Journal of Statistical Planning and Inference 88 (2000) 339–353

journal of statistical planning and inference

www.elsevier.com/locate/jspi

A mixed-type test for linearity in time series

Hong-Zhi An *, Li-Xing Zhu, Run-Ze Li

Institute of Applied Mathematics, Academia Sinica, Beijing 100080, People's Republic of China

Received 18 February 1998; received in revised form 15 November 1998

Abstract

We propose a new test for linearity of time-series model by introducing a mixed-type statistic. It consists of a Cramer–Von Mises-type statistic and a goodness-of-fit statistic concerned only with fitting a linear autoregressive model. The computation involved in the new test is considerably simple and the curse of dimensionality is partly avoided. It is shown that the test is consistent against all fixed alternatives to linearity in stationary autoregressive series. The simulation results show that the test has good performance of power. © 2000 Elsevier Science B.V. All rights reserved.

MSC: primary 62M10; 62G10; secondary 62E10; 62H15

Keywords: Nonlinearity; Test for linearity; Stationary AR(p) series

1. Introduction

Consider a stationary time series x_t. Denote by $E\{x_t|x_{t-1},x_{t-2},\ldots\}$ the conditional expectation of x_t given $x_{t-1},x_{t-2},\ldots$. If

$$\begin{aligned} E\{x_t|x_{t-1},x_{t-2},\ldots\} &= E\{x_t|x_{t-1},\ldots,x_{t-p}\} \\ &\equiv \phi(x_{t-1},\ldots,x_{t-p}) = \phi(\boldsymbol{x}_{t-1}), \end{aligned} \tag{1.1}$$

where $\boldsymbol{x}_{t-1} = (x_{t-1},x_{t-2},\ldots,x_{t-p})^\tau$ and τ denotes the transposition of a vector, then the series x_t is called an autoregressive series with order p (AR(p) for short). Hence an AR(p) series x_t satisfies the following AR(p) model, i.e.,

$$x_t = \phi(\boldsymbol{x}_{t-1}) + \varepsilon_t, \tag{1.2}$$

where $\varepsilon_t = x_t - \phi(\boldsymbol{x}_{t-1})$ is a stationary martingale difference sequence, and

$$E\{\varepsilon_t|x_{t-1},\ldots,x_{t-p}\} = 0 \quad \text{a.s.} \tag{1.3}$$

* Corresponding author. Fax: +86-1062-541689.

E-mail address: hzan@sun.ihep.ac.cn (H.-Z. An).

PII: S0378-3758(00)00087-2

In particular, if ϕ is linear (1.2) becomes the following linear AR(p) model:

$$\begin{aligned} x_t &= \beta_0 + \beta_1 x_{t-1} + \cdots + \beta_p x_{t-p} + \varepsilon_t \\ &= \beta_0 + \beta^\tau \boldsymbol{x}_{t-1} + \varepsilon_t, \end{aligned} \tag{1.4}$$

where the parameter vector $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)^\tau$ belongs to the stationary domain of the linear AR(p) models to ensure the stationarity of series x_t (e.g. see Box and Jenkins, 1970).

This paper aims to study testing for linearity of model (1.2). That is to test the null hypothesis as follows:

H_0: ϕ is linear

against the alternative

H_1: ϕ is not linear

with given order p of the stationary AR(p) series.

If order p is unknown, it may be replaced by an estimator obtained by making use of an order selection method, such as the AIC method (see An et al., 1982). Therefore, the test proposed in this paper may be approximately used to test for linearity in stationary time series, not only in AR(p) series. The same treatment is also considered by other authors (e.g. see An and Cheng, 1991). Thus we restrict our study within the test of H_0 against H_1 in the case where order p is known.

In the literature there is a number of parametric tests, for instance Keenan (1985), Chan and Tong (1986), Tsay (1986), Petruccelli (1990) and Terästirta et al. (1993). A parametric test is to test the linearity of (1.2) against a specific alternative. i.e. H_1 is replaced by that ϕ which belongs to a specific subclass of nonlinear functions with finite unknown parameters. A parametric test may be constructed by the Lagrange multiplier principle as outlined for example in Saikkonen and Luukkonen (1988), and Luukkonen et al. (1988). The advantage of the parametric approaches is that relatively few observations are required to get reasonable power for a correctly specified nonlinear function ϕ in (1.2). It is however such tests may not be sensitive to the above alternative H_1 being a nonparametric version. In other words, a parametric test may not be consistent against all fixed alternatives.

Taking into account such a problem, nonparametric approaches have been developed. One is based on the higher-order spectral analysis (e.g. Subba Rao and Gabr, 1980; Hinich, 1982; Ashley et al., 1986; Brockett et al., 1987), in particular the bispectrum. Spectrum based tests are easy to use. They are however not consistent for all fixed alternatives either.

Another way is concerned with nonparametric estimate of the function ϕ of (1.2). The use of nonparametric estimates is first considered in testing for linearity of regression model which is similar to model (1.2). Some local smoothing methods such as the kernel method are used in constructing such tests, for example, Cox and Koh (1989), Eubank and Spiegelman (1990), Eubank and Hart (1993). Under regularity conditions on the function ϕ the consistency of the nonparametric tests may be expected, as shown in regression models by Chen and Krishnaiah (1990). But the calculation involved in nonparametric tests is considerably larger than that in parametric tests. Besides "the curse of dimensionality" of $\boldsymbol{x}_{t-1}=(x_{t-1},\ldots,x_{t-p})$ is a big trouble when order p is large. Hjellivik and Tjϕstheim (1993) have recently developed a nonparametric test in time series based on the conditional mean and conditional variance which are estimated by a nonparametric procedure. They use $E\{x_t|x_{t-i}\}$ for $i=1,\ldots,p$ instead of $E\{x_t|x_{t-1},\ldots,x_{t-p}\}$ in the test proposed to avoid "the curse of dimensionality". In such a way, however the consistency of the resulting test no longer holds.

The third kind of nonparametric approach is based on a Kolmogorov–Smirnov-type statistic proposed by An and Cheng (1991), further developed by Zhu and An (1992). The consistency of the test has been shown only in the AR(1) model. The main difficulty is that the asymptotic distribution of the Kolmogorov–Smirnov-type statistic for AR(p) series is concerned with a multivariate index Brownian motion, and is not distribution free of $(x_{t-1},\ldots,x_{t-p})$, and then of the parameters $\beta_1,\beta_2,\ldots,\beta_p$ in model (1.4) either.

