Statistica Sinica 1(1991), 453-464

POWER PROPERTIES OF A TIME SERIES LINEARITY TEST AGAINST SOME SIMPLE BILINEAR ALTERNATIVES

Pentti Saikkonen and Ritva Luukkonen

University of Helsinki and Institute of Occupational Health

Abstract: The power of a score or Lagrange multiplier test to discriminate between a linear time series model and some simple bilinear alternatives is examined. It is assumed that the exact form of the bilinear model is unknown so that, instead of the true linear part of the model, the best fitting linear approximation is used as a null model in the test. Theoretical results based on Pitman's asymptotic relative efficiency suggest that selecting the null model in this way may have an adverse effect on the power of the test. According to the simulation results of the paper the problem may have practical significance if the model has a non-zero intercept and the sample size is not large.

Key words and phrases: Asymptotic relative efficiency, bilinear model, Lagrange multiplier test, linear time series, non-linear time series, score test.

1. Introduction

As the papers by Tong (1986) and Davies and Petruccelli (1986) indicate, the interest in applying non-linear time series models has considerably increased recently. However, in a majority of applications the need of a non-linear model is a priori rather uncertain. Therefore it seems reasonable to start the model building by applying a linearity test which does not require the usually rather cumbersome estimation of a non-linear alternative. Such tests have been proposed both in the time domain (McLeod and Li (1983), Keenan (1985), Tsay (1986), Weiss (1986), Petruccelli and Davies (1986), Saikkonen and Luukkonen (1988), Luukkonen et al. (1988a)) and in the frequency domain (Subba Rao and Gabr (1980), Hinich (1982), Ashley et al. (1986)). Properties of the time domain tests have been studied by Chan and Tong (1986) and Luukkonen et al. (1988b). From these studies one can conclude that none of the tests is effective as a general linearity test so that, if the type of the relevant non-linearity is unknown, it is not advisable to apply only one of them. However, even if one knows the type of the possible non-linearity one seldom knows the exact form or lag structures of the relevant non-linear model. Using some simple bilinear models as examples

the effect of this ignorance on the power of a time domain linearity test will be discussed in this paper.

If the alternative to linearity is a bilinear model it is natural to use a score or Lagrange multiplier (LM) test designed for bilinear models (see Pagan (1978), Weiss (1986) and Saikkonen and Luukkonen (1988)). To apply this test it is necessary to specify a bilinear model and estimate the corresponding linear null model. When the form of the bilinear model is not known it seems reasonable to start by specifying an adequate linear model and use it as a null model in the linearity test. A problem with this approach is that the specification of the linear model is most often data-based. Some of the parameters incorporated in the specified linear model may therefore act as substitutes for the true but omitted bilinear parameters. This means that in practice the linearity test may be based on an overspecified model which contains redundant linear parameters, not needed in the true bilinear alternative. If the omitted bilinear parameters can effectively be substituted by adding linear parameters the power of the linearity test can be expected to decrease. Whether this problem has practical significance or not will be investigated in this paper both theoretically and by simulation. The idea is to compare the actually applied LM test, based on an overspecified model, with the corresponding "optimal" LM test, based on the true bilinear data generating process. In the theoretical considerations the two tests are compared by using Pitman's asymptotic relative efficiency (ARE) (see e.g. Kendall and Stuart (1979), Chapter 25). The results of these comparisons will be used to design a simulation study to find out what happens in small samples. The performance of the LM tests will be evaluated by studying their power in relation to the reduction of forecasting accuracy caused by erroneously adopting a linear model. Somewhat surprisingly it turns out that the significance of the problem is related to the value of the intercept of the model.

The paper is organized as follows: Section 2 presents some preliminaries and a detailed formulation of the problem. Theoretical properties of the tests considered in the paper are discussed in Section 3. Section 4 reports results of the simulation study and Section 5 comprises the conclusion.

