

MODEL DIAGNOSIS FOR SETAR TIME SERIES

Hira L. Koul¹, Winfried Stute² and Fang Li¹

¹Michigan State University and ²University of Giessen

Abstract: This paper discusses asymptotically distribution free (ADF) tests in self-exciting threshold autoregressive (SETAR) models. We also consider the case when the two different line segments have no jump. These tests are based on a marked empirical process of the underlying residuals. The paper also discusses the asymptotic behavior of the residual empirical process and ADF tests for the error distribution. We find that under some mild conditions, the asymptotic null behavior of both of these processes does not depend on the preliminary estimator of the change point parameter. Moreover, somewhat surprisingly, the asymptotic behavior of the residual empirical process in these models is the same as in the one-sample location model, as long as the residuals are based on an asymptotically linear estimator of the line segment parameters. The paper also includes a simulation study analyzing the finite sample behavior of some of the proposed tests.

Key words and phrases: Marked point processes, martingale type transform, model checks, time series.

1. Introduction

Let X_i , $i = 0, \pm 1, \pm 2, \dots$ be a real valued strictly stationary time series having finite expectation. Denote with $m(x) = \mathbb{E}[X_i | X_{i-1} = x]$ the associated autoregressive (AR) function of order 1. The time series is then called a self exciting threshold model of order 1 (SETAR(1)), if

$$X_i = m(X_{i-1}) + \varepsilon_i,$$

where $(\varepsilon_i)_i$ are i.i.d. $\sim F$ and m is piecewise linear over two different ranges of x . The importance of this model and its extensions and some statistical inference about the underlying parameters in these models have been discussed in Tong (1990) and references therein, and in Chan (1993) and Qian (1998).

The present paper discusses some model checks for SETAR(1) and some goodness-of-fit tests for the noise distribution function (d.f.) F . To be more precise about m , under SETAR(1), we have

$$m(x) = (a_0 + a_1x)I(x \leq r) + (b_0 + b_1x)I(x > r). \quad (1.1)$$

Let $h(x, \vartheta) = (\alpha_0 + \alpha_1x)I(x \leq s) + (\beta_0 + \beta_1x)I(x > s)$, where $\vartheta = (\alpha_0, \alpha_1, \beta_0, \beta_1, s)' \in \mathbb{R}^5$, the family of all functions of SETAR(1) type. Here and in the following, the

symbol I denotes the indicator of the set in brackets and $'$ is transposition. The typical statistical analysis within the SETAR(1) model then consists in estimating or testing hypotheses about the unknown parameter $\vartheta^* = (a_0, a_1, b_0, b_1, r)'$.

In this paper we are interested in checking the assumption, whether SETAR(1) holds at all, i.e., whether the hypothesis $H_0 : m(x) = h(x, \vartheta)$, for some $\vartheta \in \mathbb{R}^5$ is true or not, when the error distribution is not necessarily known.

Another goodness-of-fit problem is concerned with the distribution F of the innovations ε_i , assuming that SETAR(1) with $\vartheta = \vartheta^*$ holds. Here the problem of interest is to test the simple hypothesis $K_0 : F = F_0$, against the alternative that $F \neq F_0$, where F_0 is a known d.f. The knowledge of the error distribution plays some role if one wants to compute tolerance intervals for future values of the time series.

Both problems are special cases of classical problems of model checking and goodness-of-fit testing, and numerous tests for other models are available in the literature. See, e.g., the review paper by MacKinnon (1992), the papers by McKeague and Zhang (1994) Hjellvik and Tjøstheim (1995, 1996), the monographs of Hart (1997), Koul (2002), and the references therein. In most of the literature, however, the autoregressive function under study was invariably assumed to be smooth in the parameter, first order differentiability at least being required. Note that under SETAR(1), the autoregressive function m is non-smooth in both the lag variable and in the parameter vector.

In this paper, we investigate the behavior of the following test processes under SETAR(1):

$$V_n(x, \vartheta) = n^{-1} \sum_{i=1}^n (X_i - h(X_{i-1}, \vartheta)) I(X_{i-1} \leq x)$$

$$\hat{F}_n(x, \vartheta) = n^{-1} \sum_{i=1}^n I(X_i - h(X_{i-1}, \vartheta) \leq x), \quad x \in \mathbb{R}.$$

The process V_n may be viewed as a normalized point process of the observed X_i marked by the $X_i - h(X_{i-1}, \vartheta)$. For $\vartheta = \vartheta^*$ these marks become the true errors ε_i . Tests about H_0 will be based on V_n while \hat{F}_n will be needed for K_0 .

Actually, since the true parameter $\vartheta^* = (a_0, a_1, b_0, b_1, r)'$ is unknown, we need to study V_n with $\hat{\vartheta}$, where $\hat{\vartheta} = \hat{\vartheta}_n$ is an estimator of ϑ^* based on the observations X_0, \dots, X_n . Similarly, for \hat{F}_n .

In the 1990's, tests based on the analogue of V_n for checking the validity of a smooth parametric autoregressive model of order one were investigated by An and Cheng (1991) and Koul and Stute (1999). Extensions to smooth higher order autoregressive models are due to Dominguez and Lobato (2003) and Stute et al. (2004). In the regression context, tests based on the analogue of V_n have been

investigated by Su and Wei (1991), Stute (1997), Stute, Thies and Zhu (1998), Diebolt and Zuber (1999) and Stute and Zhu (2002), among others. Tests for the error distribution in a smooth autoregressive model based on $\hat{F}_n(x, \hat{\vartheta})$ have been studied in detail, see, e.g., Boldin (1982) and Koul (2002). Under discontinuities and structural changes, these methods are not now well understood.

As our two main results, we obtain, in Section 2, expansions of $V_n(\cdot, \hat{\vartheta})$ and $\hat{F}_n(\cdot, \hat{\vartheta})$ which could be used to obtain convergence in distribution in the Skorokhod space $D[-\infty, \infty]$ to centered Gaussian processes with a specified covariance structure.

A major role is played by the jump size $d \equiv b_0 - a_0 + r(b_1 - a_1)$. Theorems 2.1 and 2.2 cover in detail the case

$$d \neq 0. \tag{1.2}$$

An informal discussion of $d = 0$ with $a_1 \neq b_1$, i.e., when m is continuous but the slopes differ will be presented in Remark 2.2.

In both cases it will turn out that estimation of the change point r has a negligible effect on the asymptotic distributions of the involved processes, while the effect of the estimation of ϑ_1^* is non-negligible.

As in many cases when parameters are estimated, critical values for standard goodness-of-fit tests are, however, difficult to obtain. Therefore, in Section 3, we propose martingale transformations of $V_n(\cdot, \hat{\vartheta})$ and $\hat{F}_n(\cdot, \hat{\vartheta})$. The weak limit of the transformed $V_n(\cdot, \hat{\vartheta})$ under H_0 is a Brownian motion in proper time. For testing H_0 , we then apply known scaling properties to finally make our tests applicable to real data. The martingale transformation for $V_n(\cdot, \hat{\vartheta})$ is the same as given in Koul and Stute (1999) for the smooth case.

As to K_0 , somewhat surprisingly, we find that if $\hat{\vartheta}_1$ is asymptotically linear in probability, the asymptotic expansion of $\hat{F}_n(\cdot, \hat{\vartheta})$ is similar to the one appearing in a one-sample location model. Hence the martingale transformation here is the same as in Khmaladze (1981).

In Section 4 we report on some simulation results which indicate that the nominal level of our test for H_0 is well attained with high power against the selected alternatives. This section also includes a finite sample simulation of the bias and mean square error of the least square estimators of SETAR(1) parameters. Proofs are deferred to Section 5.

2. Main Results

This section describes the main results of the paper and some of their implications. Throughout we assume that the sequence $(X_i)_i$ is strictly stationary and ergodic under H_0 . Some conditions under which this holds are given in Tong (1990). Denote with G the (unknown) distribution function of X_0 . To proceed

further, it will be convenient to write $\vartheta = (\vartheta'_1, s)'$ and $h(x, \vartheta) = h_s(x, \vartheta_1)$, and to refer to ϑ_1 and s as the coefficient and the change-point parameters, respectively. Throughout the paper, $\vartheta'_1 := (\alpha_0, \alpha_1, \beta_0, \beta_1)$. Let

$$\begin{aligned} \dot{h}_s(x) &= (\partial/\partial\vartheta_1)h_s(x, \vartheta_1) \\ &= (I(x \leq s), xI(x \leq s), I(x > s), xI(x > s))', \end{aligned}$$

for $s \in \mathbb{R}^* \equiv \mathbb{R} \cup \{-\infty, \infty\}$ and $x \in \mathbb{R}$ denote the vector of partial derivatives of $h_s(x, \vartheta_1)$ w.r.t. ϑ_1 . Also, let

$$J_r(x) = \mathbb{E}\dot{h}_r(X_0)I(X_0 \leq x) = \mathbb{E} \begin{pmatrix} I(X_0 \leq x \wedge r) \\ X_0 I(X_0 \leq x \wedge r) \\ I(X_0 > r, X_0 \leq x) \\ X_0 I(X_0 > r, X_0 \leq x) \end{pmatrix}.$$

Finally put, for $x \in \mathbb{R}$ and $t \in \mathbb{R}$,

$$D_n(x, t) := n^{-1/2} \sum_{i=1}^n [\dot{h}_r(X_{i-1}) - \dot{h}_{r+tn^{-1}}(X_{i-1})]I(X_{i-1} \leq x).$$

We also need some estimator of ϑ . Chan (1993) and Qian (1998) have proved, e.g., that under $d \neq 0$ the conditional least squares and maximum likelihood estimator $\hat{\vartheta} = (\hat{\vartheta}'_1, \hat{r})'$ satisfies, under H_0 ,

$$n^{1/2}(\hat{\vartheta}_1 - \vartheta_1^*) = O_{\mathbb{P}}(1), \quad (2.1)$$

$$n(\hat{r} - r) = O_{\mathbb{P}}(1). \quad (2.2)$$

Here, $\vartheta_1^* = (a_0, a_1, b_0, b_1)'$ constitutes the first part of ϑ^* .

