BENIN JOURNAL OF
STATISTICS
ISSN 2682-5767
Vol. 4, pp. 75– 89 (2021)

Performance of Some Nonlinear Time Series Models on Non-Stationary Data

I. Akeyede*
Department of Statistics Federal University of Lafia, PMB 146, Lafia, Nigeria

(*Received: 10 November 2020; accepted: 12 March 2021*)

Abstract. Time series analyses are based on assumptions of linearity and stationarity, whereas many real life problems may not satisfy these assumptions. Thus, there is a need for further investigation of nonlinear time series models for cases that are non-stationary coupled with the features of nonlinearity. This study examines the Autoregressive (AR), Self Exiting Threshold Autoregressive (SETAR), Smooth Transition Autoregressive (STAR) and Logistic Smooth Transition Autoregressive (LSTAR) models. Mont-Carlo simulations are conducted using the R statistical package, to investigate the relative performances of these models at sample sizes of 50, 80, 100, 130, 150, 180, 200, 250, 300 and 400 based on the Mean Square Error (MSE), the Residual Variance (RV), the Akaike Information Criteria (AIC) and the Mean Absolute Percentage Error (MAPE). Thereafter, the models were fitted to data on exchange rate and their performances were evaluated. The study found that the LSTAR model outperformed others in all forms of the generated nonlinear autoregressive cases, except for the polynomial models (where SETAR is preferred to the others).

Keywords: SETAR, STAR, LSTAR, stationarity, nonlinear autoregressive model.

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

1. Introduction

Time series models are usually employed to fit and analyse the dynamic behaviour of time series data; for example, the linear models such as autoregressive (AR) models, moving average (MA) models, and mixed autoregressive moving average (ARMA) models (David, 2011). The incorporation of the linear time series models into several statistical and econometric packages makes them more prominent than their nonlinear counterparts. Although these linear

*Corresponding author. Email: imamakeyede@gmail.com

models are quite successful in many applications, they are unable to represent many nonlinear dynamic patterns such as waves, jump, asymmetry, thresholds and amplitude dependence. For instance, inflation rate and gross domestic product (GDP) typically fluctuate around a higher level and are more persistent during expansions, but they stay at a relatively lower level and are less persistent during contractions (Hansen, 2000). For such data, it would not be reasonable to expect a single, linear model to capture these distinct behaviours (Chung-Ming, 2002). Indeed, in some situations, variations in the data do not exhibit simple regularities and are difficult to analyse and predict accurately. Linear recurrence and their combinations for describing the behaviour of such data are often found to be inadequate. Nonlinear time series models have the advantage of being able to capture asymmetries, jumps, and time irreversibility, which are mostly observed in financial and economic time series (Akeyede et al., 2016). They provide a much wider range of possible dynamics for the economic and financial time series data than linear models.

Aslan et al. (2018) proposed Temporal clustering of time series via threshold autoregressive models and applied it on application to commodity prices. The proposed clustering approach is mainly based on feature vectors derived from models estimates. The pricing of Bitcoin options with a view to incorporating both conditional heteroscedasticity and regime switching in Bitcoin returns was studied by Tak and Robert (2020). They adoped a nonlinear time series model combining both the self-exciting threshold autoregressive (SETAR) model and the generalized autoregressive conditional heteroscedastic (GARCH) model for modeling Bitcoin return dynamics. Specifically, the SETAR model is used to model regime switching and the Heston-Nandi GARCH model is adopted to model conditional heteroscedasticity. Shankar et al (2019) examined the determinants of mispricing in single stock futures traded in the National Stock Exchange of India, the second largest global trading venue for such contracts. Selahattin Güri and Burak GLS detrending in nonlinear unit root test The Monte Carlo simulations made indicate that the proposed test is more powerful than the Kruse (2011) test. Using the proposed test, it was examined whether the consumer price index permanent or transitory for 25 countries. According to the results obtained, by using the Kruse test, we find that unit root hypothesis was rejected only in 5 countries while using the GLS Kruse test, the unit root hypothesis was rejected in 15 countries.

The motivation for this study is that many real world problems may not satisfy the assumptions of linearity and/or stationarity. It is essential to investigate data that are non-stationary and nonlinear. This can be done using nonlinear time series models such as the SETAR, the STAR and the LSTAR models. These nonlinear time series models are discussed in the subsequent subsections.

1.1 *Self-Exciting Threshold Autoregressive (SETAR) model*

For the SETAR model, the threshold variable is a certain lagged value of the process itself; hence the adjective; self-exciting. (More generally, the threshold variable may be some vector covariate process or even some latent process, but this extension will not be pursued here. The simplest class of TAR models is the Self Exciting Threshold Autoregressive (SETAR) models of order p introduced by Tong(1983) and specified by equation (1). The popularity of SETAR models is due to their being relatively simple to specify, estimate, and interpret as compared to many other nonlinear time series models. A k-regime SETAR($d; p_1, p_2, \cdots, p_k$) model has the form

$$Y(t) = \begin{cases} \phi_0^1 + \sum_{j=1}^{p_1} \phi_j^1 Y_{t-i} e_t^1, \ if\, Y_{t-d} \leq r_1 \\ \phi_0^2 + \sum_{j=1}^{p_2} \phi_j^2 Y_{t-j} + e_t^2, \ if\, r_1 \leq Y_{t-d} \leq r_2 \\ \phi_0^k + \sum_{j=1}^{p_k} \phi_j^k Y_{t-j} Y_{t-j} + e_t^k, \ if\, r_{k-1} \leq Y_{t-d} \end{cases} \tag{1}$$

where k is the number of regimes in the model, d is the delay parameter, and p_i is the autoregressive order in the ith regime of the model. The threshold parameters satisfy the constraint $-\infty = r_0 < r_1 < \cdots < r_{(k-1)} < r_k = \infty$. In this model, the threshold variable is a lagged value of the process itself; hence the adjective; self-exciting. The threshold variable may be some vector covariate process or even some latent process, but this extension will not be pursued here.