2. Preliminaries

In this paper we shall always assume that the observed time series $y_1, \ldots, y_T$ is generated by one of the simple bilinear models

$$y_t = \mu + c\varepsilon_{t-1}y_{t-i} + \varepsilon_t \tag{2.1}$$

where $i \geq 1$ is a known integer, μ and c are unknown parameters and $\varepsilon_t \sim \text{NID}(0, \sigma^2)$. For more general bilinear models and their properties, see Granger and Andersen (1978), Subba Rao (1981) and Subba Rao and Gabr (1984). If we

knew that the "true" model is (2.1) we could easily construct an LM test for the linearity hypothesis

$$H_0 : c = 0 \quad \text{against} \quad H_1 : c \neq 0.$$

In this context the null model is the simple white noise process

$$y_t = \mu + \varepsilon_t \tag{2.2}$$

so that the resulting LM statistic, denoted by S_i, is also very simple (see Weiss (1986) or Saikkonen and Luukkonen (1988)). However, a practical problem with this reasoning is that even if we can choose the relevant value of i to be used in the test, we can very rarely know that the only linear parameter needed in the model is the intercept μ. Therefore we are not very likely to use (2.2) as the null model and consequently apply test statistic S_i. In order to demonstrate this the autocovariance function of (2.1) will be considered.

Let $\gamma_k = \text{cov}(y_t, y_{t-k})$ be the autocovariance function of y_t. Assuming $\mu = 0$ Granger and Andersen (1978, Chapters V and VI) and Subba Rao (1981) show that $\gamma_k = 0$ for all $k \geq 2$. Although this result also holds when $\mu \neq 0$, the value of the intercept affects the expressions of γ_0 and γ_1. It also affects the second moment of y_t and thereby the condition $c^2 E y_t^2 < 1$ which is sufficient for y_t to be invertible (see Subba Rao (1981)). However, the intercept is not involved in the stationarity condition $|c\sigma| < 1$ which is easily seen to be relevant even if $\mu \neq 0$. Unless otherwise stated it will always be assumed that the models considered here are both stationary and invertible. The autocovariance function of (2.1) is identical for all $i > 1$, but for $i = 1$ the situation is different. Therefore these two cases have to be discussed separately.

Case $i = 1$. Following the arguments used by Granger and Andersen (1978) and Subba Rao (1981) for the case $\mu = 0$ yields

$$\gamma_k = \begin{cases} \sigma^2[c^2\mu(\mu + 2c\sigma^2) + c^2\sigma^2(1 + c^2\sigma^2) + 1]/(1 - c^2\sigma^2), & k = 0 \\ c\sigma^2(\mu + c\sigma^2), & k = 1 \\ 0, & k \geq 2. \end{cases} \tag{2.3}$$

As can be seen from (2.3), $\mu = -c\sigma^2$ implies that $\gamma_k = 0$ for all $k > 0$. If $\mu = 0$ this can happen only when $c = 0$. In general, since $Ey_t = \mu + c\sigma^2$, the condition $Ey_t = 0$ is always sufficient for $\gamma_k = 0$, $k > 0$, to hold.

Case $i > 1$. In the same way as in the case $\mu = 0$ in the above citations one obtains

$$\gamma_k = \begin{cases} [c^2\sigma^2\mu^2 + \sigma^2]/(1 - c^2\sigma^2), & k = 0 \\ c\sigma^2\mu, & k = 1 \\ 0, & k \geq 2. \end{cases} \tag{2.4}$$

Now the influence of the intercept is even more pronounced than in the case $i = 1$. If $\mu = 0$ we always have $\gamma_k = 0$ for all $k > 0$. Since $Ey_t = \mu$, the condition $Ey_t = 0$ again implies that the autocovariances with non-zero lags vanish.

At this point we note that if $\varepsilon_{t-1}y_{t-i}$ in (2.1) is replaced by $\varepsilon_{t-i}y_{t-1}$ $(i > 1)$, the autocovariance function of the resulting model generally becomes more complicated than above (see Granger and Andersen (1978), Chapter VII). Therefore these models are not considered in this paper. In the present context it follows from (2.3) and (2.4) that, for all values of i, y_t has a simple first order moving average representation

$$y_t = \nu + a_t + \psi a_{t-1}, \tag{2.5}$$

where $\nu = Ey_t$ and a_t is a sequence of uncorrelated (but generally not independent) random variables with mean zero and variance σ_a^2. Model (2.5) is the best fitting linear approximation of (2.1) in the sense that the innovation variance of any (genuinely) other linear model is larger than σ_a^2. Denote $\rho_1 = \gamma_1/\gamma_0$. As is well known, ψ is the solution of

$$\rho_1 = \psi/(1 + \psi^2), \ |\psi| < 1,$$

whereas

$$\sigma_a^2 = \gamma_0/(1 + \psi^2).$$

By using (2.3) and (2.4) one can thus determine ψ and σ_a^2 from the parameters of (2.1).