Theorem 2.1. *Assume that $(X_i)_i$ is a strictly stationary time series with autoregressive function (1.1). Assume $d \neq 0$ and that ϑ_1^* and r admit estimators $\hat{\vartheta}_1$ and \hat{r} satisfying (2.1) and (2.2). Additionally, suppose G is continuous at r . Then, uniformly in $x \in \mathbb{R}^*$,*

$$n^{1/2}[V_n(x, \hat{\vartheta}) - V_n(x, \vartheta^*)] = \vartheta_1^{*'} D_n(x, n(\hat{r} - r)) - n^{1/2}(\hat{\vartheta}_1 - \vartheta_1^*)' J_r(x) + o_{\mathbb{P}}(1). \quad (2.3)$$

Some implications of this theorem are discussed in Remark 2.1 below.

To state our second result, we need to introduce

$$\begin{aligned} \hat{F}_n(x) &= n^{-1} \sum_{i=1}^n I(\varepsilon_i \leq x), \\ \mu(x) &= \mathbb{E}X_0 I(X_0 \leq x), \quad \bar{\mu}(x) = \mathbb{E}X_0 I(X_0 > x), \quad \bar{G}(x) = 1 - G(x). \end{aligned} \quad (2.4)$$

Recall $J_r(x)$ and put $\Gamma_r \equiv J_r(\infty) = (G(r), \mu(r), \bar{G}(r), \bar{\mu}(r))'$.

Theorem 2.2. *Suppose the assumptions of Theorem 2.1 hold. In addition, assume that the error d.f. F has a uniformly continuous Lebesgue density f , and the stationary d.f. G satisfies*

$$n^{1/2} [G(r + bn^{-1}) - G(r - bn^{-1})] = o(1), \quad \forall 0 < b < \infty. \tag{2.5}$$

Then, uniformly in $x \in \mathbb{R}^*$,

$$n^{1/2} [\hat{F}_n(x, \hat{\vartheta}) - \hat{F}_n(x)] = n^{1/2} (\hat{\vartheta}_1 - \vartheta_1^*)' \Gamma_r f(x) + o_{\mathbb{P}}(1). \tag{2.6}$$

Remark 2.1. Consider the first term in the approximation (2.3). Uniformly in $x \in \mathbb{R}^*$ and $|t| \leq K$ we have

$$\begin{aligned} & \|\dot{h}_r(X_{i-1}) - \dot{h}_{r+tn^{-1}}(X_{i-1})\| I(X_{i-1} \leq x) \\ & \leq 2(1 + |X_{i-1}|) I(r - Kn^{-1} < X_{i-1} \leq r + Kn^{-1}), \end{aligned}$$

cf. (5.4). Hence on the event $|n(\hat{r} - r)| \leq K$ we get

$$\sup_x \|D_n(x, n(\hat{r} - r))\| \leq Cn^{1/2} [G_n(r + Kn^{-1}) - G_n(r - Kn^{-1})].$$

Under (2.5), the expectation of the last quantity, however, tends to zero. Conclude that if, in Theorem 2.1, we in addition require (2.5), then the first term on the right hand side of (2.3) tends to zero uniformly in x , in probability. For a later reference we summarize this discussion in

Corollary 2.1. *Under the assumptions of Theorem 2.1 and (2.5), uniformly in $x \in \mathbb{R}^*$, $n^{1/2} [V_n(x, \hat{\vartheta}) - V_n(x, \vartheta^*)] = -n^{1/2} (\hat{\vartheta}_1 - \vartheta_1^*)' J_r(x) + o_{\mathbb{P}}(1)$.*

Observe that in the above result the effect of estimating the jump point on the process V_n is not reflected in the same way as that of estimating the coefficients parameter vector ϑ_1^* . The primary reason for this is that the processes involved are of the magnitude $O_{\mathbb{P}}(n^{-1/2})$, while \hat{r} converges to r at the rate n^{-1} , in probability.

Remark 2.2. In the case of no jump, i.e., $d = 0$, $a_1 \neq b_1$, the two line segments have different slopes but they meet at the change-point r . In this case, \hat{r} converges to r at the rate $n^{-1/2}$, in probability, so that in (2.2) n is replaced by $n^{1/2}$. See, e.g., Chan and Tsay (1998) about the least squares estimator. In such a situation, Theorem 2.1 needs to be modified as follows. First, observe that $d = 0$ is equivalent to $\vartheta_1^{*'} a(r) = 0$ with $a(x) = (-1, -x, 1, x)'$. In addition to the

conditions of Theorem 2.1, assume G has a continuous Lebesgue density g at r with $g(r) > 0$. Then (2.3) changes to, uniformly in $x \in \mathbb{R}^*$,

$$\begin{aligned} & n^{1/2}[V_n(x, \hat{\vartheta}) - V_n(x, \vartheta^*)] \\ &= \vartheta_1^{*'} a(r) g(r) \left\{ \left(n^{1/2}(x - r) I(r < x \leq \hat{r}) + n^{1/2}(\hat{r} - r) I(x > \hat{r}) \right) I(\hat{r} > r) \right. \\ &\quad \left. - \left(n^{1/2}(x - \hat{r}) I(\hat{r} < x \leq r) + n^{1/2}(r - \hat{r}) I(x > r) \right) I(\hat{r} \leq r) \right\} \\ &\quad - n^{1/2}(\hat{\vartheta}_1 - \vartheta_1^*)' J_r(x) + o_{\mathbb{P}}(1). \end{aligned}$$

Similarly, under the conditions of Theorem 2.2 and the above conditions on g , the conclusion (2.6) is replaced with the following: uniformly in $x \in \mathbb{R}^*$,

$$\begin{aligned} n^{1/2}[\hat{F}_n(x, \hat{\vartheta}) - \hat{F}_n(x)] &= \left[-|n^{1/2}(\hat{r} - r)| \vartheta_1^{*'} a(r) g(r) + n^{1/2}(\hat{\vartheta}_1 - \vartheta_1^*)' \Gamma_r \right] f(x) \\ &\quad + o_{\mathbb{P}}(1). \end{aligned} \tag{2.7}$$

These results will not be proved here, but can be deduced by an analysis similar to that appearing in the proofs of Theorems 2.1 and 2.2 below. Note, however, that because $\vartheta_1^{*'} a(r) = 0$ in the case of no jump, these approximations reduce to those given in the above theorems when $d \neq 0$. Thus, even in this case, the effect of estimating the threshold parameter r is not reflected in the limiting behavior of these processes.

Remark 2.3. In many cases, the estimator $\hat{\vartheta}$ of ϑ^* is such that $\hat{\vartheta}_1$ is asymptotically linear, i.e., for some constant $c \neq 0$ and for some function ψ with $\mathbb{E}\psi(\varepsilon) = 0$ and $\mathbb{E}\psi^2(\varepsilon) < \infty$, we have

$$n^{1/2}(\hat{\vartheta}_1 - \vartheta_1^*) = c \Sigma_r^{-1} n^{-1/2} \sum_{i=1}^n \dot{h}_r(X_{i-1}) \psi(\varepsilon_i) + o_{\mathbb{P}}(1), \tag{2.8}$$

$$\Sigma_r = \mathbb{E} \dot{h}_r(X_0) \dot{h}_r(X_0)' = \begin{pmatrix} G(r) & \mu(r) & 0 & 0 \\ \mu(r) & \tau(r) & 0 & 0 \\ 0 & 0 & \bar{G}(r) & \bar{\mu}(r) \\ 0 & 0 & \bar{\mu}(r) & \bar{\tau}(r) \end{pmatrix}.$$

Here $\tau(r) = \mathbb{E} X_0^2 I(X_0 \leq r)$ and $\bar{\tau}(r) = \mathbb{E} X_0^2 I(X_0 > r)$. For example, the least squares estimator satisfies this condition with $\psi(x) = x$, $c = 1$, cf. Chan (1993). Or if F_0 has an absolutely continuous density f_0 with a.e. derivative f_0' such that $0 < \int (f_0'/f_0)^2 dF_0 < \infty$, the ratio $\psi := -f_0'/f_0$ is differentiable and ψ' is Lipschitz (1) with $\int \psi' dF_0 \neq 0$, then the MLE under K_0 satisfies this condition with $c = 1/\int \psi' dF_0$; see Qian (1998).

Now, combining (2.6) with (2.8) we obtain, under K_0 , that

$$\begin{aligned} & n^{1/2}[\hat{F}_n(x, \hat{\vartheta}) - F_0(x)] \\ &= n^{1/2}[\hat{F}_n(x) - F_0(x)] + c \Gamma'_r \Sigma_r^{-1} n^{-1/2} \sum_{i=1}^n \dot{h}_r(X_{i-1}) \psi(\varepsilon_i) f_0(x) + o_{\mathbb{P}}(1). \end{aligned}$$

In view of the fact $\Gamma'_r \Sigma_r^{-1} = (1, 0, 1, 0)$, the coefficient of $f_0(x)$ in the above approximation is $c n^{-1/2} \sum_{i=1}^n \psi(\varepsilon_i)$. In other words, under the conditions of Corollary 2.1 and under (2.8), we obtain that uniformly in $x \in \mathbb{R}$,

$$\begin{aligned} & n^{1/2}[\hat{F}_n(x, \hat{\vartheta}) - F_0(x)] \\ &= n^{1/2}[\hat{F}_n(x) - F_0(x)] + c n^{-1/2} \sum_{i=1}^n \psi(\varepsilon_i) f_0(x) + o_{\mathbb{P}}(1). \end{aligned} \tag{2.9}$$

The structure of the leading term on the right-hand side above is similar to the expansion of the classical residual empirical process in the one-sample location model. Hence any methodology leading to asymptotically distribution-free tests there will work here also. In the next section we focus on this methodology to test for H_0 and K_0 .