The innovation within the ith regime e_t^i is a sequence of identically independent normal random variables with zero mean and constant variance $\sigma_i^2 < \infty (i = 1, 2, \cdots k)$. The overall process Y_t, is non-linear when there are at least two regimes with different linear models. The common variance σ_e^2 can be estimated by the sample pooled variance in the data. The superscripts in the model (1) indicate states of the world or regimes. Within each regime, it is assumed that the dynamical behaviour of the time series variable follows a linear autoregressive process. The regime that is operative at any time t depends on the observable past history of $\{Yt\}$ itself, in particular, on the value of Y_{t-d}. Therefore, Tong and Lim (1980, p. 249) therefore calls the process in equation (1) a self-exciting threshold autoregressive model.

The class of self-exciting threshold autoregressive (SETAR) model (Tong 1983, 1990) has been widely employed in the literature to explain various empirical phenomena in an observed time series. See the work of Tong and Yeung (1991) for beach water pollution, Yadav, Pope, and Paudyal (1994) for future markets, Watier and Richardson (1995) for epidemiological applications, Lewis and Ray (1997) for sea surface temperatures, Montgomery et al. (1998) for U.S. unemployment, Fuecht et al. (1998) for medical studies, and Clements and Smith (2001) for exchange rate variables. Tong (1990) lists many more examples from diverse fields. The statistical properties and forecasting performance of SETAR models have been extensively examined (Tong, 1990; Hansen, 1996;

1999; 2000; Clements and Smith, 1999; Kapetanios, 2000; De Gooijer and Ray, 2001). The model has certain features, such as limit cycles, amplitude dependent frequencies, and jump phenomena that cannot be captured by a linear time series model. Tong and Lim (1980) showed that the threshold model is capable of producing asymmetric, periodic behavior exhibited in the annual Wolf's sunspot and Canadian lynx data.

The Threshold Autoregressive model with two regimes is considered in this study as an extension of the autoregressive models. This allows for the parameters changing in the model according to the value of an exogenous threshold variable St-d If it is substituted by the past value of which means $S_{t-d} = Y_{t-d}$ then we call it Self-Exciting Threshold Autoregressive model (SETAR). Some simple cases are considered in this study.

1.2 *Smooth Transition AR (STAR) model*

A criticism of the SETAR model is that its conditional mean equation is not continuous (Tsay 2010). The thresholds (r_j) are the discontinuity points of the conditional mean function μt . In response to this criticism, smooth TAR models have been proposed (Chan and Tong, 1986; Terasvirta, 1994). A time series Y_t follows a 2-regime STAR(p) model of the form

$$Y_t = c_0 + \sum_{j=1}^{p} \phi_{0,i} Y_{t-d} + F\left(\frac{Y_{t-d} - \triangle}{s}\right)\left(c_1 + \sum_{j=1}^{p} \phi_{1,i} Y_{t-i}\right) + e_t \qquad (2)$$

where d is the delay parameter and s are parameters representing the location and scale of model transition, and $F(.)$ is a smooth transition function. In practice, $F(.)$ often assumes one of three forms: logistic, exponential, or a cumulative distribution function. The conditional mean of a STAR model is a weighted linear combination between the following two equations:

$$\mu_1 t = c_0 + \sum_{j=1}^{p} \phi_{0,i} Y_{t-d}$$

$$\mu_2 t = c_0 + \sum_{j=1}^{p} (c_0 + c_1) + (\phi_{0,i} + \phi_{1,i})\, Y_{t-d}$$

The weights are determined in a continuous manner by $F[\left(\frac{Y_{t-d} - \triangle}{s}\right)]$. The prior two equations above also determine properties of a STAR model. For instance, a prerequisite for the stationarity of a STAR model is that all zeros of both AR polynomials are outside the unit circle. An advantage of the STAR model over the TAR model is that the conditional mean function is differentiable. However, experience shows that the transition parameters ?and s of a STAR model are hard to estimate. In particular, most empirical studies show that standard errors of the estimates of ? and s are often quite large, resulting in t ratios of about 1.0 (Terasvirta, 1994; Tsay, 2010 and Avdoulas et al, 2018).

1.3 *Logistic Smooth Transition AR (LSTAR) model*

The logistic smooth transition autoregressive of order p [LSTAR(P)] model is:

$$Y_t = F(\gamma, c; Y_{t-d}) = (1 + \exp - \{\gamma (Y_{t-d} - d)\})^{-1} \tag{3}$$

The coefficient$\gamma, \gamma > 0$ is the smoothness parameter and the scalar c is the location parameter and d is known as the delay parameter, the variable $Y_(t - d)(d > 0)$ is then called the transition variable.

This study is aimed at evaluating the performance of nonlinear time series models in (1), (2) and (3) that can be fitted to the data generated from classes of nonlinear second order autoregressive model. Their performances are evaluated by simulations under violation of the stationarity assumption.

2. Materials and Method

The data were generated through simulation in R statistical software using the procedure carried out by Akeyede et al (2016) and Fayçal et al (2019), the parameters of the linear and nonlinear autoregressive functions in models 1-4 are fixed $\alpha_1 = 0.5, \alpha_2 = 0.6$ as to ensure the condition for non-stationarity in the data generated. The sample sizes simulated from each of autoregressive cases are; 50, 80, 100,130, 150, 180, 200, 250, 300 and 400. At a particular choice of sample size, the simulation study was carried out 1000 times for different forms of the autoregressive functions. Thereafter, a number of steps ahead were simulated to sample sizes of 50, 150 and 300 only, representing small, moderate and large sample sizes, respectively, to predict h-steps ahead (where h = 5, 10, . . ., 50) from the data generated from different classes of linear and nonlinear autoregressive time series stated earlier. Each model from (1) to (3) was fitted to the different sizes of the simulated data from (4) to (7) and this was used to forecast the future values. The effect of the sample size on the performance and the predictive ability of the linear and nonlinear models under study were determined. The relative performance of the models were examined using, Mean Square Error (MSE) and Akaike Information Criteria (AIC). The model with lowest criteria is the best among the models for the simulated data.