Except for the parametric and nonparametric tests, a mixed type test was proposed by Zhu et al. (1993) for testing linearity of regression models. Their test can also be developed for autoregressive models. They suggest a test statistic ξ_n, which is denoted by

$$\xi_n=u_n+v_n, \tag{1.5}$$

where u_n and v_n are nonnegative statistics, and u_n is a goodness of fit-type statistic based on the least-squares procedures, and v_n is based on nonparametric estimation procedures. Neither u_n nor v_n is sufficient by itself as a test for linearity, while ξ_n is sufficient (see Zhu et al., 1993). The test may be extended to autoregressive models. The consistency of the test can be expected following the same arguments as in Zhu et al. (1993), provided that

$$\lim_{n\to\infty} u_n \sim \chi^2(1) \quad \text{under } \mathrm{H}_0, \tag{1.6}$$

$$\lim_{n\to\infty} v_n \overset{\mathrm{p}}{=} 0 \quad \text{under } \mathrm{H}_0, \tag{1.7}$$

$$\lim_{n\to\infty} v_n \overset{\mathrm{p}}{=} \infty \quad \text{under } \mathrm{H}_1, \tag{1.8}$$

where $\chi^2(1)$ denotes the chi-squared distribution with one freedom degree, and "$\overset{\mathrm{p}}{=}$" denotes the convergence in probability.

In this paper a new mixed-type test is suggested based on a Cramer–von Mises-type statistic and a goodness-of-fit-type statistic. The structure of the new test statistic is similar to the mixed-type statistic proposed by Zhu et al. (1993). We use the same form of (1.5), and the same definition of u_n, but replace v_n in (1.5) by a Cramer–Von Mises-type statistic. As long as (1.6), (1.7) and (1.8) hold, a consistent mixed-type test will be obtained. The new test is computationally simple, and "the curse of dimensionality" may be avoided partly as well. The simulation results show that the test performs well in all of the linear cases and most of the nonlinear cases we considered.

In Section 2 we construct the statistics u_n and v_n and then ξ_n, and show the asymptotic behaviors. In Section 3 some simulations are presented. Section 4 provides further consideration and conclusion. Technical proofs are put in the appendix.

2. A mixed-type test

2.1. Fitting a linear AR(p) model

Consider an AR(p) series generated by model (1.2). Whether ϕ is linear or not, let $L(\boldsymbol{x}_{t-1})$ be the projection of x_t on the linear space spanned by $x_{t-1}, x_{t-2}, \ldots, x_{t-p}$, i.e.,

$$L(\boldsymbol{x}_{t-1}) = b_0 + b_1 x_{t-1} + \cdots + b_p x_{t-p} = b_0 + \boldsymbol{b}^\tau \boldsymbol{x}_{t-1}, \tag{2.1}$$

such that

$$E\{x_t - L(\boldsymbol{x}_{t-1})\}^2 = \inf_{a_0,\boldsymbol{a}} E\{x_t - a_0 - \boldsymbol{a}^\tau \boldsymbol{x}_{t-1}\}^2, \tag{2.2}$$

where $\boldsymbol{a} = (a_1, \ldots, a_p)^\tau$, and $\boldsymbol{b} = (b_1, \ldots, b_p)^\tau$. Let $e_t = x_t - L(\boldsymbol{x}_{t-1})$, be the error of the projection. It is obvious that the null hypothesis

$$\mathrm{H}_0 \text{ is true iff } e_t = \varepsilon_t, \quad \text{a.s.} \tag{2.3}$$

for all t (see An and Cheng, 1991). The null hypothesis H_0 is true iff

$$Ee_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s}) = 0, \tag{2.4}$$

for each $\boldsymbol{s}$, where $I(\cdot)$ denotes the indicator function, and

$$I(\boldsymbol{x}_{t-1} < \boldsymbol{s}) = \prod_{i=1}^{p} I(x_{t-i} < s_i), \quad \boldsymbol{s} = (s_1, \ldots, s_p)^\tau. \tag{2.5}$$

From (2.4) it follows that the null hypothesis H_0 is true iff

$$Ee_t w(\boldsymbol{x}_{t-1}) = 0 \tag{2.6}$$

for each $w(\cdot)$, where function $w(\cdot)$ is measurable with respect to $x_{t-1}, \ldots, x_{t-p}$, and has finite second moment.

Let $x_1, x_2, \ldots, x_n$ be the sample generated by model (1.2). We first fit an AR(p) model (1.2) with the sample by making use of the least-squares procedure and then obtain the residuals

$$\hat{e}_t = x_t - \hat{L}(\boldsymbol{x}_{t-1}), \quad t = p+1,\ p+2, \ldots, n, \tag{2.7}$$

where $\hat{L}(\boldsymbol{x}_{t-1}) = \hat{b}_0 + \hat{\boldsymbol{b}}^\tau \boldsymbol{x}_{t-1}$, and $(\hat{b}_0, \hat{\boldsymbol{b}})$ is the least-squares estimate of $(b_0, \boldsymbol{b})$ in (2.1). By the property of projection, it is true that

$$\sum_{t=p+1}^{n} \hat{e}_t w(\boldsymbol{x}_{t-1}) = 0 \tag{2.8}$$

for each $w(\cdot)$. Similar to Zhu et al. (1993), we use $\hat{e}_t$ to construct an improved u_n in the following subsection. Further we follow An and Cheng (1991) in constructing a Cramer–von Mises-type statistic in Section 2.3.

2.2. Construction of u_n in (1.5)

In (2.6) we replace e_t by $\hat{e}_t$, and choose a nonlinear function $w(\cdot)$ to define

$$u_n = \frac{c_u^2}{n}\left\{\sum_{t=p+1}^{n} \hat{e}_t w_1(\boldsymbol{x}_{t-1})\right\}^2, \tag{2.9}$$

which is similar to the definition given by Zhu et al. (1993), where c_u and $w_1(\cdot)$ are determined below.

By (2.6) u_n can be used to test the necessity of linearity of model (1.2) in terms of using the limiting distribution of u_n of (2.9). Here it should be noted that because of (2.8) we have to choose $w_1(\cdot)$ which should be remarkably different from linear functions. In this paper we use

$$w_1(\boldsymbol{x}_{t-1}) = \{(2\exp\{-q_1||\boldsymbol{x}_{t-1}||\} - \exp\{-2q_1||\boldsymbol{x}_{t-1}||\})q_1/3\}, \tag{2.10}$$

where $||\cdot||$ denotes the squared norm, q_1 is a positive number. Here we use $w_1(\boldsymbol{x}_{t-1})$ rather than the simple form $\exp\{-q_1||\boldsymbol{x}_{t-1}||\}$ since w_1 may have better performance than the latter. The simulation results supported such a choice.