3. ADF tests for H_0 and K_0

Write $\hat{V}_n(x)$ and $V_n(x)$ for $V_n(x, \hat{\vartheta})$ and $V_n(x, \vartheta^*)$, respectively. Now take $\hat{\vartheta}$ as the least squares estimator of ϑ . This section will discuss a transformation T_n of \hat{V}_n , so that under the conditions of Corollary 2.1 and under H_0 , the processes $n^{1/2} T_n \hat{V}_n$ converge weakly to $\sigma B \circ G$, where B is a Brownian motion on $[0, 1]$. Hence, many tests of H_0 based on a continuous function of $n^{1/2} T_n \hat{V}_n$ will be ADF. Remark 3.1 outlines the situation for \hat{F}_n , while Remark 3.2 discusses the transformation for more general SETAR-models.

Because under the conditions of Corollary 2.1 the asymptotic behavior of \hat{V}_n does not depend on the estimator \hat{r} , the transformation T_n is the same as in Koull and Stute (1999) (K-S) studied under smoothness conditions. To describe it, let $H_r(y) := \mathbb{E} \dot{h}_r(X_0) \dot{h}_r(X_0)' I(X_0 \geq y), y \in \mathbb{R}$. In view of the definition of \dot{h}_r , we have

$$H_r(y) = \begin{pmatrix} \mathbb{E} I(X_0 \leq r, X_0 \geq y) & \mathbb{E} X_0 I(X_0 \leq r, X_0 \geq y) & 0 & 0 \\ \mathbb{E} X_0 I(X_0 \leq r, X_0 \geq y) & \mathbb{E} X_0^2 I(X_0 \leq r, X_0 \geq y) & 0 & 0 \\ 0 & 0 & \bar{G}(r \vee y) & \bar{\mu}(r \vee y) \\ 0 & 0 & \bar{\mu}(r \vee y) & \bar{\tau}(r \vee y) \end{pmatrix}.$$

Thus, because $I(x \leq r, x \geq y) = 0$ for all $y > r$, and because of the continuity of the distribution of X_0 , the matrix $H_r(y)$ is invertible only for $y < r$. Write

$H_r^{-1}(y)$ for its inverse and define the operator T through

$$T\ell := \ell(x) - \int \dot{h}_r(y)' H_r^{-1}(y) \left[\int \dot{h}_r(z) I(z \geq y) \ell(dz) \right] I(y \leq x) G(dy), \quad x < r.$$

In this definition ℓ is either a function of bounded variation or a Brownian motion. As mentioned in K-S, it preserves the Brownian motion and

$$n\sigma^{-2} \text{Cov}(TV_n(x), TV_n(y)) = G(x \wedge y) = n\sigma^{-2} \mathbb{E}(V_n(x)V_n(y)), \quad x, y < r,$$

the covariance function of the time transformed Brownian motion $B \circ G$. Moreover, if additionally $\mathbb{E}\varepsilon^4 < \infty$, then, for any $\delta > 0$, $TV_n \Leftarrow \sigma B \circ G$ in the Skorokhod space $D[0, r - \delta]$. But this result is of little use as T depends on the unknown parameters r and G . Let T_n and H_n denote T and H_r after r and G are replaced by their estimates \hat{r} and G_n , respectively. For numerical issues, see the next section. We now have

Theorem 3.1. *Suppose the assumptions of Corollary 2.1 hold with $\mathbb{E}\varepsilon^4 < \infty$. Then under H_0 , for any $\delta > 0$, and with $\tau := r - \delta$,*

$$n^{1/2} \sup_{x \leq \tau} |T_n \hat{V}_n(x) - TV_n(x)| = o_{\mathbb{P}}(1), \quad (3.1)$$

$$n^{1/2} \hat{\sigma}_n^{-1} T_n \hat{V}_n \Leftarrow B \circ G \quad \text{in distribution} \quad (3.2)$$

in the space $D[-\infty, \tau]$, where $\hat{\sigma}_n^2$ is a consistent estimator of σ^2 .

It readily follows from the derivations in the previous section that $\hat{\sigma}_n^2 := n^{-1} \sum_{i=1}^n (X_i - \hat{\vartheta}'_1 \dot{h}_{\hat{r}}(X_{i-1}))^2$ is a consistent estimator of σ^2 . By letting $\delta \rightarrow 0$, and using the consistency of \hat{r} for r , from the above theorem we are thus able to conclude that if the support of G contains $(-\infty, r)$, then under H_0 ,

$$\tilde{\mathcal{D}}_n := \sup_{x < \hat{r}} \left| \frac{n^{1/2} T_n \hat{V}_n(x)}{\hat{\sigma}_n G_n^{1/2}(\hat{r}-)} \right| \Leftarrow \sup_{0 \leq t \leq 1} |B(t)| \quad \text{in distribution.}$$

Thus the test that rejects H_0 when $\tilde{\mathcal{D}}_n > b_\alpha$ will be ADF of the asymptotic size α , provided b_α is determined so that $\mathbb{P}(\sup_{0 \leq t \leq 1} |B(t)| > b_\alpha) = \alpha$, $0 < \alpha < 1$.

Remark 3.1. To obtain ADF tests for K_0 , recall (2.9). Following Khmaladze (1981), assume F_0 has an absolutely continuous density f_0 with its a.e. derivative \dot{f}_0 satisfying $0 < \int (\dot{f}_0/f_0)^2 dF_0 < \infty$. Put

$$\varphi_0(u) := \frac{\dot{f}_0}{f_0}(F_0^{-1}(u)), \quad q(u) := f_0(F_0^{-1}(u)), \quad \gamma(u) := (1, \varphi_0(u))',$$

$$B_u := \begin{pmatrix} 1 - u & -q_0(u) \\ -q_0(u) & \int_u^1 \varphi_0^2 \end{pmatrix}, \quad 0 \leq u \leq 1.$$

Additionally, assume that the constant function 1 and the score function $\varphi_0(u)$ are linearly independent on the set $u > 1 - \eta$, for all sufficiently small $\eta > 0$. Then B_u^{-1} exists for all $0 \leq u < 1$. Define

$$\begin{aligned} \hat{\varepsilon}_i &:= X_i - h(X_{i-1}, \hat{\vartheta}), & L(u) &:= \int_0^u \gamma(s)' B_s^{-1} ds, \\ \zeta_i(u) &:= L(u)\gamma(\hat{\varepsilon}_i), & & 1 \leq i \leq n. \end{aligned}$$

Finally, set

$$n^{1/2}T_n\hat{F}_n(u, \hat{\vartheta}) := n^{-1/2} \sum_{i=1}^n \left\{ [1 - \zeta_i(\hat{\varepsilon}_i)]I(\hat{\varepsilon}_i \leq u) - \zeta_i(u)I(\hat{\varepsilon}_i > u) \right\}, \quad 0 \leq u \leq 1.$$

Similar to Khmaladze (1981) it follows then that, e.g., the Kolmogorov-Smirnov test based on the process $n^{1/2}T_n\hat{F}_n(\cdot, \hat{\vartheta}_n)$ will be ADF for testing for K_0 .

Remark 3.2. The above results also hold for some general stationary and ergodic SETAR models where the AR function may consist of more than two linear or nonlinear segments. More specifically, consider the following set up. Let k, p, q, d be known positive integers with $1 \leq d \leq p$, $Y_{i-1} := (X_{i-1}, X_{i-2}, \dots, X_{i-p})'$, $-\infty = r_0 < r_1 < \dots < r_k < r_{k+1} = \infty$ be a partition of \mathbb{R} , $\vartheta_1 \in \mathbb{R}^q$, and $g_j(Y_{i-1}, \vartheta_1)$, $j = 1, \dots, k + 1$, be some known functions. Let $r := (r_1, \dots, r_k)$. Consider the problem of testing for the general SETAR model where the AR function is given by

$$h(Y_{i-1}, \vartheta) := \sum_{j=1}^{k+1} g_j(Y_{i-1}, \vartheta_1)I(r_{j-1} < X_{i-d} \leq r_j), \quad \vartheta := (\vartheta_1', r)'. \quad (3.3)$$

Tong (1990) discusses some sufficient conditions for the stationarity and ergodicity of some AR models of type (3.3).

The appropriate analogue of the process \hat{V}_n on which tests are to be based here is

$$\hat{V}_n(x) := n^{-1} \sum_{i=1}^n (X_i - h(Y_{i-1}, \hat{\vartheta}))I(X_{i-d} \leq x),$$

where $\hat{\vartheta}$ is the least squares estimator of ϑ^* , and can be shown to satisfy the conditions (2.1) and (2.2) provided $\Sigma_r := \sum_{j=1}^{k+1} \mathbb{E}g_j(Y_0, \vartheta_1^*)g_j'(Y_0, \vartheta_1^*)I(r_{j-1} < X_{1-d} \leq r_j)$ is positive definite. In fact one can show that the least squares estimator of ϑ_1^* satisfies

$$n^{1/2}(\hat{\vartheta}_1 - \vartheta_1^*) = \Sigma_r^{-1}n^{-1/2} \sum_{i=1}^n \dot{h}_r(Y_{i-1}, \vartheta_1^*)\varepsilon_i + o_{\mathbb{P}}(1), \quad (3.4)$$

with $\dot{h}_r(Y_{i-1}, \vartheta_1^*) := \sum_{j=1}^{k+1} \dot{g}_j(Y_{i-1}, \vartheta_1^*) I(r_{j-1} < X_{i-d} \leq r_j)$.