2.1 *Models selected for simulation*

Data are generated from several linear and nonlinear second orders of general classes of autoregressive functions given as follows. These are time series models and they are considered to investigate the performance of the model study on different forms of nonlinear time series data (linear, Trigonometry, exponential and polynomial)

$$\textbf{Model 1.}\ AR(2) : Y_{ti} = 0.3Y_{ti-1} - 0.6Y_{ti-2} + e_t, \tag{4}$$

with the source code to simulate data of sample size 50 given as $x < -e < -rnorm(50)$
for $(t\ in\ 3:50)x[t] < -0.3 * x[t-1] - 0.6 * x[t-2] + e[t]$

$$\textbf{Model 2.}\ TR(2) : Y_{ti} = 0.5sin(Y_{ti-1}) + 0.6cos(Y_{ti-2}) + e_t, \tag{5}$$

with the source code to simulate data of sample size 50 given as $x < -e < -rnorm(50)$
for $(t\ in\ 3:50)x[t] < -0.5 * sin(x[t-1]) + 0.6 * cos(x[t-2]) + w[t]$

$$\textbf{Model 3:}\ EX(2) : Y_{ti} = 0.5Y_{ti-2} + exp(+0.6Y_{ti-2}) + e_t, \tag{6}$$

with $x < -e < -rnorm(50)$
for $(t\ in\ 3:50)x[t] < -0.5 * x[t-1] + exp(+0.6 * x[t-2]) + e[t]$

$$\textbf{Model 4:}\ PL(2) :\ Y_{ti} = 0.5Y^2_{ti-1} + 0.6Y_{ti-2} + e_t, \tag{7}$$

$$t = 1, 2, \cdots, 50, 80, 100, 130, 150, 180, 200, 250, 300\ and\ 400.i = 1, 2, \cdots, 1000.$$

where, Y_{ti} are present responses simulated from random samples of normal distribution and Y_{ti-1} and Y_{ti-2} are past responses of first and second order respectively. e_{ti} is a random error which is known as white noise which is also distributed from normal distributions as follows: $Y_{ti} \sim N(2000, 20)\ and\ e_{ti} \sim N(1000, 10)$ for non-stationary data structures. The model 4, 5, 6 are trigonometry, exponential and polynomial autoregressive functions respectively with coefficients of Y_{t-1} being 0.5 and Y_{t-2} being +0.6. Results on the effect of sample size and the stationarity of the models are examined in the next section.

2.2 *Forecast performance comparison using Theil's U statistic*

The predictive ability of the four models were examined using MSE and AIC criteria and forecast comparison were carried using Theil's U statistic. Theil's U statistic is a relative accuracy measure that compares the forecasted results with the results of forecasting with minimal historical data. It also squares the deviations to give more weight to large errors and to exaggerate errors, which can help eliminate methods with large errors. $U > 1$ indicate that he forecasting technique is better than guessing, $U = 1$ indicates that he forecasting technique is as good as guessing and $U < 1$ indicates that he forecasting technique is worse than guessing

$$U = \sqrt{\frac{\sum_{i=1}^{n-1}\left(\frac{\hat{Y}_{t+1}-Y_{t+1}}{Y_t}\right)^2}{\sum_{i=1}^{n-1}\left(\frac{Y_{t+1}-Y_t}{Y_t}\right)^2}}$$

3. Results and Discussion

In this section, the performance of the time series models (AR, SETAR, STAR and LSTAR) are examined under non-stationarity assumptions and then fitted. The mean square error and the residual variance of each model at various sample sizes are recorded in Tables 3-5, while that of AIC and MAPE are recorded in Tables 3-6.

Table 1: Performances of the fitted models on the basis of MSE and RV criteria for model 2 under violation of stationarity assumption: $TR(2) : Y_{ti} = 0.5sin(Y_{ti-1}) + 0.6cos(Y_{ti-2}) + e_t$

Sample	MSE				RV			
Size(n)	AR	SETAR	STAR	LSTAR	AR	SETAR	STAR	LSTAR
50	208.98	102.19	125.05	99.65	202.34	117.03	125.05	101.66
80	197.90	97.24	14.17	97.01	199.10	109.12	114.13	97.37
100	175.95	96.29	108.57	95.82	175.12	99.44	108.67	95.83
130	161.78	95.51	106.48	93.59	153.65	96.63	106.54	93.6
150	138.26	94.19	104.23	91.95	138.07	95.85	104.25	91.96
180	121.03	93.72	103.39	91.67	121.42	94.07	103.44	91.67
200	109.57	92.37	102.83	91.20	111.74	94.01	102.84	91.21
250	103.65	91.27	102.54	91.00	105.36	93.36	102.52	89.62
300	102.73	90.61	101.16	89.62	105.80	92.95	101.23	84.33
400	102.02	88.52	98.07	84.32	102.44	90.61	98.08	80.07

From Table 1, it can be seen that the best model is LSTAR at various sample sizes, but the SETAR model competes well with LSTAR when the sample size increases. The performances of the four fitted models increased with increase in sample size. From Table 2, the AIC values show that the best model is LSTAR at various sample sizes followed by the SETAR model. The three nonlinear models are equally good based on MAPE criterion and perform relatively better when the sample size decreases.