Now suppose we wish to test linearity and start by specifying an adequate linear ARMA model for the series. By the above discussion we can assume that a first order moving average model will often be selected. To simplify matters we assume that this is always the case. Then the LM test for H_0 against H_1 will not be based on the "correct" model (2.1) but on the "overspecified" bilinear model

$$y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1} + c\varepsilon_{t-1}y_{t-i}. \tag{2.6}$$

When the linearity hypothesis is tested within (2.6) the null model is not the simple white noise process (2.2) but the first order moving average process

$$y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}. \tag{2.7}$$

The relevant LM statistic can again be obtained from Weiss (1986) or Saikkonen and Luukkonen (1988), and will be denoted by S_i^*.

The purpose of this paper is to investigate the effect of the above overspecification on the power of the LM test. This will be done by comparing the actually applied LM statistic S_i^* with S_i which, in a sense, represents the best theoretically obtainable result. Both of these test statistics assume that an appropriate value

of the integer i can a priori be selected. At least in some cases this assumption may be questioned, so that one might also be interested in other comparisons, like those between S_1^* and S_2^* or S_1^* and S_2. Although not irrelevant, these comparisons are outside the scope of this paper. To demonstrate that even the present case has practical relevance we first note that when the specified null model is a first order moving average model the most common choice is probably $i = 1$. For instance, in a relatively wide application by de Gooijer (1989), white noise or a first order moving average model is often the null model and (2.1) or (2.6) with $i = 1$ the alternative. Therefore, there is some justification to concentrate on the case $i = 1$ and consider the values $i \geq 2$ mainly to demonstrate that the results obtained for $i = 1$ are not a peculiarity of a single model. It may also be noted that the adverse effects of misspecifying the value of i are more obvious than those caused by a mere overspecification of the linear part of the model. Therefore, this problem is probably well recognized in applications. If necessary, time series analysts can always guard against it by augmenting the model with additional bilinear terms. For example, instead of $\varepsilon_{t-1}y_{t-1}$ one can use $\varepsilon_{t-j}y_{t-i}$, $i, j = 1, 2$, in the test. This results in a model which also contains redundant bilinear parameters. However, since the effects of such an overspecification on the LM test have already been studied it is not necessary to pursue this matter here. Theoretical results can be found in Saikkonen (1989) and simulation evidence in Saikkonen and Luukkonen (1988).

3. Theoretical Properties of the LM Tests

The definitions of the LM statistics S_i and S_i^* can readily be obtained from the above-mentioned references so that they are not repeated in this paper. It is well known that the LM test has the same asymptotic properties as the corresponding likelihood ratio test but it only requires the efficient estimation of the (linear) null model. Under the linearity hypothesis both S_i and S_i^* are asymptotically distributed as χ_1^2 and their large values are critical. However, under the alternative the asymptotic equivalence of S_i and S_i^* does not generally hold. As usual in asymptotic power considerations, the non-null properties of S_i and S_i^* will be studied under a sequence of Pitman's local alternatives

$$c = \delta/T^{1/2}, \quad 0 < |\delta| < \infty. \tag{3.1}$$

Assuming conventional regularity conditions of the maximum likelihood theory it can be shown that under (3.1) the asymptotic distributions of S_i and S_i^* are non-central χ_1^2 (see e.g. Saikkonen (1989) and the references therein). The associated non-centrality parameters are generally different for the two tests and they depend on the constant δ. Instead of considering these asymptotic non-null

distributions as such we shall use them to derive the ARE of S_i^* with respect to S_i. This ARE is denoted by e_i, and it can be obtained from the general results in Saikkonen (1989). Since the null hypothesis is one-dimensional, e_i is independent of δ and only depends on the parameters of the null model (2.2), i.e., on μ and σ^2. This is a convenient feature and makes the comparison between S_i^* and S_i very easy.