Now we discuss how to choose c_u in (2.9). Let $\tilde{b}_0 = \hat{b}_0 - b_0$, $\tilde{\boldsymbol{b}} = \hat{\boldsymbol{b}} - \boldsymbol{b}$, $\tilde{e}_t = \hat{e}_t - e_t$ and $\tilde{w}_1(\boldsymbol{x}_{t-1}) = w_1(\boldsymbol{x}_{t-1}) - Ew_1(\boldsymbol{x}_{t-1})$. Then from (2.4), (2.7) and (2.8), we have

$$\begin{aligned}
\frac{c_u}{\sqrt{n}}\sum_{t=p+1}^{n} \hat{e}_t w_1(\boldsymbol{x}_{t-1}) &= \frac{c_u}{\sqrt{n}}\sum_{t=p+1}^{n} \hat{e}_t \tilde{w}_1(\boldsymbol{x}_{t-1}) \\
&= \frac{c_u}{\sqrt{n}}\sum_{t=p+1}^{n} e_t \tilde{w}_1(\boldsymbol{x}_{t-1}) + \frac{c_u}{\sqrt{n}}\sum_{t=p+1}^{n} \tilde{e}_t \tilde{w}_1(\boldsymbol{x}_{t-1}) \\
&= \frac{c_u}{\sqrt{n}}\sum_{t=p+1}^{n} e_t \tilde{w}_1(\boldsymbol{x}_{t-1}) + \frac{c_u}{\sqrt{n}}\sum_{t=p+1}^{n} (\tilde{b}_0 + \tilde{\boldsymbol{b}}^{\tau}\boldsymbol{x}_{t-1})\tilde{w}_1(\boldsymbol{x}_{t-1}) \\
&= \frac{c_u}{\sqrt{n}}\sum_{t=p+1}^{n} e_t \tilde{w}_1(\boldsymbol{x}_{t-1}) + \frac{c_u}{\sqrt{n}}\tilde{b}_0\sum_{t=p+1}^{n} \tilde{w}_1(\boldsymbol{x}_{t-1}) \\
&\quad + \frac{c_u}{\sqrt{n}}\tilde{\boldsymbol{b}}^{\tau}\sum_{t=p+1}^{n} \boldsymbol{x}_{t-1}\tilde{w}_1(\boldsymbol{x}_{t-1}).
\end{aligned} \tag{2.11}$$

In the appendix it will be proved that, under H_0,

$$\frac{c_u}{\sqrt{n}}\tilde{b}_0\sum_{t=p+1}^{n} \tilde{w}_1(\boldsymbol{x}_{t-1}) \to 0 \quad \text{in probability}, \tag{2.12}$$

$$\frac{c_u}{\sqrt{n}}\sum_{t=p+1}^{n} e_t \tilde{w}_1(\boldsymbol{x}_{t-1}) + \frac{c_u}{\sqrt{n}}\tilde{\boldsymbol{b}}^{\tau}\sum_{t=p+1}^{n} \boldsymbol{x}_{t-1}\tilde{w}_1(\boldsymbol{x}_{t-1}) \Rightarrow \mathrm{N}(0, c_u^2\sigma_u^2), \tag{2.13}$$

where "⇒" means convergence in distribution,

$$\sigma_u^2 = \sigma_\varepsilon^2\{E\tilde{w}_1^2(\boldsymbol{x}_{t-1}) + E[\boldsymbol{x}_{t-1}^{\tau}\tilde{w}_1(\boldsymbol{x}_{t-1})]\Gamma^{-1}E[\boldsymbol{x}_{t-1}\tilde{w}_1(\boldsymbol{x}_{t-1})]\}, \tag{2.14}$$

$\sigma_\varepsilon^2 = E\varepsilon_t^2$, and

$$\Gamma = (\gamma_{i,j})_{1\leqslant i,j\leqslant p}, \quad \gamma_{i,j} = E(x_i - Ex_i)(x_j - Ex_j). \tag{2.15}$$

Let

$$\gamma_u = E\boldsymbol{x}_{t-1}^{\tau}\tilde{w}_1(\boldsymbol{x}_{t-1}). \tag{2.16}$$

For practical use, we define the estimator of σ_u^2 by the following procedure, i.e.

$$\bar{x}_n = \frac{1}{n}\sum_{t=1}^{n} x_t, \quad \hat{\gamma}_{ij} = \frac{1}{n}\sum_{t=p+1}^{n}(x_{t-i}-\bar{x}_n)(x_{t-j}-\bar{x}_n), \tag{2.17}$$

$$\hat{\Gamma} = (\hat{\gamma}_{ij})_{1\leqslant i,j\leqslant p}, \tag{2.18}$$

$$\bar{w}_1 = \frac{1}{n}\sum_{t=p+1}^{n} w_1(\boldsymbol{x}_{t-1}), \tag{2.19}$$

$$\hat{\sigma}_w^2 = \frac{1}{n}\sum_{t=p+1}^{n}[w_1(\boldsymbol{x}_{t-1})-\bar{w}_1]^2, \tag{2.20}$$

$$\hat{\gamma}_u = \frac{1}{n}\sum_{t=p+1}^{n}\boldsymbol{x}_{t-1}[w_1(\boldsymbol{x}_{t-1})-\bar{w}_1], \tag{2.21}$$

$$\hat{\sigma}_e^2 = \frac{1}{n}\sum_{t=p+1}^{n}\hat{e}_t^2 = \hat{\sigma}_\varepsilon^2 \tag{2.22}$$

and then

$$\hat{\sigma}_u^2 = \hat{\sigma}_\varepsilon^2\{\hat{\sigma}_w^2 + \hat{\gamma}_u^{\tau}\hat{\Gamma}^{-1}\hat{\gamma}_u\}. \tag{2.23}$$

In the appendix we shall prove that

$$\hat{\sigma}_u^2 \to \sigma_u^2 \quad \text{a.s.} \tag{2.24}$$

Hence we take $c_u^2 = \hat{\sigma}_u^{-2}$ in (2.9), from (2.11), (2.12), (2.13) and (2.24), (1.6) follows under H_0.