Let $L_j(x) := \mathbb{E}[g_j(Y_0, \vartheta_1^*) | I(X_{1-d} \leq x)]$, $j = 1, \dots, k$, $x \in \mathbb{R}$. The analogue of condition (2.5) is

$$\sum_{j=1}^{k+1} n^{1/2} [L_j(r_j + b/n) - L_j(r_j - b/n)] = o(1), \quad \forall 0 < b < \infty. \quad (3.5)$$

Suppose, additionally, that for almost all $y \in \mathbb{R}^p$ (with respect to the distribution of Y_0), the functions $g_j(y, \cdot)$, $j = 1, \dots, k+1$, are absolutely continuous in a neighborhood of ϑ_1^* and the vector of the corresponding a.e. derivatives \dot{g}_j 's satisfies

$$\mathbb{E}[\dot{g}_j(Y_0, \vartheta_1^* + s) - \dot{g}_j(Y_0, \vartheta_1^*)] \rightarrow 0, \quad \text{as } \|s\| \rightarrow 0, \quad \forall j = 1, \dots, k+1. \quad (3.6)$$

Then the analogue of Corollary 2.1 continues to hold for \widehat{V}_n :

$$n^{1/2} [\widehat{V}_n(x) - V_n(x, \vartheta^*)] = -n^{1/2} (\hat{\vartheta}_1 - \vartheta_1^*)' J_r(x, \vartheta_1^*) + o_{\mathbb{P}}(1),$$

where now $J_r(x, \vartheta_1^*) := \mathbb{E} \dot{h}_r(Y_0, \vartheta_1^*) I(X_{1-d} \leq x)$, and $V_n(x, \vartheta^*) := n^{-1} \sum_{i=1}^n (X_i - h(Y_{i-1}, \vartheta^*)) I(X_{i-d} \leq x)$, $x \in \mathbb{R}$.

The analogue of the matrix $H_r(y)$ needed in the transformation T is

$$\begin{aligned} H_r(y, \vartheta_1^*) &:= \mathbb{E} \dot{h}_r(Y_0, \vartheta_1^*) \dot{h}_r(Y_0, \vartheta_1^*)' I(X_{1-d} \geq y) \\ &= \sum_{j=1}^{k+1} \mathbb{E} \dot{g}_j(Y_0, \vartheta_1^*) \dot{g}_j(Y_0, \vartheta_1^*)' I(r_{j-1} < X_{1-d} \leq r_j, X_{1-d} \geq y). \end{aligned}$$

This matrix is nonsingular only for $y < r_1$. Let $\widehat{h}(z) := \dot{h}_{\widehat{r}}(z, \widehat{\vartheta}_1)$, $z \in \mathbb{R}^p$, and $\widehat{H}_n(y) := n^{-1} \sum_{i=1}^n \widehat{h}(Y_{i-1}) \widehat{h}(Y_{i-1})' I(X_{i-d} \geq y)$, $y \in \mathbb{R}$. The analogue of the transformation T_n of the process $\widehat{V}_n(x)$ here is as follows. Let $\widehat{\varepsilon}_i := X_i - h(Y_{i-1}, \widehat{\vartheta})$. Then, for $x < r_1$,

$$\begin{aligned} T_n \widehat{V}_n(x) &:= n^{-1} \sum_{i=1}^n [I(X_{i-d} \leq x) \\ &\quad - n^{-1} \sum_{k=1}^n \widehat{h}(Y_{k-1})' \widehat{H}_n^{-1}(X_{k-d}) \widehat{h}(Y_{i-1}) I(X_{i-d} \wedge x \geq X_{k-d})] \widehat{\varepsilon}_i. \end{aligned}$$

An analogue of Theorem 3.1 also holds, and the analogue of the test \widetilde{D}_n here will be ADF for checking the above more general model.

Similarly, in connection with testing for an error distribution in these models, the analogues of (2.6) and (2.9) continue to hold under the above assumptions,

with Γ_r appearing in (2.6) given now by $\Gamma_r = \sum_{j=1}^{k+1} \mathbb{E}g_j(Y_0, \vartheta_1^*)I(r_{j-1} < X_0 \leq r_j)$.

In a general linear SETAR model of Tong (1990), $g_j(Y_0, \vartheta_1) = \vartheta'_{j,1} \begin{pmatrix} 1 \\ Y_0 \end{pmatrix}$, $j = 1, \dots, k + 1$, where each $\vartheta_{j,1}$ is a vector in \mathbb{R}^{p+1} , and $\vartheta_1 = (\vartheta_{1,1}, \dots, \vartheta_{k+1,1})$. In this case the assumption that the time series is stationary and ergodic, with the stationary d.f. having a bounded density that is positive at all jump points r_j , $j = 1, \dots, k$, and the error d.f. having zero mean and finite variance, imply (3.5), (3.6) and (3.4) for the LSE.

Under the same conditions, the analogues of Theorem 2.2 and Remark 2.3 also continue to hold for testing for an error d.f. in these general linear SETAR models, using the residuals based on the LSE. Note that even here with Γ_r, Σ_r as above, one has the property $\Gamma'_r \Sigma_r^{-1} = (1, 0, \dots, 1, 0)$.

4. Simulations

This section contains a simulation study of the proposed model check for a SETAR(1) model. It investigates the finite sample behavior of the level and power of the proposed test and includes the finite sample analysis of the bias and standard deviation (s.d) of the least square estimators of ϑ^* . The section also contains the graphs of the simulated densities of the standardized estimator $n(\hat{r} - r)$.

In the simulation, we chose $\vartheta^* = (0.5, 0.3, 0.6, -0.7, 0.5)'$, i.e., the true m was $h(x, \vartheta^*) = (0.5 + 0.3x)I[x \leq 0.5] + (0.6 - 0.7x)I[x > 0.5]$. Hence the jump equals $d = -0.4$. The errors were chosen to be normal $\mathcal{N}(0, 0.1)$ and logistic with mean zero and scale 0.05 (logis(0.05)), so that the s.d. is close to 0.1.

The data was simulated from the following three models:

Model 1: $X_i = h(X_{i-1}, \vartheta^*) + \varepsilon_i$

Model 2: $X_i = h(X_{i-1}, \vartheta^*) + 0.5(X_{i-1} - 0.6)^2 - 0.4(X_{i-1} - 0.6)^3 + \varepsilon_i$,

Model 3: $X_i = h(X_{i-1}, \vartheta^*) - 1.2 \exp(-X_{i-1}^2)X_{i-1} + \varepsilon_i$.

Note that Model 1 belongs to the null model while the other two are part of the alternative. The sample sizes chosen were 100, 200, 500 and 1,000, each simulation being repeated 2,000 times. The test statistic is

$$\tilde{\mathcal{D}}_n = \sup_{x < \hat{r}} \left| \frac{n^{1/2} T_n \hat{V}_n(x)}{\sigma_n \{G_n(\hat{r}-)\}^{\frac{1}{2}}} \right|, \quad \sigma_n = \left(\frac{1}{n} \sum_{i=1}^n (X_i - h(X_{i-1}, \hat{\vartheta}))^2 \right)^{\frac{1}{2}},$$

with $\hat{\vartheta}$ being the least squares estimator of ϑ^* .

We used the nominal levels $\alpha = 0.05, 0.025$ and 0.01 to implement the tests. Let b_α satisfy $\mathbb{P}(\sup_{0 \leq t \leq 1} |B(t)| > b_\alpha) = \alpha$. Using the well known fact

$$\mathbb{P}\left(\sup_{0 \leq t \leq 1} |B(t)| < b\right) = \mathbb{P}(|B(1)| < b) + 2 \sum_{i=1}^{\infty} (-1)^i \mathbb{P}((2i-1)b < B(1) < (2i+1)b),$$

we obtain the following table for some selected values of b_α . Since under H_0 , $\tilde{\mathcal{D}}_n \Rightarrow \sup_{0 \leq t \leq 1} |B(t)|$, the b_α 's are the asymptotic critical values of the proposed test. The empirical size and power are computed by using $\#\{\tilde{\mathcal{D}}_n > b_\alpha\}/2,000$.

Table 1. The critical values b_α .

α	0.05	0.025	0.01
b_α	2.24241	2.49771	2.80705

The simulation programming was done using S-plus. We first generated $(501 + n)$ error variables from $\mathcal{N}(0, 0.1)$ and $\text{logis}(0.05)$. Using these errors and Models 1-3 with the initial value of $X_0 = 0$, we generated $(501 + n)$ observations. The last $(n + 1)$ observations from the data thus generated are used in carrying out the simulation study for $n = 100, 200, 500$ and $1,000$. The density curves of the normalized \hat{r} are plotted by using density plot command with Gaussian kernel option in S-plus.

The results of the simulation study are shown in the Tables 2 and 3. Data simulated from Model 1 are used to study the empirical size and the data from Models 2 and 3 are used to study the empirical power of the test. One sees that under (null) Model 1, the empirical sizes of the tests are smaller than the true α levels for most of the moderate sample sizes, but they are much closer to the true levels when the sample size gets larger. Under Models 2 and 3, the simulated powers are seen to increase quickly with n and they are quite large for $n \geq 500$, even at α -level 0.01 , for both error distributions.

Table 2. Proportion of rejections for test $\tilde{\mathcal{D}}_n$ under models 1–3 with $\mathcal{N}(0, 0.1)$ errors.

α -level	$H_a \setminus n$	100	200	500	1,000
0.05	Model 1	0.0205	0.0320	0.0395	0.0415
	Model 2	0.2045	0.4760	0.8975	0.9975
	Model 3	0.0935	0.3870	0.8385	0.9865
0.025	Model 1	0.0085	0.0145	0.0190	0.0195
	Model 2	0.1270	0.3640	0.8265	0.9915
	Model 3	0.0515	0.3040	0.7890	0.9800
0.01	Model 1	0.0035	0.0070	0.0080	0.0085
	Model 2	0.0560	0.2375	0.7105	0.9770
	Model 3	0.0235	0.2185	0.7345	0.9605

Table 3. Proportion of rejections for test \tilde{D}_n under models 1–3 with logis(0.05) errors.

