9th National Convention on Statistics (NCS)
EDSA Shangri-La Hotel
October 4-5, 2004

A Procedure for Detecting SETAR-Nonlinearity

by
Joselito C. Magadia

For additional information, please contact:

Author's name:	Joselito C. Magadia
Designation:	Professor
Agency:	School of Statistics
	University of the Philippines
Address:	Diliman, Quezon City, Philippines
Telefax:	(632) 928-0881
E-mail:	joselito.magadia@up.edu.ph

A Procedure for Detecting SETAR-Nonlinearity

by
Joselito C. Magadia [1]

ABSTRACT

The inadequacy of the Box-Jenkins ARMA model to address nonlinear phenomena, such as time-irreversibility, chaos, amplitude-frequency dependence, limit cycles, to name just a few, stems from ARMA's intrinsically linear form. Thus, a number of nonlinear time series models have been developed. Among these is a class of models known as Self-exciting Threshold Autoregression (SETAR) models.

This paper proposes a test for detecting nonlinearity of the SETAR type. The performance of this test is compared to that of Tsay's TAR-F and that of Petrucelli and Davies' reverse CUSUM test.

Keywords: nonlinearity, ARMA, frequency-domain, bilinear models

I. Introduction

A number of nonlinear time series models have been proposed to address the limitations posed by the more popular Box-Jenkins ARIMA models. This paper focuses on one such nonlinear time series model, the self-exciting threshold autoregressive (SETAR) model.

A SETAR model is a piecewise linear autoregressive model. It is piecewise linear, not in time, but in the space of the threshold variable. A simple SETAR model is as follows.

$$X_n = \begin{cases} \phi_0^{(1)} + \phi_1^{(1)} X_{n-1} + \dots + \phi_{p_1}^{(1)} X_{n-p_1} + \varepsilon_n^{(1)}, & X_{n-d} \geq r \\ \phi_0^{(2)} + \phi_1^{(2)} X_{n-1} + \dots + \phi_{p_2}^{(2)} X_{n-p_2} + \varepsilon_n^{(2)}, & X_{n-d} < r \end{cases}$$

where $\begin{pmatrix} \varepsilon_n^{(1)} \\ \varepsilon_n^{(2)} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \right)$.

Aside from its ability to exhibit and/or address nonlinear behavior and its competence in approximating other classes of nonlinear models, SETAR has other attractive features, among them ease of interpretation. The popularity of the autoregressive process of order p, or AR(p), process as a time series model is due to its simple interpretation: the present observation is just a linear combination of previously observed values plus an error term. The extension of AR(p) to SETAR involves only some modifications, viz. the thresholds and the delay parameter.

The notion of 'thresholds' can easily be incorporated in most analyses. After all, it is reasonable to presuppose that the dynamics of a system (whether it

---

[1] Professor, School of Statistics, University of the Philippines, Diliman, Quezon City, Philippines

be biological, economic, social, environmental, etc.) may change as a result of attaining a certain value in the state space. For example, the population of an organism in an isolated environment can experience change in its growth pattern as it exceeds, or goes below, some bounds. The delay parameter, on the other hand, can be explained by the fact that it may take some time before a system changes its dynamics.

A list of real applications given by Tong(1990) shows that SETAR has been considered by researchers in diverse fields. These include topics that range from economics, finance, population dynamics, neural science, geophysics, meteorology, automatic control, medicine and even earthquake prediction. It could be argued that an appropriate SETAR model would give more insight into a system's dynamics than a properly fitted ARMA model

All these point to the utility of the SETAR model. The question that naturally arises then is: when is SETAR appropriate? It is in the hope of answering this question that this paper proposes a test procedure for detecting SETAR nonlinearity.

II. Objective and significance of the study

The primary objective of the study is to develop a testing procedure using the frequency-domain approach for detecting SETAR. The frequency-domain approach is adopted here to overcome the difficulties presented by the time-domain approach. The time-domain approach makes use of the correlogram for exploring the nature of the internal structure of a time series. The correlogram, denoted by $\{\rho_k\}$, however, is such that its sampling distribution not only involves unknown parameters but the estimators $\hat{\rho}_k, \hat{\rho}_{k'}$ for $k \neq k'$ are also correlated with one another (Kendall, *et al.*, 1983).

