Journal of Statistical and Econometric Methods, Vol. 11, No.3, 2022, 1-12
ISSN: 2241-0384 (print), 2241-0376 (online)
https://doi.org/10.47260/jsem/1131
Scientific Press International Limited

A New Approach to the Threshold Autoregressive Models

Nuri Celik[1]

Abstract

Time series analyzing is very important tool for economic and financial system. However, recent developments show that financial systems are known in a structural change. Therefore, nonlinear time series have been analyzed for past decades because of these changes. In this paper, we consider Threshold Autoregressive (TAR) model. The most popular method for estimating the parameters and threshold value is least square (LS) method. However, LS method is not robust to the outliers and departures from normality. Therefore, we propose a robust version of estimation in order to provide robust results.

JEL classification numbers: C01, C22, C13.
Keywords: Threshold Autoregressive Model, Iterated Weighted Least Square, Skew Normal, Long Tailed Symmetric Distribution, Robustness.

[1] Gebze Technical University, Department of Mathematics.

Article Info: Received: June 29, 2022. Revised: August 3, 2022.
Published online: August 17, 2022.

1. Introduction

In Statistical applications, linear time series models have been widely used over the past decades. However, in real world, nonlinear cases give much better modeling and solutions. Because of this reason, nonlinear time series models have been studied in economic and statistics literature. Threshold Autoregressive (TAR) models are quite popular in nonlinear time series modeling. This popularity comes from its easy calculation and estimation according other nonlinear alternatives. At first TAR models had not been widely used because of its adversity to identify and estimate threshold value and modeling. However, having been proposed more simple procedures, they have been become the most popular nonlinear models used in economic and statistical literature. Now it becomes a more or less standard model in nonlinear time series and has been widely used in diverse areas, including biological sciences, econometrics, environmental sciences, finance, hydrology, physics, and population dynamics (Li et al., 2011).

The class of TAR models was firstly introduced to literature by Tong (1978, 1983), and Tong and Lim (1980) as an alternative model for describing time models. TAR models have good properties that cannot be captured by linear time series models, like limit cycles, amplitude dependent frequencies and jump phenomena (Tsay, 1989). A time series $\{y_t\}$ is said to be TAR model with *k* regimes if it satisfies,

$$y_t = \Phi_0^{(j)} + \sum_{i=1}^{p_i} \Phi_i^{(j)} y_{t-i} + \varepsilon_t^{(j)} \quad r_{j-1} \leq y_{t-d} < r_j, \;\; j = 1,2,\dots,k \qquad (1)$$

where, ε_t is identically and independently distributed error term, *k* is the number of regimes, *d* is the delay parameter and p_i is the order of AR process in i^{th} regime. It should be also noted that the orders *(p)* of AR models may differ from regime to regime. Additionally, the TAR model becomes a nonhomogeneous linear AR model when the variance of error terms is different in each regime and it reduces to random level shift model if the constants in each regime $\left(\Phi_0^{(j)}\right)$ are different.

In application, it must firstly be detected the nonlinearity in data. For this reason, the nonlinearity tests have been proposed for TAR models; see Tong and Lim (1980), Tsay (1989) and Hansen (1997). In this paper Tsay's approach is used. Because, it is based on simple linear regression techniques and simpler than the method proposed by Tong and Lim (1980). On the other hand, Hansen's method supports only two regime TAR models. Tsay's approach is a combined version of nonlinearity test introduced by Keenan (1985) and Petruccelli and Davies (1986). Its asymptotic distribution under the linear assumption is central *F* distribution.

After detecting nonlinearity, threshold values and delay parameter are estimated in order to go on modeling. There are various forms of detecting threshold values but in this paper, we use again Tsay's approach to detect the delay parameters and threshold value. This method is based on simple scatter diagram and *F* statistics mentioned before. Then, the unknown model parameters are estimated with Least Squares (LS) or Maximum Likelihood (ML) methods when the error terms are

normally distributed with mean 0 and variance σ^2. However, when the normality assumption is not satisfied, LS estimators of parameters and the test statistics based on them lose their efficiency, see Tukey (1960). There are lots of studies in the literature pointing out that nonnormal distributions are more prevalent than normal distribution in practice, see for example, Geary (1947), Huber (1981), Pearson (1932) and Tan and Tiku (1999). For this reason, to solve this problem, robust procedures that are not unduly affected by small departures from normality assumptions are proposed. In this paper, we propose robust method to estimate the model parameters for TAR models.
The rest of the paper organizes as follows; In Section 2, we give details about the robust estimation methods and features of it. We propose a simple algorithm for modeling TAR models with a robust version of the algorithm proposed by Tsay (1989). In Section 3, Monte Carlo simulation study is done in order to compare the proposed method with traditionally used method. A real-life example is given in Section 4 just for illustration. Conclusion is given at the end of the paper.

