Series</th>
<th>T-TAR Model</th>
<th>Model due to Priestley</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MAPE</td>
<td>MSE</td>
</tr>
<tr>
<td>1</td>
<td>Sunspot</td>
<td>0.7901</td>
<td>4.4814</td>
</tr>
<tr>
<td>2</td>
<td>Stock Exchange</td>
<td>1.12</td>
<td>0.48</td>
</tr>
<tr>
<td>3</td>
<td>IBM Stock Price</td>
<td>0.0401</td>
<td>27.696</td>
</tr>
</tbody>
</table>
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:

Date of Submission: 27-11-2009