2.3. Construction of v_n in (1.5)

First, we rewrite (2.4) in the following equivalent form, i.e. the null hypothesis H_0 is true iff

$$\sup_{\boldsymbol{s}}|Ee_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s})| > 0 \tag{2.25}$$

or equivalently the null hypothesis H_0 is true iff

$$\int\{Ee_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s})\}^2 w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} > 0, \tag{2.26}$$

where $w_2(\boldsymbol{s})$ is a positive kernel density to be chosen below.

Again replacing e_t by $\hat{e}_t$ of (2.7), we define

$$S_n(\boldsymbol{s}) = \frac{c_v}{n}\sum_{t=p+1}^{n}\hat{e}_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s}), \tag{2.27}$$

where c_v is determined below. According to (2.25), An and Cheng (1991) suggested to use

$$\xi_n = \sqrt{n} \sup_{\boldsymbol{s}} |S_n(\boldsymbol{s})| \tag{2.28}$$

as the test statistic for the linearity of (1.2), and showed the consistency of the test for linear AR(1) model. In this paper, we shall suggest a new statistic v_n based on $S_n(\boldsymbol{s})$ of (2.27). According to (2.26) we shall obtain (1.7) and (1.8) rather than obtaining the limiting distribution of v_n which is not distribution-free of x_t and therefore not free of the parameter (β_0, β) in (1.4) under H_0.

Now, we define

$$v_n = c_n \int \{S_n(\boldsymbol{s})\}^2 w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s}, \tag{2.29}$$

where $\{c_n\}$ is a constant sequence to be specified in (2.35) below. In order to simplify calculation, we take

$$w_2(\boldsymbol{s}) = \prod_{i=1}^{p} \{(2\mathrm{e}^{-q|s_i|} - \mathrm{e}^{-2q|s_i|})q/3\}, \tag{2.30}$$

where q is a positive number. We choose $w_2(\cdot)$ to have the same form as $w_1(\cdot)$, but with different parameter q. Substituting (2.27) in (2.29) we obtain

$$\begin{aligned} v_n &= \frac{c_n c_v^2}{n^2} \sum_{k,j=p+1}^{n} \hat{e}_k \hat{e}_j \int I(\boldsymbol{x}_{k-1} < \boldsymbol{s}) I(\boldsymbol{x}_{j-1} < \boldsymbol{s}) w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \\ &= \frac{c_n c_v^2}{n^2} \sum_{k,j=p+1}^{n} \hat{e}_k \hat{e}_j \psi(\boldsymbol{x}_{k-1}, \boldsymbol{x}_{j-1}), \end{aligned} \tag{2.31}$$

where

$$\begin{aligned} \psi(\boldsymbol{x}_{k-1}, \boldsymbol{x}_{j-1}) &= \int I(\boldsymbol{x}_{k-1} < \boldsymbol{s}) I(\boldsymbol{x}_{j-1} < \boldsymbol{s}) w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \\ &= \prod_{i=1}^{p} \int I(x_{k-i} < s_i) I(x_{j-i} < s_i)(2\mathrm{e}^{-q|s_i|} - \mathrm{e}^{-2q|s_i|})q/3\,\mathrm{d}s_i \\ &= \prod_{i=1}^{p} \int_{-\infty}^{x_{k-i} \wedge x_{j-i}} (2\mathrm{e}^{-q|s_i|} - \mathrm{e}^{-2q|s_i|})q/3\,\mathrm{d}s_i, \end{aligned} \tag{2.32}$$

where $x \wedge y = \min(x, y)$.

Now we discuss the choice of c_v in (2.27). We should take

$$c_v^2 = \{\sigma_\varepsilon^2 \operatorname{Var} I(\boldsymbol{x}_{t-1} < \boldsymbol{s})\}^{-1} \tag{2.33}$$

as a standardization factor. However $\operatorname{Var} I(\boldsymbol{x}_{t-1} < \boldsymbol{s})$ depends on $\boldsymbol{s}$, and σ_ε^2 is unknown, thus we use the maximum value of $\operatorname{Var} I(\boldsymbol{x}_{t-1} < s)$ instead of $\operatorname{Var} I(\boldsymbol{x}_{t-1} < s)$, and use $\hat{\sigma}_\varepsilon^2$ instead of σ_ε^2, i.e.,

$$c_u^2 = 4\hat{\sigma}_\varepsilon^{-2}, \tag{2.34}$$

where $\hat{\sigma}_\varepsilon^2$ is defined as in (2.24), and $\operatorname{Var} I(\boldsymbol{x}_{t-1} < \boldsymbol{s}) = P(\boldsymbol{x}_{t-1})[1 - P(\boldsymbol{x}_{t-1} < \boldsymbol{s})] \leqslant \frac{1}{4}$ for each $\boldsymbol{s}$.

Finally in order to ensure (1.7) and (1.8), in the appendix we shall prove that, as long as $c_n = o(n)$, (1.7) and (1.8) shall be true. According to our simulation results, we prefer to take

$$c_n = n/\log\log n. \tag{2.35}$$

Combining (2.31) and (2.35) we obtain a Cramer–von Mises-type statistic v_n.

Remark 2.1. The choice of the weight functions $w_1(\cdot)$ and $w_2(\cdot)$ may be important for improving the power of the test. In the simulation we tried the normal density function for $w_1(\cdot)$. The nonlinearity of threshold models 4 and 5 was not detected by the test. This may be due to that u_n is only sensitive to smooth alternatives. We then choose an exponential-type weight function which is less smooth than the normal one. The power of the test was significantly improved. Clearly, how to choose optimal weight functions deserves further study, but it may be beyond the scope of this paper.