In the frequency-domain approach, the correlogram is (Fourier-)transformed to a function*, f,* called the *spectral density function* (sdf), defined on the interval $(-\pi,\pi]$ and has the desirable properties that functional values are asymptotically independent and can subsequently be transformed to have a distribution which is not dependent on unknown parameters.

III. Review of relevant tests

Many authors have proposed tests for detecting nonlinearities. A few relevant tests are mentioned here.

In 1980, Subba Rao and Gabr(1984) proposed a frequency domain approach for detecting nonlinearity. Hinich (1982) pointed out that Subba Rao and Gabr's test (SRG) statistic can be sensitive to outliers. He proposed an improved and robustified version of the test. Chan and Tong(1986) suggest that great skill is necessary in applying the SRG test because of the large number of parameters involved, which are often chosen subjectively in the absence of extraneous information. In their paper, they also tested Hinich's 'improvement' which performed poorly as a test for SETAR nonlinearity.

More recent tests include the Brock, Dechert, and Scheinkman (BDS) test and the Ramsey and Rothman time reversibility (TR) test. Each of these tests was developed to be run against a general unspecified alternative hypothesis. However, according to Rothman(1993), other tests were specifically designed to have greater power against SETAR alternatives. Cromwell, *et al.*(1994a and 1994b) present other test procedures not only for detecting nonlinearities but also stationarity, serial independence, cointegration, *etc.*

Tong (1990) reviews a number of tests for detecting nonlinearity. These tests include a number of graphical methods and portmanteau tests as well as several tests which deal specifically with SETAR nonlinearity. The literature provides two classes of tests that are SETAR-specific: one class is based on the (conditional) likelihood ratio (see Chan(1990), Hansen(1996) and Carrasco(1997)) and essentially consists of tests of the goodness-of-fit of the hypothesized SETAR model. Analytic expressions for these distributions are difficult to obtain, however. (Tong, 1990) Another class is based on arranged autoregression (see Petrucelli and Davies(1986) , Petrucelli (1988) and Tsay(1989)).

For other nonlinear models, the Lagrange Multiplier approach for the first class of tests yields tests statistics that are asymptotically chi-square distributed under the null hypothesis of a linear model. The usual theory does not give the correct null distribution.

IV. Lag and spectral windows

Given N observations $X_1, X_2, \ldots, X_N$ from a zero mean stationary process, we form the sample analogue of (3.3) which we shall call the 'modified periodogram,' or the 'spectrogram,' or the 'periodogram,' for short, given by

$$\hat{f}^*(\omega) = \frac{1}{2\pi}\left[\hat{R}(0) + 2\sum_{r=1}^{N-1}\hat{R}(r)\cos\omega r\right]$$

where $\hat{R}(r) = \frac{1}{N}\sum_{n=1}^{N-|r|} X_n X_{n+|r|}$.

But since this estimator of the sdf has a large variance, it is modified (see Priestley, 1962) to

$$\hat{f}(\omega) = \frac{1}{2\pi}\left[\lambda(0)\hat{R}(0) + 2\sum_{r=1}^{M_N}\lambda(r)\hat{R}(r)\cos\omega r\right] \tag{1}$$

where $\{\lambda_r)$ are a set of weights called the **lag window** and M_N (<N-1) is some integer, called the **truncation point**.

V. Amplitude-modulated processes

Parzen (1963) and Jones (1971) considered the problem of estimating the sdf of a zero mean stationary process X_n with missing observation. To do this, we consider the process

$$b_n = \begin{cases} 1, & \text{if } X_n \text{ is observed at time } n \\ 0, & \text{otherwise.} \end{cases}$$

We then call $Y_n = b_n X_n$ the *amplitude-modulated version of* X_n *, i.e.*, Y_n is just the original process X_n wherein missing observations are recorded as zeros.

