ISTITUTO DI STUDI E ANALISI ECONOMICA

Forecasting Using Functional Coefficients Autoregressive Models

by

Giancarlo Bruno

ISAE, Institute for Studies and Economic Analyses, P.zza dell'Indipendenza, 4
00185 Rome, Italy - Tel: +39-06-4448 2719, fax: +39-06-4448 2249
email: g.bruno@isae.it

Working paper n. 98
June 2008

Stampato presso la sede dell'Istituto

ISAE - Piazza dell'Indipendenza, 4 – 00185 Roma.
Tel. +39-06444821; www.isae.it

ABSTRACT

The use of linear parametric models for forecasting economic time series is widespread among practitioners, in spite of the fact that there is a large evidence of the presence of non-linearities in many of such time series. However, the empirical results stemming from the use of non-linear models are not always as good as expected. This has been sometimes associated to the difficulty in correctly specifying a non-linear parametric model. I this paper I cope with this issue by using a more general non-parametric approach, which can be used both as a preliminary tool for aiding in specifying a suitable parametric model and as an autonomous modelling strategy. The results are promising, in that the non-parametric approach achieve a good forecasting record for a considerable number of series.

Keywords: Non-linear time-series models, non-parametric models.

JEL Classification: C22, C53.

CONTENTS

1 INTRODUCTION

In this paper I explore the usefulness of a class of non-linear non-parametric time series models in producing multi-step ahead forecasts for a set of economic time series. The results are compared with those derived from applying a linear parametric model belonging to the well known class of the autoregressive, moving average (ARMA) models. The main finding is that the use of non-linear non-parametric models can improve the forecasting performance, irrespective of the linearity of the series examined (as measured by some usual tests). Such an improvement results to be substantial in our dataset if one main focus is the median absolute error or the directional error rather than the mean absolute or the mean square error.

The widespread use of linear parametric models, in particular ARMA models, is based on the assumption that a time series can be expressed as the realisation of a stochastic process as the following:

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j} \qquad Z_t \sim IID(0, \sigma^2) \tag{1}$$

with $\sum_{j=0}^{\infty} |\psi_j| < \infty$.

In this case the best mean square predictor is equal to the best linear predictor and process (1) is usually approximated by a so called $ARMA(p, q)$ model:

$$(1 - \varphi_1 L - \ldots \varphi_p L^p) X_t = (1 + \theta_1 L + \cdots + \theta_q L^q) \varepsilon_t \tag{2}$$

where L is the lag operator such that $L^k X_t = X_{t-k}$ and p and q are typically of low order.

It is necessary to stress (Brockwell and Davis, 1991) that the Wold decomposition only insures that any zero-mean covariance stationary process can be expressed in a similar way as in (1) but with $Z_t \sim WN(0, \sigma^2)$; in this case one has to add the hypothesis of gaussianity to the sequence Z_t in order to have that the best linear predictor is the best predictor in mean square sense, otherwise a non-linear model should be used.

Indeed, many time series exhibit non-linear features, such as non-normality, asymmetric cycles, bi-modality, time irreversibility, prediction performance depending on the starting point, etc., thus making the linear hypothesis hard to maintain. Therefore, in such cases it could be appropriate to use more general (non-linear) models to describe those series as well as to forecast them.

An issue with this approach is represented by the large number of non-linear parametric models which can be constructed; moreover, in practical applications, the superior forecasting performance of non-linear models has been hardly observed. An extensive analysis of the forecasting performance of linear *vs.* non-linear parametric models was carried out by Stock and Watson (1998): the authors find that linear autoregressive models with unit root pretesting outperform the non-linear parametric models considered.

In a recent contribution, Teräsvirta (2006) argues that a combination of a large number of non-linear models can obtain point forecasts superior to linear ones. This can be explained by the fact that the presence of non-linear features in actual time series is often coupled with the difficulty in specifying a correct parametric non-linear model.

This last observation can lead to the alternative approach of letting the data specify the unknown non-linear functional form, i.e. using a non-parametric approach. In this paper I will analyse the forecasting performance, for a large number of time series, of the application of a non-parametric, non-linear forecasting model.