3. Simulation study

We choose seven models as shown in Table 1 to generate simulated data and use the above approach as a test for linearity of the models. For convenience the sequence of residuals $\{\varepsilon_t\}$ is independent identically distributed and generated from N(0,1). Among these models, Models 1, 2 and 4 have been simulated by An and Cheng (1991). Similar to their choice, we choose the length of sample $n = 100$, 200 and 400 respectively. The number of replications is 1000, and the level of significance $\beta = 0.05$. The results are summarized in Table 2. We also present the simulation results for An and Cheng's test in Table 2 with bold face. In An and Cheng's test, we take $m = 90$ for $n = 100$, $m = 180$ for $n = 200$ and $m = 360$ for $n = 400$. (for the definition of m, see An and Cheng, 1991)

Table 1
Seven models used to generate simulated data

Model 1 [Linear AR(1)]	$x_t = 0.5x_{t-1} + \varepsilon_t$
Model 2 [Linear AR(2)]	$x_t = 0.4x_{t-1} - 0.3x_{t-2} + \varepsilon_t$
Model 3 [Threshold AR]	$x_t = \begin{cases} 2 + 0.5x_{t-1} + \varepsilon_t, & x_{t-1} < 1 \\ 0.5 - 0.4x_{t-1} + \varepsilon_t, & x_{t-1} \geqslant 1 \end{cases}$
Model 4 [Threshold AR]	$x_t = \begin{cases} 2 + 0.5x_{t-1} + \varepsilon_t, & x_{t-2} < 1 \\ 0.5 - 0.4x_{t-1} + \varepsilon_t, & x_{t-2} \geqslant 1 \end{cases}$
Model 5 [Threshold AR]	$x_t = \begin{cases} 2 + 0.5x_{t-1} + \varepsilon_t, & x_{t-3} < 1 \\ 0.5 - 0.4x_{t-1} + \varepsilon_t, & x_{t-3} \geqslant 1 \end{cases}$
Model 6 [Exponential AR]	$x_t = (0.8 + 4x_{t-1}^2 \mathrm{e}^{-x_{t-1}^2})x_{t-1} + \varepsilon_t$
Model 7 [Rational AR]	$x_t = \frac{1.9x_{t-2}^2}{(1+x_{t-1}^2)} + \varepsilon_t$

Table 2
Frequency (percent) of correct decisions

Model no.	Model's type	$n = 100$	$n = 200$	$n = 400$
1	Linear	93 (100)	93 (100)	94 (100)
2	Linear	92 (100)	95 (100)	96 (100)
3	Nonlinear	100 (83)	100 (100)	100 (100)
4	Nonlinear	14 (96)	38 (100)	73 (100)
5	Nonlinear	77 (56)	100 (98)	100 (100)
6	Nonlinear	65 (4)	95 (20)	99 (19)
7	Nonlinear	65 (0)	96 (0)	100 (0)

Table 3
The medians of v_n and u_n in the simulation

Model no.	Model's type	$n = 100$		$n = 200$		$n = 400$	
		u_n	v_n	u_n	v_n	u_n	v_n
1	Linear	0.16	0.60	0.13	0.53	0.13	0.82
2	Linear	0.15	0.66	0.13	0.52	0.12	0.44
3	Nonlinear	7.75	14.32	13.05	29.16	24.71	60.24
4	Nonlinear	0.98	0.81	1.65	1.32	3.09	1.97
5	Nonlinear	4.57	0.38	7.37	0.54	13.44	0.40
6	Nonlinear	1.33	4.50	2.81	8.59	5.70	5.04
7	Nonlinear	3.88	0.34	8.64	0.51	16.15	0.25

Table 4
The variance of v_n and u_n in the simulation

Model no.	Model's type	$n = 100$		$n = 200$		$n = 400$	
		u_n	v_n	u_n	v_n	u_n	v_n
1	Linear	0.08	1.77	0.07	3.48	0.07	2.19
2	Linear	0.12	2.94	0.11	1.28	0.07	2.20
3	Nonlinear	1.77	61.38	3.41	80.17	6.81	144.25
4	Nonlinear	0.77	1.88	1.22	3.67	1.44	7.96
5	Nonlinear	4.57	0.38	7.37	0.54	13.44	0.39
6	Nonlinear	0.92	59.37	3.23	167.6	5.57	689.9
7	Nonlinear	4.09	1.78	6.44	0.48	11.68	0.30

It is found that the test holds well the significance level (0.05) well, but not the case with An and Cheng's test in Table 2. The rejection proportion is close to the norminal level of 0.05, while An and Cheng's test seems to be oversensitive to the null hypothesis. In Table 2 we also find that the test is sensitive to most of the alternative models we considered. Although model 4 may be detected, the frequency of correct decision significantly increases with increasing the sample size n. Overall, the new test outperform the An and Cheng's test. From our experience on simulations, the new test may be easier to calculate even for higher order p. We also investigate the role of v_n and u_n. From Tables 3 and 4, we find that under H_0 v_n is the main determining factor, and under the alternatives, both are important for the power of the test.

4. Conclusion

The new mixed-type test introduced in Section 2, is showed to be consistent against any fixed alternative, and is easy to calculate without the curse of dimensionality involved as mentioned in Section 1. The new mixed type test for testing of nonlinearity of a time series may be suitable in the stage of exploratory data analysis on which no specific class of nonlinear function $\phi(\cdot)$ is assumed as the alternative. On the other hand, if we have some information on the form of the nonlinear function $\phi(\cdot)$, we may use a suitable test for linearity against that specific alternative (e.g. see Tong, 1990, Chapter 5.3). In this case, we may obtain higher power of the test. Finally, we point out that the approaches for test of nonlinearity in autoregressive model may readily be extended to test whether the underlying model is parametric. Such a topic will be discussed further.

Acknowledgements

The authors thank two referees for their comments which improved the presentation of the manuscript.

Appendix A.

A.1. Some notations and conditions

In addition to conditions (1.1) and (1.3) on the stationary time series x_t, we assume that ε_t is a stationary ergodic martingale difference sequence with the following moment conditions:

$$E\{\varepsilon_t^2|x_{t-1},\dots,x_{t-p}\} = \sigma_\varepsilon^2 = E\varepsilon_{t-1}^2 \quad \text{a.s.} \tag{A.1}$$

$$E|\varepsilon_t|^{4+\lambda} < \infty,\ E|x_t|^{4+\lambda} < \infty \quad \text{for some } \lambda > 0 \tag{A.2}$$

and x_t is an uniformly strong mixing series with

$$\sum_{k=1}^{\infty} k^2 \alpha_k^r < \infty, \quad \alpha_k = \mathrm{O}((k\log^4 k)^{-1}), \tag{A.3}$$

where

$$\alpha_k = \sup_{A\in\mathscr{F}^0_{-\infty}} |P(AB) - P(A)P(B)|/P(A), \tag{A.4}$$

$$r < \lambda/4(4+\lambda), \tag{A.5}$$

$\mathscr{F}^0_{-\infty} = \sigma\{x_s\colon s\leqslant 0\}$ and $\mathscr{F}_k^\infty = \sigma\{x_s\colon s\geqslant k\}$. The function $\phi(\cdot)$ in (1.1) is piecewise continuous. For convenience we call all of the above conditions together with (1.1), (1.3) and (A.1)–(A.3) Condition A.