From Table 3, it can be seen that the SETAR and the LSTAR models fit the exponential function very well than the other two models between sample sizes 50 and 250. The LSTAR model performs better when the sample size is 300 and above. The AR model is worse at all levels. The performance of the four models increase when sample the size increases.

From table 4 it can be seen that the SETAR model fits the exponential func-

Table 2: Performance of the fitted models on the basis of AIC and MAPE criteria for model 2 under violation of stationarity assumption: $TR(2) : Y_{ti} = 0.5sin(Y_{ti-1}) + 0.6cos(Y_{ti-2}) + e_t$

Sample	AIC				MAPE			
Size(n)	AR	SETAR	STAR	LSTAR	AR	SETAR	STAR	LSTAR
50	6391.49	2033.27	2033.72	1936.12	17.6258	0.0092	0.0097	0.0089
80	6255.74	1843.86	1880.96	1847.41	14.0640	0.0079	0.0089	0.0079
100	5059.64	1383.17	1395.92	1377.70	13.9491	0.0078	0.0084	0.0077
130	4263.29	1152.78	1160.19	1139.89	12.5065	0.0078	0.0083	0.0076
150	3456.85	922.04	923.14	920.27	12.2648	0.0078	0.0083	0.0076
180	3130.92	830.25	840.93	828.36	11.3035	0.0077	0.0082	0.0076
200	2637.79	691.87	706.19	681.21	11.0280	0.0077	0.0081	0.0074
250	2305.74	599.69	607.94	609.13	10.3361	0.0077	0.0081	0.0074
300	1801.72	461.22	470.66	467.82	10.7869	0.0076	0.0081	0.0071
400	1460.82	374.52	384.96	366.56	9.5906	0.0075	0.0079	0.0070

Table 3: Performances of the fitted models on the basis of MSE and RV criteria for model 3 under violation of stationarity assumption: EX(2):$Y_{ti} = 0.5Y_{ti-2} + exp(+0.6Y_{ti-2}) + e_t$

Sample	MSE				RV			
Size(n)	AR	SETAR	STAR	LSTAR	AR	SETAR	STAR	LSTAR
50	331.75	107.29	114.78	102.43	230.07	110.02	119.79	103.44
80	288.90	97.08	113.72	97.35	226.15	96.34	113.74	97.36
100	267.34	96.10	108.32	96.13	223.59	94.65	108.37	96.14
130	249.43	95.33	105.71	95.75	215.29	93.99	105.71	95.75
150	205.00	94.09	104.03	93.40	200.94	92.04	104.29	93.41
180	186.75	93.54	102.83	92.22	183.30	92.03	102.82	92.22
200	180.20	92.25	102.38	90.89	180.03	90.63	102.45	90.89
250	173.22	91.18	101.74	89.54	160.26	88.31	101.73	89.55
300	162.74	88.45	100.67	89.13	158.31	86.27	100.62	80.13
400	141.91	85.89	97.53	71.78	140.10	85.58	97.51	71.78

Table 4: Performances of the fitted models on the basis of AIC and MAPE criteria for model 3 under violation of stationarity assumption: EX(2):$Y_{ti} = 0.5Y_{ti-2} + exp(+0.6Y_{ti-2}) + e_t$

Sample	AIC				MAPE			
Size(n)	AR	SETAR	STAR	LSTAR	AR	SETAR	STAR	LSTAR
50	7002.68	2030.17	2033.58	2036.60	23.5937	0.0058	0.0083	0.0059
80	6326.19	1843.20	1880.07	1847.36	15.3888	0.0054	0.0061	0.0056
100	4833.21	1370.77	1394.63	1377.09	12.5204	0.0054	0.0058	0.0055
130	4074.78	1122.32	1158.81	1138.53	9.7816	0.0054	0.0058	0.0055
150	3306.28	901.85	921.98	917.94	8.3678	0.0054	0.0057	0.0053
180	2995.53	809.91	839.97	825.06	7.0466	0.0053	0.0057	0.0053
200	2525.15	691.67	705.11	700.26	6.4899	0.0053	0.0057	0.0053
250	2208.26	599.55	606.92	609.55	5.1483	0.0053	0.0056	0.0053
300	1727.01	461.13	470.47	426.42	4.3578	0.0052	0.0056	0.0052
400	1401.28	369.08	384.71	347.89	3.2149	0.0051	0.0054	0.0051

tion better than the other models for the sample sizes considered, followed by the LSTAR model based on the AIC criterion. But their performance is not significantly different judging by the MAPE criterion.

From Table 5, the SETAR model gives the best fit at the different sample sizes followed by the LSTAR as shown by the MSE and the residual variance values. The STAR and AR models performed poorly compared to the SETAR

Table 5: Performance of the fitted models on the basis of MSE and RV criteria for model 4 under violation of stationarity assumption: $PL(2) : Y_t = 0.5Y_{t-1}^2 + 0.6Y_{t-2} + e_t$

Sample	MSE				RV			
Size(n)	AR	SETAR	STAR	LSTAR	AR	SETAR	STAR	LSTAR
50	8353.43	587.33	6757.12	608.99	6745.37	395.0563	5755.987	409.23
80	7393.56	477.34	6043.36	593.13	5485.34	337.4387	5643.068	393.10
100	6837.19	398.01	5934.70	590.88	5203.12	320.3332	5593.099	390.80
130	6381.20	387.48	5849.25	566.86	5067.56	304.987	5584.023	366.79
150	6321.07	218.89	5696.23	365.33	5017.78	299.4432	4569.077	335.20
180	5916.31	197.45	5472.05	349.18	4814.17	282.897	4472.003	309.04
200	5156.09	193.79	4835.30	192.11	4160.15	270.007	4196.76	172.10
250	5314.74	190.09	4248.68	132.01	4115.34	190.0007	3835.09	122.00
300	5003.34	190.01	4197.30	108.00	3487.23	189.3435	3244.877	107.99
400	5001.97	189.88	4192.13	101.85	2937.17	179.6745	3190.078	101.80

and LSTAR models.

