ESKİŞEHİR TECHNICAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY
A- APPLIED SCIENCES AND ENGINEERING

2022, 23(1), pp. 48-69, DOI: 10.18038/estubtda.881251

RESEARCH ARTICLE

SOME REGIME-SWITCHING MODELS FOR ECONOMIC TIME SERIES: A COMPARATIVE STUDY

Dursun AYDIN [1], **Selman MERMİ** [2,*]

[1] Department of Statistics, Faculty of Science, Muğla Sıtkı Koçman University, Muğla, Turkey
[2] Department of Statistics, Faculty of Science, Muğla Sıtkı Koçman University, Muğla, Turkey

ABSTRACT

This paper mainly discusses some regime-switching models and explore their usefulness in modeling the economic time series. In recent years, several time series models have been proposed which shape the idea of the existence of different regimes produced by a stochastic process. Especially, nonlinear time series models have gained more attention because linear time series models faced various limitations. The purpose of this study is to establish the methodology of the Self-Exciting Threshold Autoregressive (SETAR) model, Smooth Transition Autoregressive (STAR) model and Markov-Switching (MSW) model from parametric nonlinear time series models in the mean and to compare these models with each other through two financial data sets. For this purpose, some theoretical information on the subject models are given without going into too much detail. In the light of the obtained theoretical information, all models are modeled by using two financial data sets. The obtained models are compared with the help of some performance criteria, measurement of relative efficiency and graph showing the relation of the actual-fitted values of the models.

Keywords: Markov-Switching Model, Nonlinear Time Series Models, SETAR, STAR

1. INTRODUCTION

Time series analysis can be examined in two groups as linear and nonlinear time series analysis. Although linear time series analysis is more preferred because of its ease of theory and application, in some cases the use of nonlinear time series analysis is inevitable. The analysis of econometric time series with nonlinear models implies that some characteristics such as average variance and autocorrelation of the series are variable over time. Since many econometric time series exhibit these characteristics, it necessitates the use of nonlinear time series analysis. In other words, many econometric time series show severe breaks in their behavior from time to time, usually in response to events such as financial crises or rapid changes in government policy. In order to reflect the state of the underlying economy, the general attitude of market investors and other economic aspects, regime switching models became necessary at this stage. Several nonlinear time series models are presented in the literature of statistics. These models are exemplified as bilinear models by Granger and Andersen [1], the Threshold Autoregressive model (TAR) by Tong [2], Smooth Transition Autoregressive (STAR) Models by Terasvirta and Anderson [3] and the Markow-Switching (MSW) model by Hamilton [4]. Tong and Lim [5], Chan and Tong [6], Tsay [7] and Tong [8] have argued that a single econometric equation is not sufficient to model a financial series. Hence, they argued that linearization of the series section by section is more appropriate by using certain threshold values in the nonlinear time series.

The nonlinear time series literature is born as TAR, STAR, SETAR, LSTAR and MSW models and these models are enriched by various studies. These regime-switching models are divided into two basic groups according to the determination of the regime. In the first group, TAR and STAR models are available. Regimes in TAR and STAR models are determined utilizing a variable that can be observed. It is known exactly where the regimes determined by statistical methods coincide over time. The second

*Corresponding Author:selmanmermi@mu.edu.tr
Received: 16.02.2021 Published:30.03.2022

group of models is the MSW model. In the MSW model, the transition between regimes is not observable but it can be determined through a stochastic process that cannot be observable. In this case, it is not known exactly where the regime is at the time, only probabilities are assigned for different regime occurrences [9, 10].

Studies in the literature contain a gap in comparison of regime switching models in nonlinear time series models. Nonlinear time series models are often compared with linear time series models. In this respect, this study is one of the pioneering studies in the literature in terms of comparing nonlinear time series models including SETAR, STAR and MSW models through nonlinear financial data sets. Accordingly, the primary purpose of this paper is to examine the theoretical framework of nonlinear time series models based upon the regime-switching in time series analysis from the current literature and the application of these nonlinear time series models to two financial data sets and as a result to compare the efficiencies of the models among themselves. Thus, researchers will be able to find an answer to the question of which model may be more effective in nonlinear time series modeling.

The paper is structured as follows: we firstly describe the SETAR model and STAR model based on the deterministic transition between regimes and Markov Switching model (MSW) based on the stochastic process in the transition between regimes. Secondly, we apply these models to sales prices of gold bar (TL/gr) and sales prices of US dollar (USD/TL). Then we use scale-dependent measures and measures based upon percentage errors in order to compare the obtained models. Measurement of relative efficiency is also used for comparing the models. Finally, in the light of obtained results, we interpret the efficiencies of estimated models and compare findings.

2. REGIME-SWITCHING MODELS

In this part of the paper, the two types of regime-switching models are discussed, as well as their basic characteristics. We restrict our attention to models covering only two regimes to make the paper easier to understand. Some explanations on the extension of the models to allow more than one regime are given in the next part. It should also be noted that, throughout this article, $\{y_t\}$ symbolizes observable univariate time series of interest which is assumed to be ergodic and stationary. But it does not require these models to be stationary in each regime. Besides, $\{\varepsilon_t\}$ are i.i.d. random variables with zero mean and unit variance.

2.1. Models with Regimes Determined by Observable Variables

The Threshold Autoregressive (TAR) model is the most important member of these models, which assumes that the regime occurs at a time t that can be determined by an observable variable. This model initially developed by Tong [2], and Tong and Lim [5] and it is also discussed detailedly in the study of Tong [8].

When considering a TAR model, it is assumed that regimes can be established by a threshold variable (or transition variable) associated with the threshold value c. If the threshold variable z_t is taken to be lagged values of the time series y_t (i.e., $z_t = y_{t-d}$) for a certain positive integer d, then the time series itself determines the regime. This type of model is called the SETAR model.
The SETAR model with two regimes for a stationary time series y_t can be defined as

$$y_t = \left[\emptyset_{1,0} + \sum_{i=1}^{p_1} \emptyset_{1,i}\, y_{t-i}\right]\left(1 - I(y_{t-d} > c)\right) + \left[\emptyset_{2,0} + \sum_{j=1}^{p_2} \emptyset_{2,j}\, y_{t-j}\right] I(y_{t-d} > c) + \varepsilon_t \quad (1)$$

where I is an indicator function expressed as $I = \{1 \; if \; z_t = y_{t-d} > c \; and \; 0 \; otherwise\}$ and $\varepsilon_t's$ are the error terms sequence of independent and identically distributed (i.i.d.) variables with mean zero and constant variance, p_1 and p_2 are the degrees of the autoregressive (AR) model in lower and upper regimes, respectively. It also should be noted that z_{t-d} is the threshold variable by assuming that $z_{t-d} = y_{t-d}$, d and c are the delay parameter and threshold value, respectively in Equation (1) (see [11] for a detailed discussion). Here it is understood in Equation (1) that a SETAR model is also a piecewise linear AR model in the two regimes defined by the threshold variable y_{t-d} and c.

To estimate the unknown parameters in Equation 1, Tong [12] constructed maximum likelihood estimation method by using Akaike Information Criteria (AIC) when there are a finite number of possible threshold values and Gaussian errors are assumed. Chan et al. [13] obtained the consistency and asymptotic normality property of the least-squares estimators of the coefficient parameter $\emptyset_c = (\emptyset_1', \emptyset_2')'$ under some regularity conditions for the case of first-order autoregression models in each regime. However, in practice, the threshold parameter c is undefined and takes infinite values in $\Re$. In this case, Petruccelli [14] proved that the conditional least squares estimator of $\emptyset$ for the SETAR (2,1,1) model is strongly consistent. Afterward, Chan [15] revealed that the conditional least squares estimator of a stationary ergodic TAR model is strongly consistent and the estimator of the threshold parameter is N consistent for the SETAR (2, p, p) model. From this point forth, this study is motivated by [15] in point of the estimation of unknown parameters in the SETAR model. After important results of [15], Hansen [16] showed that the threshold effect (the disparity between two regimes' slopes) becomes small while the sample size goes to infinity.