2. Estimation

Although TAR models are one of the most important nonlinear time series models, in estimation unknown model parameters, LS method can be applied since they are locally linear models. The details and features LS techniques used in TAR models have been argued in different papers; see Tsay (1989), Chan (1993) and Qian (1998). The bad performance of LS estimators for contaminated data shows the necessity of robust estimation methods, methods which are robust toward outliers and wrong specification of the model (Stockinger and Dutter, 1983). In modeling time series data, different type of outlier may be dealt with such as additive outliers (AOs), replacement outliers (ROs) and innovations outliers (IOs). Time series outliers can have an arbitrarily adverse influence on parameter estimates for time series models, and the nature of this influence depends on the type of outlier (Maronna and Zamar, 2002). For this reason, a robust estimation technique is introduced in order to get robust estimators in TAR models.

2.1 M estimation method for AR(p) models

AR(p) models can be represented linear regression models like,

$$\boldsymbol{y} = \boldsymbol{Z\beta} + \boldsymbol{\varepsilon} \tag{2}$$

where $\boldsymbol{\beta} = \left(\Phi_0, \Phi_1, \dots, \Phi_p\right)^T$ parameter vector, $\boldsymbol{y} = \left(\mathrm{y}_{p+1}, \mathrm{y}_{p+2}, \dots, \mathrm{y}_n\right)^T$ observations, $\boldsymbol{\varepsilon} = \left(\varepsilon_{p+1}, \varepsilon_{p+2}, \dots, \varepsilon_n\right)^T$ error terms and $\boldsymbol{Z} = \left(\mathrm{Z}_{p+1}, \mathrm{Z}_{p+2}, \dots, \mathrm{Z}_n\right)^T$.

An M-estimator of $\widehat{\boldsymbol{\beta}}$ is defined by

$$\sum_{\mathrm{i=p+1}}^{\mathrm{n}} \mathrm{Q}\left(\frac{y_i - z_i^T \beta'}{\hat{\sigma}}\right) = min \tag{3}$$

where *Q(.)* is a loss function and $\hat{\sigma}$ is an estimate of the scale parameter. The loss function is often used in the form of its first derivative $\psi(t) = \frac{dQ(t)}{dt}$. There is various form of $\psi(t)$ function in the literature, in this paper we use Tukey's $\psi(t)$ function introduced in Huber, 1981.

$$\psi(t) = \begin{cases} t\left(1 - \left(\frac{t}{c}\right)^2\right)^2 & |t| \leq c \\ t & |t| > c \end{cases} \tag{4}$$

According to its purpose, a $\psi(t)$ function should be odd, bounded and continuous. The following iterated weighted least squares (IWLS) algorithm can be used for estimating $\hat{\beta}$ and $\hat{\sigma}$ simultaneously. A convergence proof for the estimation of linear models is given by Dutter (1975).

2.2 IWLS Algorithm

- Identify the initial values $\beta^{(0)}$,$\sigma^{(0)}$ and tolerance level ϵ,
- Set the iteration *m=0*,
- Denote $r_i^{(m)} = y_i - Z_i^T \boldsymbol{\beta}^{(m)} \quad i = p+1, \dots, n$
- Compute σ using

$$\left(\sigma^{(m+1)}\right)^2 = \frac{1}{c} \sum_{i=p+1}^{n} \chi\left(\frac{r_i^{(m)}}{\sigma^{(m)}}\right) \left(\sigma^{(m)}\right)^2$$

where $\chi(t) = t\psi(t) - Q(t)$.