Table 6: Performances of the fitted models on the basis of AIC and MAPLE criteria for model 4 under violation of stationarity assumption: $PL(2) : Y_t = 0.5Y_{t-1}^2 + 0.6Y_{t-2} + e_t$

Sample	AIC				MAPE			
Size(n)	AR	SETAR	STAR	LSTAR	AR	SETAR	STAR	LSTAR
50	13840.79	980.67	9995.28	2397.56	18.0715	0.0008	10.017	0.0009
80	11050.46	943.79	9875.59	2358.24	17.6511	0.0008	9.1632	0.0009
100	8288.85	898.01	7386.54	1787.50	17.2596	0.0008	8.7866	0.0008
130	6908.31	552.99	6187.38	1508.10	16.7551	0.0008	8.7016	0.0008
150	5527.61	521.88	4964.43	1210.83	16.5065	0.0008	8.6115	0.0008
180	4974.94	396.45	4474.47	998.80	16.1365	0.0007	8.5015	0.0007
200	4146.26	389.99	3727.01	586.82	16.0938	0.0007	8.4335	0.0007
250	3593.67	382.00	3225.54	360.91	15.9967	0.0007	8.2214	0.0006
300	2765.58	367.89	2453.25	353.62	15.6539	0.0007	8.1825	0.0006
400	2213.05	367.87	1962.73	320.41	15.3027	0.0007	8.1074	0.0006

In Table 6, the SETAR model gives the best fit at the different sample sizes followed by LSTAR on the basis of AIC and MAPLE criteria. The STAR and AR model have the lowest AIC and MAPE values.

4. Predictive Ability of Fitted Models for Non-stationary Data Structure

Tables 7-9 show the comparison among the predictive ability of the four fitted models to model 5, 6 and 7 (trigonometric, exponential and polynomial functions) when the sample size of 300 is used in the simulation and the stationarity assumptions are violated. The values of the mean square error and the AIC of the forecast values of each model at various horizons (H), i.e., the number of steps ahead and sample sizes are recorded in the tables. The MSE and AIC criteria are selected so as to be consistent with the criteria used in fitting the data. The model with lowest criteria has the best forecast at every horizon.

Table 7: Forecast performances of the fitted models on the basis of MSE and AIC for model TR(2): $Yti = 0.5sin(Y_{ti-1}) + 0.6cos(Y_{ti-2}) + et$

Step Ahead	MSE				AIC			
(H)	AR	SETAR	STAR	LSTAR	AR	SETAR	STAR	LSTAR
5	0.0079	9.45E-05	1.78E-04	9.73E-05	-122.103	-910.441	-655.87	-613.101
10	0.0065	7.89E-05	8.01E-04	8.71E-05	-215.187	-1060.56	-712.08	-701.044
15	0.0056	6.34E-05	8.91E-04	7.70E-05	-301.561	-1331.97	-940.08	-933.397
20	0.0043	3.90E-05	7.10E-04	1.00E-05	-613.990	-1615.41	-1260.15	-1447.53
25	0.0031	3.32E-05	5.86E-04	1.15E-06	-766.231	-2007.56	-1581.60	-2061.73
30	8.72E-04	3.01E-05	4.98E-04	8.95E-06	-912.561	-2213.71	-1904.16	-2288.61
35	7.12E-04	2.91E-05	4.32E-04	9.65E-06	-987.002	-2576.55	-2227.49	-2407.09
40	4.32E-04	2.11E-05	3.81E-04	4.72E-07	-991.671	-2610.75	-2551.69	-2716.63
45	4.10E-04	1.71E-05	3.41E-04	1.21E-07	-1031.77	-2710.91	-2876.51	-2970.89
50	3.82E-04	1.91E-05	3.09E-04	2.86E-07	-1078.89	-2910.91	-3001.86	-3524.65

The best forecast was observed from SETAR when number of steps ahead is less than 25 followed by LSTAR model but LSTAR model forecasts better than others when the number of steps ahead increased.

Table 8: Forecast performances of the fitted models on the basis of MSE and AIC for model EX(2): $Y_{ti} = 0.5Y_{ti-2} + exp(0.6Y_{ti-2}) + et$

Step Ahead	MSE				AIC			
(H)	AR	SETAR	STAR	LSTAR	AR	SETAR	STAR	LSTAR
5	0.0067	4.67E-06	5.62E-04	8.05E-06	-271.89	-782.82	-690.71	-872.39
10	0.0042	3.56E-06	3.35E-04	6.56E-06	-402.35	-989.77	-812.23	-1034.53
15	0.0040	1.07E-06	2.77E-04	4.95E-06	-582.36	-1271.63	-941.43	-1320.42
20	0.0037	8.77E-07	1.06E-04	3.75E-06	-631.41	-1572.35	-1284.38	-1651.61
25	0.0035	6.13E-07	8.85E-05	1.03E-06	-725.18	-2125.72	-1632.40	-2223.73
30	0.0031	3.51E-07	7.42E-05	7.71E-07	-742.48	-2461.80	-1843.23	-2653.17
35	0.0028	3.11E-07	5.50E-05	5.90E-07	-867.86	-2782.02	-2312.56	-2603.89
40	0.0026	1.66E-07	2.77E-05	2.99E-07	-989.48	-2796.54	-2671.58	-2731.37
45	0.0022	1.23E-07	1.99E-05	1.87E-07	-1100.00	-2831.42	-2762.82	-2872.38
50	0.0017	1.04E-07	1.63E-05	1.32E-07	-1156.39	-2892.91	-2813.77	-2894.27

It was observed that the LSTAR models have the best forecast followed by SETAR based on the two criteria. But both models have close performances when the numbers of steps ahead increase.