The SETAR model assumes that the threshold value c determines the boundary between the two regimes. The gradual transition between the regimes can be provided by the following way: If the indicator function $I(y_{t-d} > c)$ in SETAR model given in Equation (1) is replaced by a continuous function $GF(s_t, \gamma, c)$, which changes from 0 to 1 in a smooth manner, the resulting model is called the STAR model and is given by

$$y_t = \left[\varphi_{1,0} + \sum_{i=1}^{p_1} \varphi_{1,i} y_{t-i}\right]\left(1 - GF(s_t, \gamma, c)\right) + \left[\varphi_{2,0} + \sum_{j=1}^{p_2} \varphi_{2,j} y_{t-j}\right] GF(s_t, \gamma, c) + \varepsilon_t \quad (2)$$

where s_t is a stationary threshold variable, γ is the smoothness parameter, c is the threshold value, $GF(s_t, \gamma, c)$ is a transition function, and ε_t's are the error terms, as defined in Equation (1).
As expressed in Equation (2), a STAR model is defined by two autoregressive regimes connected to each other by a smooth transition function. For notational convenience, the STAR model can also be expressed in a closed-form taking $p_1 = p_2 = p$ as follows.

$$y_t = \varphi_1' x_t \left(1 - GF(s_t, \gamma, c)\right) + \varphi_2' x_t GF(s_t, \gamma, c) + \varepsilon_t \quad (3)$$

where $\varphi_i = \left(\varphi_{i,0}, \varphi_{i,1}, \cdots, \varphi_{i,p}\right)$ are the parameters to be estimated for $i = 1,2$ and $x_t = \left(1, y_{t-1}, \cdots, y_{t-p}\right)$ contains constant term and lagged values of the endogenous variable. Here, p is the degree of autoregressive construction. Note that the smoothing parameter γ characterizes the speed of transition between regimes. As indicated above, the transition function $GF(s_t, \gamma, c)$ is a continuous function that varies from 0 to 1. It controls the transition between regimes, and finally, the transition function can be a linear trend function $(s_t = t)$.

In this paper, we are inspired by the study of [10] which proposed the application of two types of STAR model, the logistic (LSTAR) and exponential (ESTAR) autoregressive models. This also includes some specification tests of the model such as linearity testing against smooth transition autoregression, determining the delay parameter, and choosing between LSTAR and ESTAR models. Van Dijk et al. [17] surveyed a recent development for STAR models. Liew et al. [18] investigated the exchange rates adjustment behavior with STAR models.

In the STAR model, the Logistic Smooth Transition Autoregressive (LSTAR) model and Exponential Smooth Transition Autoregressive (ESTAR) model are obtained when the transition function *GF*(.) is used as the logistic function and exponential function, respectively [10]. There are many STAR type nonlinear models in the literature according to the type of transition function. However, this part of the study will focus primarily on the LSTAR and ESTAR models. A popular selection for the transition function $GF(s_t, \gamma, c)$ is the first-order logistic function and determined as:

$$GF(s_t, \gamma, c) = (1 + exp\{-\gamma(s_t - c)\})^{-1}, \quad \gamma > 0 \tag{4}$$

The resultant model in Equation (4) is called a logistic STAR or LSTAR model. Here, c can be interpreted as the threshold value that separates the two regimes, as in SETAR models. In this context, the logistic function varies monotonically from 0 to 1 while s_t is increasing. The speed and smoothness of transition from one regime to another is determined by the parameter γ. When $\gamma \to 0$, the transition between regimes hardens. While $\gamma \to 0$, the transition between regimes smooths [19]. Another popular selection for the transition function $GF(s_t, \gamma, c)$ is the exponential function and specified as:

$$GF(s_t, \gamma, c) = 1 - exp^{-\gamma(s_t - c)^2}, \quad \gamma > 0 \tag{5}$$

The resultant model in Equation (5) is referred to as the exponential STAR or ESTAR model. As in the LSTAR model, c is the threshold value and γ can be expressed as the speed and smoothness of the transition between regimes. In the ESTAR approach, as $\gamma \to 0$, the transition function approaches to 0. Thus, ESTAR (p) becomes AR (p) model. As γ approaches to ∞, $GF(s_t, \gamma, c) = 1$. In this case, the ESTAR (p) model becomes another AR (p) model.

2.2. Models with Regimes Determined by Unobservable Variables

The second class most popular model, which suppose that the regime occurs at a time t cannot be observed but is determined by an unobservable stochastic process, is the Markov Switching (MSW) model. It is stressed in this case that we can only assign probabilities to the occurrence of different regimes. In the MSW model, the transition between the regimes is determined by an unobservable state or regime variable s_t that cannot be observed, unlike the TAR and STAR models. The MSW model of [4], also known as the regime-switching model, is one of the most widely used nonlinear time series models. It has been applied for modeling and forecasting business cycles, volatility in financial and economic variables, and the term structure of interest rates. This model involves many equations in different regimes to characterize the time series structure. The current value of the state variable in the Markov chain used for transitions between regimes depends on the previous cycle's variable.

In the case of only two regimes, s_t can simply be supposed to have the values 1 and 2, such that the MSW model with an AR(p) model in both regimes is given by

$$y_t = \begin{cases} \emptyset_{0,1} + \sum_{i=1}^{p} \emptyset_{i,1} y_{t-i} + \varepsilon_{t1} \ if \ s_t = 1 \\ \emptyset_{0,2} + \sum_{i=1}^{p} \emptyset_{i,2} y_{t-i} + \varepsilon_{t2} \ if \ s_t = 2 \end{cases} \tag{6}$$

where $\emptyset_{1i}$ and $\emptyset_{2i}$ are the autoregressive delay parameters of each regime, ε_{t1} and ε_{t2} are mutually independent "white noise" series and s_t specifies the first-degree Markov chain, which has the probabilities in Equation (7). If the serial is in the lower regime $s_t = 1$, if it is in the upper regime $s_t = 2$. In the MSW model, the process s_t is supposed to be a first-order Markov-process. This means that the current regime s_t is only dependent on the previous regime, s_{t-1}. It's worth noting that the MSW model is completed by determining the transition probabilities matrix

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} \{[s_t = 1|s_{t-1} = 1] = p\} & \{[[s_t = 2|s_{t-1} = 1] = 1 - p]\} \\ \{[s_t = 1|s_{t-1} = 2] = q\} & \{[[s_t = 2|s_{t-1} = 2] = 1 - q]\} \end{bmatrix} \tag{7}$$

where each row sums up to one, and p_{11} denotes the probability of transition from lower regime to lower regime; p_{12} gives the probability of transition from lower regime to the upper regime; p_{21} shows the probability of transition from upper regime to lower regime; p_{22} presents the probability of transition from upper regime to upper regime.

3. ESTIMATION OF THE MODELS

In this section, we now provide the estimates of the parameters in the different regime-switching models. For further discussions, we refer to the study of [8, 20] for the SETAR model, [21] for the STAR model and [22] for the MSW model. It is also noted that we discuss the estimation problem for two-regime models.

3.1. Estimation of SETAR Models

Suppose that we consider the SETAR model given in Equation (1). The parameters of this model can be commonly estimated by conditional least squares. To see that least squares is the best estimation process, model (1) can be rewritten as

$$y_t = \emptyset_1' x_t I(y_{t-d} \leq c) + \emptyset_2' x_t I(y_{t-d} > c) + \varepsilon_t \tag{8}$$

or more compactly as

$$y_t = x_t(c)\phi + \varepsilon_t \tag{9}$$

where $x_t = (1, y_{t-1}, \cdots, y_{t-p})$, $x_t(c) = (x_t' I(y_{t-d} \leq c), x_t' I(y_{t-d} > c))$, and $\phi = (\emptyset_1' \ \emptyset_2')'$ with $\emptyset_1 = (\emptyset_{1,0}, \emptyset_{1,1}, \cdots, \emptyset_{1,p})$ and $\emptyset_2 = (\emptyset_{2,0}, \emptyset_{2,1}, \cdots, \emptyset_{2,p})$. For a given value of c, estimates of ϕ are then computed by the least-squares method, defined as

$$\hat{\phi}(c) = (\sum_{i=1}^{n} x_t(c)' x_t(c))^{-1} (\sum_{i=1}^{n} x_t(c)' y_t) \tag{10}$$

The notation $\hat{\phi}(c)$ is known as the conditional least square estimator that depends on conditional c. It should be noted that for the model (9), the corresponding residuals are expressed as $\hat{\varepsilon}_t(c) = y_t - x_t(c)'\hat{\phi}(c)$ with residual variance $\hat{\sigma}^2(c) = \frac{1}{n}\sum_{t=1}^{n} \hat{\varepsilon}_t^2 (c)$. This idea shows that the least square estimate of c can be defined by minimizing the residual variance. In other words, the estimate of threshold value c is

$$\hat{c} = argmin\ \hat{\sigma}^2(c), c \in C \tag{11}$$

where C indicates the set of all possible values that threshold variable takes [16].