α -level	$H_a \setminus n$	100	200	500	1,000
0.05	Model 1	0.0180	0.0310	0.0475	0.0520
	Model 2	0.1670	0.4065	0.8490	0.9935
	Model 3	0.0855	0.3725	0.8440	0.9870
0.025	Model 1	0.0090	0.0140	0.0260	0.0300
	Model 2	0.1070	0.3030	0.7610	0.9850
	Model 3	0.0465	0.2800	0.7915	0.9750
0.01	Model 1	0.0025	0.0055	0.0075	0.0140
	Model 2	0.0550	0.1925	0.6390	0.9600
	Model 3	0.0205	0.1905	0.7260	0.9600

Table 4 below lists the means and standard deviations of the least squares estimator under H_0 for both normal and logistic error processes. From this table one sees very little bias for all sample sizes and that $\hat{\vartheta}$ converges to ϑ^* and standard deviations tend to decrease as sample sizes increase from 100 to 1,000.

Table 4. Means and (standard deviations) of $\hat{\vartheta}$ under model 1, i.e., H_0 .

estimate \ n	$\mathcal{N}(0, 0.1)$ errors				Logistic errors			
	100	200	500	1,000	100	200	500	1,000
\hat{a}_0	0.4997 (0.0241)	0.4999 (0.0171)	0.4997 (0.0108)	0.4999 (0.0076)	0.4996 (0.0224)	0.4999 (0.0158)	0.5001 (0.0103)	0.4998 (0.0073)
\hat{a}_1	0.3040 (0.0872)	0.3001 (0.0591)	0.3013 (0.0384)	0.3003 (0.0272)	0.3027 (0.0819)	0.3018 (0.0581)	0.3003 (0.0370)	0.2998 (0.0262)
\hat{b}_0	0.6019 (0.1250)	0.5969 (0.0888)	0.6018 (0.0537)	0.5996 (0.0387)	0.5985 (0.1163)	0.5998 (0.0812)	0.5987 (0.0513)	0.6004 (0.0351)
\hat{b}_1	-0.7029 (0.2004)	-0.6952 (0.1428)	-0.7026 (0.0859)	-0.6994 (0.0621)	-0.6978 (0.1900)	-0.6994 (0.1326)	-0.6982 (0.0834)	-0.7002 (0.0571)
\hat{r}	0.4940 (0.0065)	0.4971 (0.0033)	0.4989 (0.0013)	0.4994 (0.0007)	0.4946 (0.0061)	0.4973 (0.0030)	0.4989 (0.0012)	0.4994 (0.0006)

The simulation results of the densities of $n(\hat{r} - 0.5)$ are shown in Figures 1 and 2. The first figure contains the Monte Carlo density curves from 2,000 replications for the normal error processes and sample sizes $n = 100, 200, 500$ and 1,000, respectively, while the second figure has similar densities for the logis(0.05) process. Under both error processes, the graphs show that the distributions of the normalized estimate $n(\hat{r} - r)$ are skewed for all the sample sizes chosen. The graphs also give evidence that the convergence rate of \hat{r} to r is n^{-1} .

In the following figures, “.....” is for $n = 100$, “- . -” is for $n = 200$, “- - -” is for $n = 500$ and the solid line is for $n = 1,000$.

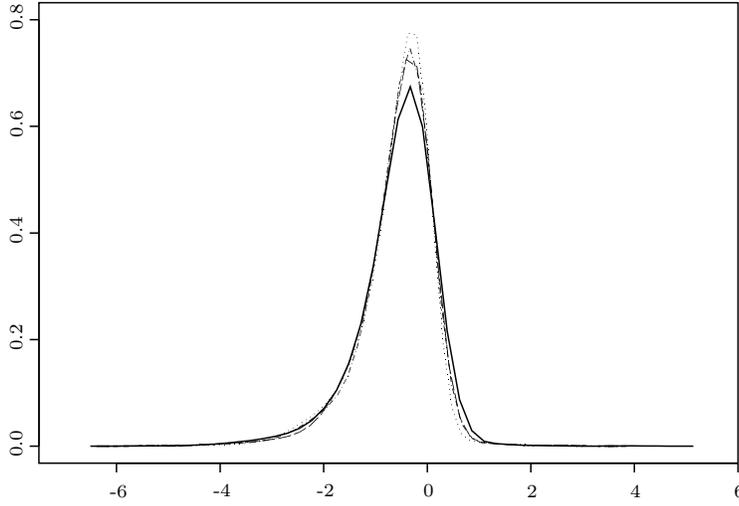


Figure 1. The density of $n(\hat{r} - 0.5)$ under H_0 with $\mathcal{N}(0, 0.1)$ errors.

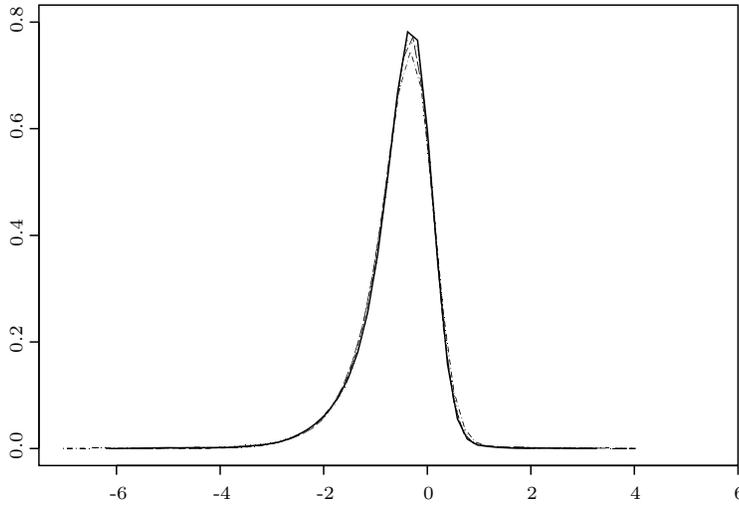


Figure 2. The density of $n(\hat{r} - 0.5)$ under H_0 with $\text{logis}(0.05)$ errors.

Computational scheme. For an interested reader we now describe the computation of $T_n \hat{V}_n$ and \tilde{D}_n used in the above simulations. As before, let $\vartheta_1 = (\alpha_0, \alpha_1, \beta_0, \beta_1)'$, $\vartheta_1^* = (a_0, a_1, b_0, b_1)'$, and $\vartheta^* = (\vartheta_1^{*'}, r)'$.

Step 1. Sort $\begin{pmatrix} X_0 & X_1 \\ \vdots & \vdots \\ X_{n-1} & X_n \end{pmatrix}$ according to the first column. Let $\begin{pmatrix} X_{(1)} & Y_1 \\ \vdots & \vdots \\ X_{(n)} & Y_n \end{pmatrix}$ denote the ordered observations, where $X_{(1)} \leq \dots \leq X_{(n)}$ are the ordered

X_0, \dots, X_{n-1} .

Step 2. For $k = 1, \dots, n$, minimize, with respect to $\vartheta_1 \in \mathbb{R}^4$,

$$M(\vartheta_1, k) = \sum_{i=1}^k (Y_i - (\alpha_0 + \alpha_1 X_{(i)}))^2 + \sum_{i=k+1}^n (Y_i - (\beta_0 + \beta_1 X_{(i)}))^2.$$

Let $\hat{\vartheta}_{1,k} = (\hat{a}_{0,k}, \hat{a}_{1,k}, \hat{b}_{0,k}, \hat{b}_{1,k})'$ denote the minimizer.

Step 3. Compute the least squares estimate $\hat{\vartheta}$ of ϑ^* under H_0 : $\hat{\vartheta} = (\hat{\vartheta}'_{1,\hat{k}}, \hat{r})'$, $\hat{r} = X_{(\hat{k})}$, $\hat{k} = \operatorname{argmin}_{1 \leq k \leq n} M(\hat{\vartheta}_{1,k}, k)$.

Step 4. Compute σ_n and $G_n(\hat{r}-)$: $\sigma_n = \sqrt{M(\hat{\vartheta}_{1,\hat{k}}, \hat{k})/n}$, $G_n(\hat{r}-) = (\hat{k} - 1)/n$.

Step 5. For $l = 1, \dots, \hat{k} - 1$, compute $T_n \hat{V}_n(X_{(l)})$:

Substep 1. Let $\hat{\varepsilon}_i := Y_i - h(X_{(i)}, \hat{\vartheta})$. Compute $n^{1/2} \hat{V}_n(X_{(l)}) = (1/\sqrt{n}) \sum_{i=1}^l \hat{\varepsilon}_i$.