The three nonlinear models have close forecast performances based on MSE and in Figure 12, the SETAR models have the best forecast at number of steps greater than 30 based on AIC while SETAR and LSTAR have equal performances at the number of steps that are less than 30.

Table 9: Forecast performance of the fitted models on the basis of MSE and AIC for model 4 PL(2): $Y_t = 0.5Yt - 1^2 + 0.6Y_{t-2} + e_t$

Step Ahead	MSE				AIC			
(H)	AR	SETAR	STAR	LSTAR	AR	SETAR	STAR	LSTAR
5	0.0057	7.83E-07	8.78E-05	8.08E-08	-385.63	-1378.29	-908.60	-1362.52
10	0.0042	5.57E-07	3.76E-05	7.34E-08	-497.39	-1712.27	-980.79	-1756.22
15	0.0035	4.05E-07	1.13E-05	5.43E-08	-322.42	-1883.23	-1003.59	-1897.76
20	0.0034	2.55E-07	9.16E-06	4.08E-08	-591.01	-1879.34	-1130.45	-1902.18
25	0.0031	9.18E-08	8.87E-06	2.44E-08	-772.18	-1909.58	-1222.72	-1923.45
30	0.0027	8.02E-08	8.89E-06	2.57E-08	-890.38	-1997.69	-1412.74	-1987.41
35	0.0025	6.42E-08	7.78E-06	1.99E-08	-967.72	-2203.48	-1567.33	-1998.74
40	0.0021	5.19E-08	6.42E-06	1.41E-08	-996.56	-2511.49	-1890.86	-2143.53
45	0.0021	3.12E-08	5.81E-06	1.20E-08	-1100.19	-2751.56	-2773.76	-2432.39
50	0.0020	3.00E-08	5.34E-06	1.01E-08	-1290.55	-2856.58	-2897.48	-2678.91

5. Application of the Fitted Models to Real Life Data (Exchange Rate)

For the purpose of the flow of the analysis, the time series data on monthly exchange rate from government for the periods of twenty are shown in Figure. Before the data were fitted to different models, tests of linearity and stationarity were carried out using Keenan and Tsay F-tests for linearity and Augmented Dickey-Fuller (ADF) and Philip Peron (PP) tests for stationarity.

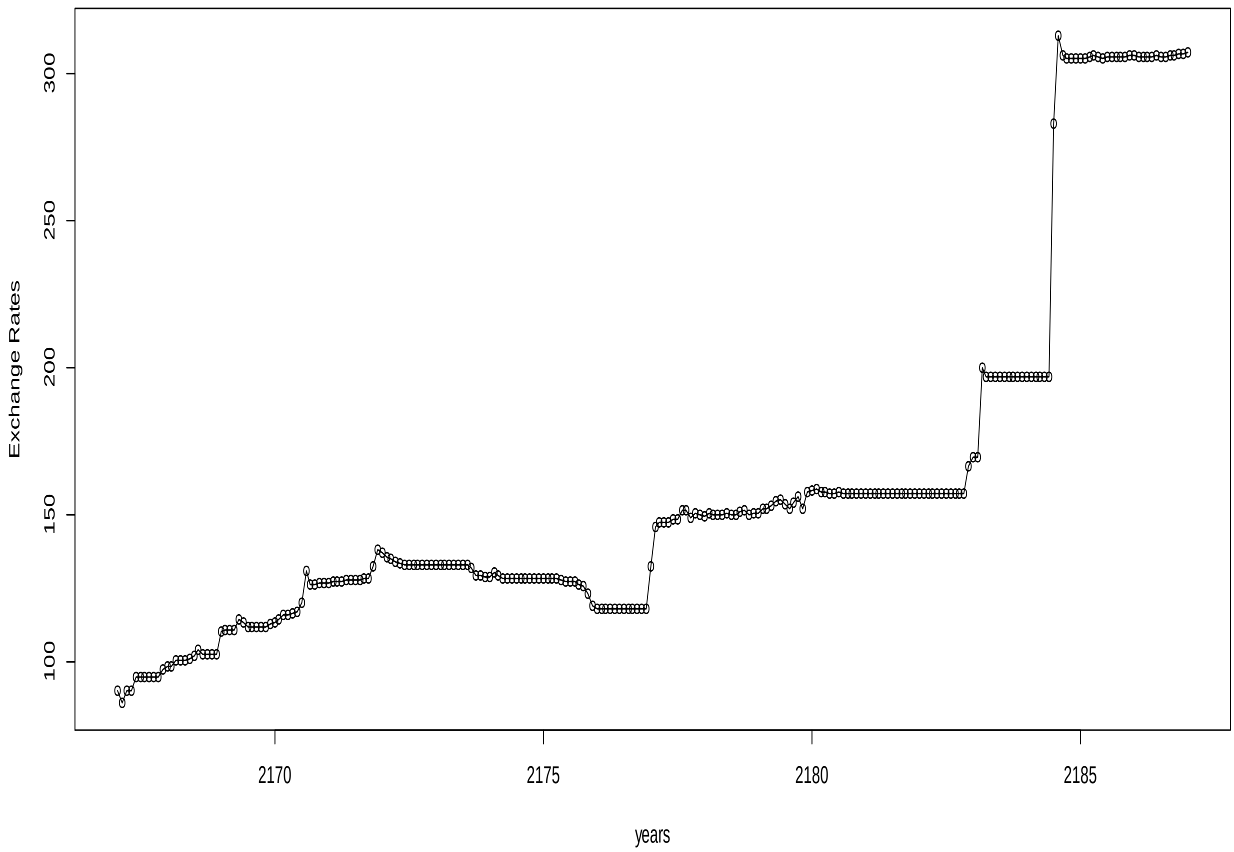

Figure 1: Nigeria's exchange rate (official)