Given observations (x_1, x_2, . . ., x_N) from a stationary process, we estimate the sdf in a manner similar to the previous section. If a nonzero mean is suspected, we first subtract the quantity $\tilde{x} = \frac{\sum_{i=1}^{N} b_i x_i}{\sum_{i=1}^{N} b_i}$ from each observation. We, then, use a modified estimator of the autocovariances given by

$$\hat{R}(r) = \frac{\sum_{n=1}^{N-|r|} b_{n+|r|}\, x_{n+|r|}\; b_n\, x_n}{\sum_{n=1}^{N} b_n^2}. \qquad (2)$$

Substituting (2) in (1) would give an estimator with the following asymptotic properties (see Jones(1971)):

$$Var[\hat{f}(\omega)] \approx f^2(\omega) \sum_{k=-M}^{M} d_k \lambda_k^2 \qquad (3)$$

and

$$\operatorname{cov}[\hat{f}(\omega_i), \hat{f}(\omega_j)] \approx f(\omega_i) f(\omega_j) \sum_{k=-M}^{M} d_k \lambda_k^2 \cos(\omega_i - \omega_j)k$$

$$= O(M/N),$$

where

$$d_k = \frac{\sum_{t=1}^{N} b_{t+k}^2 b_t^2}{\left(\sum_{t=1}^{N} b_t^2\right)^2}.$$

VI. Detecting SETAR-nonlinearity

Given two realizations (y_1, y_2, . . . , y_N) and (z_1, z_2, . . . , z_N) from stationary processes $\{Y_n\}$ and $\{Z_n\}$, respectively, suppose it is of interest to know if there is a difference between the two processes or if the realizations are derived from the same process. To address this matter, we might reason in the following manner: A stationary process may be characterized, among other things, by its spectral density function. Therefore, if we can test the equality of the sdf's of $\{Y_n\}$ and $\{Z_n\}$ then we can conclude with a certain degree of confidence, whether the two realizations come from different processes or not.

Proposition . Let $\hat{f}_Y(\omega^*)$ and $\hat{f}_Z(\omega^*)$ denote the value of the sample sdf's at a fixed frequency ω^* computed from their respective realizations using the same scale parameter lag window and the same truncation point M_N. Suppose $E(\hat{f}_Y(\omega^*)) \neq 0$ and $E(\hat{f}_Z(\omega^*)) \neq 0$. Under the null hypothesis H_0 that the observations come from the same process, then

$$M_{Z,Y} = \sum_{j=1}^{k} [M_{Z,Y}^{(1)}(\omega_j)]^2. \tag{4}$$

has an asymptotic central chi-square distribution with *k* degrees of freedom,

where $$M_{Z,Y}^{(1)}(\omega^*) = \frac{\log\left|\hat{f}_Z(\omega^*)\right| - \log\left|\hat{f}_Y(\omega^*)\right|}{\sqrt{2\frac{M_N}{N}\sum_r \lambda_r^2}}$$

and ω_j are the Fourier frequencies. (see procedure described below).

Proof: see Magadia(2000).

V. A Proposed Test Procedure

Based on the foregoing, the following procedure for testing a univariate time series for SETAR-nonlinearity of the form

$$X_n = \sum_{i=1}^{k_j} \phi_i^{(j)} X_{n-i} + \sigma_j \varepsilon_n^{(j)}$$

is proposed.

Given observations $\mathbf{X} = (X_1, X_2, \ldots, X_N)$:

1) choose ξ such that $0.5 \leq \xi < 1$.
2) Let $Y^* \equiv 100\xi th$ percentile of **X** and let $Z^* \equiv 100(1-\xi)$th percentile of **X**.
3) For $d = 1, 2, \ldots, d^*$, define the sequences $\{Y_n\}$ and $\{Z_n\}$, $n = 1, 2, \ldots, N - d$ by

$$Y_n = \begin{cases} X_{n+d}, \; if\; X_n \leq Y^* \\ 0, \quad otherwise, \end{cases}$$

and $$Z_n = \begin{cases} X_{n+d}, \; if\; X_n \geq Z^* \\ 0, \quad otherwise. \end{cases}.$$

