

On Equal Predictive Ability and Parallelism of Self-Exciting Threshold Autoregressive Model

F. I. Ugwuowo^{1, *}, E. C. Uzochukwu² and T. E. Ugah³

^{1,2,3}*Department of Statistics University of Nigeria, Nsukka, Enugu State, Nigeria*

(Received: 16 October 2023; Accepted: 10 January 2024)

Abstract. Several authors have developed statistical procedures for testing whether two models are equivalent. In this work, we not only present the notion of equivalence but also extend this to a measure of predictive ability of a time series following a stationary self-exciting threshold autoregressive (SETAR) process. A test is developed for equal predictive ability and a proposition and lemma given and proved. Illustrative examples are given to show how to conduct the test which can help the practitioner to avoid incorrect assessment of the accuracy of a forecast.

Keywords: Non-linear time series, Self-exciting threshold autoregressive model, ARIMA, SETAR process

Published by: Department of Statistics, University of Benin, Nigeria

1. Introduction

Popularization and extensive research of linear time series modeling began in 1927 with Yule's Autoregressive models, used in studying sunspot numbers. In the decades that followed, these models have been successfully applied in different fields because as far as one-step ahead prediction is concerned, linear time series models are often adequate. However, this is not so as can be seen from the Canadian lynx data. The causes of this are mentioned later herein. Nonlinear time series analysis gained attention in the 1970's. The interest grew due to the need to model nonlinear changes in everyday time series data exhibiting non-linearity. Autoregressive Integrated Moving Average (ARIMA) models cannot describe adequately limit cycles, time-irreversibility, amplitude-frequency dependency and jump phenomena (the Canadian lynx data mentioned earlier is a good example). As a result, Tong (1978) came up with a procedure for modeling nonlinear changes in time series data in which different Autoregressive (AR) processes are functioning, and the switch between these AR

*Corresponding author. Email: ngfedelis.ugwuowo@unn.edu.ng

models depends on the delay parameter and threshold value(s), which has certain time lag values from the given time series. Tong and Lim (1980) and Tong (1983) followed up the work with an extensive description of the procedure. Tsay (1989) proposed a much simpler procedure. He noted that the Tong's procedure is not statistically adequate for formally determining if a given data can be described using a threshold model.

Several nonlinear time series models have been proposed over the years including the Threshold Autoregressive (TAR) models. This involves a piece-wise linearization of nonlinear models over the state space by the introduction of the thresholds $\{r_0, \dots, r_k\}$. TAR models have been of significant interest because of its ability to model nonlinear data adequately. Common notions were employed by Priestly (1965), and Ozaki and Tong (1975) in the analysis of non-stationary time series and time dependent systems. They used local stationarity as the counterpart of our present local linearity. The overall process is nonlinear when there are at least two regimes with different parameters and/or process order. Tong and Lim (1980) proposed the requirements for modeling of nonlinear time series. Asikgil (2018) used five different SETAR models to generate the autocorrelated disturbances in a multiple regression study. In a recent work by Yang et al (2018), they used empirical likelihood method to analyze the SETINAR models which captured the various characteristics of time series of counts.

Hamdi and Khalfi (2019) used a variant of the predictive Density criterion (PDC) for joint determination of the delay parameter and autoregressive orders of the SETAR models. The result was shown to be very effective. Blasques *et al.* (2019) introduced a new class of nonlinear autoregressive models from their representation as linear autoregressive models with time-varying coefficients. Using some illustration, they compared the dynamic properties of the new model with those of well-known nonlinear dynamic models such as the threshold and smooth transition autoregressive models and found it to perform better. Guris and Guris (2019) proposed a new unit root test against ESTAR model based on a class of modified statistics by applying the GLS detrending method for the unit root test procedure. Siu and Elliot (2020) adopted a nonlinear time series model for modeling Bitcoin return dynamics. The approach involved combining both the self-exciting threshold autoregressive (SETAR) model and the generalized autoregressive conditional heteroscedastic (GARCH) model. Liu and Zhang (2020) obtained the law of the iterated logarithm for error density estimators in first-order autoregressive models and threshold models as corollaries.

Predictive ability in time series indicates the degree to which the past can be used in ascertaining the future. This is a fundamental measurement in time series analysis. Assessing whether there is predictability among macroeconomic variables has always been a central issue for applied researchers. For example, much effort has been devoted to analyzing whether money has predictive content for output. This question has been addressed by using both simple linear Granger Causality (GC) tests (Stock and Watson, 1989) as well as tests that allow for non-linear predictive relationships (Amato and Swanson, 2001 and Stock and Watson, 1999). Several authors have studied this measure and proposed extensions to it in several fields, for instance tourism and finance. However, not much has been done to investigate whether more than one series have

equal predictive ability. Testing whether the models provide similar forecast performance represents a test of equal predictive ability. Examining equal predictive ability is essential in risk management. Furthermore, it could be interesting to establish if time series which have the same variables (economic, climate, etc), recorded in different spatial areas or calculated with different methodologies, have equal predictive abilities.

The issue of equivalence have not been thoroughly addressed, yet there are many instances in practice where it is necessary. For instance, the task of forecasting demand for common items in different markets maybe facilitated if it can be shown that the models characterizing demand are equivalent in the various market. As an extension of this application on a broader scale, the equivalence of economic indicator series in separate geographic areas could suggest that the economic condition are common in the studied regions, (Steece and Wood, 1985).

Steece and Wood (1985) introduced a simple method for assessing the equivalence of k time series when the series are not necessarily independent. By equivalence, the k series exhibit parallelism and concurrence. They introduced a test of equivalence which can be easily computed from information routinely reported by software packages and can be easily interpreted.

Triacca (2004) analyzed the relationship between a measure of dissimilarity between ARIMA models and a condition of parallelism of two ARIMA processes. Otranto (2004) extended well known results developed for the ARIMA models to GARCH models by considering financial time series characterized by similar volatility structures. The argument being that the selection of series having similar behavior could be important for the analysis of the transmission mechanisms of volatility and to forecast the time series, using the series with more similar structure. Otranto and Triacca (2007) used a measure of predictive ability of a time series following a stationary Autoregressive Moving Average (ARMA) process to develop a test of equal predictive ability of two series. The test is derived by a set of propositions which links the structure of the AR and MA coefficients to the predictability measure.