- Calculate weights

$$w_i^{(m)} = \begin{cases} \psi\left(\frac{r_i^{(m)}}{\sigma^{(m+1)}}\right) / \left(\frac{r_i^{(m)}}{\sigma^{(m)}}\right) & r_i^{(m)} \neq 0 \\ 1 & otherwise \end{cases}$$

- Define a diagonal matrix $W^{(m)}$ with w_i as its *(i-p)th* diagonal element
- Solve $\sum_{i=p+1}^{n} \left(r_i^{(m)} - Z_i^T \tau^{(m)}\right)^2 w_i^{(m)}$=*min*

where

$$\tau^{(m)} = \left(Z^T W^{(m)} Z\right)^{-1} Z^T W^{(m)} y - \boldsymbol{\beta}^{(m)}$$

- Compute new value of $\boldsymbol{\beta}$

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} + w\tau^{(m)}$$

where *0<w<2* is an arbitrary relaxation vector.

- Stop if $|\sigma^{(m)} - \sigma^{(m+1)}| < \epsilon\sigma^{(m+1)}$
- Augment *m=m+1* and go to step 3.

Martin (1978) showed that under regularity conditions M estimators obtained by using IWLS are consistent and asymptotic normal.

2.3 Robust autocorrelation and partial autocorrelation

The sensitivity of the traditional estimators, the sample autocorrelation functions (acf) and partial autocorrelation functions (pacf), to outliers is well known (see Chan, 1993, Deutsch et al., 1990, or Maronna et al., 2006). The acf and pacf of time series (Y_t) simply can be notated $Cov(Y_{t+h}, Y_h) = \gamma(h)$ and $Cor(Y_{t+h}, Y_h) = \rho(h)$ for all $t, h \in Z$. It should be also noted that $\rho(h) = \frac{\gamma(h)}{\gamma(0)}$ where $\gamma(0)$ is the variance. Several robust alternatives have been proposed in the literature. Dürre et al., (2014) have resulted that the estimation approach based on robust scale estimators give better solutions. The estimation simply can be found like

$$\rho(h) = \frac{Var(Y_{t+h}+Y_t) - Var(Y_{t+h}-Y_t)}{Var(Y_{t+h}+Y_t) + Var(Y_{t+h}-Y_t)} \tag{5}$$

Maronna and Zamar (2002) recommended the τ estimator which is obtained by

$$\hat{\sigma}^2 = \frac{\hat{\sigma}_0^2}{n} \sum_{i=1}^{n} d_{c2}\left(\frac{y_i - \hat{\mu}}{\sigma_o}\right) \tag{6}$$

where $\hat{\mu}$ is a weighted mean of the observations, σ_o is the ordinary MAD and $d_c(x) = min(x^2, c^2)$. In Maronna and Zamar (2002) tuning constants $c_1 = 4.5$ (for $\hat{\mu}$) and $c_2 = 3$.

As a result, we propose a robust version of the algorithm given by Tsay (1989).

2.4 Algorithm for modeling TAR models

Step 1: Select the AR order p by using robust pacf and the set of possible thresholds *S*.

Step 2: Perform a nonlinearity test for a given *p* and every element d of *S*, and if nonlinearity is detected select the delay parameter by using *F* statistics described in Tsay (1983).

Step 3: For a given *p* and *d*, locate the threshold values by using the scatter diagram of *t* ratios.

Step 4: Estimate the unknown model parameters via IWLS.

3. Simulation Study

In simulation study, among nonnormal distributions used in statistics, we use long tailed symmetric distribution and Azzalini's skew normal distribution (Azzalini, 1985, 1986). Because, both distribution cover various type of distribution used in application. Long tailed symmetric distribution which has probability density function (pdf),

$$f(e)=\frac{1}{\sigma}\left\{1+\frac{e^2}{k\sigma^2}\right\}^{-\lambda} \qquad -\infty<e<\infty,\ \ k=2\lambda-3, \lambda\geq 2\ . \tag{7}$$

It may be noted that $t=\sqrt{v/k}\left(e/\sigma\right)$ has Student's t distribution with $v=2\lambda-1$ degrees of freedom. The variance of the distribution (2) is σ^2 for all $(\lambda\geq 2)$.