4) Compute the autocovariance sequences $\{\hat{R}_Y(r)\}$ and $\{\hat{R}_Z(r)\}$, for $r = 1, 2, \ldots, M$.
5) Choose a lag window $\lambda(r)$, set $\Delta\omega = B_W/2$ (see Table 1, below) and compute $\hat{f}_Y(\omega)$ and $\hat{f}_Z(\omega)$ at the Fourier frequencies $\omega = \Delta\omega, 2(\Delta\omega), 3(\Delta\omega), \ldots, k(\Delta\omega)$ where $k = [\pi/\Delta\omega]$.
6) Obtain the value of the test statistic given by eqn. (4) with the value of the denominator adjusted by the quantity suggested by eqn (3) and conclude that a SETAR-nonlinearity is indicated by the test if, for some

d, $M_{Z,Y} > \chi^2_{k,\alpha}$, where $\chi^2_{k,\alpha}$ is such that if χ^2 is distributed as a central chi-square random variable with k degrees of freedom then $\Pr(\chi^2 > \chi^2_{k,\alpha}) = \alpha$.

Table 1 . Values of B_W for selected windows

Estimator	B_W
Fejèr window	4.9/M
Parzen window	12/M
Bartlett-Priestley window	1.55π/M

An explanation is in order. The sequences $\{Y_n\}$ and $\{Z_n\}$ can be interpreted as "observations generated by the lower 100ξth percentile of **X**," and "observations generated by the upper 100ξth percentile of **X**,", respectively, given d. If the process that generated **X** is SETAR, then, for some d, the test statistic eqn. (4) would indicate a difference, provided a threshold value lies between $X_{(1)}$ and $X_{(N)}$, the smallest and largest observations, respectively. On the other hand, if the process that generated **X** is of the form $X_n = f(X_{n-1}, X_{n-2}, \ldots, \varepsilon_n, \varepsilon_{n-1}, \ldots)$ where f is a continuous, analytic function (which can be linear, as in ARMA models, or nonlinear, as in ARCH models), then eqn. (4) is expected to give a small value, indicating that there is not much difference between the spectra of the sequences generated by the lower 100ξth and upper 100ξth percentiles. This is so, because under the latter case, there is one and only one function that generated both sequences $\{Y_n\}$ and $\{Z_n\}$.

The performance of the proposed test procedure will be compared with two other tests which were designed specifically for SETAR. These are Tsay's TAR-F (TF) test and the Petrucelli-Davies' reverse CUSUM (RC) test. (see Magadia,2000 for a description of these tests.)

Before the results of the comparisons are presented, some remarks should be made:

(a) In practice, ξ should be chosen to be between 0.6 and 0.75. This would generate Y and Z processes with approximately 100(1-ξ)% zeroes.

(b) The choice for the lag window should be one which has a nonnegative spectral window. The Fejèr, Parzen and Bartlett-Priestley windows all have this property. (see Chatfield,1989).

(c) The proposed test cannot detect SETAR-nonlinearity if the model is of the form : SETAR(d ; p, p, . . ., p) with $\phi_0^{(i)} \neq \phi_0^{(j)}$ for $i,j = 1, 2, \ldots, l$, $i \neq j$ and for $k = 1, 2, \ldots, p$, $\phi_k^{(1)} = \phi_k^{(2)} = \ldots = \phi_k^{(l)}; \sigma_1 = \sigma_2 = \ldots = \sigma_l$.

VI. A Comparison of Tests using Simulation

Tsay (1989) pointed out that "(s)ince the number and location of the thresholds are unknown, there exists no (global) most powerful test for threshold nonlinearity." The performance of the proposed test is thus evaluated using real and simulated data. All previous tests for nonlinearity used the same approach and as much as possible, the same data sets and models used here are the same as those in the literature.

First, the proposed test is applied to some data sets which have been extensively analyzed in the literature. These are i) the annual sunspot series data from 1700 - 1989, ii) the Canadian lynx series, iii) the blowfly population data (observations 48 to 206), and iv) Series A and Series C from Box and Jenkins (1976). For the TF and RC tests, the delay parameter *d* and the AR order *p* used are the same as those adopted by Tsay (1989, 1991). The delay parameter is chosen from 1, 2 , 3 and/or 8. For the Canadian lynx data, Tsay used *p* = 11 in the 1989 article, while the 1991 article used *p* = 2. Also, r_{min} = [N/10] + *p*. The proposed test used ξ= 0.67 and all the nonnegative windows. For the Fejèr and Bartlett (BP) windows, the truncation point M is given by a value near [$\xi N^{1/2}$]. For comparative purposes, the truncation point for the Parzen window was chosen to come up with a bandwidth close to that of the Fejèr and BP windows.