Substep 2. Compute $U(y, z, X_{(l)}) := \dot{h}_{\hat{r}}(y)' H_n^{-1}(y) \dot{h}_{\hat{r}}(z) [z \geq y] [y \leq X_{(l)}]$, where $[A] := I(A)$ for any event A :

$$\begin{aligned} U(y, z, X_{(l)}) = & \frac{1}{D_n(y)} \left(\frac{1}{n} \sum_{i=1}^{\hat{k}} X_{(i)}^2 [X_{(i)} \geq y] [y \leq X_{(l)} \wedge z] [z \leq \hat{r}] \right. \\ & - \frac{1}{n} \sum_{i=1}^{\hat{k}} X_{(i)} [X_{(i)} \geq y] y [y \leq X_{(l)} \wedge z] [z \leq \hat{r}] \\ & - \frac{1}{n} \sum_{i=1}^{\hat{k}} X_{(i)} [X_{(i)} \geq y] [y \leq X_{(l)} \wedge z] z [z \leq \hat{r}] \\ & \left. + \frac{1}{n} \sum_{i=1}^{\hat{k}} [X_{(i)} \geq y] y [y \leq X_{(l)} \wedge z] z [z \leq \hat{r}] \right), \end{aligned}$$

$$D_n(y) = \frac{1}{n^2} \sum_{i=1}^{\hat{k}} X_{(i)}^2 [X_{(i)} \geq y] \sum_{i=1}^{\hat{k}} [X_{(i)} \geq y] - \frac{1}{n^2} \left(\sum_{i=1}^{\hat{k}} X_{(i)} [X_{(i)} \geq y] \right)^2.$$

Substep 3. Compute

$$\begin{aligned} \int U(y, z, X_{(l)}) G_n(dy) = & \frac{1}{n} \sum_{j=1}^l \frac{1}{D_n(X_{(j)})} \left(\frac{1}{n} \sum_{i=1}^{\hat{k}} X_{(i)}^2 [X_{(i)} \geq X_{(j)}] [X_{(j)} \leq z] [z \leq \hat{r}] \right. \\ & - \frac{1}{n} \sum_{i=1}^{\hat{k}} X_{(i)} [X_{(i)} \geq X_{(j)}] X_{(j)} [X_{(j)} \leq z] [z \leq \hat{r}] \\ & \left. - \frac{1}{n} \sum_{i=1}^{\hat{k}} X_{(i)} [X_{(i)} \geq X_{(j)}] [X_{(j)} \leq z] z [z \leq \hat{r}] \right) \end{aligned}$$

$$+\frac{1}{n} \sum_{i=1}^{\hat{k}} [X_{(i)} \geq X_{(j)}] X_{(j)} [X_{(j)} \leq z] z [z \leq \hat{r}].$$

Substep 4. Compute

$$\begin{aligned} & n^{1/2} \int \int U(y, z, X_{(l)}) G_n(dy) \widehat{V}_n(dz) \\ &= \frac{1}{\sqrt{n}} \sum_{k=1}^{\hat{k}} \hat{\varepsilon}_k \frac{1}{n} \sum_{j=1}^{l \wedge k} \frac{1}{D_n(X_{(j)})} \left[\frac{1}{n} \sum_{i=j}^{\hat{k}} X_{(i)}^2 \right. \\ & \quad \left. - \frac{1}{n} \sum_{i=j}^{\hat{k}} X_{(i)} X_{(j)} - \frac{1}{n} \sum_{i=j}^{\hat{k}} X_{(i)} X_{(k)} + \frac{1}{n} \sum_{i=j}^{\hat{k}} X_{(j)} X_{(k)} \right]. \end{aligned}$$

Substep 5. From the definition of $T_n \widehat{V}_n$, and the above substeps 1-4, we obtain

$$\begin{aligned} & n^{1/2} T_n \widehat{V}_n(X_{(l)}) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^l \hat{\varepsilon}_i - \frac{1}{\sqrt{n}} \sum_{k=1}^{\hat{k}} \hat{\varepsilon}_k \frac{1}{n} \sum_{j=1}^{l \wedge k} \frac{1}{D_n(X_{(j)})} \left[\frac{1}{n} \sum_{i=j}^{\hat{k}} X_{(i)}^2 \right. \\ & \quad \left. - \frac{1}{n} \sum_{i=j}^{\hat{k}} X_{(i)} X_{(j)} - \frac{1}{n} \sum_{i=j}^{\hat{k}} X_{(i)} X_{(k)} + \frac{1}{n} \sum_{i=j}^{\hat{k}} X_{(j)} X_{(k)} \right]. \end{aligned}$$

Step 6. Compute the test statistic

$$\tilde{\mathcal{D}}_n = \sup_{1 \leq l \leq \hat{k}-1} \left| \frac{n^{1/2} T_n \widehat{V}_n(X_{(l)})}{\sqrt{\frac{M(\hat{\vartheta}_{1, \hat{k}}, \hat{k})(\hat{k}-1)}{n^2}}} \right|.$$

5. Proofs

We first summarize some facts about the AR function h and its derivatives \dot{h}_s . For $x \in \mathbb{R}$ and $\vartheta, \vartheta^* \in \mathbb{R}^5$ we have

$$\|\dot{h}_s(x)\| \leq 1 + |x|, \tag{5.1}$$

$$|h_s(x, \vartheta_1)| \leq \|\vartheta_1\| (1 + |x|), \tag{5.2}$$

$$|h_s(x, \vartheta_1) - h_s(x, \vartheta_1^*)| \leq \|\vartheta_1 - \vartheta_1^*\| (1 + |x|). \tag{5.3}$$

Moreover, for all real numbers x and $s \leq t$,

$$\|\dot{h}_s(x) - \dot{h}_t(x)\| \leq 2(1 + |x|) I(s < x \leq t). \tag{5.4}$$

Finally, note that

$$h_s(x, \vartheta_1) = \vartheta_1' \dot{h}_s(x). \tag{5.5}$$

Proof of Theorem 2.1. We obviously have, cf. (5.5),

$$n^{1/2}[V_n(x, \hat{\vartheta}) - V_n(x, \vartheta^*)] = n^{-1/2} \sum_{i=1}^n [\vartheta_1^{*'} \dot{h}_r(X_{i-1}) - \hat{\vartheta}_1' \dot{h}_{\hat{r}}(X_{i-1})] I(X_{i-1} \leq x).$$

Write

$$\hat{\vartheta}_1 = \vartheta_1^* + un^{-1/2} \text{ and } \hat{r} = r + tn^{-1}. \tag{5.6}$$

By (2.1) and (2.2), for a given $\varepsilon > 0$, we may find a (large) constant K so that for $n \geq n_0$, say, we have up to an event of probability $\leq \varepsilon$, that $\|u\| \leq K$ and $|t| \leq K$. We therefore have to study the processes

$$\Delta_n(x, t, u) := n^{-1/2} \sum_{i=1}^n [\vartheta_1^{*'} \dot{h}_r(X_{i-1}) - (\vartheta_1^* + un^{-1/2})' \dot{h}_{r+tn^{-1}}(X_{i-1})] I(X_{i-1} \leq x)$$

uniformly in $x \in \mathbb{R}^*$ and $\|u\|, |t| \leq K$. Expand Δ_n as

$$\begin{aligned} \Delta_n(x, t, u) &= n^{-1/2} \sum_{i=1}^n \vartheta_1^{*'} [\dot{h}_r(X_{i-1}) - \dot{h}_{r+tn^{-1}}(X_{i-1})] I(X_{i-1} \leq x) \\ &\quad - u'n^{-1} \sum_{i=1}^n [\dot{h}_{r+tn^{-1}}(X_{i-1}) - \dot{h}_r(X_{i-1})] I(X_{i-1} \leq x) \\ &\quad - u'n^{-1} \sum_{i=1}^n \dot{h}_r(X_{i-1}) I(X_{i-1} \leq x). \end{aligned}$$

Now apply Cauchy-Schwarz and (5.4) to get, uniformly in x and $\|u\| \leq K, 0 < t \leq K$, that the second sum in absolute values is less than or equal to $2Kn^{-1} \sum_{i=1}^n (1 + |X_{i-1}|) I(r \leq X_{i-1} \leq r + Kn^{-1})$. From the continuity of G at r and the Ergodic Theorem, the last term is easily seen to converge to zero with probability one. Similarly, for $-K \leq t \leq 0$.

Finally, from a Glivenko-Cantelli result for strictly stationary time series (see, e.g., Stute and Schumann (1980)), we have with probability one, uniformly in $x \in \mathbb{R}^*$,

$$\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n \dot{h}_r(X_{i-1}) I(X_{i-1} \leq x) = \mathbb{E} \dot{h}_r(X_0) I(X_0 \leq x) = J_r(x).$$

Recalling (5.6), we thus obtain, uniformly in $x \in \mathbb{R}^*$,

$$\begin{aligned} &n^{1/2}[V_n(x, \hat{\vartheta}) - V_n(x, \vartheta^*)] \\ &= \vartheta_1^{*'} n^{-1/2} \sum_{i=1}^n [\dot{h}_r(X_{i-1}) - \dot{h}_{\hat{r}}(X_{i-1})] I(X_{i-1} \leq x) - n^{1/2}(\hat{\vartheta}_1 - \vartheta_1^*)' J_r(x) + o_{\mathbb{P}}(1), \end{aligned}$$

as desired.

Proof of Theorem 2.2. By (5.5), the residuals may be written as

$$\begin{aligned} X_i - \hat{\vartheta}'_1 h_{\hat{r}}(X_{i-1}) &= \varepsilon_i + \left[\vartheta_1^{*'} \dot{h}_r(X_{i-1}) - (\vartheta_1^* + un^{-1/2})' \dot{h}_{r+tn^{-1}}(X_{i-1}) \right] \\ &= \varepsilon_i + \xi_n(X_{i-1}, t, u), \text{ say,} \end{aligned}$$

where with large probability, t and u satisfy $|t| \leq K$ and $\|u\| \leq K$, for some $K < \infty$.

Hence for u and t from (5.6), $\hat{F}_n(x, \hat{\vartheta}) = n^{-1} \sum_{i=1}^n I(\varepsilon_i \leq x - \xi_n(X_{i-1}, t, u)) \equiv \hat{F}_n(x, t, u)$. Recall the definition of \hat{F}_n from (2.4) and let $F_n(x, t, u) = n^{-1} \sum_{i=1}^n F(x - \xi_n(X_{i-1}, t, u))$, $\alpha_n(x, t, u) = n^{1/2} [\hat{F}_n(x, t, u) - F_n(x, t, u)]$. Let $\mathbf{0}$ denote the zero vector in \mathbb{R}^4 . Note that we have $\xi_n(x, 0, \mathbf{0}) \equiv 0$, so that $F_n(x, 0, \mathbf{0}) = F(x)$ and $\hat{F}_n(x, 0, \mathbf{0}) = \hat{F}_n(x)$.