The pdf of the skew normal distribution, on the other hand, is given by

$$h(z)=2\phi(z)\Phi(\lambda z),\quad -\infty<z<\infty,\quad -\infty<\lambda<\infty \tag{8}$$

where $\phi(z)$ and $\Phi(z)$ are the pdf and the cumulative distribution function (cdf) of the standard normal distribution, respectively. λ is the skewness parameter. It determines the shape of the distribution. We compare the M estimators and the LS estimators of the model parameters in terms of means, variances and mean square errors (MSE) for some representative alternative models. All the simulations are based on $[1000,000/n]$ Monte Carlo runs. In the simulation study, we take TAR model with 2 regimes and in each regime we take AR(1) models for the sake of brevity. Threshold value is determined 0 and the delay parameter is 1.

Without loss of generality, we choose the following setting in our simulation:

$$\Phi_0^{(1)}=\Phi_0^{(2)}=0, \qquad \Phi_1^{(1)}=0.5, \Phi_1^{(2)}=0.4, \quad \sigma=1$$

3.1 Alternative Models

Model (1): Dixon's outlier model: (n-1) observations come from N(0,1) but one observation (we do not know which one) comes from N(0,10).
Model (2): Dixon's outlier model: (n-1) observations come from N(0,1) but one observation (we do not know which one) comes from N(0,10).
Model (3): All observation comes from SN(0,1,1).
Model (4): All observation comes from LTS(3,3,1).
Model (5): Mixture model: 0.90 SN $(0,1,1)$+0.10 SN $(0,1,0.4)$
Model (6): Contamination model: 0.90 SN $(0,1,1)$+0.10 N$(0,1)$.

Simulation results are given in Table 1. Relative Efficiencies (RE) are calculated with the division of two MSE's.

Table 1: Means, variances and MSEs for the LS and M estimators for $\boldsymbol{\Phi}_1^{(1)}$, $\boldsymbol{\Phi}_1^{(2)}$ and $\boldsymbol{\sigma}$.

	Mean ($\boldsymbol{\Phi}_1^{(1)}$)		Mean ($\boldsymbol{\Phi}_1^{(2)}$)		Mean($\boldsymbol{\sigma}$)		RE ($\boldsymbol{\Phi}_1^{(1)}$)	RE ($\boldsymbol{\Phi}_1^{(2)}$)	RE ($\boldsymbol{\sigma}$)
n	LS	M	LS	M	LS	M			
Model 1									
25	0.522	0.515	0.438	0.413	1.313	1.242	88	89	94
50	0.519	0.508	0.429	0.409	1.201	1.138	87	88	90
100	0.508	0.502	0.421	0.405	1.153	1.094	83	81	88
Model 2									
25	0.615	0.554	0.501	0.439	2.178	1.851	79	76	84
50	0.571	0.531	0.478	0.421	1.738	1.522	75	69	83
100	0.546	0.519	0.462	0.418	1.565	1.367	68	64	80
Model 3									
25	0.426	0.447	0.321	0.377	1.024	0.988	91	90	90
50	0.454	0.487	0.358	0.382	1.044	0.990	87	89	90
100	0.479	0.509	0.378	0.389	1.045	0.992	81	84	90
Model 4									
25	0.567	0.539	0.481	0.456	1.067	1.049	89	90	89
50	0.546	0.521	0.467	0.423	1.054	1.041	88	87	88
100	0.532	0.511	0.444	0.419	1.048	1.028	82	83	81
Model 5									
25	0.586	0.546	0.501	0.467	1.046	0.994	81	82	91
50	0.579	0.537	0.477	0.451	1.063	1.009	78	77	89
100	0.565	0.523	0.453	0.439	1.071	1.016	70	70	88
Model 6									
25	0.554	0.529	0.458	0.431	1.074	1.015	86	89	90
50	0.543	0.511	0.441	0.428	1.081	1.025	82	87	88
100	0.528	0.507	0.438	0.419	1.088	1.034	77	82	86

It can be seen from Table 1; M estimators are more efficient and robust from the LS estimator for simulated TAR models. It should be also noted that the same results are obtained for different true models (3 regimes, different delay parameters and AR orders) and different true parameter values.

4. Application (Help Wanted Index Data)

As an example, it is used Help Wanted Index (HWI) data from January 1997 to January 2020 in USA (https://www.conference-board.org/data/). The scatter plot of the data with the transformation of $\log\left(Y_{t-1}/Y^{t}\right)$ is shown in Figure 1.