Table 10: Test of nonlinearity and non-stationarity on monthly exchange rate

Stationarity Test	Test of Nonlinearity			Test of Unit root
	Keenan Test	Tsay F-Test	ADF	PP
			p-value	
Test-Stat	1.4837	1.787	-1.1283	4.1779
DF	24	24	24	24
P-value	0.2239	0.09424	0.9162	0.01
Decision	Accept Nonlinearity	Accept Nonlinearity	Accept Unit Root	Accept Unit Root
Remarks	Nonlinear	Nonlinear	Not Stationary	Not Stationary

Table 11: Performance of the fitted models on exchange rate data

Model	MSE	AIC	MAPLE	RV
AR	6022.00	5546.21	0.9756	2.6507
SETAR	5873.27	4179.52	0.8833	2.4009
STAR	5998.32	4181.63	0.8687	2.3017
LSTAR	5873.29	4181.52	0.8634	2.3002

Table 12: Forecast performance of the models using Theil U statistic

Steps Ahead	AR	SETAR	STAR	LSTAR
5	1.33013	2.84114	2.57219	3.1075
10	1.30324	2.81424	2.54529	3.09641
15	0.27634	2.78735	2.5184	3.0435
20	0.24944	2.76045	1.4915	3.0286
25	0.22255	2.73356	1.46461	3.0026
30	0.19565	2.70666	1.43771	2.9785
35	0.16876	2.67977	1.41082	2.94872
40	0.14186	2.65287	1.38392	2.92182
45	0.11497	2.62598	1.35703	2.89493
50	0.08807	2.59908	1.33013	2.86803

Table 10 shows that the null hypothesis of nonlinearity was accepted for the exchange data using the two statistics. Also, the null hypothesis of unit root was accepted in favour of alternative of stationarity in both ADF and PP tests for the real data and therefore the data is nonlinear and non-stationary.

Table 11 shows that LSTAR performs better than others based on the two criteria. The 12-Months Forecast plots of the Exchange rate using the four fitted models are presented in Figures 2

Based on the Theil's analysis above, the LSTAR has the highest forecasting power due to their values greater than 1 and also greater than other values of the models across the steps ahead; this is followed by SETAR and STAR. However the Theil values of AR, at higher steps ahead are close to zero, hence it is not as good as other models in forecasting. Indeed, the forecasting ability of all the models decrease as steps ahead increase.