Table 2(a) - (g) P-values of nonlinearity tests on real data, where TF, RC, and B-P denote Tsay's TAR-F test, Petrucelli and Davies' reverse CUSUM and Bartlett-Priestley lag window, respectively, ".000" indicates a p-value less than .001 and " * " denotes a p-value greater than 0.100. For computational expediency the algorithm for the RC test was not programmed to compute the attained p-value once 0.100 is exceeded.

(a) Lynx , n=114

Test	d=1	d=2	d=3
TF (p = 11)	.184	.161	.000
TF (p = 2)	.002	.000	.000
RC (p = 11)	.001	.024	.009
RC (p = 2)	.001	.006	.028
Proposed Test, B-P Window(M = 7)	.000	.023	.836
Proposed Test, Fejèr Window(M = 7)	.000	.000	.899
Proposed Test, Parzen Window(M = 18)	.000	.071	.021

(b) Logged lynx , n=114

Test	d=1	d=2	d=3
TF (p = 9)	.012	.012	.040
RC (p = 9)	*	.001	.040
Proposed Test, B-P Window (M = 7)	.000	.000	.000
Proposed Test, Fejèr Window (M = 7)	.002	.000	.000
Proposed Test, Parzen Window (M = 18)	.000	.000	.003

(c) Blowfly , n=159

Test	d=1	d=2	d=3	d=8
TF (p = 2)	.000	.000	.000	.000
RC (p = 2)	.001	.001	.008	.001
Proposed Test, B-P Window (M = 8)	.000	.001	.047	.075
Proposed Test, Fejèr Window (M = 8)	.004	.883	.948	.000
Proposed Test, Parzen Window (M = 20)	.048	.386	.030	.006

(d) Logged Blowfly , n=159

Test	d=1	d=2	d=3	d=8
TF (p = 11)	.000	.000	.015	.008
RC (p = 11)	.041	.041	.041	.016
Proposed Test, B-P Window (M = 7)	.139	.000	.000	.000
Proposed Test, Fejèr Window (M = 7)	.000	.045	.000	.028
Proposed Test, Parzen Window (M = 18)	.005	.004	.000	.000

(e) Series A, n=197

Test	d=1	d=2	d=3
TF (p = 7)	.455	.366	.746
RC (p = 7)	*	*	*
Proposed Test, B-P Window (M = 8)	.998	.920	.908
Proposed Test, Fejèr Window (M = 8)	.998	.977	.978
Proposed Test, Parzen Window (M = 23)	.998	.894	.843

(f) Series C , n=226

Test	d=1	d=2	d=3
TF (p = 2)	.905	.880	.735
RC (p = 2)	*	*	*
Proposed Test, B-P Window (M = 7)	.000	.000	.000
Proposed Test, Fejèr Window (M = 7)	.000	.000	.002
Proposed Test, Parzen Window (M = 18)	.935	.928	.952

(g) Sunspot , n=289

Test	d=1	d=2	d=3	d=4
TF (p = 11)	.000	.000	.000	.023
RC (p = 11)	.001	.065	.014	.014
Proposed Test, B-P Window (M = 10)	.000	.000	.000	.000
Proposed Test, Fejèr Window (M=10)	.000	.000	.000	.000
Proposed Test, Parzen Window(M = 23)	.000	.000	.003	.000

Table 2(a) - (g) give the following results. All tests show that Series A is linear and the lynx, sunspot and blowfly data are nonlinear. The Tsay's TAR-F, Petrucelli-Davies' RC and the proposed test using the Parzen window show Series C to be linear, while the proposed test using the BP and Fejèr windows indicate otherwise. An interesting point here is that Tsay (1989) mentions Series C as linear, but his more recent New-F (1991) test indicates nonlinearity (not necessarily SETAR-nonlinearity) for d = 2,3. In addition, the latter article reported that the BDS test also indicates nonlinearity at a 5% asymptotic critical value.