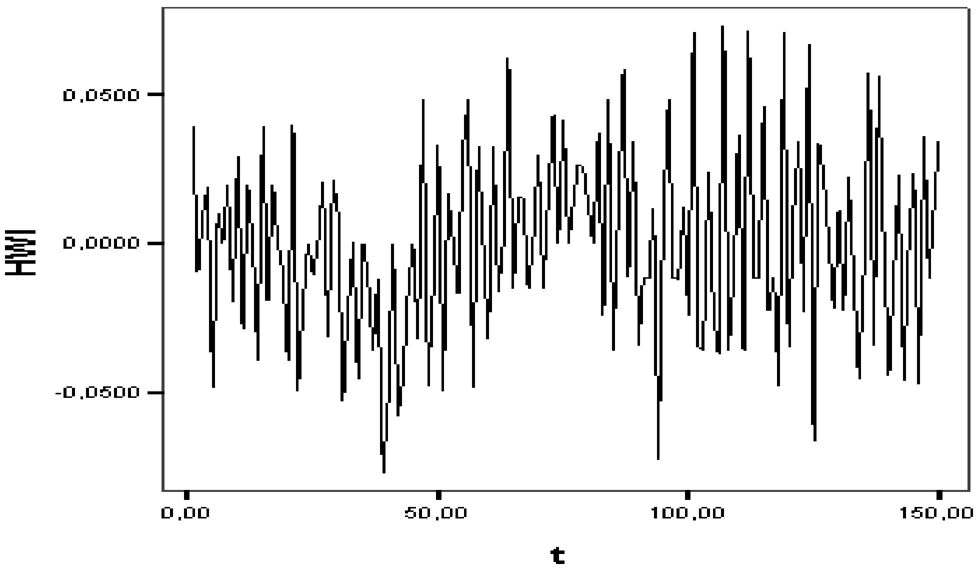

Figure 1: Transformed Help Wanted Index Data

In order to perform nonlinearity, test we must detect the order of AR model. We use robust PACF for detecting the order *p*. According to the robust PACF values of the data *p=2* is found as an order. Then, the *F* statistics proposed by Tsay (1989) confirms that the process is nonlinear. The second step is detecting delay parameter. We use again Tsong's technique in order to determine d. As proposed in the paper, maximum *F* statistics provided by the combination of *(p,d)* should be selected as a delay parameter. We select Y_{t-2} as the threshold value. Therefore, the scatter plot *t* ratios versus Y_{t-2} values suggest that there is one significant change which means we have TAR model with two regimes. From the plot it can be easily seen that the threshold is 0.

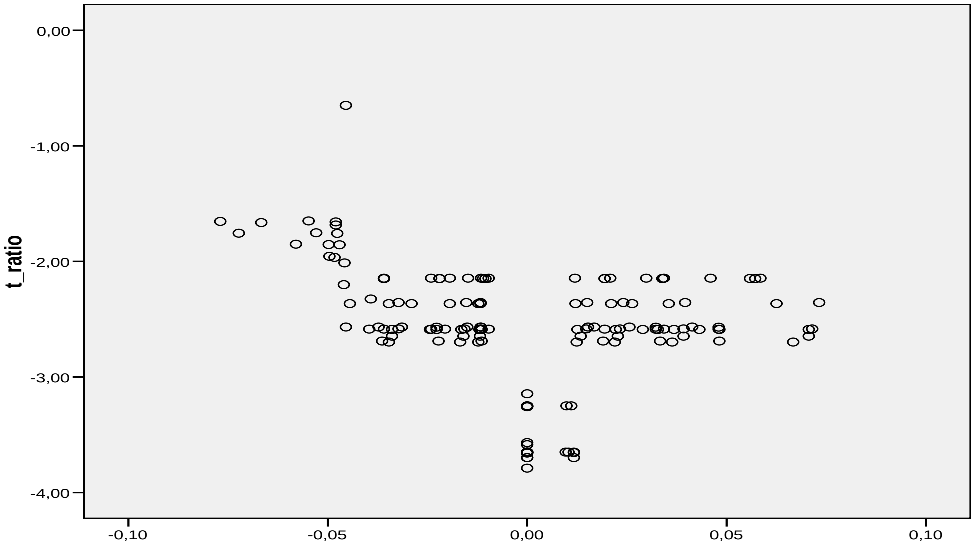

Figure 2: Scatter plot of t ratios versus Y_{t-2}

For the next step, we estimate the model parameters by using robust technique as shown in Section 2. To identify the distribution of the error terms, we use Q-Q plot technique which is one of the well-known and widely used graphical techniques. On the other hand, among the Q-Q plots of the residuals obtained for various different values of the skewness parameter λ, $SN(\mu,\sigma,\lambda=1)$ adequately models the residuals, since the observations do not deviate very much from the straight line, see Figure 3.