This work presents a test of equal predictive ability and its relationship with parallelism in Self-Exciting Threshold Autoregressive (SETAR) model. Three illustrative examples using Wolf yearly sunspot numbers, Abuja Ozone depletion data and simulated data are presented. It is organized as follows; section 2 discusses materials and method; section 3 presented the results and discussions of the numerical examples. The conclusion is presented in section 4.

2. Materials and Method

The R^2 statistic is a measure of the proportion of the total variation in the data that is explained by the time series model. It can be stated in the context of the self-exciting autoregressive (SETAR) time series models as one minus the ratio of the residual variance to the total variance of the time series. It makes available a measure of the relative predictive ability of a time series given its past history. Its understanding in the time series context gains much from the similarity to linear regression. However, the special character of time-series models offers the opportunity of additional insight which can be of considerable value

in actual data analysis. R^2 can be related to the underlying parameters of a time series and to its autocorrelation structure, which shall be presented shortly.

2.1 R^2 Defined for SETAR Time Series Models

We are interested in time series $\{Z_t\}$ which have a self-exciting threshold autoregressive (SETAR) representation of the form

$$Z_t = \phi_0^j + \sum_{i=1}^p \phi_i^j Z_{t-i} + \epsilon_t^j \quad \text{if } r_{j-1} \leq Z_{t-d} < r_j, \quad (1)$$

where ϕ_i^j , $i = 1, 2, 3, \dots, p$ and $j = 1, 2, 3, \dots, k$ are fixed parameters for different regime satisfying conditions for stationarity and ϵ_t^j are a sequence of uncorrelated random variables with common variance $\sigma_\epsilon^{2(j)}$. The fraction of the variation in Z , $V(Z)^j$ which cannot be predicted is therefore

$$1 - R^2 = \frac{\sigma_\epsilon^{2(j)}}{V(Z)^{(j)}}$$

and fraction that can be predicted is define as

$$R^2 = 1 - \frac{\sigma_\epsilon^{2(j)}}{V(Z)^{(j)}}$$

For a series which cannot be forecasted, such as white noise, this ratio is equal to zero. If the series can be forecasted without error, the ratio is 1. The interpretation of R^2 gives extra meaning when it can be associated with the parameters of the time series and to its autocorrelation. This tool can be very useful at the identification stage when it is the desire of an analyst to assess the predictive ability of a time series. Theoretical R^2 can be associated with the coefficients and autocorrelation coefficients in the following way. Denoting the autocovariances of the process by γ_i^j , the variance of Z_t is given by

$$V(Z_t)^{(j)} = E(Z_t^2)^{(j)} = \phi_1^j \gamma_1^j + \dots + \phi_p^j \gamma_p^j + \sigma_\epsilon^{2(j)}$$

We then have $R^{2(j)}$ to be

$$\begin{aligned} 1 - R_p^{2(j)} &= \frac{\sigma_\epsilon^{2(j)}}{V(Z_t)^{(j)}} \\ &= 1 - \phi_1^j \rho_1^j - \dots - \phi_p^j \rho_p^j \end{aligned}$$

$$\begin{aligned}
R_p^{2(j)} &= \phi_1^j \rho_1^j + \dots + \phi_p^j \rho_p^j \\
&= \underline{\rho}_p^{(j)} \underline{\phi}_p^{(j)}
\end{aligned} \tag{2}$$

$$\underline{\phi}_p^{(j)} = R_p^{2(j)} \left(\underline{\rho}_p^{(j)} \right)^{-1} \tag{3}$$

where $\underline{\rho}_p^{(j)}$ denote the column vector $(\rho_1^j, \dots, \rho_p^j)$ and $\underline{\phi}_p^{(j)}$ is $(\phi_1^j, \dots, \phi_p^j)$. Yule-Walker equations are given by,

$$\underline{\rho}_p^{(j)} = \Gamma_p^{(j)} \underline{\phi}_p^{(j)}$$

where $\Gamma_p^{(j)}$ is the correlation matrix for $(Z_{t-1}, \dots, Z_{t-p})$. Using Equation (3),

$$\underline{\phi}_p^{(j)} = \left[\Gamma_p^{(j)} \right]^{-1} \underline{\rho}_p^{(j)}$$

we can then obtain $R_p^{2(j)}$ as

$$R_p^{2(j)} = \underline{\rho}_p^{(j)} \left[\Gamma_p^{(j)} \right]^{-1} \underline{\rho}_p^{(j)}$$

Proposition:

Let $\{Y_t; t = 0, \pm 1, \dots\}$ and $\{Z_t; t = 0, \pm 1, \dots\}$ be two SETAR processes. The processes $\{Y_t; t = 0, \pm 1, \dots\}$ and $\{Z_t; t = 0, \pm 1, \dots\}$ have equal predictive ability if

$$\underline{\rho}_p^{(j)} \underline{\phi}_{pY}^j - \underline{\rho}_q^{(j)} \underline{\phi}_{qZ}^j = 0$$

Proof:

If $\{Y_t; t = 0, \pm 1, \dots\}$ and $\{Z_t; t = 0, \pm 1, \dots\}$ have equal predictive ability, then

$$\begin{aligned}
R_{pY}^{2(j)} &= R_{qZ}^{2(j)} \\
\underline{\rho}_p^{(j)} \underline{\phi}_{pY}^j - \underline{\rho}_q^{(j)} \underline{\phi}_{qZ}^j &= 0
\end{aligned}$$

$$\rho_1^j \phi_{1Y}^j + \rho_2^j \phi_{2Y}^j + \dots + \rho_p^j \phi_{pY}^j = \rho_1^j \phi_{1Z}^j - \rho_2^j \phi_{2Z}^j - \dots - \rho_q^j \phi_{qZ}^j$$

These measures are the same if

$$\rho_i^j \phi_{iY}^j - \rho_i^j \phi_{iZ}^j = 0, \quad i = 1, \dots, p, q \text{ for each } i.$$

Hence,

$$\rho_1^j \phi_{1Y}^j + \rho_2^j \phi_{2Y}^j + \dots + \rho_{pY}^j \phi_{pY}^j - \rho_1^j \phi_{1Z}^j - \rho_2^j \phi_{2Z}^j - \dots - \rho_{qz}^j \phi_{qZ}^j = 0$$

$$\rho_{py}^{(j)} \phi_{PY}^j - \rho_{qz}^{(j)} \phi_{qZ}^j = 0$$

2.2 Parallelism and Predictive Ability

Let $\{Y_t; t = 0, \pm 1, \dots\}$ and $\{Z_t; t = 0, \pm 1, \dots\}$ be two SETAR processes.

$$\phi_Y^j(B) Y_t = \epsilon_t^j, \epsilon_{Y,t}^j \sim WN(0, \sigma_Y^{j2})$$

$$\phi_Z^j(B) Z_t = \epsilon_t^j, \epsilon_{Z,t}^j \sim WN(0, \sigma_Z^{j2})$$

where $\phi_Y^j(L)$ and $\phi_Z^j(L)$ are finite in L of degree PY and PZ respectively. If $\{Y_t; t = 0, \pm 1, \dots\}$ and $\{Z_t; t = 0, \pm 1, \dots\}$ are parallel then $\phi_Y^j(L) = \phi_Z^j(L)$. This definition is given in Steeco and Wood (1985), Guo(1999) and Otranto and Triacca (2007) for the case of AR models.

Lemma:

Let $\{Y_t; t = 0, \pm 1, \dots\}$ and $\{Z_t; t = 0, \pm 1, \dots\}$ be two SETAR processes. If $\{Y_t; t = 0, \pm 1, \dots\}$ and $\{Z_t; t = 0, \pm 1, \dots\}$ are parallel then they have equal predictive ability

Proof:

If $\{Y_t; t = 0, \pm 1, \dots\}$ and $\{Z_t; t = 0, \pm 1, \dots\}$ are parallel then

$$Y_t \phi_Y^j(L) = \epsilon_t^j = \epsilon_t^j = \phi_Z^j(L) Z_t$$

$$Y_t = \frac{\epsilon_t^j}{\phi_Y^j(L)} = \frac{\epsilon_t^j}{\phi_Z^j(L)} = Z_t$$

with $\phi_Y^j(L) \neq 0$ and $\phi_Z^j(L) \neq 0$ for $|L| < 1$

$$Y_t = \frac{\epsilon_t^j}{\phi_Y^j(L)} = \left(1 + \psi_{1y}^j L + \psi_{2y}^j L^2 + \dots\right) \epsilon_t^j = \left(1 + \psi_{1z}^j L + \psi_{2z}^j L^2 + \dots\right) \epsilon_t^j = \frac{\epsilon_t^j}{\phi_Z^j(L)} = Z_t \quad (4)$$

From the LHS, we obtain

$$\frac{\epsilon_t^j}{\phi_Y^j(L)} = \left(1 + \psi_{1y}^j L + \psi_{2y}^j L^2 + \dots\right) \epsilon_t^j \Rightarrow \phi_Y^j(L) \left(1 + \psi_{1y}^j L + \psi_{2y}^j L^2 + \dots\right) = 1 \quad (5)$$

$$\left(1 - \phi_{1,y}^j L - \phi_{2,y}^j L^2 - \dots\right) \left(1 + \psi_{1,y}^j L + \psi_{2,y}^j L^2 + \dots\right) = 1$$

$$1 + \psi_{1,y}^j L + \psi_{2,y}^j L^2 + \dots - \phi_{1,y}^j L - \psi_{1,y}^j \phi_{1,y}^j L^2 - \psi_{2,y}^j \phi_{1,y}^j L^3 - \dots - \phi_{2,y}^j L^2 - \psi_{1,y}^j \phi_{2,y}^j L^3 - \dots = 1 \quad (6)$$

and from RHS

$$\left(1 + \psi_{1,z}^j L + \psi_{2,z}^j L^2 + \dots\right) \epsilon_t^j = \frac{\epsilon_t^j}{\phi_Z^j(L)} \Rightarrow \phi_Z^j(L) \left(1 + \psi_{1,z}^j L + \psi_{2,z}^j L^2 + \dots\right) = 1$$

$$\left(1 - \phi_{1,z}^j L - \phi_{2,z}^j L^2 - \dots\right) \left(1 + \psi_{1,z}^j L + \psi_{2,z}^j L^2 + \dots\right) = 1$$

$$1 + \psi_{1,z}^j L + \psi_{2,z}^j L^2 + \dots - \phi_{1,z}^j L - \psi_{1,z}^j \phi_{1,z}^j L^2 - \psi_{2,z}^j \phi_{1,z}^j L^3 - \dots - \phi_{2,z}^j L^2 - \psi_{1,z}^j \phi_{2,z}^j L^3 - \dots = 1 \quad (7)$$

For the model of order p , we obtain

$$\phi_{p,y}^j = \psi_{p,y}^j - \phi_{1,y}^j \psi_{p-1,y}^j - \phi_{2,y}^j \psi_{p-2,y}^j - \dots - \phi_{p-1,y}^j \psi_{1,y}^j \quad (8)$$

$$\phi_{p,y}^j = \psi_{p,y}^j - \sum_{i=1}^{p-1} \phi_{i,y}^j \psi_{p-i,y}^j \quad (9)$$

$$\psi_{l,y}^j = \sum_{i=1}^p \phi_{i,y}^j \psi_{l-i,y}^j \text{ for } l > p \quad (10)$$

The same applies for Z process as follows

$$\phi_{p,z}^j = \psi_{p,z}^j - \phi_{1,z}^j \psi_{p-1,z}^j - \phi_{2,z}^j \psi_{p-2,z}^j - \dots - \phi_{p-1,z}^j \psi_{1,z}^j \quad (11)$$

$$\phi_{p,z}^j = \psi_{p,z}^j - \sum_{i=1}^{p-1} \phi_{i,z}^j \psi_{p-i,z}^j \quad (12)$$

$$\psi_{l,z}^j = \sum_{i=1}^p \phi_{i,z}^j \psi_{l-i,z}^j \text{ for } l > p \quad (13)$$

equating Equations (7) and (10) we have,

$$\phi_{p,y}^j = \psi_{p,y}^j - \sum_{i=1}^{p-1} \phi_{i,y}^j \psi_{p-i,y}^j = \psi_{p,z}^j - \sum_{i=1}^{p-1} \phi_{i,z}^j \psi_{p-i,z}^j = \phi_{p,z}^j \quad (14)$$

From Equation (3)

$$\begin{aligned}\phi_{k,y}^j &= R_y^{2(j)} \left(\rho_{k,y}^{(j)} \right)' = R_z^{2(j)} \left(\rho_{k,z}^{(j)} \right)' = \phi_{k,z}^j \\ R_y^2 &= \rho_p^{(j)} \phi_{py}^j = \rho_q^{(j)} \phi_{qz}^j = R_z^2 \\ R_y^2 - R_z^2 &= 0\end{aligned}$$

2.3 Testing for Equal Predictive Ability

The theoretical results of the previous section suggest the possibility of testing the equal predictive ability of two SETAR processes. Verifying the hypothesis of equal predictive ability is an important task in several fields. For example, it could help to choose among several procedures of seasonal adjustment; it could help in investment decisions and in establishing the different degrees of predictive ability among several returns (Otranto and Triacca, 2007).

The above provides the way to confirm the equivalence of the equal predictive ability and parallelism between two SETAR processes. The null hypothesis can be expressed in terms of the parameters of the time series.

$$H_0 : \phi_{k,y}^j = \phi_{k,z}^j \quad (15)$$

The null hypothesis can be expressed as a set of linear restrictions:

$$\mathbf{A}\Theta = \mathbf{0} \quad (16)$$

where $\Theta = \left(\phi_{1y}^j, \phi_{2y}^j, \dots, \phi_{py}^j, \phi_{1z}^j, \phi_{2z}^j, \phi_{1y}^j, \dots, \phi_{pz}^j \right)'$ and $\mathbf{A}$ is a $m \times m$ matrix of the form

$$\begin{pmatrix} I_m - I_m & 0_m & \dots & 0_m & 0_m \\ 0_m & I_m & -I_m & \dots & 0_m & 0_m \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_m & 0_m & 0_m & \dots & I_m - I_m \end{pmatrix} \quad (17)$$

with I_m representing $m \times m$ identity matrix and 0_m an $m \times m$ matrix with elements equal to zero, $m = p$. The Wald test statistic for the null hypothesis is of the form:

$$W = (\mathbf{A}\hat{\Theta})' (\mathbf{A}\hat{\Lambda}\mathbf{A}')^{-1} (\mathbf{A}\hat{\Theta})$$

where $\hat{\Theta}$ is the maximum likelihood estimator of Θ whereas $\hat{\Lambda}$ is the maximum likelihood estimator of the covariance matrix of Θ . W is asymptotically distributed as a central chi-square random variable with m degrees of freedom.

3. Results and Discussion

Example 1: Consider different transformations of the Wolf Yearly Sunspots numbers collecte from 1700 - 2001. The series consists of 302 observations. From Figure 1, we see that the plot of sunspot differs from the log. transform and the reciprocal but it is considerably similar to the square root transform.

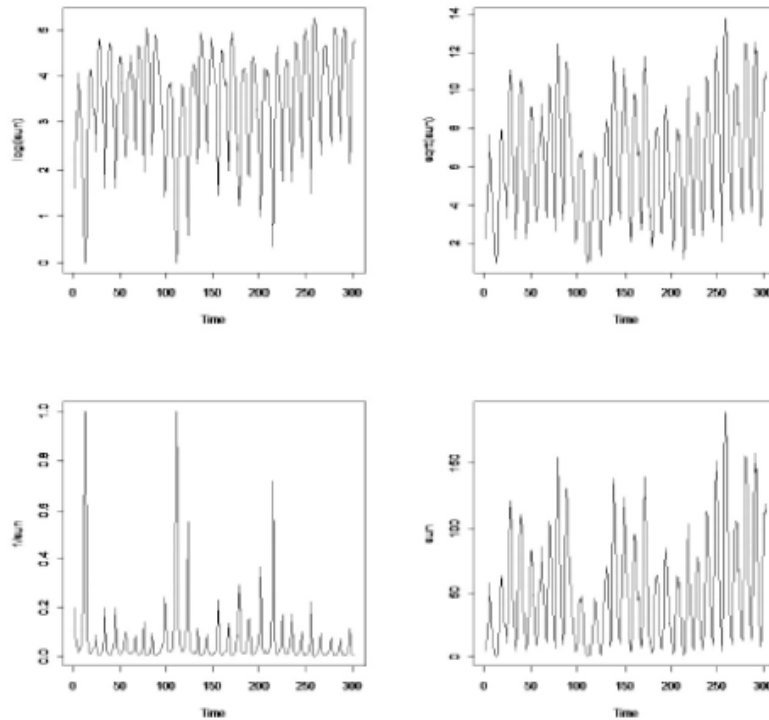


Figure 1: Plot of square root, $\log_e$, reciprocal and actual data of sunspot numbers

In Table 1 we present the summary of the estimated parameters(with standard error in parenthesis) of the sunspot numbers and the logarithm, square root and reciprocal transformations. Threshold values depict the point of change from lower regime to upper regime. The measure of predictive ability R^2 are shown for the two regimes.

Table 1: Estimates of the parameter for sunspot numbers and other transformations

	Sunspot Numbers SETAR(2;2,2),d=1 Estimated threshold=100.1		$\log_e$ Sunspot Numbers SETAR(2;2,2),d=1 Estimated threshold=2.632		Sq.root Sunspot Numbers SETAR(2;2,2),d=1 Estimated threshold=3.317		Reciprocal of Sunspot Numbers SETAR(2;2,2),d=1 Estimated threshold=0.125	
Lower Regim	Estimates	t-value	Estimates	t-value	Estimates	t-value	Estimates	t-value
Intercept	13.358(1.729)	7.723	1.098(0.319)	3.446	0.809(0.709)	1.144	0.0023(0.004)	0.608
lag 1	1.1591(0.053)	29.888	1.098(0.187)	5.837	1.439(0.326)	4.414	1.536(0.1)	15.31
lad 2	-0.82(0.047)	-17.578	-0.371(0.138)	-2.692	-0.469(0.176)	-2.665	-0.252(0.039)	-6.483
RSE	15.68		0.851		1.359		0.0016()	
n	260		66		52		267	
R^2	0.914		0.891		0.877		0.665	
Upper Regim	Estimates	t-value	Estimates	t-value	Estimates	t-value	Estimates	t-value
Intercept	-9.985(15.86)	-0.63	0.132(0.137)	0.963	1.599(0.223)	7.164	0.083(0.071)	1.179
lag 1	1.247(0.125)	9.938	1.617(0.047)	33.26	1.471(0.044)	33.613	0.843(0.208)	4.045
lad 2	-0.356(0.081)	-4.37	-0.7(0.036)	-19.262	-0.732(0.04)	-18.124	-0.425(0.197)	-2.155
RSE	16.958		0.333		1.099		0.249	
n	40		234		248		33	
R^2	0.979		0.993		0.979		0.635	

The different models from Table 1 are given below;
The sunspots numbers,

$$z_t = \begin{cases} 13.358 + 1.591z_{t-1} - 0.82z_{t-2}, & z_{t-1} \leq 100.1 \\ -9.985 + 1.247z_{t-1} - 0.356z_{t-2}, & z_{t-1} > 100.1 \end{cases}$$

The $\log_e$ of sunspots numbers,

$$z_t = \begin{cases} 1.098 + 1.089z_{t-1} - 0.37z_{t-2}, & z_{t-1} \leq 2.632 \\ 0.132 + 1.617z_{t-1} - 0.732z_{t-2}, & z_{t-1} > 2.632 \end{cases}$$

The square root of sunspots numbers,

$$z_t = \begin{cases} 0.809 + 1.439z_{t-1} - 0.469z_{t-2}, & z_{t-1} \leq 3.317 \\ 1.599 + 1.471z_{t-1} - 0.732z_{t-2}, & z_{t-1} > 3.317 \end{cases}$$

The reciprocal of sunspots numbers,

$$z_t = \begin{cases} 0.0023 + 1.536z_{t-1} - 0.252z_{t-2}, & z_{t-1} \leq 0.125 \\ 0.083 + 0.843z_{t-1} - 0.425z_{t-2}, & z_{t-1} > 0.125 \end{cases}$$

The tests for parallelism and equal predictive ability for sunspot numbers are as follows:

$$H_0 : \phi_{k,z_1}^j = \phi_{k,z_2}^j \text{ vs } \phi_{k,z_1}^j \neq \phi_{k,z_2}^j \quad (18)$$

$\alpha = 0.05$.

Test statistics:

$$W = (\hat{A}\hat{\Theta})'(\hat{A}\hat{\Lambda}\hat{A}')^{-1}(\hat{A}\hat{\Theta})$$

Computation:

Using the estimates of the parameters

$$\hat{\Theta} = (1.591, -0.82, 1.089, -0.371)$$

$$\text{and } \hat{\Lambda} = \begin{pmatrix} 0.0013 & -0.0011 & 0 & 0 \\ -0.001 & 0.0013 & 0 & 0 \\ 0 & 0 & 0.013 & -0.01 \\ 0 & 0 & -0.01 & 0.013 \end{pmatrix},$$

we obtain the values in Table 2 which shows the test statistic and p -values in parenthesis for comparison of sunspot numbers(S) with square root(SS), $\log_e$ transforms(LS) and reciprocal (RS) of sunspot numbers.

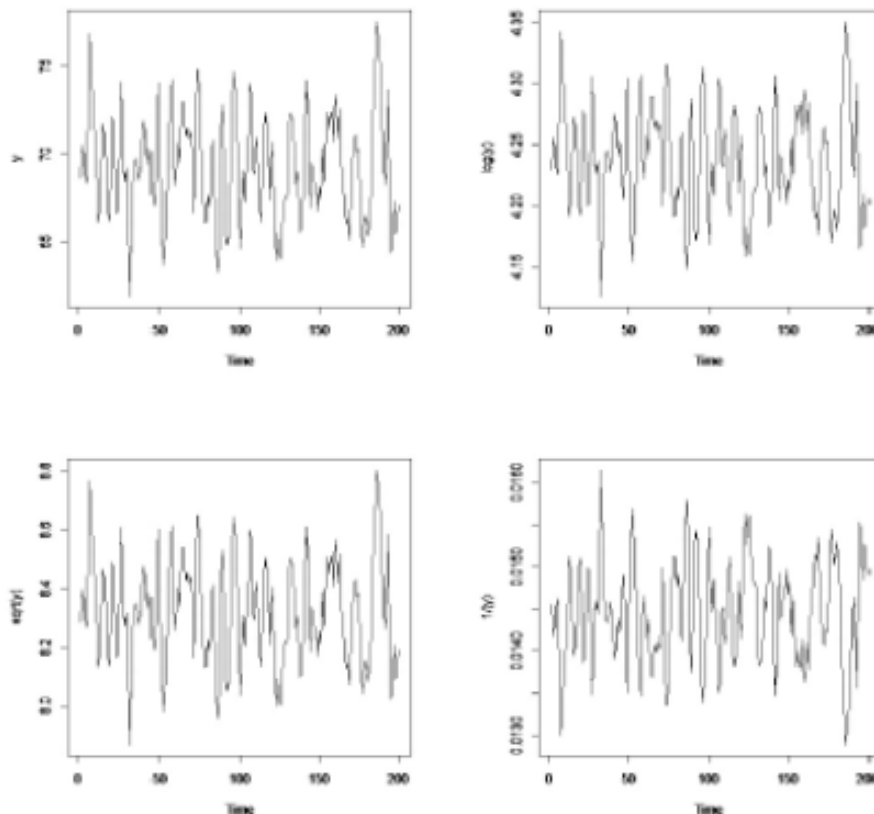
Table 2: Result on Sunspot Numbers

	S	LS	SS
RS	676(0)		
	10.006(0.0004)		
SS	20.529(0.0001)	49.49(0)	
	366.835(0)	15(0.0006)	
LS	18556(0.0001)	55.128(0)	12.012(0.0025)
	5.666(0.058)	24.747(0.00001)	32.528(0)

Inference:

Since when $0.05 > p - \text{value}$ we reject the null hypothesis of parallelism and equal predictive ability. We conclude as follows; the sunspots numbers is not parallel to any of these transformations and neither is any of the transformations to the others, even though they possess similar predictive abilities. Nevertheless, for this reason of no parallelism, using these transformations may be inappropriate for modeling and inference.

Example 2: A set of 200 observations was simulated using the following parameters $\hat{\phi}_0^1 = 22, \hat{\phi}_1^1 = 1.39, \hat{\phi}_2^1 = -0.24, \hat{\phi}_0^2 = 29, \hat{\phi}_1^2 = 1.081, \hat{\phi}_2 = -0.5$ and $r = 4$. Figure 2 represents different transformations of the simulated data.

Figure 2: Plot of square root, $\log_e$, reciprocal and actual data of simulated data

In Table 3, we present the parameter estimates(with standard errors in parenthesis) obtained using the simulated data and logarithm, square root and reciprocal transformations.

Table 3: Estimates of the parameter for simulated data and other transformations

	Simulate data SETAR(2;2,2),d=2 Estimated threshold=71.88		log _e Simulate data SETAR(2;2,2),d=1 Estimated threshold=2.632		Sq.root Simulate data SETAR(2;2,2),d=1 Estimated threshold=3.317		Reciprocal of Simulate data SETAR(2;2,2),d=1 Estimated threshold=0.125	
Lower Regim	Estimates	t-value	Estimates	t-value	Estimates	t-value	Estimates	t-value
Intercept	29.993(4.4)	9.96951	1.974(0.251)	7.871	3.849(0.495)	7.785	0.005(0.004)	3.467
lag 1	1.222(0.067)	18.354	1.209(0.064)	18.888	1.216(0.064)	18.958	1.254(0.0898)	13.959
lad 2	-0.656(0.079)	-8.2	-0.676(0.074)	151.9	-0.680(0.075)	-9.122	-0.595(1.116)	-5.143
RSE	2.016		0.0292		0.1216		0.0004	
n	157		163		163		102	
R ²	0.999		1		0.999		0.999	
Upper Regim	Estimates	t-value	Estimates	t-value	Estimates	t-value	Estimates	t-value
Intercept	10.699(11.59)	0.932	1.179(0.821)	1.436	2.345(1.603)	1.462	0.008(0.0013)	6.447
lag 1	0.992(0.111)	8.944	0.965(0.129)	7.489	0.951(0.127)	7.481	1.0722(0.0751)	14.286
lad 2	-0.169(-0.169)	-0.952	-0.246(0.216)	-1.139	-0.239(0.210)	-1.138	-0.6626(0.0915)	-7.243
RSE	1.811		0.0256		0.108		0.0004	
n	42		35		35		96	

The different models are given below;

The simulated data,

$$z_t = \begin{cases} 29.993 + 1.222z_{t-1} - 0.656z_{t-2}, & z_{t-2} \leq 71.88 \\ 10.699 + 0.992z_{t-1} - 0.169z_{t-2}, & z_{t-2} > 71.88 \end{cases}$$

The $\log_e$ of simulated data,

$$z_t = \begin{cases} 1.974 + 1.209z_{t-1} - 0.070z_{t-2}, & z_{t-2} \leq 4.281 \\ 1.179 + 0.965z_{t-1} - 0.24z_{t-2}, & z_{t-2} > 4.281 \end{cases}$$

The square root of simulated data,

$$z_t = \begin{cases} 3.849 + 1.216z_{t-1} - 0.680z_{t-2}, & z_{t-2} \leq 8.506 \\ 2.345 + 0.951z_{t-1} - 0.239z_{t-2}, & z_{t-2} > 8.506 \end{cases}$$

The reciprocal of simulated data,

$$z_t = \begin{cases} 0.005 + 1.254z_{t-1} - 0.595z_{t-2}, & z_{t-2} \leq 0.0145 \\ 0.0087 + 1.0722z_{t-1} - 0.413z_{t-2}, & z_{t-2} > 0.0145 \end{cases}$$

The tests for parallelism and equal predictive ability for simulated data are as follows:

$$H_0 : \phi_{k,z_1}^j = \phi_{k,z_2}^j \text{ vs } \phi_{k,z_1}^j \neq \phi_{k,z_2}^j \quad (19)$$

$$\alpha = 0.05.$$

Test statistics:

$$W = (A\hat{\Theta})'(A\hat{\Lambda}A')^{-1}(A\hat{\Theta})$$

<http://www.bjs-uniben.org/>

Table 4: Result on Simulated data

	S	LS	SS
RS	2.186(0.335)		
	17.408(0.0002)		
SS	0.463(0.793)	3.866(0.145)	
	0.703(0.704)	8.274(0.016)	
LS	0.274(0.874)	4.01(0.135)	0.0098(0.995)
	0.634(0.728)	8.343(0.015)	0.0044(0.998)

Table 4 shows the values of the test statistic and p-values (in parenthesis) for comparison of simulated data (S) with square root(SS), $\log_e$ transforms(LS) and reciprocal (RS) of simulated data.

Inference: The reciprocal of the simulated data is not parallel to the simulated data, the square root transform and the $\log_e$ as figure 2 shows, but they all have equal predictive ability. For the purpose of forecasting if we wish to transform the data it will be appropriate to use either the square root or $\log_e$ transforms of the data since they possess similar structure with the actual simulated data.

Example 3: Ozone depletion is measured by reduction in the total column ozone above a point on the earth's surface. The series consists of 114 observations. Figure 3 displays plots of Ozone depletion data and other transformations of the data.

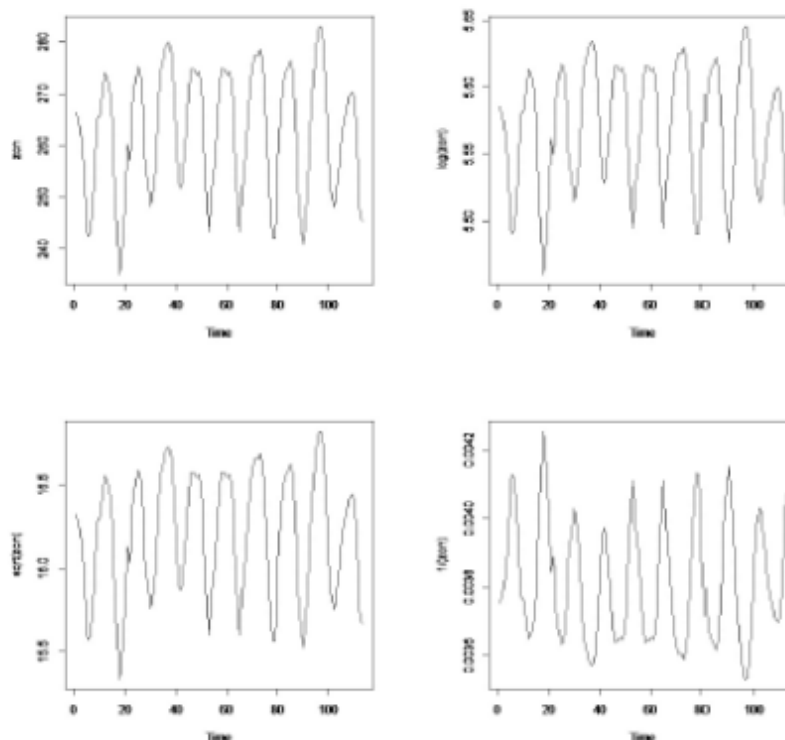


Figure 3: Plot of square root, $\log_e$, reciprocal and actual data of ozone depletion data

In Table 5 we present the parameter estimates(with the standard error in parenthesis) obtained using the ozone depletion data and logarithm, square root and reciprocal transformations.

Table 5: : Estimates of the parameter for Ozone Depletion data, and other transformations

	Ozone Depletion data SETAR(2;3,3),d=3 Estimated threshold=260.7		log _e Ozone Depletion data SETAR(2;3,3),d=3 Estimated threshold=5.563		Sq. root Ozone Depletion data SETAR(2;3,3),d=3 Estimated threshold=16.14		Reciprocal of Ozone Depletion data SETAR(2;3,3),d=3 Estimated threshold=0.004	
Lower Regim	Estimates	t-value	Estimates	t-value	Estimates	t-value	Estimates	t-value
Intercept	76.143(21.136)	3.603	1.609(0.451)	3.567	4.674(1.312)	3.562	0.0008(0.0004)	2.328
lag 1	0.974(0.109)	8.934	0.963(0.107)	9.019	0.968(0.108)	8.976	1.799(0.144)	12.523
lad 2	-0.063(-0.191)	-0.384	-0.071(0.160)	-0.443	-0.067(0.162)	-0.413	-1.076(0.263)	-4.097
lad 3	-0.191(0.115)	-1.662	-0.179(0.111)	-1.606	-0.185(0.113)	-1.634	0.064(0.180)	0.355
RSE	3.213		0.012		0.0998		0.0004	
n	47		47		47		63	
R ²	0.999		1				0.999	
Upper Regim								
Intercept	59.359(24.001)	2.348	1.194(0.514)	2.325	3.477(1.483)	2.344	0.0003(3.433)	3.433
lag 1	1.772(0.140)	12.628	1.786(0.142)	12.583	1.779(0.141)	12.608	0.951(0.104)	9.106
lad 2	-1.053(0.251)	-4.188	-1.065(0.257)	-4.147	-1.059(0.254)	-4.169	-0.079(0.155)	-0.507
lad 3	0.065(0.170)	0.385	0.065(0.175)	0.372	0.065(0.172)	0.379	-0.167(0.108)	-1.547

The different models are given below;

The Ozone depletion data,

$$z_t = \begin{cases} 76.143 + 0.974z_{t-1} - 0.06z_{t-2} - 0.191z_{t-3}, & z_{t-3} \leq 260.7 \\ 56.359 + 1.772z_{t-1} - 1.053z_{t-2} + 0.065z_{t-3}, & z_{t-3} > 260.7 \end{cases}$$

The square root of Ozone depletion data,

$$z_t = \begin{cases} 4.674 + 0.968z_{t-1} - 0.067z_{t-2} - 0.185z_{t-3}, & z_{t-3} \leq 16.14 \\ 3.477 + 1.779z_{t-1} - 1.059z_{t-2} + 0.065z_{t-3}, & z_{t-3} > 16.14 \end{cases}$$

The log_e of Ozone depletion data,

$$z_t = \begin{cases} 1.609 + 0.963z_{t-1} - 0.071z_{t-2} - 0.179z_{t-3}, & z_{t-3} \leq 5.563 \\ 1.194 + 1.780z_{t-1} - 1.065z_{t-2} + 0.065z_{t-3}, & z_{t-3} > 5.563 \end{cases}$$

The reciprocal of of Ozone depletion data,

$$z_t = \begin{cases} 0.008 + 1.799z_{t-1} - 1.076z_{t-2} - 0.064z_{t-3}, & z_{t-3} \leq 0.004 \\ 0.001 + 0.9510z_{t-1} - 0.079z_{t-2} + 0.167z_{t-3}, & z_{t-3} > 0.004 \end{cases}$$

The tests for parallelism and equal predictive ability for ozone depletion data are as follows:

$$H_0 : \phi_{k,z_1}^J = \phi_{k,z_2}^J \text{ vs } \phi_{k,z_1}^J \neq \phi_{k,z_2}^J, \quad (20)$$

$$\alpha = 0.05.$$

Test statistics:

$$W = (A\hat{\Theta})'(A\hat{\Lambda}A')^{-1}(A\hat{\Theta})$$

In Table 6, we present the values of the test statistic and the *p*-values (in paren-

thesis) for comparison of Abuja ozone depletion data (Z) with square root(SZ), $\log_e$ transforms(LZ) and reciprocal (RZ) of Abuja ozone depletion data (Z) .

Table 6: Result on Ozone Depletion data

	Z	RZ	SZ
RZ	21.02(0.0001)		
	20.994(0.0001)		
SZ	0.0035(0.9999)	21.56(0.0001)	
	0.0127(0.9996)	21.37(0.0001)	
LZ	0.0122(0.9996)	21.516(0.0001)	0.0027(0.9999)
	0.0506(0.9970)	21.75(0.0001)	0.0127(0.9996)

Inference: In conclusion, the reciprocal of the ozone depletion data is not parallel to the ozone depletion data, the square root transform and the $\log_e$ as figure 3 shows. However, they all have equal predictive ability. For the purposes of forecasting, it will be appropriate to use either the square root or $\log_e$ transforms of the data since they possess similar structure with the actual data.

4. Conclusion

In this work, we considered the parallelism and equal predictive ability between different Self-Exciting Threshold Autoregressive models in which the comparison was between different transformations of the given time series data. We established conditions for the relationship between equal predictive ability and parallelism. From the work we see that in some cases the equal predictive ability of different models can be tested for, by using the parameters of the model in which we simply consider if they are parallel. In other cases where the estimated R^2 shows equal predictive ability for all, the test will still indicate which model(s) (or data transformation(s)) are more appropriate for forecasting by considering the parallelism between the series (which indicates similarities in the structure of the time series).

References

- Amato, J. and Swanson, N. R. (2001). The real time predictive content of money for output. *Journal of Monetary Economics*, 48, 3-24.
- Asikgil, B. (2018). An Adapted Approach for Self-Exciting Threshold Autoregressive Disturbances in Multiple Linear Regression. *GU J Sci*, 31(4), 1268-1282.
- Aslan, S., Yozgatligil, C. and Iyigun, C.(2018) Temporal clustering of time series via threshold autoregressive models: application to commodity prices. *Annals Operations Research*, 260, 51–77.
- Avdoulas, C., Bekiros, S. and Boubaker, S. (2018). Evolutionary-based return forecasting with nonlinear STAR models: evidence from the Eurozone peripheral stock markets. *Annals Operation Research*, 262, 307–333.
- Blasques, F., Loopman, S. J. and Lucas, A. (2020). Nonlinear autoregressive models with optimality properties. *Econometric Reviews*, 39(6), 559-578.
- Box, G. E. P. and Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
- Guo, J. H . (1999). A Nonparametric Test for Parallelism of Two First-Order Autoregres-

- sive Processes. The Australian and Newzealand Journal of Statistics, 41, 59-65.
- Güris, S. and Güris, B. (2019). GLS detrending in nonlinear unit root test. Communications in Statistics- Simulation and Computation. <https://doi.org/10.1080/03610918.2019.1662442>
- Hamdi, F. and Khalfi, A. (2019). Predictive density criterion for SETAR models. Communications in statistics-simulation and computation. <https://doi.org/10.1080/03610918.2019.1653915>
- Hathaway, D. H. (2014). Royal Observatory, Greenwich-USA/NOAASunspot Data, at <http://solarscience.msfc.nasa.gov/greenwich.shtml>.
- Liu, T. and Zhang, Y. (2020). Law of the iterated logarithm for error density estimators in nonlinear autoregressive models. Communications in Statistics- Theory and Methods, 49(5), 1082-1098.
- Otranto, E. (2004). Classifying the Market Volatility with ARMA Distance Measures. Quaderni di Statistica, 6, 1-19.
- Otranto, E. and Triacca, U. (2007). Testing for Equal Predictability of Stationary ARMA Processes, Journal of Applied Statistics, 34(9), 1091-1108.
- Ozaki, T. and Tong, H. (1975). On Fitting Non-stationary Autoregressive Models in Time Series Analysis. In Proc. 8th Hawaii Int. Conf. on System Sciences, North Hollywood: Western Periodicals, 225-226.
- Siu, T.K. Elliott, R.J. (2020). Bitcoin option pricing with a SETAR-GARCH model. The European Journal of Finance. <https://doi.org/10.1080/1351847X.2020.1828962>
- Steece, B. and Wood, S. (1985). A Test For The Equivalence of k ARMA Models, Empirical Economics, 10, 1-11.
- Stock, J. H. and Watson, M. W. (1989). Interpreting the evidence on money income causality. Journal of Econometrics 40, 161-181.
- Stock, J. H. and Watson, M. W. (1999). Forecasting Inflation. Journal of Monetary Economics, 44, 293-335.
- Tong, H. (1978). On a Threshold Model. In Pattern Recognition and Signal Processing (ed. C. H. Chen). Amsterdam.
- Tong, H. and Lim, K. S. (1980). Threshold Autoregressive, Limit Cycles and Cyclical Data. Journal Royal Statistical Society, B, 42(3), 245-292.
- Tong, H. (1983). Threshold Models in Nonlinear Time Series Analysis. In Lecture Notes in Statistics 21.
- Tsay, R.S. (1989). Testing and modeling threshold autoregressive processes, Journal of American Statistical Association, 84: 231-240.
- Triacca, U. (2004). A Note on Distance and Parallelism Between Two ARIMA Processes Quaderni di Statistica, 6, 1-9.
- Yang, K., Li, H. and Wang, D. (2018). estimation of parameters in the self-exciting threshold Autoregressive process for nonlinear time series of counts. Applied Mathematical Modelling, Doi: 10.1016/j.apm.2018.01.003