The next set of tables give the empirical frequencies of rejecting the null hypothesis of a non-SETAR process based on 1000 replications and 1% and 5% critical values. The model used to generate the data used in constructing Table

3(a) - (c) is $X_n = \begin{cases} 0.5X_{n-1} + \varepsilon_n, \ X_{n-1} \le 0 \\ \phi X_{n-1} + \varepsilon_n, \ otherwise \end{cases}$ as originally proposed by Tsay (1989), where ϕ = -2, -1, -0.5, 0, 0.5 and $\varepsilon_n \sim$ iid N(0,1).

For Table 4(a) - (c), the following models were used to generate the data. Thus the first two models generate non-SETAR processes, in which case the tests should exhibit low empirical frequencies of rejecting the null hypothesis. The next three models, models 3 to 5, generate EXPAR and bilinear processes which are closely related to SETAR per the result of Petrucelli, 1992. The last two models generate SETAR processes; hence the empirical frequencies of rejecting the null hypotheses should be high for these two models.

Model 1**: linear AR(2) $X_n = 0.4X_{n-1} - 0.3X_{n-2} + \varepsilon_n$

Model 2**: linear MA(2) $X_n = \varepsilon_n - 0.4\varepsilon_{n-1} + 0.3\varepsilon_{n-2}$

Model 3*:EXPAR(1) $X_n = (0.3 - 0.8\exp(-X_{n-1}^2))X_{n-1} + \varepsilon_n$

Model 4*: bilinear $X_n = 0.5 - 0.4X_{n-1} + 0.4X_{n-1}\varepsilon_{n-1} + \varepsilon_n$

Model 5**:bilinear $X_n = 0.4X_{n-1} - 0.3X_{n-2} + 0.5X_{n-1}\varepsilon_{n-1} + 0.8\varepsilon_{n-1} + \varepsilon_n$

Model 6*:SETAR $X_n = \begin{cases} 2 + 0.5X_{n-1} + \varepsilon_n, \ X_{n-1} < 1 \\ 0.5 - 0.4X_{n-1} + \varepsilon_n, otherwise \end{cases}$

Model 7 :SETAR $X_n = \begin{cases} 0.6X_{n-1} + \varepsilon_n, \quad X_{n-1} \ge 0 \\ 0.6X_{n-1} + 1.5\varepsilon_n, \ otherwise \end{cases}$

* ≡ from Chan and Cheung (1986).
** ≡ from Keenan (1945), also analyzed in Chan and Cheung (1986).

The sample sizes used are 150, 300 and 500. For each realization of sample size N in the simulation, 5000 + N observations were generated and the first 5000 values are discarded to reduce any effect of the zero starting values, *i.e.,* setting Y_n and ε_n equal to zero for t ≤ 0. The ε_n are standardized normal variates. For the TF and RC tests, $p = d^* = 1$. For the proposed test, ξ = 0.67 and the Parzen window with M = 4, 7, 11 are used.

Table 3(a) - (c). Empirical frequencies of rejecting a non-SETAR model based on 1000 replications of a SETAR(2;1,1) model

(a) N = 150

					Proposed Test (Parzen Window)					
	TF		RC		M=4		M=7		M=11	
φ	α=.05	α=.01	α=.05	α=.01	α=.05	α=.01	α=.05	α=.01	α=.05	α=.01
- 2	999	998	968	945	985	958	946	850	927	835
- 1	975	923	938	903	740	524	622	375	621	416
- 0.5	793	568	800	622	350	176	285	111	304	152
0	286	113	305	121	76	19	66	16	77	33
0.5	38	10	31	5	16	1	6	1	20	2

(b) N = 300

φ	TF α=.05	TF α=.01	RC α=.05	RC α=.01	Proposed Test (Parzen Window) M=4 α=.05	M=4 α=.01	M=7 α=.05	M=7 α=.01	M=11 α=.05	M=11 α=.01
- 2	1000	1000	983	969	1000	1000	1000	1000	1000	999
- 1	1000	1000	965	943	966	905	978	850	945	866
- 0.5	982	932	937	908	724	478	688	424	691	460
0	577	327	597	358	161	49	140	46	174	48
0.5	45	11	31	3	13	2	3	0	6	0

(c) N= 500

φ	TF α=.05	TF α=.01	RC α=.05	RC α=.01	Proposed Test (Parzen Window) M=4 α=.05	M=4 α=.01	M=7 α=.05	M=7 α=.01	M=11 α=.05	M=11 α=.01
- 2	1000	1000	984	972	1000	1000	1000	1000	1000	1000
- 1	1000	1000	972	960	999	996	999	988	999	990
- 0.5	999	995	953	935	928	821	950	782	928	811
0	832	631	818	662	298	112	442	124	359	166
0.5	54	13	42	10	9	1	25	0	9	4

Table 4(a) - (c). Empirical frequencies of rejecting a non-SETAR model based on 5% and 1% critical values. The generating models are given in page 52 and 53.

(a) N = 150

Model	TF (p = 1) α=.05	TF (p = 1) α=.01	RC (p = 1) α=.05	RC (p = 1) α=.01	Proposed Test (Parzen Window) M=4 α=.05	M=4 α=.01	M=7 α=.05	M=7 α=.01	M=11 α=.05	M=11 α=.01
1	22	5	15	3	12	0	4	0	29	8
2	43	8	26	2	35	9	11	4	19	6
3	143	36	125	30	16	4	6	0	19	5
4	997	979	973	943	727	526	602	388	602	406
5	168	59	417	166	984	945	974	906	957	904
6	997	992	969	965	849	624	456	172	420	233
7	75	18	17	2	443	252	256	98	215	74

(b) N = 300

Model	TF (p = 1) α=.05	TF (p = 1) α=.01	RC (p = 1) α=.05	RC (p = 1) α=.01	Proposed Test (Parzen Window) M=4 α=.05	M=4 α=.01	M=7 α=.05	M=7 α=.01	M=11 α=.05	M=11 α=.01
1	22	6	12	3	17	3	4	0	9	2
2	40	5	29	0	41	8	18	2	24	2
3	267	103	281	89	22	3	4	0	14	0
4	1000	1000	976	972	958	889	928	788	910	787
5	322	137	665	412	1000	1000	1000	1000	1000	999
6	1000	1000	973	972	994	970	904	674	921	581
7	79	19	79	1	821	596	574	320	657	253

(c) N = 500

					Proposed Test (Parzen Window)					
	TF (p = 1)		RC (p = 1)		M=4		M=7		M=11	
Model	α=.05	α=.01	α=.05	α=.01	α=.05	α=.01	α=.05	α=.01	α=.05	α=.01
1	24	4	13	2	19	2	4	0	10	1
2	48	8	28	6	33	10	20	4	21	5
3	459	213	510	228	19	3	6	0	8	0
4	1000	1000	979	972	995	983	996	968	992	962
5	513	281	868	687	1000	1000	1000	1000	1000	1000
6	1000	1000	972	972	1000	1000	997	971	984	985
7	95	27	16	4	962	897	876	694	812	793

Based on results from Tables 3 and 4, we note the following.
(i) The power of all the tests increases as the sample size n increases. (ii) The proposed test has less power than the TF and RC tests when the AR processes that comprise the SETAR model are "close". However, the proposed test appears to compensate for this characteristic, since it has smaller Type I errors when the data generating process is indeed non-SETAR, or ARMA to be precise. This is evident from values in the bottom rows of Tables 3 and the first two rows of Tables 4.3.

It may also be noted that the TF, RC, and the proposed tests have rather high empirical frequencies of rejecting the null hypothesis when the data generating process is bilinear. This would support Petrucelli's claim regarding the close relationship between bilinear and SETAR models. The proposed test, however, apparently does not associate SETAR-like properties with the particular EXPAR data generating model chosen.

When the data generating process is indeed SETAR without marked volatility, the power of the proposed test is comparable to that of the TF and RC tests when the sample size is sufficiently large. The fact that the power of the proposed test is lower for moderate sample sizes may be due to the effect of effectively setting 33% of the observations to zero when $\xi = 0.67$ is used.

When the SETAR model is characterized by marked volatility, the simulation results indicate greater power for the proposed test than either the TF or RC test, whatever the sample size.

VII. On the choice of M

It is worthy to note that all of the algorithms presented above require the user to specify the value of d^*, the maximum value that a delay could attain. Tsay's and Petrucelli-Davis' tests require that the maximum value of the order of the AR process be specified, while the proposed test requires the user to specify the value for M. The suggested value of M above albeit *ad hoc,* nevertheless, satisfies theoretical considerations. Several prominent authors have tried to tackle the problem The general consensus is to choose M to achieve the proper balance between the variance and bias of the spectral estimate. A large value of M increases the variance and decreases the bias, while a small value of M has the opposite effect, that of decreasing the variance but increasing the bias. But, to make good use of this principle, we need to have an idea of the process since the variance and bias are functions of the true spectral values.

VIII. Limitations

One limitation of the proposed testing procedure is that it will be applicable to data sets that are relatively lengthy. Estimation of spectral densities, which the procedure uses, requires a larger number of observations than the usual time-domain approach.

In addition, the testing procedures developed in this study are suited to nonstationary processes that can be rendered stationary by variance-stabilizing transformations or differencing, or are piecewise or locally stationary. This is because the assumption of stationarity is essential to the existence of the spectrum, which is the cornerstone of the spectral analysis approach used to construct the proposed test.

References

Box, G.E.P and Jenkins, G. (1976). *Time Series Analysis forecasting and control.* Holden-Day, San Francisco.

Carrasco, M. (1997). *Misspecified Structural Change, Threshold and Markov Switching Models.* Ohio State University.

Chan, K.S. (1990). "Testing for Threshold Autoregression," *Ann. Stat*, 18, 1886-1894.

Chan, W.S. and Tong, H. (1986). "On tests for non-linearity in time series analysis." *Journal of Forecasting,* **5**, 217-228.

_______ (1990). "On Likelihood Ratio Tests for Threshold Autoregression," *J. R. Statist. Soc. B*, 52, No.3, 469-476.

Chatfield, C. (1989). *The Analysis of Time Series.-4th ed.* Chapman and Hall, Ltd.

Cromwell, J. B., Labys, W. C., and Terraza, M. (1994a)**.** *Univariate Tests for Time Series Models.* Sage Publications, Inc., Thousand Oaks, Ca.

Cromwell, J. B., Hannan, M. J., Labys, W. C., and Terraza, M. (1994b)**.** *Multivariate Tests for Time Series Models.* Sage Publications, Inc., Thousand Oaks, Ca.

Hansen, B. E. (1996). *Inference in TAR models.* Working paper, Boston College.

Hinich, M. (1982). "Testing for Gaussinity and linearity of a stationary time series. *Journal of Time Series Analysis,* **3**, 169-176.

Jones, R.H. (1971). "Spectrum estimation with missing observations," *Ann. Inst. Statist. Math.,* 23, 387 - 398.

Kendall, M. G., Stuart, A., and Ord, J. K. (1983). *The Advanced Theory of Statistics, Vol. 3: Design and Analysis, and Time Series. (4th ed.).* Charles Griffin & Co., England.

Magadia, J.C. (2000). *A Frequency-domain Approach for Detecting Self-Exciting Threshold Autoregression (SETAR) Nonlinearity.* Unpublished PhD theses, School of Statistics, University of the Philippines.

Parzen, E. (1963a). "On spectral analysis with missing observations and amplitude modulation;" *Sankhya, A,* 25, 383 -392.

_______ (1963b). Notes on Fourier analysis and spectral windows. Technical Report No. 48, Department of Statistics, Stanford University. Published in *Time Series Analysis Papers* by E. Parzen. Holden-day, San Francisco.

Petrucelli, J. D.(1992). "On the approximation of time series by threshold autoregressive models." *Sankhya*, *B,**54*** , 106-113.

Petrucelli, J.D. and Davies, N. "A Portmanteau test for self-exciting threshold autoregressive-type nonlinearity in time series." *Biometrika,* **73,** 687-694.

Priestley, M. B.(1962). "Basic considerations in the estimation of power spectra." *Technometrics*, **4,** 551 - 564.

Subba Rao, T. and Gabr, M.M. (1980). " A test for linearity of stationary time series." *Journal of Time Series Analysis,* **1**, 145-158.

___________. (1984). *An Introduction of Bispectral Analysis and Bispectral Time series Models.* Lecture Notes in Statistics, **24**, Springer, New York.

________ (1981). *Spectral Analysis And Time Series, Vols. I, II.* Academic Press, New York.

Tong, Howell (1990) *Non-linear Time Series: A Dynamic System Approach*. Oxford University Press, Oxford.