Denote $z_{1t} = (\partial\varepsilon_t/\partial c)\varepsilon_t$, $z_{2t} = (\partial\varepsilon_t/\partial\theta)\varepsilon_t$ and $z_{3t} = (\partial\varepsilon_t/\partial\mu)\varepsilon_t$ where the partial derivatives are obtained from (2.6) and evaluated under the null model (2.2). It follows from Saikkonen (1989) that $1-e_i$ equals the square of the partial correlation coefficient between z_{1t} and z_{2t} conditional on z_{3t}. Since

$$z_{1t} = -(\mu + \varepsilon_{t-i})\varepsilon_{t-1}\varepsilon_t,\ z_{2t} = -\varepsilon_{t-1}\varepsilon_t \text{ and } z_{3t} = -\varepsilon_t,$$

it is straightforward to verify that

$$e_i = \begin{cases} 2/(2+\mu^2/\sigma^2), & i = 1 \\ 1/(1+\mu^2/\sigma^2), & i > 1. \end{cases} \tag{3.2}$$

In the same way as with the autocovariance function of the model, the cases $i = 1$ and $i > 1$ again differ from each other. It follows from (3.2) that $e_i \leq e_1 \leq 1$ with equalities if and only if $\mu = 0$. This implies that when $\mu = 0$, test statistics S_i and S_i^* are asymptotically equivalent, not only under the null hypothesis, but also under the local alternatives (3.1). Hence, if $\mu = 0$, the performance of test statistic S_i relative to S_i^* in finite samples is likely to be good. On the other hand, the value of (3.2) can always be made arbitrarily close to zero by taking μ^2/σ^2 large enough. When $\mu \neq 0$, the test statistic S_i^* may therefore be very inefficient and its power may also be rather low. We shall take a closer look at these possibilities in the next section.

In practice one often starts the model building by mean correcting the observations. Since the mean value of the series depends on the intercept of the model, the above discussion gives a reason to consider the effect of a mean correction on the properties of test statistics S_i and S_i^*. In most cases the mean correction is also data-based; so that, suppose the sample mean is subtracted from the observations. It is well known that this has no effect on the asymptotic null distribution of an LM test used to check the order of a linear ARMA model. However, this is not the case in the present context. It is not difficult to show that analogues of test statistics S_i and S_i^* based on such mean corrected observations do not have an asymptotic χ_1^2 distribution under the null hypothesis (details of the required derivations are available upon request). A similar result can also be obtained for other LM type linearity tests, so that these tests should not be used with mean corrected observations.

4. Simulation Study

We shall now present simulation results of the ability of test statistics S_i^* and S_i to reveal the non-linearity in series generated by (2.1) with $i = 1, 2$. Our interest is focused on fairly short time series because they are frequently encountered in practice, and because in very large samples even a relatively poor test can have high power. The examined sample sizes are $T = 50, 75, 100, 150$. A low power, as such, may not be a serious drawback if the test fails to reject linearity only when a linear model approximates the true data generating process so well that improvements achieved by a bilinear model are only marginal. Since univariate time series models are mainly used for forecasting, it seems reasonable to follow Granger and Andersen (1978, p.53) and use the ratio σ_a^2/σ^2 to measure the loss caused by erroneously accepting (2.5) instead of the true model (2.1). The idea is to consider models with various values of the loss function σ_a^2/σ^2. For each value we try to find a model for which the ARE (3.2) is small, implying difficulties for test statistic S_i^*. We can make (3.2) small by fixing σ^2 and choosing $|\mu|$ large. However, since the value of μ affects the invertibility condition of the bilinear alternative there is an upper bound of $|\mu|$ which can be considered for a given value of c.

The observations were obtained as follows. The normally distributed innovations ε_t were first generated by using a random number generator in the NAG subroutine library on a VAX 8800 computer of the University of Helsinki. The value of the innovation variance σ^2 is always $\sigma^2 = 1$. Observations from model (2.1) with various values of μ and c were obtained recursively. The initial value ε_0 was simulated and y_0 chosen to be $y_0 = Ey_t$. The first 100 observations were discarded to avoid initialization effects. The number of replications was 1000 and between successive replications 700 observations of the innovation series were discarded to better guarantee the independence of the replications.