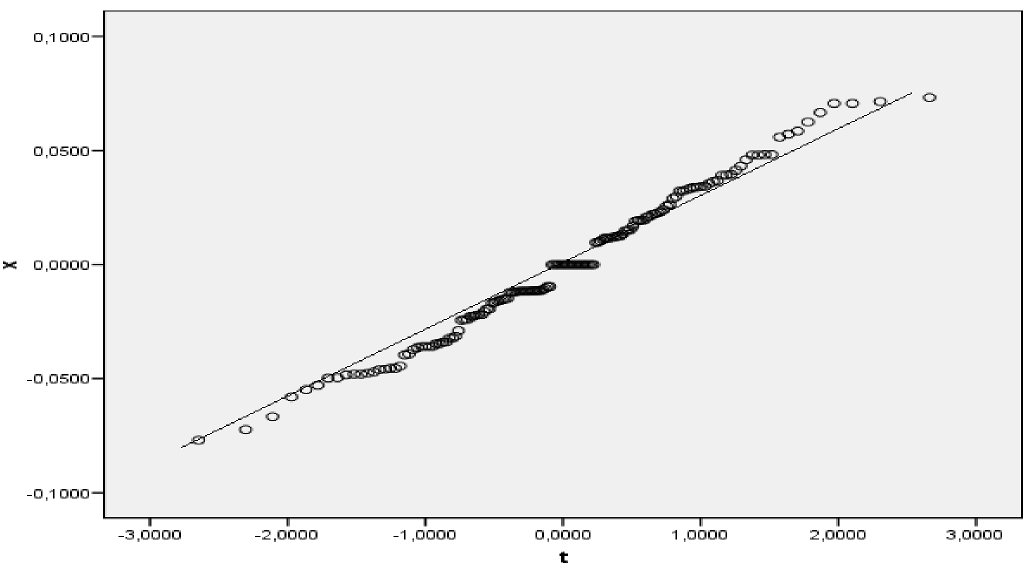

Figure 3: Q-Q plot for transformed help wanted index data

When we take the skewness parameter λ as 1, the parameter estimates are obtained by using LS estimation method and M-estimation method as shown below respectively.

$$Y_t = \begin{cases} 0.775Y_{t-1} + 0.087Y_{t-2} + \varepsilon_t & Y_{t-2} < 0 \\ -0.543Y_{t-1} + 0.264Y_{t-2} + \varepsilon_t & Y_{t-2} > 0 \end{cases}$$

and

$$Y_t = \begin{cases} 0.678Y_{t-1} + 0.181Y_{t-2} + \varepsilon_t & Y_{t-2} < 0 \\ -0.457Y_{t-1} + 0.342Y_{t-2} + \varepsilon_t & Y_{t-2} > 0 \end{cases}$$

AIC information the R^2 values for model estimation are shown in Table 2.

Table 2: AIC information the R^2 values calculated by the LS and M estimation methods

	R^2	**Akaike Information**
LS	0.842	-8.454
M	0.891	-9.567

In model checking, the ACF and PACF of the standardized residuals and squared standardized residual of the model all fail to suggest any model inadequacy. The coefficient of determination value obtained by the M-estimation method is higher than the value obtained by the LS estimation method. On the other hand, the absolute value of AIC value is higher corresponding to the normal theory. In addition, it should be noted that the parameter estimations calculated by using M-estimation have small standard errors than the estimations calculated by LS method. Therefore, the model obtained by M-estimation method is more reliable than the model obtained by LS estimation method.