A.2. Proof of (1.6)

Because (1.6) holds only under H_0, in this section we consider only the linear AR(p) series generated by model (1.4). Therefore under H_0 with Condition A, the following results concerned with the central limit theory and the law of large numbers are well known (e.g. see Hall and Heyde, 1980; An et al., 1982):

$$\lim_{n\to\infty} \hat{\sigma}_w^2 = E\{w_1(\boldsymbol{x}_{t-1}) - Ew_1(\boldsymbol{x}_{t-1})\}^2 \quad \text{a.s.,} \tag{A.6}$$

$$\lim_{n\to\infty} \bar{w}_1 = Ew_1(\boldsymbol{x}_{t-1}) \quad \text{a.s.,} \tag{A.7}$$

$$\lim_{n\to\infty} \hat{\Gamma} = \Gamma \quad \text{a.s.,} \tag{A.8}$$

$$\lim_{n\to\infty} \hat{\gamma}_u = \gamma_u \quad \text{a.s.,} \tag{A.9}$$

$$\lim_{n\to\infty} \hat{\sigma}_\varepsilon^2 = \sigma_\varepsilon^2 \quad \text{a.s.,} \tag{A.10}$$

$$\lim_{n\to\infty} \sqrt{n}\tilde{b}_0 \sim \mathrm{N}(0,*), \quad \text{and} \quad \lim_{n\to\infty} \sqrt{n}\tilde{\boldsymbol{b}} \sim \mathrm{N}(\boldsymbol{0}, \sigma_\varepsilon^2 \boldsymbol{\Gamma}^{-1}), \tag{A.11}$$

where $\hat{\sigma}_w$, $\bar{w}_1, \hat{\boldsymbol{\Gamma}}$, $\hat{\gamma}_u, \hat{\sigma}_\varepsilon^2$ are defined by (2.19) to (2.24), $\sigma_\varepsilon^2, \boldsymbol{\Gamma}$, and γ_u are denoted by (2.16)– (2.18) respectively, and $\tilde{b}_0 = \hat{b}_0 - b_0$ and $\tilde{\boldsymbol{b}} = \hat{\boldsymbol{b}} - \boldsymbol{b}$ are defined in (2.13). Hence from (A.6)– (A.11) and (2.25), (2.26) follow, i.e.,

$$\begin{aligned}\lim_{n\to\infty} \hat{\sigma}_u^2 &= \lim_{n\to\infty} \hat{\varepsilon}^2\{\hat{\sigma}_w^2 + \hat{\gamma}_u^\tau \hat{\boldsymbol{\Gamma}}^{-1} \hat{\gamma}_u\} \\ &= \sigma_\varepsilon^2\{\sigma_w^2 + \gamma_u^\tau \boldsymbol{\Gamma}^{-1} \bar{\gamma}_u\} \quad \text{a.s.}\end{aligned} \tag{A.12}$$

Theorem A.1. *Under* H_0 *with Condition A*, (1.6) *holds.*

Proof. By (2.13) we only need to prove (2.14) and (2.15). By (A.7) and (A.11) it is easy to see that

$$\begin{aligned}\frac{1}{\sqrt{n}}\tilde{b}_0 \sum_{t=p+1}^{n} \tilde{w}_1(\boldsymbol{x}_{t-1}) &= \sqrt{n}\tilde{b}_0 \cdot \frac{1}{n} \sum_{t=p+1}^{n} \{w(\boldsymbol{x}_{t-1}) - Ew_1(\boldsymbol{x}_{t-1})\} \\ &= (\sqrt{n}\tilde{b}_0) \cdot \{\bar{w}_1 - Ew_1(\boldsymbol{x}_{t-1})\} \to 0 \quad \text{in probability}\end{aligned} \tag{A.13}$$

which implies (2.14).

By (A.9) and (A.11), it is easy to see that

$$\begin{aligned}\frac{1}{\sqrt{n}}\tilde{\boldsymbol{b}}^\tau \sum_{t=p+1}^{n} \boldsymbol{x}_{t-1}\tilde{w}_1(\boldsymbol{x}_{t-1}) &= (\sqrt{n}\tilde{\boldsymbol{b}}^\tau) \cdot \left\{\frac{1}{n} \sum_{t=p+1}^{n} \boldsymbol{x}_{t-1}\tilde{w}_1(\boldsymbol{x}_{t-1})\right\} \\ &= (\sqrt{n}\tilde{\boldsymbol{b}}^\tau) \cdot \hat{\gamma}_u \sim \mathrm{N}(0, \sigma_\varepsilon^2 \gamma_u^\tau \boldsymbol{\Gamma}^{-1} \gamma_u).\end{aligned} \tag{A.14}$$

By (1.3), (A.1), (2.5) and Corollary 5.1 in Hall and Heyde (1980), we have

$$\frac{1}{\sqrt{n}}\sum_{t=p+1}^{n} e_t\tilde{w}_1(\boldsymbol{x}_{t-1}) = \frac{1}{\sqrt{n}}\sum_{t=p+1}^{n} \varepsilon_t\tilde{w}_1(\boldsymbol{x}_{t-1}) \sim \mathrm{N}(0, \sigma_\varepsilon^2 E\tilde{w}_1^2(\boldsymbol{x}_{t-1})). \tag{A.15}$$

It is easy to check (or see the proof of Theorem 1 in Zhu et al., 1993) that the statistics in (A.14) and (A.15) are asymptotically independent of each other. Finally from (A.14) and (A.15), (2.14) follows immediately.