For test statistic S_i^* we have to fit a linear first order moving average model to the generated series. This was done by using the conditional least squares method with the maximum number of iterations 30 (see e.g. Box and Jenkins (1976), Chapter 7). Realizations for which the estimation algorithm fails to converge within 30 iterations were discarded. Since the proportion of such realizations was always small (mostly a few percents), the number of discarded realizations is not reported. It should be noted, however, that the discarded realizations were not used to compute the empirical power of test statistic S_i so that in this respect the results of test statistics S_i^* and S_i are comparable. The reason for using the conditional least squares method is that the LM tests for bilinearity are derived from a corresponding conditional likelihood function. Moreover, in the models considered below the value of the moving average parameter is always

reasonably far from the boundary of the invertibility region ($|\psi| \leq 0.67$) so that the performance of the conditional least squares method should be quite good (see Dent and Min (1978)).

Tables 1 and 2 present empirical rejection frequencies of test statistics S_i^* and S_i for $i = 1, 2$, respectively, at the nominal 5% significance level when $\mu \neq 0$. The results are fairly similar in both tables. Since $\sigma^2 = 1$ the variance σ_a^2 directly expresses the loss of not rejecting the linearity hypothesis. The value of σ_a^2 increases when one moves from the top of the tables to the bottom. In each case the estimated power of test statistic S_i is very high so that if we knew the right alternative we could almost always reveal the non-linearity in the series.

Now consider the performance of test statistic S_i^* $(i = 1, 2)$. When $T = 50$ its power is much weaker than that of S_i. For instance, in the fifth entries of Tables 1 and 2 the use of the best fitting linear model results in an increase of about 20% in the one step ahead forecast error variance of the model, but the empirical rejection frequencies of test statistics S_1^* and S_2^* are below 70% and 50%, respectively, whereas those of S_1 and S_2 are over 90%. When the sample size increases to 75 and 100 the performance of S_i^* improves although the probability of type II error can still be nonnegligible for σ_a^2/σ^2 up to 1.15 or so. At this point it may be noted that in an economic application in Maravall (1983) a reduction of 8% in the one step ahead forecast error variance is deemed worthwhile. When T = 150 the performance of S_i^* is further improved and the test fails only rarely unless σ_a^2/σ^2 is less than 1.10.

In the examples of Tables 1 and 2 the value of the ARE (3.2) is always very low but sometimes the power of S_i^* is nevertheless fairly high. As discussed in Luukkonen et al. (1988b) the ARE may not be a very good indicator of the power of a linearity test if the true model is bilinear and the sufficient condition of invertibility is nearly violated. Even though our models are always clearly invertible, the invertibility condition of the bilinear model is usually much easier to violate than that of the linear moving average model. Despite this remark the ARE calculations have been useful in helping to design the simulation study.

In the models of Tables 1 and 2 the intercept μ is always non-zero because in this way it is possible to construct alternatives against which test statistic S_i^* is likely to have low power. In order to demonstrate the behaviour of S_i^* in the favourable situation $\mu = 0$ a few results of S_1 and S_1^* are presented in Table 3. Now the ARE equals one and S_1^* is practically as powerful as the "optimal" test statistic S_1. The necessity of specifying an adequate linear model using data before applying an LM test against a bilinear alternative is thus not a drawback if one, a priori, knows that the model has no intercept. On the other hand, the results of Table 3 indicate that when $\mu = 0$ and the sample size is not large even the "optimal" test statistic S_1 may have rather low power.

5. Conclusions

Of the time domain tests considered in Luukkonen et al. (1988b) the LM test designed for bilinear models is the most powerful one to discriminate between a linear model and a bilinear alternative. However, a practical problem with the LM test, as well as any other time domain test, is that the structure of the relevant bilinear model is unknown and the test is applied by first specifying an adequate linear model using the observed time series. As the results of this paper show, this can have an adverse effect on the power of the LM test. Somewhat surprisingly, the power is related to the intercept of the model and it can be substantially reduced if the model has a non-zero intercept and the sample size is not large. As compared with linear time series models this is very peculiar and implies that in non-linear models the role of the intercept is rather different from that in linear models in which it can simply be eliminated before inference on other parameters of the model is performed. Unfortunately we have not been able to find any intuitive interpretation for this phenomenon.

It is not surprising that the performance of the LM test improves when the sample size or the non-linearity of the process increases. We would expect that for sample sizes larger than 150 or 200 the power of the LM test is quite good even if a data-based technique is used to specify the null model.