5. Conclusion

Time series analyzing are very important tool for economic and financial system. However, recent developments show that financial systems are known in a structural change. Therefore, nonlinear time series analyzing have been analyzed for past decades because of modeling these structural changes. The most popular nonlinear time series models are threshold autoregressive models. However, despite of its popularity, there is only small development in TAR models. Tsay (1989) proposed an easy way for estimating the regimes, delay parameters and threshold values. The method is easy to understand and applicable to various types of data. In this paper, we proposed a robust way to estimate the model parameters, and then we combine the technique with robust methodology. The simulation study shows that the robust method is more robust and efficient than the traditionally used LS method. Therefore, in order to get rid of the normal distribution assumption difficulties and the problems of outlier, the proposed technique is more feasible.

References

[1] Azzalini, A. (1985). A class of distributions which includes the normal ones. Scand. Journal of Statistics. 12, pp. 171-178.

[2] Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones. Statistica, 46, pp. 199-208.

[3] Chan, K. S. (1993). Consistency and limiting distribution of the least squares' estimator of a threshold autoregressive model. Ann. Statist. 21, pp. 520–533.

[4] Deutsch, S. J., Richards, J. E., and Swain, J. J. (1990). Effects of a single outlier on ARMA identification. Communications in Statistics: Theory and Methods, 19(6), pp. 2207–2227.

[5] Dutter R. (1975): Robust Regression: Different Approaches to Numerical Solutions and Algorithms. Res. Rep. 6, Fachgruppe f. Statist., Eidgen. Techn. Hochsch., Zürich.

[6] Dürre, A., Fried, R. and Liboschik, T. (2014). Robust estimation of (partial) autocorrelation, SFB discussion paper, TU Dortmund

[7] Geary, R.C. (1947). Testing for normality, Biometrika. 34, pp. 209-242.

[8] Hansen, B.E. (1997). Inference in TAR models. Studies in Nonlinear Dynamics and Econometrics 2: 1–14. Berkeley Electronic Press, Berkeley.

[9] Huber, P.J. (1981). Robust Statistics, Jonh Wiley, New York, 1981.

[10] Keenan, D.M. (1985). A Tukey Nonadditivity-Type Test for Time Series Nonlinearity. Biometrika, 72, pp. 39-44.

[11] Li, D., Li, W.K. and Ling, S. (2011), On the least square estimation of threshold autoregressive and moving average models, Statistics and Inferences, 4, pp. 183- 196

[12] Maronna, R. A., Martin, R. D., and Yohai, V. J. (2006). Robust statistics. J. Wiley,Chichester.

[13] Maronna, R. A. and Zamar, R. H. (2002). Robust estimates of location and dispersion for high-dimensional datasets. Technometrics, 44(4), pp. 307–317.

[14] Martin, R.D (1978). Asymptotic Properties of M-estimates for p-th order Autoregressions. Techn. Rep. 212 Dept. Electrical Engineering. Univ Washington, Seattle

[15] Pearson, E.S (1932). The analysis of variance in cases of nonnormal variation, Biometrika, 23, pp.114-133.

[16] Petruccelli, J., and Davies, N. (1986). A Portmanteau Test for Self- Exciting Threshold Autoregressive-Type Nonlinearity in Time Series. Biometrika, 73, pp. 687-694.

[17] Qian, L. (1998). On maximum likelihood estimators for a threshold autoregression. J. Statist. Plann. Inference 75, pp. 21–46.

[18] Stockinger N. and R. Dutter (1983): Robust Time Series Analysis — An Overview. Res. Rep. 9, Inst. Statist., Techn. Univ. Graz.

[19] Tan, W.Y. and Tiku, M.L. (1999). Sampling distributions in terms of Laguerre Polynomials with applications, New Age International (formerly, Wiley Eastern), New Delhi.

[20] Tong, H. (1978). On a threshold model, in pattern recognition and signal processing, edited by C.H Chen, Amsterdam Kluwer.
[21] Tong, H., and Lim, K. S. (1980). Threshold Autoregression, Limit Cycles and Cyclical Data (with discussion). Journal of the Royal Statistical Society, Ser. B, 42, pp. 245-292.
[22] Tong, H. (1983). Threshold models in non-linear time series analysis, New York: Springer.
[23] Tsay, R.C. (1989). Testing and Modeling Threshold Autoregressive Process, Journal of American Statistical Association, 84, pp. 231-240.
[24] Tukey, J.W. (1960). A survey of sampling from contaminated distributions, Contributions to Probability and Statistics, I. Olkin (ed.), Stanford, CA: Stanford University Press.