A.3. Proof of (1.7)

In this section, we still consider the linear case of AR(p) series x_t. Replacing $w_1(\boldsymbol{x}_{t-1})$ in (2.13) by $I(\boldsymbol{x}_{t-1} < \boldsymbol{s})$, we obtain the following inequality which is similar to (2.13), i.e.,

$$\begin{aligned}\left|\frac{c_v}{\sqrt{n}}\sum_{t=p+1}^{n} \hat{e}_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s})\right| &= \left|\frac{c_v}{\sqrt{n}}\sum_{t=p+1}^{n} e_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s}) + \frac{c_v}{\sqrt{n}}\tilde{\boldsymbol{b}}^\tau \sum_{t=p+1}^{n} \boldsymbol{x}_{t-1} I(\boldsymbol{x}_{t-1} < \boldsymbol{s})\right. \\ &\quad \left. + \frac{c_v}{\sqrt{n}}\tilde{b}_0 \sum_{t=p+1}^{n} I(\boldsymbol{x}_{t-1} < \boldsymbol{s})\right| \\ &\leqslant \left|\frac{c_v}{\sqrt{n}}\sum_{t=p+1}^{n} \varepsilon_t I(x_{t-1} < \boldsymbol{s})\right| + \frac{2c_v}{\sqrt{n}}||\tilde{\boldsymbol{b}}|| \sum_{t=p+1}^{n} ||\boldsymbol{x}_{t-1}|| + c_v|\tilde{b}_0|\sqrt{n}.\end{aligned} \tag{A.16}$$

By (2.32), (A.11) and the law of large numbers

$$\begin{aligned}&\frac{1}{\log\log n}\int\left\{\frac{1}{\sqrt{n}}\tilde{\boldsymbol{b}}^\tau \sum_{t=p+1}^{n} \boldsymbol{x}_{t-1} I(\boldsymbol{x}_{t-1} < \boldsymbol{s})\right\}^2 w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \\ &\leqslant \frac{1}{\log\log n}\int 2\sqrt{n}||\tilde{\boldsymbol{b}}||^2 \cdot \frac{1}{n}\sum_{t=p+1}^{n} ||\boldsymbol{x}_{t-1}||^2 w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \\ &= \frac{2}{\log\log n}||\sqrt{n}\tilde{\boldsymbol{b}}||^2 \cdot \frac{1}{n}\sum_{t=p+1}^{n} ||\boldsymbol{x}_{t-1}||^2 \to 0 \quad \text{in probability}\end{aligned} \tag{A.17}$$

and

$$\begin{aligned}&\frac{1}{\log\log n}\int\left\{\frac{1}{\sqrt{n}}\tilde{b}_0 \sum_{t=p+1}^{n} I(\boldsymbol{x}_{t-1} < \boldsymbol{s})\right\}^2 w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \\ &\leqslant \frac{1}{\log\log n}|\sqrt{n}\tilde{b}_0|^2 \to 0 \quad \text{in probability.}\end{aligned} \tag{A.18}$$

By (1.3) and (2.34) for $k > j$, we derive that

$$\begin{aligned}E\varepsilon_k\varepsilon_j\psi(\boldsymbol{x}_{k-1}, \boldsymbol{x}_{j-1}) &= E\{\varepsilon_j\psi(\boldsymbol{x}_{k-1}, \boldsymbol{x}_{j-1})E[\varepsilon_k|x_{k-1}, \ldots, x_{k-p}]\} \\ &= E\{\varepsilon_j\psi(\boldsymbol{x}_{k-1}, \boldsymbol{x}_{j-1})E[\varepsilon_k|x_{k-1}, \ldots, x_{k-p}]\} = 0, \quad \text{a.s.}\end{aligned} \tag{A.19}$$

and

$$|\psi(\boldsymbol{x}_{k-1}, x_{j-1})| \leqslant 1. \tag{A.20}$$

From (A.19) and (A.20) it follows that

$$E\int\left\{\frac{1}{\sqrt{n}}\sum_{p+1}^{n}\varepsilon_t I(x_{t-1} < \boldsymbol{s})\right\}^2 w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s}$$

$$= \frac{1}{n}\sum_{k=p+1}^{n}\sum_{j=p+1}^{n} E\varepsilon_k\varepsilon_j\psi(\boldsymbol{x}_{k-1}, \boldsymbol{x}_{j-1})$$

$$\leqslant \frac{1}{n}\sum_{k=p+1}^{n}\sigma_\varepsilon^2 E\psi(\boldsymbol{x}_{k-1}, \boldsymbol{x}_{k-1}) \leqslant \sigma_\varepsilon^2. \tag{A.21}$$

Consequently by (A.21) and the Chebyshev inequality

$$\frac{1}{\log\log n}\int\left\{\frac{1}{\sqrt{n}}\sum_{t=p+1}^{n}\varepsilon_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s})\right\}^2 w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \to 0 \quad \text{in probability,} \tag{A.22}$$

hence, from (A.16)– (A.18) and (A.22), the following theorem follows immediately.

Theorem A.2. *Under* H_0 *with Condition A*, (1.7) *holds.*

A.4. Proof of (1.8)

To prove (1.8) under H_1, we have to consider general nonlinear AR(p) series generated by (1.2), hence we need the uniformly mixing condition (A.3). In the nonlinear case of stationary AR(p) series, under conditions (1.1), (1.3), (A.1), (A.2) and (A.3), (A.11) still hold as showed by An and Cheng (1991), or directly by using Corollary 5.1 in Hall and Heyde (1980). Consequently (A.17) and (A.8) can be shown by the argument used in the previous subsection. Note that $e_t \not\equiv \varepsilon_t$ under H_1. Then by (A.16) –(A.18), we have

$$\left|\frac{c_v}{\sqrt{n}}\sum_{t=p+1}^{n}\hat{e}_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s})\right| = \left|\frac{c_v}{\sqrt{n}}\sum_{t=p+1}^{n}e_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s})\right| + \mathrm{O_p}(\log\log n), \tag{A.23}$$

where $\mathrm{O_p}(\log\log n)$ denotes divergence order of $\log\log n$ in probability.

Recalling (2.3) and (1.1), we have

$$e_t = y_t - L(\boldsymbol{x}_{t-1}) = \phi(\boldsymbol{x}_{t-1}) - L(\boldsymbol{x}_{t-1}) + \varepsilon_t. \tag{A.24}$$

Note that $\{x_t\}$ is a uniformly mixing series, hence $\{\phi(\boldsymbol{x}_{t-1}) - L(\boldsymbol{x}_{t-1})\}$ is uniformly mixing as well. Then by Example 4 in Hall and Heyde (1980, p. 19) $\{[\phi(\boldsymbol{x}_{t-1}) - L(\boldsymbol{x}_{t-1})]I(\boldsymbol{x}_{t-1} < \boldsymbol{s})\}$ is a mixing series with $E\{\phi(\boldsymbol{x}_{t-1}) - L(\boldsymbol{x}_{t-1})\}^2 < \infty$. Consequently, by Theorem 2.21 in Hall and Heyde (1980) with condition (A.3), we have

$$\lim_{n\to\infty}\frac{1}{n}\sum_{t=p+1}^{n}\{\phi(\boldsymbol{x}_{t-1}) - L(\boldsymbol{x}_{t-1})\}I(\boldsymbol{x}_{t-1} < \boldsymbol{s})$$

$$= E[\phi(\boldsymbol{x}_{t-1}) - L(\boldsymbol{x}_{t-1})]I(\boldsymbol{x}_{t-1} < \boldsymbol{s}) \quad \text{a.s.} \tag{A.25}$$