We are going to show that, for any finite $K > 0$,

$$\sup_{x \in \mathbb{R}, |t| \leq K, \|u\| \leq K} |\alpha_n(x, t, u) - \alpha_n(x, 0, \mathbf{0})| = o_{\mathbb{P}}(1), \quad (5.7)$$

$$\sup_{x \in \mathbb{R}, |t| \leq K, \|u\| \leq K} |n^{1/2} [F_n(x, t, u) - F(x)] - u' \Gamma_r f(x)| = o_{\mathbb{P}}(1). \quad (5.8)$$

To show (5.8), assume $t > 0$ w.l.o.g. Decompose

$$\begin{aligned} F_n(x, t, u) &= n^{-1} \sum_{i=1}^n F(x - \xi_n(X_{i-1}, t, u)) I(X_{i-1} \in (r, r + tn^{-1})) \\ &\quad + n^{-1} \sum_{i=1}^n F(x - \xi_n(X_{i-1}, t, u)) I(X_{i-1} \notin (r, r + tn^{-1})) \\ &\equiv A_1(x, t, u) + A_2(x, t, u), \quad \text{say.} \end{aligned}$$

Similarly, write

$$\begin{aligned} F(x) &= F(x) n^{-1} \sum_{i=1}^n I(X_{i-1} \in (r, r + tn^{-1})) + F(x) n^{-1} \sum_{i=1}^n I(X_{i-1} \notin (r, r + tn^{-1})) \\ &\equiv B_1(x, t, u) + B_2(x, t, u). \end{aligned}$$

Since $\sup_{x; t, u} |A_1(x, t, u) - B_1(x, t, u)| \leq 2n^{-1} \sum_{i=1}^n I(X_{i-1} \in (r, r + Kn^{-1}))$ and the expectation of the last term equals $2[G(r + Kn^{-1}) - G(r)] = o(n^{-1/2})$, by (2.5), we obtain

$$n^{1/2} \sup_{x; t, u} |A_1(x, t, u) - B_1(x, t, u)| = o_{\mathbb{P}}(1). \quad (5.9)$$

Next, we discuss $A_2 - B_2$. Since now we sum over $X_{i-1} \notin (r, r + tn^{-1}]$, we have

$$n^{1/2} \xi_n(X_{i-1}, t, u) = -u' \dot{h}_r(X_{i-1}). \quad (5.10)$$

From this, uniformly in x and $|t| \leq K, \|u\| \leq K$,

$$\begin{aligned} & A_2(x, t, u) - B_2(x, t, u) \\ &= n^{-1} \sum_{i=1}^n [F(x - \xi_n(X_{i-1}, t, u)) - F(x)] I(X_{i-1} \notin (r, r + tn^{-1})) \\ &= n^{-1} \sum_{i=1}^n [F(x + u'n^{-1/2}\dot{h}_r(X_{i-1})) - F(x)] I(X_{i-1} \notin (r, r + tn^{-1})). \end{aligned}$$

Now we can use an argument similar to the one leading to (5.9) to show that the last sum equals, uniformly in $x \in \mathbb{R}$,

$$n^{-1} \sum_{i=1}^n [F(x + u'n^{-1/2}\dot{h}_r(X_{i-1})) - F(x)] + o_{\mathbb{P}}(n^{-1/2}). \tag{5.11}$$

Apply Taylor's formula, the integrability of $\dot{h}_r(X_0)$ and the Ergodic Theorem to obtain that (5.11) equals, uniformly in $x \in \mathbb{R}$, $f(x) u'n^{-3/2} \sum_{i=1}^n \dot{h}_r(X_{i-1}) + o_{\mathbb{P}}(n^{-1/2})$. The assertion (5.8) now follows from another application of the Ergodic Theorem.

We now sketch a proof of (5.7). As before, we may also decompose sums into X_{i-1} satisfying $r < X_{i-1} \leq r + tn^{-1}$ and the rest. We then obtain, uniformly in x and $0 \leq t \leq K, \|u\| \leq K$,

$$\begin{aligned} & n^{-1/2} \sum_{i=1}^n [I(\varepsilon_i \leq x - \xi_n(X_{i-1}, t, u)) - F(x - \xi_n(X_{i-1}, t, u))] \\ & \qquad \qquad \qquad \times I(r < X_{i-1} \leq r + tn^{-1}) = o_{\mathbb{P}}(1). \end{aligned}$$

Conclude from (5.10) that, uniformly in x and $|t|, \|u\| \leq K$,

$$\begin{aligned} & \alpha_n(x, t, u) - \alpha_n(x, 0, \mathbf{0}) \\ &= n^{-1/2} \sum_{i=1}^n \left[I(\varepsilon_i \leq x + n^{-1/2}u'\dot{h}_r(X_{i-1})) - F(x + n^{-1/2}u'\dot{h}_r(X_{i-1})) \right. \\ & \qquad \qquad \qquad \left. - I(\varepsilon_i \leq x) + F(x) \right] + o_{\mathbb{P}}(1). \end{aligned} \tag{5.12}$$

Let $U_n(x, u)$ stand for the leading term on the right hand side of (5.12). Note that for each index i the random variable ε_i is independent of $\dot{h}_r(X_{i-1})$. Hence the summands of $U_n(x, u)$ form a martingale difference array. For fixed u , we may therefore apply Theorem 1.1 of Koul and Ossiander (1994) to obtain that the process $U_n(x, u)$ is C -tight in x , provided $n^{-1/2} \max_{1 \leq i \leq n} \|\dot{h}_r(X_{i-1})\| = o_{\mathbb{P}}(1)$ and $n^{-1} \sum_{i=1}^n \|\dot{h}_r(X_{i-1})\| = O_{\mathbb{P}}(1)$. The first condition, however, follows since X_0 has a finite second moment, while the second follows from the Ergodic Theorem.

Altogether this shows that for a fixed u , $\sup_x |U_n(x, u)| = o_{\mathbb{P}}(1)$. To obtain this result uniformly in u over a compact set, we use a standard argument. First cover the cube $\|u\| \leq K$ by finitely many small cubes and let u_0 be the center of such a cube, say $C(u_0)$. To compare $U_n(x, u)$ with $U_n(x, u_0)$ over $C(u_0)$, use the monotonicity of the indicator function and of F , and observe that after telescoping, there will appear error terms of the form

$$n^{-1/2} \sum_{i=1}^n \left[F(x + n^{-1/2} u' \dot{h}_r(X_{i-1})) - F(x + n^{-1/2} u_0' \dot{h}_r(X_{i-1})) \right],$$

which, in turn, are bounded from above in absolute value by

$$\|f\| \left| n^{-1} \sum_{i=1}^n (u - u_0)' \dot{h}_r(X_{i-1}) \right|.$$

Such terms can be made arbitrarily small, uniformly in $u \in C(u_0)$, provided that $C(u_0)$ has been chosen small enough. Hence the oscillations of $U_n(x, u)$ are uniformly small over $C(u_0)$. This completes the proof of (5.7) and, together with (5.8), also of Theorem 2.2.

Proof of Theorem 3.1. The claim (3.2) follows from (3.1) and the fact $n^{1/2} TV_n \Leftarrow \sigma B \circ G$ in $D[-\infty, \tau]$, proved in K-S.

The basic details of the proof of (3.1) are similar to those appearing in K-S, but the discontinuity of h_r makes some details necessarily different. We briefly indicate the differences. Let $\Delta_n := n^{1/2}(\hat{\vartheta}_1 - \vartheta_1^*)$. We have

$$\begin{aligned} T_n \widehat{V}_n(x) &= \widehat{V}_n(x) - \int \dot{h}_{\hat{r}}(y)' H_n^{-1}(y) \left[\int \dot{h}_{\hat{r}}(z) I(z \geq y) \widehat{V}_n(dz) \right] I(y \leq x) G_n(dy), \\ TV_n(x) &= V_n(x) - \int \dot{h}_r(y)' H_r^{-1}(y) \left[\int \dot{h}_r(z) I(z \geq y) V_n(dz) \right] I(y \leq x) G(dy). \end{aligned}$$

Let $\widehat{U}_n(y) := \int \dot{h}_{\hat{r}}(z) I(z \geq y) \widehat{V}_n(dz)$, $U_n(y) := \int \dot{h}_r(z) I(z \geq y) V_n(dz)$. From Corollary 2.1, we obtain $n^{1/2} \widehat{V}_n(x) = n^{1/2} V_n(x) - \Delta_n' J_r(x) + o_{\mathbb{P}}(1)$. Thus we have

$$\begin{aligned} & n^{1/2} [T_n \widehat{V}_n(x) - TV_n(x)] \\ &= -\Delta_n' J_r(x) + o_{\mathbb{P}}(1) + n^{1/2} \int \dot{h}_r(y)' H_r^{-1}(y) U_n(y) I(y \leq x) G(dy) \\ &\quad - n^{1/2} \int \dot{h}_{\hat{r}}(y)' H_n^{-1}(y) \widehat{U}_n(y) I(y \leq x) G_n(dy) \\ &= -\Delta_n' J_r(x) + o_{\mathbb{P}}(1) + T_{n1}(x) - T_{n2}(x), \quad \text{say.} \end{aligned} \tag{5.13}$$

Next, we have the bound

$$\|H_n(x) - H_r(x)\| \leq 2n^{-1} \sum_{i=1}^n \left\| \dot{h}_r(X_{i-1}) \right\| \left\| \dot{h}_{\hat{r}}(X_{i-1}) - \dot{h}_r(X_{i-1}) \right\|$$

$$\begin{aligned}
 &+ n^{-1} \sum_{i=1}^n \sum_{i=1}^n \left\| \dot{h}_{\hat{r}}(X_{i-1}) - \dot{h}_r(X_{i-1}) \right\|^2 \\
 &+ \left\| n^{-1} \sum_{i=1}^n \dot{h}_r(X_{i-1}) \dot{h}_r(X_{i-1})' I(X_{i-1} \geq x) - H_r(x) \right\|, \\
 &= B_1 + B_2 + B_3(x), \quad \text{say.}
 \end{aligned}$$

By (5.1) and (5.4), for a finite constant C ,

$$\begin{aligned}
 B_1 + B_2 &\leq C n^{-1} \sum_{i=1}^n \left(1 + |X_{i-1}|\right)^2 I(|X_{i-1} - r| \leq |\hat{r} - r|) \\
 &\leq C \left[|\hat{r} - r|^2 + (1 + r)^2 \left\{ G_n(r + |\hat{r} - r|) - G_n(r - |\hat{r} - r|) \right\} \right] \\
 &= o_{\mathbb{P}}(1).
 \end{aligned}$$

The Ergodic Theorem and a Glivenko-Cantelli type argument show that $\sup_{x \in \mathbb{R}} \|B_3(x)\| = o_{\mathbb{P}}(1)$. Hence we have

$$\sup_{x \in \mathbb{R}} \|H_n(x) - H_r(x)\| = o_{\mathbb{P}}(1), \tag{5.14}$$

$$\sup_{x \leq \tau} \|H_n^{-1}(x) - H_r^{-1}(x)\| = o_{\mathbb{P}}(1). \tag{5.15}$$

Next, rewrite

$$\begin{aligned}
 n^{1/2} \widehat{U}_n(y) &= n^{-1/2} \sum_{i=1}^n \dot{h}_{\hat{r}}(X_{i-1}) \left(X_i - \hat{\vartheta}'_1 \dot{h}_{\hat{r}}(X_{i-1}) \right) I(X_{i-1} \geq y) \\
 &= n^{-1/2} \sum_{i=1}^n \dot{h}_{\hat{r}}(X_{i-1}) \left[\varepsilon_i + \vartheta_1^{*'} \dot{h}_r(X_{i-1}) - \hat{\vartheta}'_1 \dot{h}_{\hat{r}}(X_{i-1}) \right] I(X_{i-1} \geq y) \\
 n^{1/2} U_n(y) &= n^{-1/2} \sum_{i=1}^n \dot{h}_r(X_{i-1}) \varepsilon_i I(X_{i-1} \geq y).
 \end{aligned}$$

Then

$$\begin{aligned}
 n^{1/2} [\widehat{U}_n(y) - U_n(y)] &= n^{-1/2} \sum_{i=1}^n \left[\dot{h}_{\hat{r}}(X_{i-1}) - \dot{h}_r(X_{i-1}) \right] \varepsilon_i I(X_{i-1} \geq y) \\
 &\quad + n^{-1/2} \sum_{i=1}^n \dot{h}_{\hat{r}}(X_{i-1}) \left[\vartheta_1^{*'} \dot{h}_r(X_{i-1}) - \hat{\vartheta}'_1 \dot{h}_{\hat{r}}(X_{i-1}) \right] I(X_{i-1} \geq y) \\
 &= A_n(y) + D_n(y), \quad \text{say.}
 \end{aligned}$$

Fix an $\epsilon > 0$. By (2.2), there is a $b_\epsilon = b < \infty$, $N_\epsilon < \infty$, such that $\mathbb{P}(|\hat{r} - r| \leq$

$b/n) > 1 - \epsilon$, for all $n > N_\epsilon$. Now, on the event $\{|\hat{r} - r| \leq b/n\}$, by (5.4),

$$\begin{aligned} & \sup_{y \in \mathbb{R}^*} \|A_n(y)\| \\ & \leq C n^{-1/2} \sum_{i=1}^n \left[1 + |X_{i-1}|\right] I(|X_{i-1} - r| \leq |\hat{r} - r|) |\varepsilon_i| \\ & \leq C [1 + |r| + b/n] n^{-1/2} \sum_{i=1}^n I(|X_{i-1} - r| \leq b/n) |\varepsilon_i|, \quad \forall n > N_\epsilon. \end{aligned}$$

But the expected value of the second factor in this upper bound is proportional to $n^{1/2}[G(r + b/n) - G(r - b/n)]$, which tends to zero under the conditions of Corollary 2.1. Hence $\sup_{y \in \mathbb{R}^*} \|A_n(y)\| = o_{\mathbb{P}}(1)$.

The term $D_n(y)$ can be written as $D_{n1}(y) - D_{n2}(y)$, where $D_{n1}(y) \equiv n^{-1/2} \sum_{i=1}^n \dot{h}_{\hat{r}}(X_{i-1}) \left[\vartheta_1^{*'}(\dot{h}_r(X_{i-1}) - \dot{h}_{\hat{r}}(X_{i-1})) \right] I(X_{i-1} \geq y)$, $D_{n2}(y) \equiv H_n(y) \Delta_n$. Arguing as above, one obtains $\sup_{y \in \mathbb{R}^*} \|D_{n1}(y)\| = o_{\mathbb{P}}(1)$, while by (2.1) and (5.14), $\sup_{y \in \mathbb{R}^*} \|D_{n2}(y) - H_r(y) \Delta_n\| = o_{\mathbb{P}}(1)$. We have thus proved $n^{1/2} \sup_{y \in \mathbb{R}^*} \|\hat{U}_n(y) - U_n(y) + H_r(y) \Delta_n\| = o_{\mathbb{P}}(1)$. Use this, (5.14), (5.15), and an argument like in the proof of Theorem 2.4 of K-S, to obtain uniformly in $x \leq \tau$,

$$T_{n2}(x) = \int_{-\infty}^x \dot{h}_r H_r^{-1} U_n dG - \int_{-\infty}^x \dot{h}_r dG \Delta_n + o_{\mathbb{P}}(1) = T_{n1}(x) - \Delta_n' J_r(x) + o_{\mathbb{P}}(1),$$

This, together with (5.13), completes the proof of Theorem 3.1.

References

- An, H. Z. and Cheng, B. (1991). A Kolmogorov-Smirnov type statistic with application to test for nonlinearity in time series. *Internat. Statist. Rev.* **59**, 287-307.
- Boldin, M. V. (1982). Estimation of the distribution of noise in autoregression scheme. *Theory Probab. Appl.* **27**, 866-871.
- Chan, K. S. (1993). Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *Ann. Statist.* **21**, 520-533.
- Chan, K. S. and Tsay, R. S. (1998). Limiting properties of the least squares estimator of a continuous threshold autoregressive model. *Biometrika* **85**, 413-426.
- Diebolt, J. and Zuber, J. (1999). Goodness-of-fit tests for nonlinear heteroscedastic regression models. *Statist. Probab. Lett.* **42**, 53-60.
- Dominguez, M. A. and Lobato, I. N. (2003). Testing the martingale difference hypothesis. *Econom. Rev.* **22**, 351-377.
- Hart, J. D. (1997). *Nonparametric Smoothing and Lack-of-Fit Tests*. Springer Series in Statistics. Springer-Verlag, New York.
- Hjellvik, V. and Tjøstheim, D. (1995). Nonparametric test of linearity for time series. *Biometrika* **82**, 351-368.
- Hjellvik, V. and Tjøstheim, D. (1996). Nonparametric statistics for testing of linearity and serial independence. *J. Nonparametric Statist.* **6**, 223-251.
- Khmaladze, E. V. (1981). Martingale approach in the theory of goodness-of-fit tests. *Theory Probab. Appl.* **26**, 240-257.

- Koul, H. L. (2002). *Weighted Empirical Processes in Dynamical Nonlinear Models*. Second edition. Springer Lecture Notes Ser. in Statistics **166**, New York.
- Koul, H. L. and Ossiander, M. (1994). Weak convergence of randomly weighted dependent residual empiricals with applications to autoregression. *Ann. Statist.* **22**, 540-562.
- Koul, H. L. and Stute, W. (1999). Nonparametric model checks for time series. *Ann. Statist.* **27**, 204-236.
- MacKinnon, J. G. (1992). Model specification tests and artificial regressions. *J. Economic Literature* **30**, 102-146.
- McKeague, I. W. and Zhang, M. J. (1994). Identification of nonlinear time series from the first order cumulative characteristics. *Ann. Statist.* **22**, 495-514.
- Qian, L. (1998). On maximum likelihood estimators for a threshold autoregression. *J. Statist. Plann. Inference* **75**, 21-46.
- Stute, W. (1997). Nonparametric model checks for regression. *Ann. Statist.* **25**, 613-641.
- Stute, W., Presedo Quindimil, M., González Manteiga, W. and Koul, H. L. (2004). Model checks of higher order time series. In press.
- Stute, W. and Schumann, G. (1980). A general Glivenko-Cantelli theorem for stationary sequences of random observations. *Scand. J. Statist.* **7**, 102-104.
- Stute, W., Thies, S. and Zhu, L.X. (1998). Model checks for regression: an innovation process approach. *Ann. Statist.* **26**, 1916-1934.
- Stute, W. and Zhu, L. X. (2002). Model checks for generalized linear models. *Scand. J. Statist.* **29**, 535-545.
- Su, J. Q. and Wei, L. J. (1991). A lack-of-fit test for the mean function in a generalized linear model. *J. Amer. Statist. Assoc.* **86**, 420-426.
- Tong, H. (1990). *Non-Linear Time Series: A Dynamical System Approach*. Oxford Univ. Press, Oxford.

Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824-1027, U. S. A.

E-mail: koul@stt.msu.edu

Mathematical Institute, University of Giessen, Arndtstr. 2, D-35392 Giessen, Germany.

E-mail: winfried.stute@math.uni-giessen.de

Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824-1027, U. S. A.

E-mail: fli@math.iupui.edu

(Received August 2002; accepted August 2004)