As in the case of the threshold value, the threshold variable y_{t-d} is unknown and it is an important problem to determine this variable. In practice, the delay parameter d and the threshold variable y_{t-d} are chosen by minimizing the residual variance. In the context of the SETAR model, we need an efficient grid search in Equation (11) to identify values of d. This means that the minimization problem in Equation (11) becomes

$$(\hat{c}, \hat{d}) = argmin\ \hat{\sigma}^2(c, d),\ \ c \in C\ and\ d \in D \tag{12}$$

where D denotes the set of all possible delay integers (d<p1, p2).

Note that for the sake of convenience, we consider the method proposed by [8]. For each possible delay and threshold parameter, AIC is used to estimate these parameters and determine appropriate autoregressive orders, p_1 and p_2, in model (1). Both delay parameter d and threshold value c are then selected by minimizing the AIC, given by

$$AIC = n_1 ln\hat{\sigma}_1^2 + n_2 ln\hat{\sigma}_2^2 + 2(p_1 + 1) + 2(p_2 + 1) \tag{13}$$

where $n_k, k = 1{,}2$ is the number of observations in the k^{th} regime, $\hat{\sigma}_k^2, k = 1{,}2$ is the variance of the residuals in the k^{th} regime, p_1 and p_2 are the selected lag orders in the lower and upper regimes, respectively.

3.2. Estimation of STAR Models

Assume that we consider the STAR model expressed in equation (3). The estimation of parameters $\varphi = (\varphi_{1,p}', \varphi_{2,p}', \gamma, c)$ in this model can be estimated as

$$\hat{\varphi} = \underbrace{argmin}_{m} \sum_{t=1}^{n} \left[y_t - \left(\hat{\varphi}_{1,p}' x_t \left(1 - GF(y_{t-p}, \gamma, c) + \hat{\varphi}_{2,p}' x_t GF(y_{t-p}, \gamma, c) \right) \right) \right]^2 \tag{14}$$

It is also noted that for fixed values of (γ, c), the STAR model is linear in terms of autoregressive parameters $(\varphi'_{1,p}, \varphi'_{2,p})$. Hence, conditional upon (γ, c), the estimates of $\varphi = (\varphi'_{1,p}, \varphi'_{2,p})'$ can be estimated by ordinary least squares as

$$\hat{\varphi}(\gamma, c) = \left(\sum_{t=1}^{n} x_t(\gamma, c)\, x_t(\gamma, c)'\right)^{-1}\left(\sum_{t=1}^{n} x_t(\gamma, c)\, y_t\right) \quad (15)$$

where $x_t(\gamma, c) = \left(x'_t\left(1 - GF(y_{t-p}; \gamma, c)\right), x'_t\left(GF(y_{t-p}; \gamma, c)\right)\right)'$. Note that the parameters (γ, c) are obtained through a two-dimensional grid search by minimizing the residual variance $\hat{\sigma}^2(\gamma, c) = \frac{1}{n}\sum_{t=1}^{n} \hat{\varepsilon}_t\,(\gamma, c)$ where $\varepsilon_t = y_t - \hat{\varphi}(\gamma, c)'\hat{\varphi}(\gamma, c)$ is the vector of the model residuals. Ordinary least squares method is equivalent to maximum likelihood estimation with the additional assumption that errors ε_t are normally distributed (see [10,23] for a more detailed discussion).

3.3. Estimation of MSW Models

Maximum likelihood method can be used to estimate the coefficients in the MSW model. However, the estimation procedure is highly nonstandard in the literature, since the Markov process S_t is not observed. The basic idea of the estimation problem is not only to obtain estimates of the parameters and the transition probabilities but also to obtain an estimate of the state that occurs at each point of the sample.

We now consider the two-regime MSW model with an AR(p) specification in both regimes, given in Equation (6). If the parameters (i.e., coefficients) of this model are relaxed to be dependent on a hidden state variable S_t, it becomes:

$$y_t = \emptyset_{0,s_t} + \emptyset_{1,s_t} y_{t-1} + \cdots + \emptyset_{p,s_t} y_{t-p} + \varepsilon_t = \phi'_j X_t + \varepsilon_t\,, \quad t = 1 \quad (16)$$

where $s_t = j$ are the unobserved state variables with the transition probabilities matrix in Equation (7), $x_t = (1, y_{t-1}, \cdots, y_{t-p})'$, $\emptyset'_j = (\emptyset_{0,j}, \emptyset_{1,j}, \cdots, \emptyset_{p,j})'$, for $j = 1{,}2$ and ε_t are normally distributed random variables with mean zero and variance σ^2. As shown in Equation (16), the parameters of the MSW model are based on the value of a discrete-valued state variable s_t. In the MSW model, s_t is presumed to follow a particular stochastic process, namely an *N*-state (here, *N*=2) Markov chain or process. The evolution of the Markov process is defined by their transition probabilities expressed in Equation (7).

Here, we are interested mainly in estimating the parameters of the model in Equation (16). Note that the primary item of interest is the regime indicator variable s_t to estimate the parameter vector in Equation (17). We are interested in constructing the estimates of which regime is in effect at each point in time t since s_t is unobserved. These estimates take the form of smoothing probabilities that $s_t = j, (j = 1{,}2)$.

Let $Z_t = (z_t, z_{t-1}, \cdots, z_1)$ indicate the collection of all the observed variables up to time t, which denotes the information set we have at time t. The density of y_t conditional on regime s_t, Z_{t-1} and $s_t = j, (j = 1{,}2)$ is a normal distribution with mean $\emptyset'_j X_t$ and variance σ^2,

$$f(y_t | s_t = j, Z_{t-1}; \theta) = \frac{1}{\sqrt{2\pi\sigma}} exp\left[\frac{-\left(y_t - \emptyset'_j X_t\right)^2}{2\sigma^2}\right] \quad (17)$$

where $\theta = (\emptyset_0, \emptyset_1, \cdots, \emptyset_p, \sigma_\varepsilon^2, p_{11}, p_{22})$ is a vector containing all of the parameters in Equation (16). It should be noted that all transition probabilities are completely defined by p_{11} and p_{22} because, for example, $p_{12} = 1 - p_{11}$. The key idea is to estimate the parameter vector θ by the maximum likelihood method. As shown in [24], the maximum likelihood estimates of the transition probabilities are given by

$$\hat{p}_{ij} = \frac{\sum_{t=2}^{n} P(s_t = j, s_{t-1} = i \mid Z_n; \hat{\theta})}{\sum_{t=2}^{n} P(s_{t-1} = i \mid Z_n; \hat{\theta})} \quad (18)$$

where $\hat{\theta}$ denotes the maximum likelihood estimates of θ defined in (17). It is also seen in [24] that these satisfy the first-order conditions

$$\sum_{t=1}^{n}\left(y_t - \hat{\emptyset}_j x_t\right)x_t\, P\left(s_t = j \mid Z_n; \hat{\theta}\right) = 0,\ j = 1,2 \tag{19}$$

which indicate that the log-likelihood function $L(\theta) = \sum_{t=1}^{n} f(y_t \mid s_t = j, Z_{t-1}; \theta)$ is maximized at the actual value of $\hat{\theta}$. It is also noted that equation (18) provides the following equation.

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n}\sum_{j=1}^{2}\left(y_t - \hat{\emptyset}_j x_t\right)^2 P\left(s_t = j \mid Z_n;\ \hat{\theta}\right) \tag{20}$$

It is understood from equation (20) that $\hat{\emptyset}_j$ can be estimated by the method of weighted least squares regression of y_t on x_t, with weights given by the square root of the smoothed probability, $P\left(s_t = j \mid Z_n;\ \hat{\theta}\right)$. Thus, the estimates $\hat{\emptyset}_j$ can be computed as

$$\hat{\emptyset}_j = \left(\sum_{t=1}^{n} x_t(j)x_t(j)'\right)^{-1}\left(\sum_{i=1}^{n} x_t(j)y_t(j)\right) \tag{21}$$

where $y_t(j) = y_t\sqrt{P\left(s_t = j \mid Z_n;\ \hat{\theta}\right)}$ and $x_t(j) = x_t\sqrt{P\left(s_t = j \mid Z_n;\ \hat{\theta}\right)}$. See [25] for a more detailed discussion.