The non-linear non-parametric model I consider here is the so called functional coefficients autoregressive model (FCAR), which in practice is an autoregressive model where the coefficients are allowed to vary as a function of a lag of the modelled variable. The non-parametric nature of the model lies in the fact that the functional form of the coefficients is left unspecified. While other papers have examined some aspects of the forecasting performance of FCAR models, such as in Chen and Tsay (1993) and the recent work of Harvill and Ray (2005), nonetheless, to my knowledge, it lacks a more extensive study of the forecasting usefulness of such models with real data sets.

The paper is organised as follows: section 2 presents the model adopted; section 3 reviews the data used; section 4 presents the set-up of the empirical exercise as well as the main results, while the final section concludes and presents some issues for future research.

2 THE FUNCTIONAL COEFFICIENTS AUTOREGRESSIVE (FCAR) MODEL

A very general non-parametric setting for time series modelling can be specified as a non-parametric autoregressive (NAR) model:

$$X_t = g(X_{t-1}, X_{t-2}, \dots, X_{t-p}) + \varepsilon_t, \qquad t = p+1, \dots, T \tag{3}$$

where ε_t is a martingale difference process and $\{X_t, \dots, X_{t-p}\}$ is a strictly stationary β-mixing process.

Estimation of the unknown function $g(\cdot)$ can be carried out e.g. by means of a kernel estimator. More generally a local polynomial approach can be used: in that case a kernel estimator can be seen to be equivalent to a local constant estimator. Efficiency reasons show that in this setting a local linear estimator should be preferred (Fan and Gijbels, 1996).

Nevertheless, the generality of model (3) comes with a cost: the so called *curse of dimensionality*, that is the sample size required for having a performance comparable to the case where $p = 1$ grows exponentially fast (Fan and Yao, 2003, page 317), which in practice means that, for the usual sample sizes observed in economic time series, p can be at most one or two.

Different means have been proposed to overcame this problem, restricting in some way the behaviour of the function $g(\cdot)$ in model (3). A usual manner to accomplish this, for example, is by using so called *additive* models, that is models like the following:

$$X_t = a_1(X_{t-1}) + \dots + a_p(X_{t-p}) + \varepsilon_t \quad t = p+1, \dots, T. \tag{4}$$

In the same spirit other solutions have been proposed, among them the functional coefficient autoregressive (FCAR) model:

$$X_t = a_1(X_{t-d})X_{t-1} + \dots + a_p(X_{t-d})X_{t-p} + \varepsilon_t \quad t = p+1, \dots, T \tag{5}$$

which has been introduced by Chen and Tsay (1993), while Chen and Liu (2001) and Cai et al. (2000) further address the issues of estimation, testing and smoothing parameter selection. This kind of model has some appealing features, in that it nests the usual linear AR model, as well as some popular non-linear parametric models, such as threshold autoregressive (TAR) and exponential autoregressive (EXPAR) models; also SETAR models can be considered as nested in this framework. Moreover, it has a nice interpretation, as the coefficients depend on the "state" of the variable X_{t-d} in a smooth way, differently from what happens in the TAR model, where the autoregressive parameters shift discontinuously following the discrete number of states associated to the variable X_{t-d}. Such a model remains sufficiently general to handle many kinds of non-linearities often found in macroeconomic time series, while reducing considerably the problem of model complexity: the unknown functions, in fact, depend only on one variable in this set-up.

2.1 Estimation

Estimation of model (5) consists in the estimation of the unknown functions $a_i(\cdot)$. Provided suitable conditions on their smoothness, this can be carried out by local averaging techniques, such as kernel estimation or local polynomial estimation; following the efficiency reasons showed in Fan and Gijbels (1996) we will use a local linear estimator which can be shown to be uniformly better than the local constant (kernel) estimator.