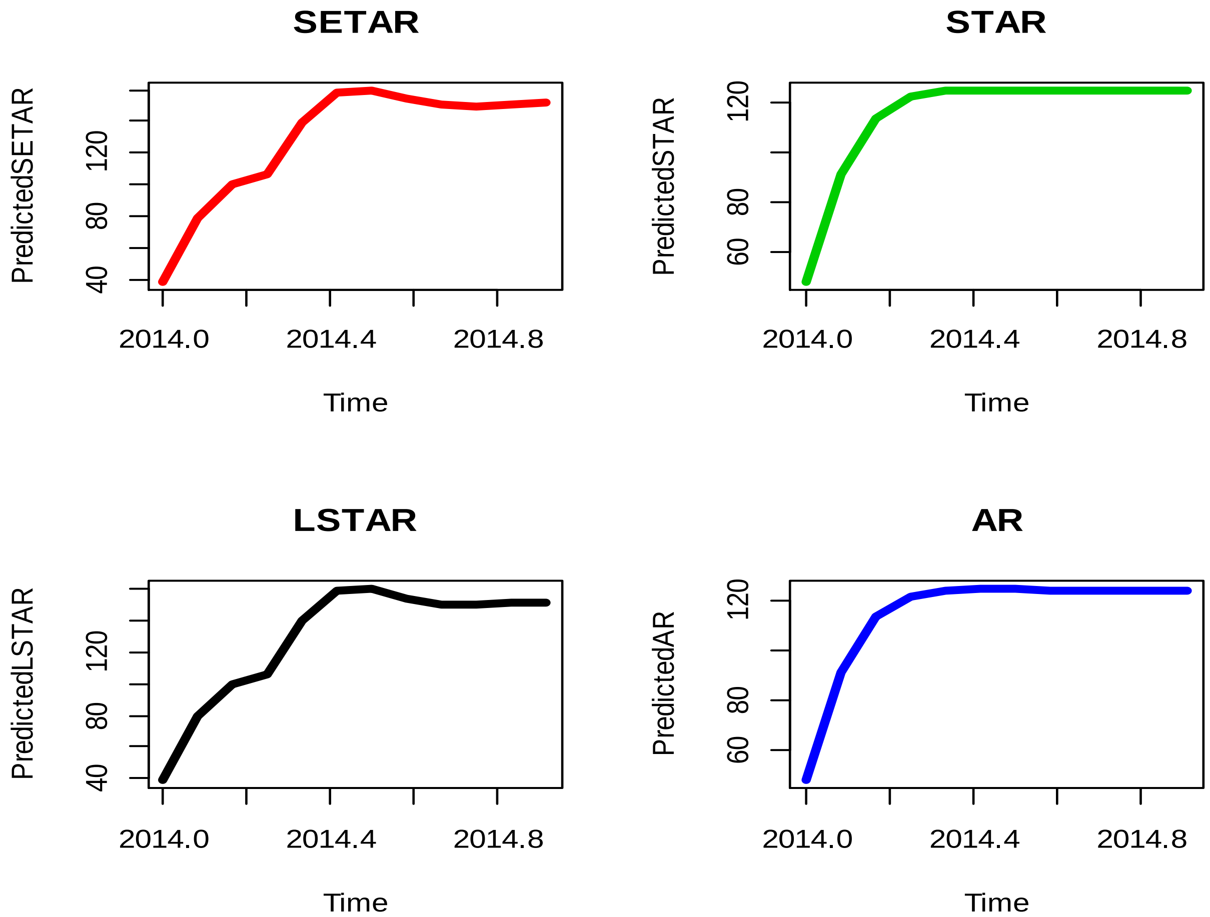

Figure 2: 12-months forecasts plots of the monthly exchange rate

6. Conclusion

The performance of LSTAR model superseded other models in nonlinear non-stationary autoregressive functions as the sample size increases except in polynomial function where SETAR model performs better than others. The three nonlinear models SETAR, LSTAR and STAR have close performances in exponential autoregressive function when sample size increases and performances of the four fitted models increase in all forms of autoregressive functions. The SETAR models had the best fit to linear autoregressive, exponential and polynomial functions at sample sizes that were less than 250 while LSTAR performed relatively better at 250 and above for non-stationary data structures. Furthermore, The LSTAR model outperformed others in forecasting as number of steps ahead increases. The predictive ability of all fitted models increase as sample sizes and numbers of steps ahead increase. Finally, it was observed that LSTAR model fits best to the data on official exchange rate (Naira to Dollar).

Acknowledgement

The author is grateful to the anonymous reviewers for their careful reading of the manuscript and their insightful comments and suggestions. Their valuable suggestions has enhanced the presentation and ideas of the article.

References

Akeyede, I., Adeleke, B.L. and Yahya, W.B. (2016). Linear and nonlinear time series model selection for stationary data structure: application to monthly rainfall data in Nigeria. Journal of Physical Sciences, **21**: 1–13.

Aslan, S., Yozgatligil, C. and Iyigun, C. (2018) Temporal clustering of time series via threshold autoregressive models: application to commodity prices. Ann. Oper. Res., **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. Ann. Oper. Res., **262**: 307–333.

Chan, K. S., and Tong, H. (1986). On estimating thresholds in autoregressive models. Journal of Time Series Analysis, **7**: 178–90.

Chung-Ming, K. (2002). Lecture on the Markov switching model. Institute of Economics Academia Sinica, www.sinica.edu.tw/as/ssrc/ckuan

David, G. and Darfiana, N. (2011). Threshold autoregressive models in finance: a comparative approach. Applied Statistics Education and Research Collaboration (ASEARC) - Conference Papers, Faculty of Engineering and Information Sciences, University of Wollongong.

DeGooijer, J. and Ray, B. (2001). Using POLYMARS to model vector time series,preprint, Department of Mathematical Sciences, New Jersey Institute of Technology.

Fayçal, H. and Abderaouf, K. (2019). Predictive density criterion for SETAR models. Communications in Statistics - Simulation and Computation, **74**: 1–7. doi:10.1080/03610918.2019.1653915

Feucht, M., Moller, U., Witte, H., Schmidt, K., Arnold, M., Benninger, F., Steinberger, K. and Friedrich, M. H. (1998). Nonlinear dynamic of 3Hz spike-and-wave discharges recorded during typical absence seizures in children. Cerebral Cortex, **8**: 524–33.

Ganesh, R., Sankar, L. and Kumar, K. K. (2019). Mispricing in single stock futures: empirical examination of Indian markets. Emerging Markets Finance and Trade, **55**(7): 1619–1633.

Hansen, B. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica, **64**(2): 413-430.

Hansen, B. (1999). Testing for linearity. Journal of Economic Surveys, **13**: 551–76.

Hansen, B. (2000). Sample splitting and threshold estimation. Econometrica, **68**: 575–603.

Lewis, P., and Ray, B. (1997). Modeling nonlinearity, long-range dependence, and periodic phenomena in sea surface temperatures using TSMARS. Journal of the American Statistical Association, **92**: 881-893.

Montgomery, Zarnowitz, A., Tsay, R. S. and Tiao, G. (1998). Forecasting the U.S. unemployment rate. Journal of the American Statistical Association, **93**: 478–93.

Kapetanios, G. (2000). Small sample properties of the conditional least squares estimator in SETAR models. Economics Letters, **69**: 267–76.

Tak, K. S. and Robert J. E. (2020). Bitcoin option pricing with a SETAR-GARCH model. The European Journal of Finance, **60**, 51–77.

Selahattin, G. and Burak, G. (2019) . GLS detrending in nonlinear unit root test. Communications in Statistics - Simulation and Computation, **74**: 1–7.

TeraSvirta, T. (1994). Specication, estimation and evaluation of smooth transition autoregressive models. Journal of the American Statistical Association, **89**: 208–18.

Tong, H. (1983). Threshold models in nonlinear time series analysis. Lecture Notes in Statistics, Springer, New York.

Tong, H. (1990). Nonlinear time series: a dynamical system approach. Clarendon, Oxford.

Tong, H. and Yeung, I. (1991). On tests for self-exciting threshold autoregressive-type nonlinearity in partially observed time series. Applied Statistics, **40**: 43–62.

Tsay, R. S. (2010). Analysis of financial time series (3rd ed.). John Wiley & Sons, Inc., New Jersey.

Tong H. and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. Journal of Royal Statist. Soc. B, **42**: 245-292.

Watier, L., and Richardson, S. (1995). Modelling of an epidemiological time series by a threshold autoregressive model. Journal of the Royal Statistical Society, Series D, **44**: 35-64.

Yadav, P. K., Pope, P. F. and Paudyal, K. (1994). Threshold autoregressive modelling in finance: the price difference of equivalent assets. Mathematical Finance, **4**: 205–21.