Finally, we believe that the results of the paper have practical relevance even though the considered models are very simple. As already pointed out, simple models of this kind have been of interest in some fields of application (see de Gooijer (1989)). Furthermore, in principle, ARE's, similar to (3.2), can always be obtained although the calculations are involved except in special cases. The following general conclusions can be made, however. If a bilinear model has no intercept, then the overspecification of the linear part of the model as in (2.6) has no effect on the asymptotic (local) power of the LM test. The ARE of the "overspecified" test with respect to the corresponding "optimal" test is then always equal to one. On the other hand, for any bilinear model with a non-zero intercept this ARE is smaller than one. We have no reason to believe that in general bilinear models the problems caused by such an overspecification would be less serious than in the simple examples of this paper. The results of Tables 1 and 2 should thus indicate what can at least happen in this respect. We would also expect that the conclusions of the paper do not greatly change if the best fitting linear model is replaced by a statistically adequate approximation.

Acknowledgements

Part of this work was carried out when the first author worked as a senior research fellow at the Academy of Finland. We wish to thank Timo Teräsvirta, an associate editor and a referee for useful comments. Financial support from

the Yrjö Jahnsson Foundation is gratefully acknowledged.

Table 1. Empirical rejection frequencies (%) of test statistics S_1^* and S_1 at the nominal 5% significance level computed from time series generated by (2.1) with $i = 1$ and $\sigma^2 = 1$

μ	c	e_1	σ_a^2	Test statistic	$T = 50$	$T = 75$	$T = 100$	$T = 150$
6.0	−0.11	0.05	1.049	S_1	94.7	99.5	99.8	100.0
				S_1^*	32.8	45.2	53.7	72.1
5.5	−0.13	0.06	1.078	S_1	96.1	99.7	99.9	100.0
				S_1^*	44.2	57.6	69.6	86.4
−4.0	0.17	0.11	1.116	S_1	96.7	99.6	100.0	100.0
				S_1^*	49.3	69.0	81.8	94.3
3.5	−0.20	0.14	1.162	S_1	96.6	99.7	99.9	100.0
				S_1^*	65.5	80.9	91.4	98.0
−3.4	0.22	0.15	1.212	S_1	98.1	99.8	100.0	100.0
				S_1^*	69.8	89.1	95.3	99.5
−3.0	0.25	0.18	1.263	S_1	98.6	99.8	100.0	100.0
				S_1^*	77.0	92.1	97.2	99.6
2.3	−0.30	0.27	1.321	S_1	97.3	99.6	99.9	100.0
				S_1^*	83.6	94.1	98.7	100.0
2.0	−0.35	0.33	1.425	S_1	98.1	99.7	99.9	100.0
				S_1^*	89.1	96.9	99.7	100.0

Table 2. Empirical rejection frequencies (%) of test statistics S_2^* and S_2 at the nominal 5% significance level computed from time series generated by (2.1) with $i = 2$ and $\sigma^2 = 1$

μ	c	e_2	σ_a^2	Test statistic	$T = 50$	$T = 75$	$T = 100$	$T = 150$
4.5	−0.15	0.05	1.059	S_2	94.4	99.6	99.9	100.0
				S_2^*	33.2	45.2	57.9	79.3
−4.2	0.17	0.05	1.085	S_2	94.7	99.8	100.0	100.0
				S_2^*	44.4	60.6	73.2	87.8
3.5	0.20	0.08	1.111	S_2	90.8	98.6	99.9	100.0
				S_2^*	42.5	60.2	74.1	89.8
−3.1	0.23	0.09	1.151	S_2	94.4	99.8	100.0	100.0
				S_2^*	60.8	77.8	88.2	97.7
−3.0	−0.25	0.10	1.197	S_2	90.2	98.9	99.7	100.0
				S_2^*	46.3	66.5	81.8	93.7
2.4	0.30	0.15	1.256	S_2	89.0	97.8	99.4	100.0
				S_2^*	54.5	74.6	86.5	96.3
1.6	−0.38	0.28	1.321	S_2	92.4	99.1	99.9	100.0
				S_2^*	71.9	90.4	97.2	99.7
1.3	−0.44	0.37	1.415	S_2	90.9	99.0	99.8	100.0
				S_2^*	73.7	90.9	97.6	99.8

Table 3. Empirical rejection frequencies (%) of test statistics S_1^* and S_1 at the nominal 5% significance level computed from time series generated by (2.1) with $i = 1$, $\mu = 0$ and $\sigma^2 = 1$

c	σ_a^2	Test statistic	$T = 50$	$T = 75$	$T = 100$	$T = 150$
0.15	1.046	S_1	22.8	38.0	48.4	67.3
		S_1^*	21.6	36.9	48.3	66.9
0.25	1.134	S_1	51.7	72.9	86.2	96.5
		S_1^*	50.5	71.1	84.9	95.8
0.35	1.285	S_1	73.0	89.5	97.7	99.6
		S_1^*	71.0	89.8	97.1	99.7

References

Ashley, R. A., Patterson, D. M. and Hinich, M. J. (1986). A diagnostic test for nonlinear serial dependence in time series fitting errors. *J. Time Ser. Anal.* **7**, 165-178.

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control.* Holden-Day, San Francisco.

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

Davies, N. and Petruccelli, J. D. (1986). Detecting non-linearity in time series. *Statistician* **35**, 271-280.

Dent, W. and Min, A.-S. (1978). A monte carlo study of autoregressive integrated moving average processes. *J. Econometrics* **7**, 23-55.

Gooijer, J. G. de (1989). Testing non-linearities in world stock market prices. *Econom. Lett.* **31**, 31-35.

Granger, C. W. J. and Andersen, A. P. (1978). *An Introduction to Bilinear Time Series Models.* Vandenhoeck and Ruprecht, Göttingen.

Hinich, M. J. (1982). Testing for Gaussianity and linearity of a stationary time series. *J. Time Ser. Anal.* **3**, 169-176.

Keenan, D. M. (1985). A Tukey nonadditivity-type test for time series non-linearity. *Biometrika* **72**, 39-44.

Kendall, M. G. and Stuart, A. (1979). *The Advanced Theory of Statistics*, Vol. 2, 4th edition. Griffin, London.

Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988a). Testing linearity against smooth transition autoregressive models. *Biometrika* **75**, 491-499.

Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988b). Testing linearity in univariate time series models. *Scand. J. Statist.* **15**, 161-175.

Maravall, A. (1983). An application of nonlinear time series forecasting. *J. Bus. & Econom. Statist.* **1**, 66-74.

McLeod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations. *J. Time Ser. Anal.* **4**, 269-273.

Pagan, A. (1978). Some simple tests for non-linear time series models. CORE Discussion Paper No. 7812.

Petruccelli, J. D. and Davies, N. (1986). A portmanteau test for self-exciting threshold autoregressive-type nonlinearity in time series. *Biometrika* **73**, 687-694.

Saikkonen, P. (1989). Asymptotic relative efficiency of the classical test statistics under misspecification. *J. Econometrics* **42**, 351-369.

Saikkonen, P. and Luukkonen, R. (1988). Lagrange multiplier tests for testing non-linearities in time series models. *Scand. J. Statist.* **15**, 55-68.

Subba Rao, T. (1981). On the theory of bilinear time series models. *J. Roy. Statist. Soc. Ser. B* **43**, 244-255.

Subba Rao, T. and Gabr, M. M. (1980). A test for linearity of stationary time series. *J. Time Ser. Anal.* **1**, 145-158.

Subba Rao, T. and Gabr, M. M. (1984). *An Introduction to Bispectral Analysis and Bilinear Time Series Models.* Springer-Verlag, Berlin.

Tong, H. (1986). Non-linear time series models of regularly sampled data: A review. Invited paper presented at the First World Congress of the Bernoulli Society for Mathematical Statistics and Probability, Tashkent, September 1986.

Tsay, R. S. (1986). Nonlinearity tests for time series. *Biometrika* **73**, 461-466.

Weiss, A. A. (1986). ARCH and bilinear time series models: Comparison and combination. *J. Bus. & Econom. Statist.* **4**, 59-70.

Department of Statistics, University of Helsinki, Aleksanterinkatu 7, SF-00100 Helsinki, Finland.

Institute of Occupational Health, Topeliuksenkatu 41 a A, SF-00250 Helsinki, Finland.

(Received June 1989; accepted January 1991)