Let $U_t = X_{t-d}$. This implies that the following function must be minimized, with respect to $\{a_i, b_i\}$:

$$\sum_{t=p+1}^{T} \left\{ X_t - \sum_{i=1}^{p} \left[a_i + b_i(U_t - u) \right] X_{t-i} \right\}^2 \frac{1}{h} K\left(\frac{U_t - u}{h} \right) \tag{6}$$

where $K(\cdot)$ is a kernel function, h is a smoothing parameter (bandwidth) and u is the point where the regression function is evaluated. The local linear regression estimate of $a_i(\cdot)$ in (5) is then simply $\hat{a}_i(u)$.

Resorting to matrix notation and denoting with $\tilde{\boldsymbol{X}}$ the $2 \times np$ matrix with the t-th row given by:

$$\{X_{t-1}, \ldots, X_{t-p}, X_{t-1}K_h(X_{t-d} - u), \ldots, X_{t-p}K_h(X_{t-d} - u)\},$$

where $K_h = \frac{1}{h}K(\cdot/h)$. Letting $\boldsymbol{Y} = (X_{1+p}, \ldots, X_T)$ and $\boldsymbol{W} = diag\{(K_h(X_{p+1-d}), \ldots, K_h(X_{T-d})\}$, then the problem can written as:

$$\text{argmin}_{\beta}(\boldsymbol{Y} - \tilde{\boldsymbol{X}}\boldsymbol{\beta})'\boldsymbol{W}(\boldsymbol{Y} - \tilde{\boldsymbol{X}}\boldsymbol{\beta}) \tag{7}$$

so that the solution vector is:

$$\hat{\boldsymbol{\beta}} = (\tilde{\boldsymbol{X}}'\boldsymbol{W}\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}'\boldsymbol{W}\boldsymbol{Y} \tag{8}$$

where $\hat{\boldsymbol{\beta}} = (\hat{a}_1, \ldots, \hat{a}_p, \hat{b}_1, \ldots, \hat{b}_p)$

The approach just described treats the smoothing parameter h as a constant over the domain of u. An alternative is represented by the k-nearest neighbour (k-NN) method, where for each value u where the function is evaluated, only the k nearest observations are used, possibly weighted by a kernel function. This is equivalent to specifying a variable bandwidth, which depends on the point u where the function is evaluated. In particular, this amounts to have a larger bandwidth for the intervals of the u domain where observations are less frequent and vice versa.

While the k-nearest neighbour approach could be in theory more suited to the problem of forecasting, which is a "local" problem, its actual effectiveness must be confirmed in practice.

2.2 Model identification

In order to estimate the model (5) it is necessary to set up a procedure to identify the different elements which pertain to the estimation process itself. In particular suitable values for p, d and h (or k for the k-NN method) must be specified. In order to accomplish this task I slightly modify the procedure proposed by Cai et al. (2000), allowing to select a subset of lags between 1 and p. Such a procedure for model identification looks like as follows:

1. First, a maximum value for p is given, depending on the frequency of the time series: daily and monthly series are given a value of $p = 13$, while for quarterly and annual series I consider $p = 5$.

2. A subset of significant lags from the set $\{1, \ldots, p\}$ is then selected. This is done with a non-parametric version of the final prediction error (FPE) criterion which has been proposed by Tschernig and Yang (2000) in the case of the general model (3) (Indeed, the proposed method is valid also in case of heteroskedasticity).

3. Once a subset of $\{1, \ldots, p\}$ has been selected, a form of multifold cross-validation, as proposed by Cai et al. (2000), is used to select both the lag d of the state variable X_{t-d} and the bandwidth h (or, in the case of k-nearest neighbour, the optimal value of k), as specified in the next subsection.

2.2.1 Selection of the bandwidth and of the d parameter

A multi-fold cross-validation procedure was proposed by Cai et al. (2000) to allow the simultaneous choice of p, d and h. Here I use it only for h and d, having already selected a suitable subset of significant lags. The procedure works as follows.

Let us take two positive integers m and Q such that $T > mQ$; the idea is to use q sub-series of length $T - qm$, with $q = 1, 2, \dots, Q$, to estimate the unknown functional coefficients, then use these estimates to produce fitted values for the next m observations.

Let us denote with $\hat{a}_{j,q}$ the estimate of $a(\cdot)_j$ using $T - qm$ data points, we have that for each q the average prediction error is given by:

$$APE_q(h,d) = \frac{1}{m} \sum_{t=T-qm+1}^{T-qm+m} \left[X_t - \sum_{j=1}^{p} \hat{a}_{j,q}(X_{t-d}) X_{t-j} \right]^2 \quad . \tag{9}$$

Moreover, for given h and d define the average forecasting error:

$$APE(h,d) = Q^{-1} \sum_{q=1}^{Q} APE_q(h,d) \quad . \tag{10}$$

The value of h and d are then selected such that (10) is minimized. In the empirical exercise I use the values $Q = 4$ and $m = 0.1T$.

All the previous steps can be repeated in much the same way for the k-NN method, substituting k for h in equations (9) and (10).

3 DATA

The aim of this exercise is to test the forecasting performance of the FCAR model with respect to actual economic time series. For this reason I do not rely on simulated examples, even though these could be important for assessing the forecasting behaviour of such model for a given data generating process (DGP).

Obviously the choice of the testing dataset limits in some respect the generality of the results obtained, but this is an unavoidable limitation in an empirical exercise like this one. I tried to cope with this critical aspect by taking a set of series which are widely used in the time series and forecasting literature and that show a certain degree of heterogeneity as regards to frequency, linearity, and stochastic features in general.

The data used in this paper come mainly from the datasets contained in the software `R` (2007) and in particular in the package `fma` (Hyndman, 2007a); some series come out also from data contained in the packages `tseries` (Trapletti and Hornik, 2006), `forecast` (Hyndman, 2007b) and `mFilter` (Balcilar, 2007). Detailed information about the time series can be found in the Appendix A.

In table 1 we show the time series classified according to their length and frequency; most of the series have between 100 and 200 observations and are recorded at monthly frequency.

Table 1: Time series used by frequency and length

frequency \ length	< 100	101-200	201-300	> 300
annual	1	3		1
quarterly		6	1	
monthly		10	3	4
daily			1	1

All the series were made stationary by differencing (possibly after a log or square root transformation); seasonal differences were always imposed on seasonal series; the need for first difference was tested by means of ADF, PP and KPSS tests: a first difference is imposed when at least two of the aforementioned tests give an indication at 95% confidence level in favour of the presence of a unit root.

Transformed series were tested for linearity: the tests proposed by Hinich (1982), Keenan (1985), Lee et al. (1993) and Teräsvirta et al. (1993) were used. None of these tests propose a specific alternative. In addition, Ljung-Box test on squared residuals of the fitted ARMA model were calculated: in case of linearity the squared residuals should in fact be white noise and departure from this behaviour can be taken as evidence of the presence of non-linearities. Detailed results on the transformation used and on the results of unit root and linearity tests for each series are presented in Appendix B

The natural benchmark against which to compare the forecasting performance of the model considered is the well known ARMA model as in (2). The subset of significant lags of the autoregressive and moving average polynomials have been selected on the basis of the AICC criterion (see Brockwell and Davis, 1991, p. 302), searching

Table 2: Results of linearity tests

number of tests signf. at 5%	series
0	5
1	7
2	9
3	3
4	6
5	1

within a maximum lag of 13 both for the autoregressive and the moving average lag polynomial. A further benchmark is represented by a simple random walk model.

4 EMPIRICAL RESULTS

Each model considered was identified using the first two thirds of observations for each series, while the remaining third was left in order to carry out a true out-of-sample forecasting exercise. For all estimations I used a Gaussian kernel. Appendix C contains some detailed information about the models selected for each series. A recursive scheme was used to get forecasts up to 12 step-ahead for the evaluation period. In this paper when I refer to multi-step forecasts I mean that the s-step-ahead forecast is obtained iteratively, considering as true values the forecasts for the $1, 2, \ldots, s-1$ step-ahead obtained in the previous rounds. