STATISTICS IN TRANSITION

new series

An International Journal of the Polish Statistical Association

CONTENTS

Volume 10, Number 3, December 2009

ISSN 1234-7655

STATISTICS IN TRANSITION-new series, December 2009
Vol. 10, No. 3, pp. 359—370

NONPARAMETRIC BOOTSTRAP TEST FOR AUTOREGRESSIVE ADDITIVE MODELS

Luca Bagnato[1], Antonio Punzo[2]

ABSTRACT

Additive autoregressive models are commonly used to describe and simplify the behaviour of a nonlinear time series. When the additive structure is chosen, and the model estimated, it is important to evaluate if it is really suitable to describe the observed data since additivity represents a strong assumption. Although literature presents extensive developments on additive autoregressive models, few are the methods to test additivity which are generally applicable. In this paper a procedure for testing additivity in nonlinear time series analysis is provided. The method is based on: Generalized Likelihood Ratio, Volterra expansion and nonparametric conditional bootstrap (Jianqing and Qiwei, 2003). Investigation on performance (in terms of empirical size and power), and comparisons with other additivity tests proposed by Chen *et al.* (1995) are made recurring to Monte Carlo simulations.

Key words: Additive models; Generalized Likelihood Ratio; Volterra expansion; Bootstrap.

1. Introduction

Additive models face the trade-off between the misspecification problem and interpretability. It is not surprising if such tools are commonly used in the statistical applications to simplify data analysis. This wide family of models embodies a key simplifying assumption that, in some scale covariate, effects are separable. Furthermore, additive structures allow us to overcome the so-called problem of the "curse of dimensionality" (Bellman, 1961) by dimension reduction. In detail, additive autoregressive models, applied to time series analysis, assume that conditional expectation function of the dependent variable Y_t can be written as the sum of smooth terms in the lagged variables:

[1] Dipartimento di Metodi Quantitativi per le Scienze Economiche Aziendali Università degli Studi di Milano-Bicocca e-mail: luca.bagnato@unimib.it.
[2] Dipartimento di Impresa, Culture e Società Università di Catania e-mail: *antonio.punzo@unict.it*.

$$E\left[Y_t \mid \left(Y_{t-1},\ldots,Y_{t-p}\right)=\left(x_1,\ldots,x_p\right)\right]=m_1\left(x_1\right)+\cdots+m_p\left(x_p\right). \quad (1)$$

The advantages of using additive models for nonlinear autoregression are based on several reasons. First, they are easier to interpret because they do not involve interactions. Secondly, in many circumstances they can provide adequate approximations for many applications. Thirdly, under the additivity assumption, univariate smoothing techniques can be used directly in nonparametric estimation, resulting into a more comprehensive estimate. In addition, under additivity, the nonlinear contribution of each lagged variable to the response variable can be easily seen; it can be displayed graphically and in some cases can be interpreted. As regard the problem of computing the additive components, many algorithms like, for example, backfitting (Buja *et al.*, 1989) and marginal integration (Linton and Nielsen, 1995), have been provided and improved.

When the additive structure is chosen, and the model estimated, it is important to evaluate if it is really suitable to describe the observed data since, however, additivity represents a strong assumption. The question arises from the misspecification problem which leads to wrong conclusions and erroneous forecasting. The diagnostic checking stage is not merely to determine whether there is evidence of lack of fit but also to suggest ways in which the model may be modified when this is necessary. There are two basic methods for model validation: *overfitting* and *diagnostic checks* applied to the residuals. This paper focuses on the overfitting approach where the model is deliberately overparameterized in a way it is expected to be needed and in a manner such that the entertained model is obtained by setting certain parameters in the more general model at fixed values, usually zero (Box and Pierce, 1970). This traditional approach, mainly based on parametric assumptions, consists of using a large family of parametric models under the alternative hypothesis. The implicit assumption is that the large family of parametric models specifies the form of the true underlying dynamics correctly. However, this is not always warranted and leads naturally to a nonparametric alternative hypothesis. Naturally, the problems increase when also the null hypothesis is nonparametric (additive structure in this specific case).

Although extensive developments on nonparametric estimation techniques, there are few generally applicable methods for testing additivity (see Chen *et al.*, 1995). Our proposed procedure, that comes on top of the procedures proposed by Chen *et al.* (1995), is based on the Generalized Likelihood Ratio (GLR) which is a generally applicable tool for testing parametric hypotheses against nonparametric alternatives. An extension for using such a procedure to the nonparametric (additive) null hypothesis case will be made. Although the GLR method has been developed for independent data, the idea can be applied to time series data. In fact, it is expected that under mixing conditions, the results should also hold for the dependent data (Jianqing and Qiwei, 2003).

The paper is organized as follows: an introduction to the GLR method and nonparametric conditional bootstrap is presented in Section 2; the proposed procedure for testing additivity is described in Section 3 and, in the last section, it is applied to additive and nonadditive models often used in the time series literature.

2. The generalized likelihood ratio

Before introducing GLR it is worth to remember the classic maximum likelihood ratio test which is generally applicable to most parametric hypothesis-testing procedure. The fundamental property that contributes to the success of the maximum likelihood ratio tests is that their asymptotic distributions under the null hypothesis are independent of nuisance parameters. This property was referred to as the "Wilks phenomenon" by Fan *et al*. (2001). Assuming such a property, one can determine the null distribution of the likelihood ratio statistic by using either the asymptotic distribution or the Monte Carlo simulation by setting nuisance parameters at some fitted values. The latter is also referred to as the parametric bootstrap.

The question arises naturally whether the maximum likelihood ratio test is still applicable to the problems with nonparametric models as alternative. First, nonparametric maximum likelihood estimators (MLE) usually do not exist. Even when they exist, they are hard to compute. To mitigate these difficulties, the maximum likelihood estimator under the alternative hypothesis can be replaced by any reasonable nonparametric estimator. This is the essence of the *generalized likelihood ratio*.

Let $\mathbf{f}$ be the vector of functions of main interest and $\boldsymbol{\eta}$ be the vector of nuisance parameters. Suppose that the logarithm of the likelihood of a given set of data is $\ell(\mathbf{f},\boldsymbol{\eta})$. Given $\boldsymbol{\eta}$, a good nonparametric estimator $\hat{\mathbf{f}}_{\boldsymbol{\eta}}$ can be obtained. The nuisance parameters $\boldsymbol{\eta}$ can be estimated by the *profile likelihood* by maximizing $\ell(\hat{\mathbf{f}}_{\boldsymbol{\eta}},\boldsymbol{\eta})$ with respect to $\boldsymbol{\eta}$, resulting in the profile likelihood estimator $\hat{\boldsymbol{\eta}}$. This gives the profile likelihood $\ell(\hat{\mathbf{f}}_{\hat{\boldsymbol{\eta}}},\hat{\boldsymbol{\eta}})$, which is not the maximum likelihood since $\hat{\mathbf{f}}_{\hat{\boldsymbol{\eta}}}$ is not an MLE.

Now, suppose that we are interested in testing whether a parametric family $\mathbf{f}_\theta$ fits a given set of data. Formally, the null hypothesis is

$$H_0 : \mathbf{f} = \mathbf{f}_\theta, \qquad \theta \in \Theta, \tag{2}$$

and we use the nonparametric model $\mathbf{f}$ as alternative. Let $\hat{\theta}_0$ and $\hat{\boldsymbol{\eta}}_0$ be the maximum likelihood estimators under the null model (2) obtained by maximizing the function $\ell(\mathbf{f}_\theta,\boldsymbol{\eta})$. Then $\ell(\mathbf{f}_{\hat{\theta}_0},\hat{\boldsymbol{\eta}}_0)$ is the maximum likelihood under the null hypothesis. The GLR statistic simply compares the log-likelihood under the two competing classes of models:

$$T = \ell\left(\hat{\mathbf{f}}_{\hat{\boldsymbol{\eta}}}, \hat{\boldsymbol{\eta}}\right) - \ell\left(\mathbf{f}_{\hat{\theta}_0}, \hat{\boldsymbol{\eta}}_0\right). \tag{3}$$

Example 2.1 (Univariate nonparametric model) Let $\{(X_i, Y_i)\}_{i=1}^n$ be a sample from the nonparametric model:

$$Y_i = m(X_i) + \varepsilon_i, \qquad i = 1, \ldots, n, \tag{4}$$

where $\{\varepsilon_i\}_{i=1}^n$ are a sequence of *i.i.d.* random variables from $N(0, \sigma^2)$. Consider testing the simple linear regression model:

$$H_0 : m(x) = \beta_0 + \beta_1 x \qquad\qquad H_1 : m(x) \neq \beta_0 + \beta_1 x, \tag{5}$$

with nonparametric alternative model (4). Then, the conditional log-likelihood function given $X_1, \ldots, X_n$ is

$$\ell(m, \sigma) = -n \ln\left(\sqrt{2\pi}\sigma\right) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} [Y_i - m(X_i)]^2. \tag{6}$$

In this specific case results $\mathbf{f}$=m, $\boldsymbol{\eta}$=σ and $\theta = (\beta_0, \beta_1)$. Ultimately θ identifies a particular model contained in the linear model class. For a given σ, let $\hat{m}(\cdot)$ be, for example, the local linear estimator based on the data $\{(X_i, Y_i)\}_{i=1}^n$, which is independent of σ. Substituting it into (6), the following profile likelihood is obtained:

$$\ell(\hat{m}, \sigma) = -n \ln\left(\sqrt{2\pi}\sigma\right) - \frac{1}{2\sigma^2} \mathrm{RSS}_1, \tag{7}$$

where $\mathrm{RSS}_1 = \sum_{i=1}^n [Y_i - \hat{m}(X_i)]^2$. Maximizing (7) with respect to σ, it results that $\hat{\sigma}_1^2 = \mathrm{n}^{-1}\mathrm{RSS}_1$. Hence the profile likelihood is

$$\ell(\hat{m}, \hat{\sigma}) = -\frac{n}{2} \ln\left(\frac{\sqrt{2\pi}\mathrm{RSS}_1}{n}\right) - \frac{n}{2}. \tag{8}$$

Under H_0 the maximum likelihood estimator $\hat{\theta}_0 = \left(\hat{\beta}_0, \hat{\beta}_1\right)$ can be obtained. Then, the profile likelihood under the null hypothesis results

$$\ell\left(m_{\hat{\theta}_0}, \hat{\sigma}_0\right) = -\frac{n}{2} \ln\left(\frac{\sqrt{2\pi}\mathrm{RSS}_0}{n}\right) - \frac{n}{2}, \tag{9}$$

where $\mathrm{RSS}_0 = \sum_{i=1}^{n}\left[Y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 X_i\right)\right]^2$. Using quantities RSS_0 and RSS_1 the statistic (3) can be obtained as follows:

$$T = \frac{n}{2}\ln\left(\frac{\mathrm{RSS}_0}{\mathrm{RSS}_1}\right). \tag{10}$$

This is a GLR test statistic.

As with parametric inference, the GLR test does not have to use the true likelihood. For example, the test statistic T in Example 2.1 applies to problem (5) whether ε_i is normally distributed or not. The normality assumption is simply used to motivate the procedure. Similarly, the GLR statistic does not have to require the MLE under the null hypothesis.

Such considerations suggested using the GLR method for testing additivity. In fact, neither MLE nor error distribution assumptions are made. In the next section it will be shown how to provide the alternative structure to compare with the additive one. In order to do this, the Volterra expansion will be used. Furthermore, to utilize the GLR statistic, the distribution under the null hypothesis needs to be provided. The question arises naturally whether the asymptotic null distribution depends on the nuisance parameter under the null hypothesis, namely, whether the Wilks phenomenon continues to hold for the GLR test. For a number of models and a number of hypotheses, studied by Fan *et al.* (2001), it has been shown that the Wilks type of results continue to hold. Such authors are not able to show that Wilks type of results hold for all problems, but their results indicate that such a phenomenon holds with sufficient generality.

3. A new test for autoregressive additivity

The Additive AutoRegressive (AAR) model is defined as follows:

$$Y_t = m\left(Y_{t-1}, \ldots, Y_{t-p}\right) + \varepsilon_t, \tag{11}$$

with

$$m\left(Y_{t-1}, \ldots, Y_{t-p}\right) = c_m + m_1\left(Y_{t-1}\right) + \cdots + m_p\left(Y_{t-p}\right), \tag{12}$$

where c_m is a constant, m_j, $j=1,\ldots,p$, are univariate unknown functions, and the ε_t are independent and identically distributed (*i.i.d.*) with mean 0 and variance σ^2. Furthermore, ε_t is assumed independent of $\{Y_{t-k}\}_{k\geq 1}$ for any t. To ensure identifiability of the additive component functions m_j, it is assumed $E\left[m_j\left(Y_{t-j}\right)\right] = 0$ for all $j=1,\ldots,p$. The intercept $c_m = E\left(Y_t\right)$ is typically estimated by

$\overline{Y} = \sum_{i=1}^{n} Y_i / n$. Technically, the *i.i.d.* assumption of the errors may be weakened when other theoretical explorations are made. However, as well known, a white noise process is no longer a pertinent building block for nonlinear models, as it is important to look for measures beyond the second moments to characterize the nonlinear dependence structure.

Once estimated the (11) under the additivity assumption (12), the obvious question is whether such a model is appropriate to describe the underlying structure. In order to deal with this model validation, the additive (null) hypothesis

$$H_0 : m\left(Y_{t-1}, \ldots, Y_{t-p}\right) = c_m + \sum_{j=1}^{p} m_j\left(Y_{t-j}\right), \tag{13}$$

will be compared with the (alternative) hypothesis that the conditional mean has one more general autoregressive structure, say

$$H_1 : m\left(Y_{t-1}, \ldots, Y_{t-p}\right). \tag{14}$$

The comparison will be made through the GLR statistic which is utilized here in the most general case, that is, when also the null hypothesis is nonparametric. Obviously, the model validation procedure (overfitting technique) needs to estimate a very general model under the alternative hypothesis. The main problem relates to the definition of a model more general than the additive one but not affected by the "curse of dimensionality". The Volterra expansion allows us to overcome such a difficulty (Chen *et al.*, 1995). In particular, through the Volterra expansion, an autoregressive model can be rewritten in the following way

$$\begin{aligned} Y_t &= m\left(Y_{t-1}, \ldots, Y_{t-p}\right) + \varepsilon_t \\ &= \sum_{u=1}^{p} \phi_u Y_{t-u} + \sum_{u \le v}^{p} \phi_{uv} Y_{t-u} Y_{t-v} + \sum_{u \le v \le w}^{p} \phi_{uvw} Y_{t-u} Y_{t-v} Y_{t-w} + \cdots + \varepsilon_t \\ &= c_m + \sum_{u=1}^{p} m_u\left(Y_{t-u}\right) + \sum_{u \le v}^{p} \phi_{uv} Y_{t-u} Y_{t-v} + \sum_{u \le v \le w}^{p} \phi_{uvw} Y_{t-u} Y_{t-v} Y_{t-w} + \cdots + \varepsilon_t \end{aligned} \tag{15}$$

where

$$m_u\left(Y_{t-u}\right) = \phi_u Y_{t-u} + \phi_{uu} Y_{t-u}^2 + \phi_{uuu} Y_{t-u}^3 + \cdots.$$

Obviously, the case $u = v = w$ is excluded from the summation of the third-order term in (15). It is clear from expression (15) that if the model is additive, then all the coefficients of the higher-order terms in the equation should be zero. Furthermore, defining

$$m_{uv}\left(Y_{t-u} Y_{t-v}\right) = \phi_{uv} Y_{t-u} Y_{t-v} + \phi_{uuvv}\left(Y_{t-u} Y_{t-v}\right)^2 + \phi_{uuuvvv}\left(Y_{t-u} Y_{t-v}\right)^3 + \cdots,$$

the expression (15) can be rewritten as follows:

$$Y_t = c_m + \sum_{u=1}^{p} m_u\left(Y_{t-u}\right) + \sum_{u \le v}^{p} m_{uv}\left(Y_{t-u}Y_{t-v}\right) + \sum_{u \le v \le w}^{p} m_{uvw}\left(Y_{t-u}Y_{t-v}Y_{t-w}\right) + \cdots + \varepsilon_t . \quad (16)$$

Such a result suggests us a simple approximate form to apply for the alternative hypothesis (14). In particular if, for example, the (16) is truncated to the second summation, the hypothesis (14) can be assumed as:

$$H_1 : m\left(Y_{t-1}, \ldots, Y_{t-p}\right) = c_m + \sum_{u=1}^{p} m_u\left(Y_{t-u}\right) + \sum_{u \le v}^{p} m_{uv}\left(Y_{t-u}Y_{t-v}\right). \quad (17)$$

This formulation allows to estimate a model that contains $p + \binom{p}{2}$ univariate additive functions. Although such a test is limited to the first-order cross-product terms, it should have acceptable power against a large class of nonadditive models. In the next Subsection the method for finding the distribution of the GLR statistic will be provided.

3.1. The conditional bootstrap test

The proposed step-procedure for testing additivity, that will be called from now on as Conditional Bootstrap Test, is described in what follows.

I. The two models, respectively under the null and the alternative hypotheses (13) and (14), are estimated and the GLR statistic T is calculated.

II. The *nonparametric conditional bootstrap* (Jianqing and Qiwei, 2003) is applied:

1. Generate the bootstrap residuals $\left\{\varepsilon_t^*\right\}$ of the empirical distribution of the centred residuals $\left\{\hat{\varepsilon}_t - \bar{\hat{\varepsilon}}\right\}$ from the alternative model, where $\bar{\hat{\varepsilon}}$ is the average of $\left\{\hat{\varepsilon}_t\right\}$. An assumption about the distribution error is made; thus, for example, a kernel density estimation can be applied. Construct the bootstrap sample: $Y_{t,1}^* = Y_{t-1}, \ldots, Y_{t,p}^* = Y_{t-p}$ and

$$Y_t^* = c_m + \hat{m}_1\left(Y_{t-1}\right) + \cdots + \hat{m}_p\left(Y_{t-p}\right) + \hat{\varepsilon}_t^*$$

for t=p,…,n.

2. Estimate the additive and the alternative model based on the bootstrap sample:

$$\left\{\left(Y_{t,1}^*, \ldots, Y_{t,p}^*, Y_t^*\right)\right\}_{t=p}^{n}.$$

Calculate the GLR statistic

$$T = \frac{n-p+1}{2}\ln\left(\frac{\text{RSS}_0^*}{\text{RSS}_1^*}\right) \approx \frac{n-p+1}{2}\frac{\text{RSS}_0^* - \text{RSS}_1^*}{\text{RSS}_1^*}.$$

3. Repeat the above two steps B times, and use the empirical distribution of $\{T^*\}$ as an approximation to the distribution of the GLR statistic T under H_0.

III. The estimated p-value of the test is the percentage of $\{T^*\}$ greater than the statistic T provided at point I.

4. **Simulation study**

In order to evaluate the performance of the proposed test, in terms of its empirical size and power, a simulation study is performed. Obtained results are also compared with size and power of three different additivity tests proposed by Chen *et al*. (1995): the *conditional mean test*, the *Lagrange multiplier test*, and the *permutation test*. The first uses the local conditional mean estimator of Truong (1993) and employs a procedure similar to the analysis of variance. The second applies the alternating conditional expectation (ACE) algorithm of Breiman and Friedman (1985) to fit an additive model to the data; additivity is then tested by means of a Lagrange multiplier type test. The third procedure uses the ACE algorithm as well, but it fits permuted residuals to some cross-product terms of the explanatory variables in order to obtain a reference distribution for the test statistic.

In order to make the above-mentioned comparison easier, two subsets of models considered in Chen *et al*. (1995) are used here – one for size considerations and the other for power analysis. The first set consists of the following two additive models:

$$Y_t = 0.8Y_{t-1} - 0.3Y_{t-2} + \varepsilon_t, \tag{18}$$

$$Y_t = 0.5Y_{t-1} - \sin(Y_{t-2}) + \varepsilon_t. \tag{19}$$

These models are used to study the behaviour of the conditional bootstrap test under the null hypothesis of additivity. They represent time series models commonly used in univariate analysis. The linear *AR* (2) model in (18) is chosen to ensure that the proposed test works well for this simple case, while the slightly more complicated model (19), containing a trigonometric sine function at lag 2, is often used in the time series literature to describe periodic series (Lewis and Ray, 1993).

The second set consists of the following two nonadditive models:

$$Y_t = 2\exp\left(-0.1Y_{t-1}^2\right)Y_{t-1} - \exp\left(-0.1Y_{t-1}^2\right)Y_{t-2} + \varepsilon_t, \quad (20)$$

$$Y_t = Y_{t-1}\sin\left(Y_{t-2}\right) + \varepsilon_t. \quad (21)$$

These models are used to study the power of the conditional bootstrap test. In particular, model (21) is a functional-coefficient $AR(1)$ with a sine function of lag 2 (Chen and Tsay, 1993).

Like Chen *et al*. (1995), for each of the models in (18)-(21), we have applied the proposed test to 300 realizations, each with 300 observations. The sample size of 300 or larger is common in nonlinear time series analysis, especially when using nonparametric methods; indeed, it is often difficult to obtain a reliable estimate of the high-dimensional surface when the sample size is small. According to Chen *et al*. (1995), the innovations ε_t are independent $N(0,1)$. In applying the conditional bootstrap test, in order to make faster the procedure, we use a value B=100 and cross-product terms of degree one in the Volterra expansion. For details on the simulation factors used for the other three tests, see Chen *et al*. (1995).

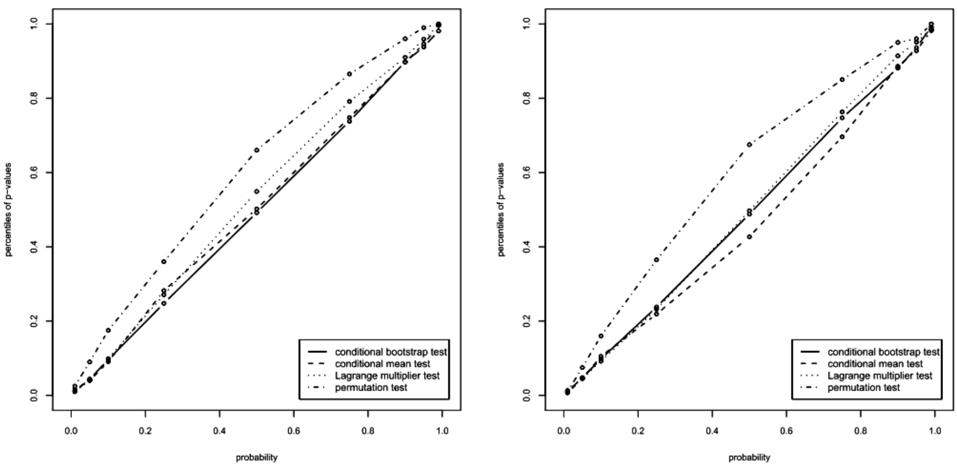
Table 1 shows the (simulated) empirical distribution function of the p-values for models (18) and (19), and for each of the four considered tests.

Table 1. Percentiles of p-values of the nonparametric bootstrap test, under the null hypothesis, in comparison with the tests proposed by Chen *et al.* (1995).

Probability	Conditional bootstrap test		Conditional mean test		Lagrange multiplier test		Permutation test	
	Model (18)	Model (19)	Model (18)	Model (19)	Model (18)	Model (19)	Model (18)	Model (19)
0.01	0.013	0.010	0.016	0.007	0.010	0.013	0.025	0.010
0.05	0.043	0.047	0.040	0.045	0.041	0.046	0.090	0.075
0.10	0.097	0.097	0.091	0.105	0.099	0.092	0.175	0.160
0.25	0.247	0.237	0.282	0.219	0.271	0.232	0.360	0.365
0.50	0.493	0.487	0.507	0.427	0.549	0.497	0.660	0.675
0.75	0.737	0.747	0.748	0.696	0.791	0.763	0.865	0.850
0.90	0.887	0.880	0.897	0.886	0.910	0.914	0.960	0.950
0.95	0.937	0.933	0.946	0.928	0.959	0.951	0.990	0.960
0.99	0.983	0.993	0.996	0.982	0.996	0.986	1.000	1.000

The graphical counterpart of Table 1 is also given in Figure 1; here, models (18) and (19) are separately considered in Figure 1(a) and Figure 1(b), respectively.

Figure 1. Percentiles of *p*-values of the nonparametric bootstrap test, under the null hypothesis, in comparison with the tests proposed by Chen *et al*. (1995).



(a) Model (18) (b) Model (19)

Figure 1 is useful because it simplifies the comparative analysis of performance; indeed, as expected, it is easy to see that the empirical distribution function of the *p*-values, for both models (18) and (19), is always close to the uniform distribution in the unit interval [0,1], which means that the nominal size equals the empirical one. This similarity is strikingly clear for the conditional bootstrap test. In these terms the conditional bootstrap test appears to be one of the best while the permutation test the worst.

Table 2 shows the percentages of rejection by the several tests under different significance levels for models (20) and (21).

Table 2. Percentages of rejection by the conditional bootstrap test, under the alternative hypothesis, in comparison with the tests proposed by Chen *et al*. (1995).

Probability	Conditional bootstrap test		Conditional mean test		Lagrange multiplier test		Permutation test	
	Model (20)	Model (21)	Model (20)	Model (21)	Model (20)	Model (21)	Model (20)	Model (21)
0.010	0.207	1.000	0.403	1.000	1.000	0.990	0.090	1.000
0.050	0.127	1.000	0.267	1.000	0.997	0.990	0.037	1.000
0.100	0.083	1.000	0.117	1.000	0.977	0.967	0.007	1.000

It can be easily noted how all the tests have a good power against the functional-coefficient autoregressive model (21). Unfortunately, above all the permutation test and the conditional bootstrap test do not have a good power against the exponential model in (20). This poor performance of the two tests is understandable because the nonadditivity of the alternative model is in higher-order terms and only the simple cross-product term $Y_{t-1}Y_{t-2}$ has been used in the tests. In practice, it may be helpful to employ several cross-product terms in using these tests.

5. Concluding remarks

In this paper a new procedure for testing additivity, which we defined conditional bootstrap test, has been proposed by means of Generalized Likelihood Ratio, Volterra expansion, and nonparametric conditional bootstrap. This procedure does not require any strong assumption on innovations. A simulated analysis of performance, in terms of empirical size and power, suggests that the conditional bootstrap test is generally reliable under the null hypothesis, even if it may result in low power when the true alternative model is nonadditive in higher-order terms and the cross-product terms considered in the Volterra expansion are small in number. This drawback, also shared by the permutation test proposed in Chen *et al*. (1995), could be however overcome by employing several cross-product terms in the above said expansion (naturally, to the detriment of the computing time).

The paper hints at some further issues; for example, a sort of "rule of thumb" in order to select the suitable number of cross-product terms in the Volterra expansion could be interesting. In detail, future works will be directed to implementing information criterion techniques to select the "correct" model inside a wide family of models resulting from the Volterra expansion.

REFERENCES

BELLMAN, R., 1961. *Adaptive Control Process*. Princeton: Princeton University Press.

BOX, G. E. P. and PIERCE, D. A., 1970. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332), pp. 1509–1526.

BREIMAN, L. and FRIEDMAN, J. H., 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80(391), pp. 580–619.

BUJA, A., HASTIE, T., and TIBSHIRANI, R., 1989. Linear smoothers and additive models. *The Annals of Statistics*, 17(2), pp. 453–510.

CHEN, R. and TSAY, R. S., 1993. Functional-coefficient autoregressive models. *Journal of the American Statistical Association*, 88(421), pp. 298-308.

CHEN, R., LIU, J. S., and TSAY, R. S., 1995. Additivity tests for nonlinear autoregression. *Biometrika*, 82(2), pp. 369–383.

FAN, J., ZHANG, C., and ZHANG, J., 2001. Generalized likelihood ratio statistics and Wilks phenomenon. *The Annals of Statistics*, 29(1), pp. 153–193.

HASTIE, T. and TIBSHIRANI, R. J., 1990. *Generalised additive models*. London: Chapman and Hall.

JIANQING, F. and QIWEI, Y., 2003. *Nonlinear time series: nonparametric and parametric methods*. New York: Springer.

LEWIS, P. and RAY, B., 1993. Nonlinear modelling of multivariate and categorical time series using multivariate adaptive regression splines. In H. Tong, ed. *Dimension Estimation and Models*. Singapore: World Scientific, pp. 136-169.

LINTON, O. and NIELSEN, J. P., 1995. A kernel method of estimating structured nonparametric regression based on marginal integration. *Biometrika*, pp. 82(1), 93–100.

TRUONG, Y., 1993. A nonparametric framework for time series analysis. *New Directions in Time Series Analysis*, pp. 371–386. New York: Springer.