By (A.2) it is obvious that $\sum_{t=1}^{k} \varepsilon_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s})$ is a martingale with finite second moment. Thus by Example 1 and Theorem 2.21 in Hall and Heyde (1980, pp. 19, 41) and (1.1) we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=p+1}^{n} \varepsilon_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s}) = E\varepsilon_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s}) = 0 \quad \text{a.s.} \tag{A.26}$$

Let

$$f(\boldsymbol{s}) = Ee_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s}) = E\{\phi(\boldsymbol{x}_{t-1} < \boldsymbol{s}) - L(\boldsymbol{x}_{t-1} < \boldsymbol{s})\} I(\boldsymbol{x}_{t-1} < \boldsymbol{s}). \tag{A.27}$$

From (A.24) to (A.27), we obtain

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=p+1}^{n} e_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s}) = f(\boldsymbol{s}) \quad \text{a.s.} \tag{A.28}$$

By Theorem 1.3.2 of Chow and Teicher (1978) under H_1

$$\sup_{\boldsymbol{s}} |f(\boldsymbol{s})| > 0. \tag{A.29}$$

Because $\phi(\cdot)$ is piecewise continuous, then there exist vectors $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ and positive number c_3 such that

$$f(\boldsymbol{s}) > c_3 > 0, \quad \boldsymbol{s} \in (\boldsymbol{c}_1, \boldsymbol{c}_2) \tag{A.30}$$

where $\boldsymbol{s} \in (\boldsymbol{c}_1, \boldsymbol{c}_2)$ means that $c_{1i} < s_i < c_{2i}$, $i = 1, 2, \ldots, p$, $\boldsymbol{c}_1 = (c_{11}, \ldots, c_{1p})^\tau$, $\boldsymbol{c}_2 = (\boldsymbol{c}_{21}, \ldots, c_{2p})^\tau$ and $\boldsymbol{s} = (s_1, \ldots, s_p)$.

By (A.28), (A.30) and the Fatou–Lebesgue Theorem (see Loeve, 1960, p. 126), it follows that

$$\begin{aligned}
&\liminf_{n\to\infty} \int \left\{ \frac{1}{n} \sum_{t=p+1}^{n} e_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s}) \right\}^2 w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \\
&\quad \geqslant \int \liminf_{n\to\infty} \left\{ \frac{1}{n} \sum e_t I(\boldsymbol{x}_{t-1} < \boldsymbol{s}) \right\}^2 w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \\
&\quad = \int f^2(\boldsymbol{s}) w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \\
&\quad \geqslant \int_{\boldsymbol{s}\in(\boldsymbol{c}_1,\boldsymbol{c}_2)} c_3 w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \\
&\quad \geqslant c_3 \int_{\boldsymbol{s}\in(\boldsymbol{c}_1,\boldsymbol{c}_2)} w_2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} > 0.
\end{aligned} \tag{A.31}$$

Eq. (1.8) is proved. Substituting (A.23) and (A.31) into (2.31), we obtain the following theorem.

Theorem A.3. *Under* H_1 *with Condition A*, (1.8) *holds.*

References

An, H.Z., Chen, Z.G., Hannan, E.J., 1982. Autocorrelation, autoregression and autoregressive approximation. Ann. Statist. 10, 926–936.

An, H.Z., Cheng, B., 1991. A Kolmogorov-Smirnov type statistic with application to test for nonlinearity in time series. Int. Statist. Rev. 59, 287–307.

Ashley, R.A., Patherson, D.M., Hinich, M.J., 1986. A diagnostic test for nonlinearity serial dependence in time series fitting errors. J. Time Ser. Anal. 7, 165–178.

Box, G.E.P., Jenkins, G.M., 1970. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.

Brockett, P.L., Hinich, M., Wilson, G.R., 1987. Nonlinear and non-Gaussianocean noise. J. Amer. Statist. Assoc. 82, 1386–1394.

Chan, W.S., Tong, H., 1986. On test for nonlinearity in time series analysis. J. Forecasting 5, 217–228.

Chen, X.R., Krishnaiah, P.R., 1990. A test of linearity of regression models. Chinese J. Appl. Probab. Statist. 6, 363–377 (in Chinese).

Chow, Y.S., Teicher, H., 1978. Probability Theory. Springer, New York.

Cox, D.D., Koh, E., 1989. A smoothing spline based test of model adequacy in polynomial regression. Ann. Inst. Statist. Math. 41, 383–400.

Eubank, R.L., Spiegelman, C.H., 1990. Testing the goodness of fit of a linear model via nonparametric regression techniques. J. Amer. Statist. Assoc. 85, 387–392.

Hall, P., Heyde, C.C., 1980. Martingale Limit Theory and its Application. Academia Press, New York.

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

Hjellivik, V., Tjϕstheim, D., 1993. Nonparametric tests of linearity for time series. Biometrika.

Keenan, P.J., 1985. A turkey nonadditivity-type test for time series. Biometrika 72, 39–44.

Loeve, M., 1960. Probability Theory, 2nd Edition. Van Nostrand Reinhold, New York.

Luukkonen, R., Saikkonen, P., Terasvirta, T., 1988. Testing linearity against smooth transition autoregression. Biometrika 75, 491–499.

Saikkonen, P., Luukkonen, R., 1988. Lagrange multiplier test for testing nonlinearity in time series models. Scand. J. Statist. 15, 55–68.

Subba Rao, T., Gabr, M.M., 1980. A test for linearity of stationary time series. J. Time Ser. Anal. 1, 145–158.

Terästirta, T, et al., 1993. Power of the neural network linearity test. J. Time Ser. Anal. 14, 209–220.

Tong, H., 1990. Non-linear Time Series. Oxford University Press, Oxford, New York.

Tsay, R., 1986. Nonlinearity tests for time series. Biometrika 73, 237–252.

Zhu, L.X., An, H.Z., 1992. A testing approach for nonlinearity in regression models. J. Math. 4, 391–397 (in Chinese).

Zhu, L.X., Hichernell, F., An, H.Z., 1993. A goodness of fit test for linearity of a stochastic regression model. Chinese J. Comtemp. 36, 199–206.