4. PERFORMANCE CRITERIA

In this section, scale-dependent measures and measures based on percentage errors are taken into consideration to compare the outputs of the models obtained in the application section.

4.1. Scale-Dependent Measures

Performance criteria calculated according to the scale of the data are called scale-based performance measures. When comparing different approaches applied to the same collection of data, these measures come in handy. However, they should not be used when comparing data sets with different scales. Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Median Absolute Error (MAE) are the most widely used scale-dependent measures (MdAE). These measures are defined in Table 1.

Table 1. Most used scale-dependent measures

$MSE = \sum_{i=1}^{n}\frac{(y_i - \hat{y}_i)^2}{n}$	$RMSE = \sqrt{\sum_{i=1}^{n}\frac{(y_i - \hat{y}_i)^2}{n}}$
$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert,$	$MdMH = median(\lvert y_i - \hat{y}_i\rvert)$

It should be noted that RMSE is preferred to MSE because it is on the same scale as the data. RMSE and MSE have historically gained popularity due to their theoretical relevance in statistical modeling. Moreover, they are more sensitive to outliers than MAE or MdAE.

4.2. Measures Based on Percentage Errors

In the performance measures based on the percentage errors, the percentage error is calculated by $p_t = 100\varepsilon_t / y_t$. Percentage errors-based performance criteria have the advantage of being able to perform calculations independent of the scale. Thus, it is frequently used for different types of data sets. Table 2 shows the most widely used measures: Mean Absolute Percentage Error (MAPE), Median Absolute Percentage Error (MdAPE), Root Mean Square Percentage Error (RMSPE), and Root Median Square Percentage Error (RMdSPE).

Table 2. Most used measures based on percentage errors

$MAPE = mean(\|p_t\|)$	$MdAPE = median(\|p_t\|)$
$RMSPE = \sqrt{mean(p_t^2)}$	$RMdSPE = \sqrt{median(p_t^2)}$

There are disadvantages such that the performance criterions based on the error percentages are infinite and undefined when the value of Y_t equals zero for t relevant to any period and have a completely skewed distribution when the value of Y_t is near zero. For example, the value of MAPE will be significantly larger than the value of MdAPE. Therefore, when the x value contains zero or near-zero values, it becomes impossible to use performance measures based on percentage errors. In this case, use of the Symmetric Mean Absolute Percentage Error (sMAPE) and the Symmetric Median Absolute Percentage Error (sMdAPE) criteria, referred to as "symmetric" measurements showed in Table 3, is required.

Table 3. Symmetric measures

$sMAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\|y_i - \hat{y}_i\|}{(\|y_i\| + \hat{y}_i)/2}$	$sMdAPE = median\left(\frac{\|y_i - \hat{y}_i\|}{(\|y_i\| + \|\hat{y}_i\|)/2}\right)$

The problem that occurs around taking a value of zero or near zero is less obvious with the criteria of sMAPE and sMdAPE. Nevertheless, when x takes a value close to zero, y tends to take a value close to zero, so that the measurement will be divided by a value close to zero. Some researchers argued that measures based on percentage errors are highly skewed and that some transformations (logarithmic transformation, etc.) may make these measurements more stable [26].

Definition 1. The relative efficiency of a model compared to another model is defined by the ratio

$$RE\left(\frac{Model_i}{Model_j}\right) = \frac{MSE(Model_i)}{MSE(Model_j)}, \quad i, j = 1, 2, 3 \tag{22}$$

where i and j correspond to obtained by SETAR, STAR and MSW models. As is known from many applications, the MSE criterion is widely used for the comparison of models. Note that i^{th} model is said to be more efficient than j^{th} model if $RE < 1$.

5. EMPIRICAL RESULTS

5.1. First Application: Data and Method

The nonlinear time series models described in the methodology section of this study are applied for the sales prices of bar gold (TL/gr) series obtained monthly between 1996-2016. A total of 241 observation values are used. The subject data set is obtained from the electronic data distribution system of the Central Bank of the Republic of Turkey (http://evds.tcmb.gov.tr/). In our study, we used the R Programming Language (3.1.3) and EViews 9 package software for time series analysis methods.

5.1.1. Results of SETAR Model

First of all, the first-order lag of the data is taken in order to ensure the stationary condition of the sales prices of the bar gold (TL/gr) series. Then, the augmented Dickey-Fuller (ADF) test proposed by Dickey-Fuller [27] was performed for the first order lagged series. The test results are given in Table 4. Test interpretation:

H_0: There is a unit root (Data is not stationary).
H_1: There is no unit root (Data is stationary).

Table 4. ADF test results by EViews 9 software

		t-Statistics	Prob.
Augmented Dickey-Fuller test statistics		-12.54035	<0.001
Test critical values:	%1 level	-3.996918	
	%5 level	-3.428739	
	%10 level	-3.137804	

According to the ADF test results, H_0 hypothesis is rejected because the ADF test statistic value (-12.54) is smaller than all critical t-statistics values (-3.9969, -3.4287, -3.1378). So, the first order lagged series is stationary.

For the SETAR model of the form (1), the parameter values for the most suitable SETAR models are obtained according to the AIC information criteria as a result of the iterations are given in Table 5, with the delay number of the threshold variable d, the number of delays of the sub regime i, the number of delays of the top regime j, and the c threshold value.

Table 5. SETAR model iteration results obtained by R programming

	Delay Number of Threshold Variable	Delay Number of Lower Regime	Delay Number of Upper Regime	Threshold Value	Information Criteria
	d (z_{t-d})	i (y_{t-i})	j (y_{t-j})	c	AIC
1	3	1	1	0.16	399.8567
2	3	1	1	0.15	399.8933
3	3	1	1	0.14	400.0392
4	3	1	1	0.13	400.0780
5	3	1	1	0.12	400.1023
6	3	1	1	0.11	400.1428
7	3	1	1	0.10	400.2160
8	3	1	1	0.09	400.2728
9	3	1	1	0.00	400.3296
10	3	1	1	0.01	400.3558

The optimal number of delays for the threshold variable according to Table 5 is 3. In the same way, the model giving the lowest AIC value comes out as a SETAR model of the first order, in which the threshold value is 0.16 and the delay number of the upper and lower regime is 1. The parameter values obtained by the least-squares method are given in Table 6.

Table 6. Estimation results of SETAR model

Model		Coefficients	Standard Deviation	t-value	p
Lower Regime	$\emptyset_{1,0}$	0.2304	0.2078	1.1086	0.2687
	$\emptyset_{1,1}$	0.5708	0.1018	5.6095	<0.001*
Upper Regime	$\emptyset_{2,0}$	0.5184	0.2225	2.3294	0.0207*
	$\emptyset_{2,1}$	0.0064	0.0776	0.0839	0.9332

*Significant Coefficients

The obtained SETAR model according to these results is as follows:

$$y_t = \begin{cases} 0.2304 + 0.5708 y_{t-1} + \varepsilon_t \ \ if \ \ y_{t-3} < 0.16 \\ 0.5184 + 0.0064 y_{t-1} + \varepsilon_t \ \ if \ \ y_{t-3} \geq 0.16 \end{cases} \tag{23}$$

129 observational values are used for the lower regime generated by the observation values below the threshold value $c = 0.16$ whereas 109 observational values are used for the upper regime of the observational values above 0.16. In other words, the number of observations corresponding to the threshold value of the threshold variable determined for the SETAR model iteratively, which is less than 0.16, is 129, as the number of observations greater than this value is 109. The graph showing the points where the upper or lower regime is active according to the observation values below or above the threshold value is given in Figure 1.

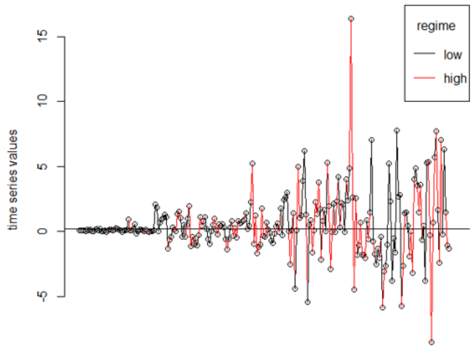

Figure 1. Switching graph between regimes of SETAR model

According to Figure 1, the values obtained by the transition variable are predominantly active in the lower regime. To investigate the homoscedasticity of residuals obtained from the SETAR model in Equation 23, a scatter plot of residuals against fitted values is obtained in Figure 2.

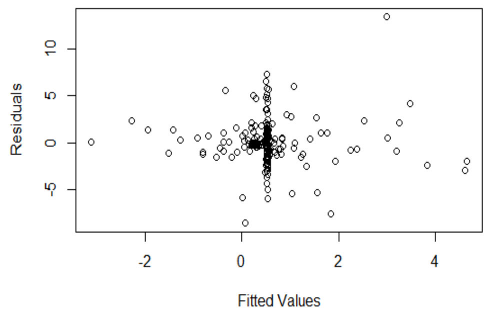

Figure 2. Scatter plot of residuals against fitted values for SETAR model

As seen in Figure 2, there is no apparent evidence for homoscedasticity of residuals. Therefore, Spearman's Rank Correlation test is carried out between the absolute values of residuals and y_t time-series values. Results of the test are given in Table 7.

Table 7. Results of Spearman's Rank Correlation Test

Spearman's Rank Correlation Test	
S-value	p-value
2165500	0.8604

According to Table 7, the null hypothesis of homoscedasticity is accepted because the p-value 0.8604 is greater than the threshold 0.05. These results indicate that the homoscedasticity assumption is provided.

5.1.2. STAR Type Nonlinearity Test and Estimation Results of STAR Model

The first step in the STAR type nonlinearity test is to determine the linear autoregressive model. The linear model suitable for sales prices of bar gold (TL/gr) series is determined as AR(2) with the help of the AIC information criterion. Probability values obtained from the AR model test against the STAR type model are given in Table 8. In the STAR type nonlinearity test, the test is repeated for $p = 1, 2, 3, 4, 5$ delays to find the threshold variable that minimizes the F probability.

Table 8. STAR type nonlinearity test and selection of transition function

Threshold Variable	Probability Values				Selection of Transition Function
	F	F_3	F_2	F_1	
y_{t-1}	0.0005494	0.1699	0.09863	0.07763	LSTAR
y_{t-2}	0.001242	0.1587	0.1345	0.1565	LSTAR
y_{t-3}*	0.00002	0.0055	0.0225	0.0043	LSTAR
y_{t-4}	0.0009	0.4546	0.5073	0.1133	LSTAR
y_{t-5}	0.0125	0.3306	0.673	0.7496	LSTAR

When the F probability values of the obtained models for the transition variables with different delay values are compared, the smallest probability value is obtained for the third delay. Comparing the probability values of F_1, F_2, F_3 for the y_{t-3} variable, F_1 has the smallest probability value with a value of 0.0043. As a result, in the process of examining the STAR type nonlinearity for the sales prices of bar gold (TL/gr) series, the third delay in which linearity is most strongly rejected is determined as the number of delays of the threshold variable while the structure of the transition function in the STAR type models is designated as LSTAR. The results of the 2-regime LSTAR model estimated by the nonlinear least-squares method are given in Table 9.

Table 9. Estimation result of LSTAR model

	Lower Regime			
	Estimation	Standard Deviation	t-value	p
Constant	0.2856	0.2103	1.3578	0.1745
y_{t-1}	0.6092	0.1022	5.9614	<0.001*
y_{t-2}	-0.1867	0.1052	-1.7743	0.076*
	Upper Regime			
	Estimation	Standard Deviation	t-value	p
Constant	0.3036	0.3127	0.9707	0.3317
y_{t-1}	-0.5731	0.1279	-4.479	<0.001*
y_{t-2}	0.0467	0.1297	0.36	0.7188

* Significant Coefficients $\gamma = 100$ $c = 0.1514$ $SSE = 1186.643$ $AIC = 400$

Mathematical representation of the estimated LSTAR model is as follows:

$$y_t = (0.2856 + 0.6092y_{t-1} - 0.1867y_{t-2}) \times \left(1 - GF(s_t, \gamma, c)\right) + [(0.3036 - 0.5731y_{t-1} + 0.0467y_{t-2}) \times GF(s_t, \gamma, c)] + \varepsilon_t \quad (24)$$

Here, the transition function belonging to the LSTAR model and given in Equation (4) can be written as in Equation (25) according to s_t threshold variable, γ smoothness parameter and c threshold value estimated in Table 9.

$$GF(s_t, \gamma, c) = GF(y_{t-3}, 100, 0.1514) = (1 + exp\{-100(y_{t-3} - 0.1514)\})^{-1} \quad (25)$$

The optimal value for the threshold parameter c in the transition function is determined as 0.1514. The smoothness value of the transition between regimes is calculated as $\gamma = 100$. The high γ value indicates that the transition between regimens is fast and rigid. To visualize this prediction, the difference between the regime transition velocities by taking $\gamma = 100$ and $\gamma = 1$ can be obtained as in Figure 3.

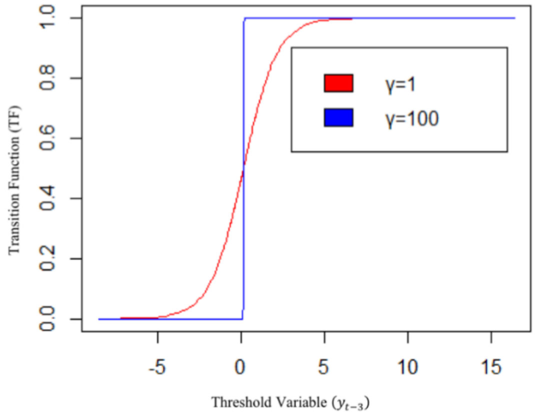

Figure 3. Transition velocities between regimes according to different smoothness parameters

As can be seen from Figure 3, the parameter that determines the speed of approach from 0 to 1 of the logistic transition function is γ. When the γ value is 100, the transition between the regimes is sudden and sharp. When γ value is 1, it is clear that the transition between the regimes is softer. In other words, as the value of γ parameter increases, the transition function converges to 1 more rapidly. As a result, it can be said that TAR-like models where the transitions between regimes are sudden and hard are more appropriate than the STAR models for the sales prices of bar gold (TL/gr) series, resulting from the fact that the transition between the regimes is hard and sudden.

5.1.3. Estimation Results of Markov Switching Model

In this part of the application, the Markov switching model which is based on a stochastic variable in the transition variable is estimated to reveal the nonlinear dynamics in the series of sales prices of bar gold (TL/gr). The Markov switching model based on 2 regimes and 2 delays in each regime obtained from the maximum likelihood method is as follows:

Table 10. Estimation result of Markov switching model

	Lower Regime			
	Estimation	Standard Error	z-value	p
Constant	8.3568	1.1796	7.0842	<0.001*
y_{t-1}	0.9529	0.4581	2.0801	0.0375*
y_{t-2}	1.2863	0.5304	2.4252	0.0153*
	Upper Regime			
	Estimation	Standard Error	z-value	p
Constant	0.3507	0.1341	2.6151	0.0089*
y_{t-1}	0.2783	0.0690	4.0318	0.0001*
y_{t-2}	-0.2640	0.0673	-3.9230	0.0001*

* Significant Coefficients

The mathematical representation of the Markov switching model obtained according to the results in Table 10 is as follows:

$$y_t = \begin{cases} 8.3568 + 0.9529y_{t-1} + 1.2863y_{t-2} + \varepsilon_{t0} & if\ s_t = 1 \\ 0.3507 + 0.2783y_{t-1} - 0.2640y_{t-2} + \varepsilon_{t1} & if\ s_t = 2 \end{cases} \tag{26}$$

The matrix of transition probabilities between the regimes is as follows:

$$P_{ij} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} 8.41 \times 10^{-09} & 0.9999 \\ 0.0246 & 0.9754 \end{bmatrix} \quad (27)$$

While the process is active in the lower regime, the probability of switching to upper regime ($P[S_t = 2 \mid S_{t-1} = 1] = 0.9999$) and while it's active in the upper regime the probability of remaining in the same regime ($P[S_t = 2 \mid S_{t-1} = 2] = 0.9754$) are quite high. This indicates that the process is predominantly active in the upper regime.

5.1.4. Comparison of Methods

In this application, SETAR, STAR, and Markov switching models which are the most used in nonlinear time series literature are applied to sales prices of bar gold (TL/gr) series. The performances of the obtained models are compared with the criteria of MSE, RMSE, MAE, and MdAE which are scale-dependent, and with the criteria of sMAPE and sMdAPE which are based on percentage errors. The reason why the symmetric criteria are preferable from the performance criteria based on error percentages is that the stationary y_t series used in the models have values of zero or near zero. Obtained results of performance criteria are given in Table 11.

Table 11. Performance indicators of models

	MSE	RMSE	MAE	MdAE	sMAPE	sMdAPE
SETAR	193.3240	13.9041	11.3961	**9.7186**	**0.4101**	44.1419
LSTAR	278.5528	16.6899	13.7689	11.6367	0.4720	53.7691
MSW	**160.8788**	**12.6838**	**10.9923**	10.1929	0.4119	**41.2045**

According to the obtained MSE, RMSE, MAE, and sMdAPE performance criteria, the Markov switching model is considered as the model with the most effective performance. All performance criteria indicate that the SETAR model in which the transition between regimes is sudden and hard is more effective than the STAR model based on a smooth transition between regimes. This supports the graph of the transition velocities of the LSTAR model between regimes according to different smoothness parameters obtained in Figure 3.

After comparing the models with performance criteria, we obtained a relative efficiency matrix from the values of MSE of each model for another comparison method. These values are obtained using (22) and the results are given in Table 12.

Table 12. Efficiency matrix of the models

SETAR	LSTAR	MSW	
1.0000	1.4409	0.8321	SETAR
0.6940	1.0000	**0.5775**	LSTAR
1.2017	1.7314	1.0000	MSW

From Table 12, it can be seen that the efficiency values of the MSW model against SETAR and LSTAR models are smaller than 1. Besides, the efficiency value of the SETAR model against the LSTAR model is smaller than 1. Hence, it can be concluded that the MSW model is more efficient than other models and also SETAR is a more effective model than the LSTAR model. The graph which compares the fitted values obtained from each model with actual values is given in Figure 4.

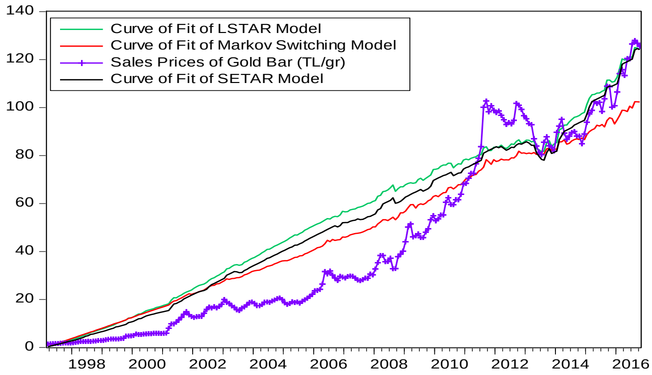

Figure 4. The curve of sales prices of bar gold (TL/gr) and curves of fit obtained from related models

According to the graph obtained in Figure 4, the Markov switching model appears to be the best performing model, followed by the SETAR and LSTAR models. It is seen that the SETAR model adapts to actual values better than the STAR model in parallel with the results obtained from performance criteria values. As a result, it can be interpreted that the Markov-Switching model indicates the best performance for the sales prices of bar gold (TL/gr).

5.2. Second Application: Data and Method

In this application, the nonlinear time series models discussed in Chapter 2 are applied for US Dollar sales prices (USD/TL) series obtained monthly between 2000 and 2016 in TL-based. A total of 204 observations are available. The subject data set is obtained from the electronic data distribution system of the Central Bank of the Republic of Turkey (http://evds.tcmb.gov.tr/). For statistical analysis, R Programming Language (3.1.3) and EViews 9 package software are used.

5.2.1. Results of SETAR Model

The first step is to check the stability of the data. Augmented Dickey-Fuller test is applied for the 1st order lagged series of US Dollar sales prices (USD/TL) data before modeling process. Results of the test are given in Table 13.

Table 13. ADF test results by EViews 9 Software

		t-Statistics	Prob.
Augmented Dickey-Fuller test statistics		-9.382990	<0.001
Test critical values:	%1 level	-3.462737	
	%5 level	-2.875680	
	%10 level	-2.574385	

According to Table 13, the *p*-value of the test statistics is smaller than the %5 significance level. Therefore, the null hypothesis (H_0), in which there is a unit root, is rejected. In other words, the 1st order lagged series is stationary.

The results of the most suitable parameter values obtained according to the AIC value as a result of the iterations made for the two regime SETAR model are given in Table 14.

Table 14. SETAR model iteration results obtained in R programming

	Delay Number of Threshold Variable	Delay Number of Lower Regime	Delay Number of Upper Regime	Threshold Value	Information Criteria
	d (z_{t-d})	i (y_{t-i})	j (y_{t-j})	c	AIC
1	5	1	1	-0.011	-1150.701
2	5	1	1	-0.012	-1150.164
3	5	1	1	0.007	-1148.851
4	5	1	1	0.012	-1148.807
5	5	1	1	0.008	-1148.745
6	1	1	1	0.008	-1148.666
7	5	1	1	0.011	-1148.602
8	1	1	1	0.006	-1148.490
9	1	1	1	0.007	-1148.489
10	5	1	1	-0.013	-1148.487

The optimal number of delays for the threshold variable according to Table 14 is 5. Likewise, the model giving the lowest AIC value comes out as a SETAR model of the first order, in which the threshold value is -0.011 and the delay number of the upper and lower regime is 1. The coefficient values of the SETAR model obtained by the conditional least-squares method are given in Table 15.

Table 15. Estimation Results of SETAR Model

Model		Coefficients	Standard Deviation	t-value	p
Lower Regime	$\emptyset_{1,0}$	-0.0043	0.0069	-0.6227	0.5342
	$\emptyset_{1,1}$	0.1996	0.1089	1.8324	0.0684*
Upper Regime	$\emptyset_{2,0}$	0.0181	0.0053	3.4013	0.0008*
	$\emptyset_{2,1}$	0.4581	0.0831	5.5103	<0.001*

*Significant Coefficients

The following equation is the mathematical representation of the resulting SETAR model:

$$y_t = \begin{cases} -0.0043 + 0.1996 y_{t-1} + \varepsilon_t \;\; if \;\; y_{t-5} < -0.011 \\ 0.0181 + 0.4581 y_{t-1} + \varepsilon_t \;\; if \;\; y_{t-5} \geq -0.011 \end{cases} \tag{28}$$

In this model, 74 observational values are used for the lower regime generated by the observation values below the threshold value $c = -0.011$ whereas 123 observational values are used for the upper regime of the observational values above -0.011. It should be noted that the total number of observations in the lower and upper regimes are determined according to the threshold variable y_{t-5} and threshold value -0.011 obtained from the SETAR model iteratively. The graph showing the points where the upper or lower regime is active according to the observation values below or above the threshold value is given in Figure 5.

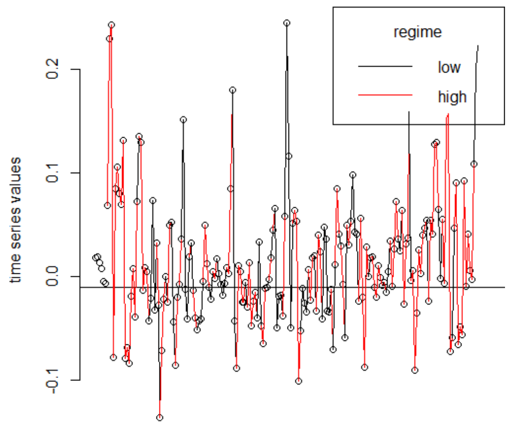

Figure 5. Graph of transition between the regimes of SETAR model

As can be seen from Figure 5, the observed values are mainly concentrated in the upper regime above the threshold value $c = -0.011$. A scatter plot is obtained in Figure 6 to provide a visual examination of the homoscedasticity assumption between the predicted dependent variable scores and the prediction errors.

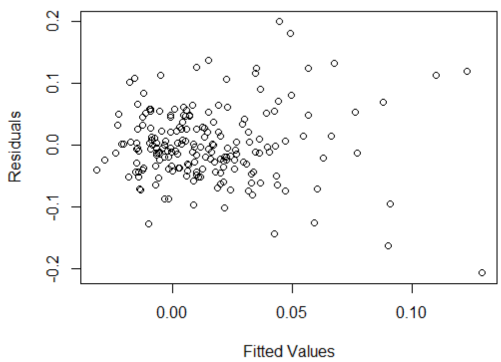

Figure 6. Scatter plot of residuals against fitted values for SETAR model

Figure 6 presents a random displacement of scores into a rectangular shape without any clustering or systematic pattern. Therefore, this figure provides the assumption of homoscedasticity of residuals. To strengthen this view, Spearman's Rank Correlation test is applied between absolute values of residuals and y_t time-series values. Test results are given in Table 16.

Table 16. Spearman's rank correlation test results

Spearman's Rank Correlation Test	
S-value	p-value
1208900	0.4745

According to Table 16, the residuals have constant variance (homoscedasticity) with a 95% level of confidence since the p-value 0.4745 is greater than 0.05.

5.2.2. STAR Type Nonlinearity Test and Estimation Results of STAR Model

The first one of the STAR model prediction phases is to determine the most suitable linear model for the stationary data set. This model is determined as AR (2) for the US Dollar sales prices (USD/TL) series with the help of the AIC information criterion. The next step after the determination of a suitable model is the test phase which consists of the STAR model test against the AR model. Although the optimum number of delays is 2, the F test statistics and probability values are calculated for each number of delays until the 5th delay and the number of delays in which the linearity is most strongly rejected is determined as the delay number of the threshold variable. The results are given in Table 17.

Table 17. STAR type nonlinearity test and selection of transition function

Threshold Variable	F^*	F_3^*	F_2^*	F_1^*	Selection of Transition Function
y_{t-1}	6.736 (<0.001)	0.1503 (0.8606)	0.6278 (0.6433)	3.0304 (0.0074)	LSTAR
y_{t-2}	6.597 (<0.001)	0.6723 (0.5117)	1.3288 (0.2607)	2.8722 (0.0106)	LSTAR
y_{t-3}	6.507 (<0.001)	3.0969 (0.047)	2.88 (0.0239)	2.7935 (0.0126)	LSTAR
y_{t-4}^{**}	7.463 (<0.001)	10.575 (<0.001)	9.1952 (<0.001)	9.8602 (<0.001)	LSTAR
y_{t-5}	4.387 (<0.001)	0.463 (0.6301)	0.5394 (0.7069)	0.4188 (0.8659)	ESTAR

*F test statistic values and probability values in parentheses are given.
**Optimal threshold variable

According to the STAR type nonlinearity test results in Table 17, the null hypothesis which indicates linearity of data, $H_0: \beta_1 = \beta_2 = \beta_3 = 0$, is rejected with the strongest 4th delay (F=7.463). Therefore, the delay number of the threshold variable can be taken as 4. When the probability values of F_1, F_2, F_3 for the y_{t-4} variable are compared, F_3 has the smallest probability value with 4.421×10^{-05}. For this reason, the structure of the transition function is defined as LSTAR.

The most favorable autoregressive model for US Dollar sales prices (USD/TL) was found to have a delay number of 2. Accordingly, for both regimes, autoregressive models that have 2 delays are calculated by using the nonlinear least-squares method. The results are given in Table 18.

Table 18. Estimation result of LSTAR model

	Lower Regime			
	Estimation	Standard Deviation	t-value	p
Constant	0.0120	0.0044	2.6902	0.0071*
y_{t-1}	0.4682	0.0730	6.4150	<0.001*
y_{t-2}	-0.1780	0.0754	-2.3577	0.0184*
	Upper Regime			
	Estimation	Standard Deviation	t-value	p
Constant	0.0002	0.0147	0.0153	0.9878
y_{t-1}	-0.5576	0.2425	-2.2994	0.0215*
y_{t-2}	0.0808	0.2212	0.3654	0.7148

*Significant Coefficients $\gamma = 4121$ $c = 0.0819$ $SSE = 0.6727$ $AIC = -1143$

The mathematical representation of the obtained model is as in (29).

$$y_t = (0.0120 + 0.4682y_{t-1} - 0.1780y_{t-2}) \times \left(1 - GF(S_t, \gamma, c)\right) + [(0.0002 - 0.5576y_{t-1} + 0.0808y_{t-2}) \times GF(S_t, \gamma, c)] + \varepsilon_t \tag{29}$$

Transition function in (29) can be written as follows:

$$GF(S_t, \gamma, c) = GF(y_{t-4}, 4121, 0.0819) = (1 + exp\{-4121(y_{t-4} - 0.0819)\})^{-1} \tag{30}$$

The optimal threshold value for the LSTAR model is calculated as 0.0819. The γ value which indicates the smoothness of the transition between the regimens is calculated as 4121. The graph showing the relationship between the threshold variable and the transition function is given in Figure 7.

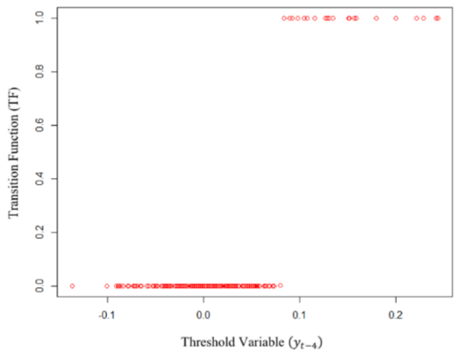

Figure 7. Graph of transition velocity between the regimes of LSTAR model

As shown in Figure 7, the high γ value indicates that the transition between regimes is sudden and sharp. As a result, it can be said that STAR models, where the transition between regimes is soft are not effective models for US Dollar sales prices (USD/TL) series.

5.2.3. Estimation Results of Markov Switching Model

Estimation results of the Markov switching model through the maximum likelihood method using the second-order autoregressive models in each regime and based on the stochastic threshold variable between regimes are given in Table 19.

Table 19. Estimation result of Markov switching model

	Lower Regime			
	Estimation	Standard Error	*z*-value	*p*
Constant	-0.0037	0.0038	-0.9606	0.3367
y_{t-1}	0.2376	0.0927	2.5631	0.0104*
y_{t-2}	-0.2037	0.1015	-2.0074	0.0447*
	Upper Regime			
	Estimation	Standard Error	*z*-value	*p*
Constant	0.1209	0.0188	6.4445	<0.001*
y_{t-1}	0.8163	0.1952	4.1812	<0.001*
y_{t-2}	-0.0967	0.2315	-0.4178	0.6761

* Significant Coefficients

The mathematical representation of the obtained model is as in (31).

$$y_t = \begin{cases} -0.0037 + 0.2376y_{t-1} - 0.2037y_{t-2} + \varepsilon_{t0} \ \ if \ s_t = 1 \\ \ \ 0.1209 + 0.8163y_{t-1} - 0.0967y_{t-2} + \varepsilon_{t1} \ \ if \ s_t = 2 \end{cases} \tag{31}$$

The matrix of transition probabilities between regimes is obtained with the help of Equation (18):

$$P_{ij} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} 0.9230 & 0.0770 \\ 0.4377 & 0.5622 \end{bmatrix} \tag{32}$$

According to all transition probabilities, the probability of being active in the lower regime appears to be more than the probability of being active in the upper regime.

5.2.4. Comparison of Methods

In this application, the TAR model, STAR model, and Markov switching model, which are the most used nonlinear time series models in the literature for US Dollar sales prices (USD / TL) series obtained monthly between 2000-2017 are estimated. To compare the performances of the models obtained, scale-dependent performance criteria (MSE, RMSE, MAE and MdAE) and the symmetric measures (sMAPE and sMdAPE) presented in Section 4 are considered. The results obtained are given in Table 20.

Table 20. Performance indicators of models

	MSE	RMSE	MAE	MdAE	sMAPE	sMdAPE
SETAR	**0.0551**	**0.2348**	**0.2061**	**0.2067**	**0.1283**	**12.5457**
LSTAR	0.1999	0.4472	0.3952	0.4398	0.2334	24.9898
MSW	0.1782	0.4222	0.3696	0.3881	0.2152	21.9987

According to all the values of performance criteria obtained in Table 20, the SETAR model is seen as the model with the most effective performance. Depending on the high value of smoothness parameter obtained in the LSTAR model, the SETAR model shows better performance than the LSTAR model by all the criteria. To compare the efficiencies of the models through relative efficiency measurement, we obtained an efficiency matrix from the values of MSE of each model in Table 21.

Table 21. Efficiency matrix of the models

SETAR	LSTAR	MSW	
1.0000	3.6279	3.2341	SETAR
0.2756	1.0000	0.8914	LSTAR
0.3092	1.1218	1.0000	MSW

According to the results in Table 21, the efficiency values of the SETAR model against LSTAR and MSW models are smaller than 1. Also, the efficiency value of the SETAR model against the LSTAR model is smaller than 1. Hence, it can be said that the SETAR model is more efficient than other models and also SETAR is a more effective model than the LSTAR model. To visually compare the US Dollar sales prices (USD/TL) series and the nonlinear time series models, a graph of the actual values and the fitted values obtained from the related models are obtained in Figure 8.

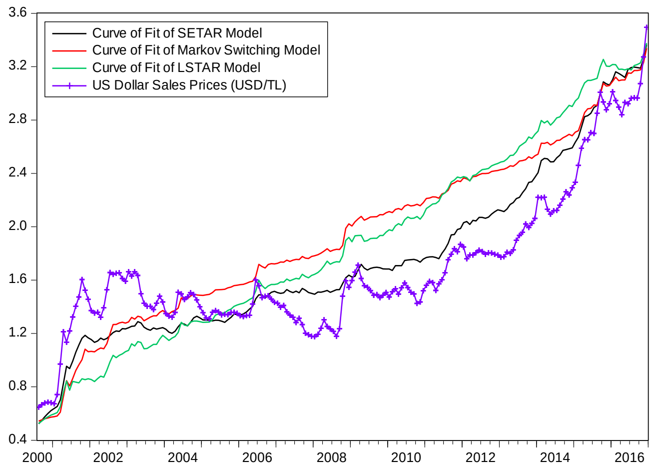

Figure 8. The curve of US dollar sales prices (USD/TL) and curves of fit obtained from related models

In parallel with the results obtained from the performance criteria, the SETAR model fit curve shows a closer fit with actual values than other parametric models.

6. CONCLUSION

Despite evidence of potential nonlinearities, linear models dominated economic time series research until recently. There are at least two explanations for the widespread use of linear models. Firstly, in many cases, linear models are thought to provide rational approximations to true nonlinear relationships, the precise form of which is often unknown. Furthermore, due to a lack of computing capabilities, it was nearly impossible to specify and estimate complex nonlinear models. In recent years, there has been a steady rise in interest in the study and actual application of nonlinear time series models [28].

In this paper, we introduced the Self-Exciting Threshold Autoregressive (SETAR) model, Smooth Transition Autoregressive (STAR) model and Markov-Switching (MSW) model, which are the most commonly used nonlinear parametric time series models to be able to model volatility in a better way in financial markets. To that end, these models were applied to two financial data sets: sales prices of bar gold (TL/gr) and US dollar (USD/TL). The efficiencies of the obtained models were compared with some performance criteria given in Section 4. The empirical results can be summarized as follows: The MSW and SETAR models, respectively, are the most effective models in both applications. In first application, it has been determined that the SETAR model is more effective than the LSTAR model in terms of all the criteria discussed. This result supports the estimation value of smoothness parameter obtained in Table 9 for the LSTAR model and the transition velocities between the regimes for this value in Figure 3. Because, as the value of smoothness parameter increases, the smoothness of transition between the regimes decreases. The same result has also obtained in the second application. In this respect, LSTAR models can be interpreted as the most ineffective models due to the high smoothness parameter values which indicates that the transition between regimes is fast and rigid in both applications. Finally, the performance of the MSW model has also found to be quite successful compared to the LSTAR model by all criteria in both applications.

It should be noted that all inferences about the models were obtained with the assumption of normally distributed of errors. Outliers, on the other hand, can distort this assumption.

CONFLICT OF INTEREST

The authors stated that there are no conflicts of interest regarding the publication of this article.

REFERENCES

[1] Granger CWJ, Andersen AP. An Introduction to Bilinear Time Series Models. Gottingen: Vandenhoek & Ruprecht, 1978.

[2] Tong, H. On a Threshold Model. In: Chen C, editor. Pattern Recognition and Signal Processing. Amsterdam, Sijthoff & Noordhoff, 1978, 575-586.

[3] Terasvirta T, Anderson H. Characterizing Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models. J Appl Econom, 1992; 7: 119-136.

[4] Hamilton JD. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. Econometrica, 1989; 57: 357-384.

[5] Tong H, Lim KS. Threshold Autoregression, Limit Cycles and Cyclical Data. J R Stat Soc, 1980; 42: 245-292.

[6] Chan KS, Tong H. On Estimating Thresholds in Autoregressive Models. J Time Ser Anal 1986; 7: 179-190.

[7] Tsay RS. Nonlinearity Tests for Time Series. Biometrika, 1986; 73: 461-466.

[8] Tong H. Non-linear Time Series: A Dynamical Systems Approach. USA: Oxford University Press, 1990.

[9] Granger CWJ, Terasvirta T. Modelling Nonlinear Economic Relationships. USA: Oxford University Press, 1993.

[10] Terasvirta T. Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models. J Am Stat Assoc, 1994; 89: 208-218.

[11] Khan MY. Advanced in Applied Nonlinear Time Series Modeling. Dissertation, Munich University, Münih, Germany, 2015.

[12] Tong H. Threshold Models in Non-Linear Time Series Analysis. Berlin: Springer-Verlag,1983.

[13] Chan KS, Petruccelli JD, Tong H, Woolford SW. A Multiple-Threshold AR(1) Model. J Appl Prob, 1985; 22: 267-279.

[14] Petruccelli JD. On The Consistency of Least Squares Estimators for a Threshold AR(1) Model. J Time Ser Anal, 1986; 7: 269-278.

[15] Chan KS. Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model. Ann Stat, 1993; 21: 520-533.

[16] Hansen BE. Inference in TAR Models. Stud Nonlinear Dyn E 1997; 2: 1-16.

[17] van Dijk DJC, Terasvirta T, Franses PH. Smooth Transition Autoregressive Models – A Survey of Recent Developments. Economet Rev, 2002; 21: 1-47.

[18] Liew VKS, Baharumshahb AZ, Lau E. Nonlinear Adjustment towards Purchasing Power Parity in ASEAN Exchange Rates, J Appl Econ, 2002; 6: 7-18.

[19] van Dijk DJC, Terasvirta T, Franses PH. Smooth Transition Autoregressive Models: A Survey of Recent Developments. Econometric Institute Research Report EI2000-23/A, 2000.

[20] Hansen BE. Sample Splitting and Threshold Estimation. Econometrica, 2000; 68: 575-603.

[21] Terasvirta T. Modelling Economic Relationships with Smooth Transition Regressions. In: Ullah A, Giles DEA, editors. Handbook of Applied Economic Statistics. New York, Marcel Dekker, 1998, 507-552.

[22] Hamilton JD. Time Series Analysis. Princeton, Princeton University Press, 1994.

[23] Bekiros SD. A Robust Algorithm for Parameter Estimation in Smooth Transition Autoregressive Models. Econ Lett ,2009; 103: 36-38.

[24] Hamilton JD. Analysis of Time Series Subject to Changes in Regime. J Econometrics, 1990; 45: 39-70.

[25] Franses PH, van Dijk DJC. Nonlinear Time Series Models in Empirical Finance. New York, Cambridge University Press, 2000.

[26] Hyndman RJ, Koehler AB. Another Look at Measures of Forecast Accuracy. Int J Forecasting, 2006; 22: 679-688.

[27] Dickey DA, Fuller WA. Distribution of the Estimates for Autoregressive Time Series with a Unit Root. J Am Stat Assoc, 1979; 74: 427-431.

[28] van Dijk DJC. Smooth Transition Models: Extensions and Outlier Robust Inference. PhD, Erasmus School of Economics, Rotterdam, Netherlands, 1999.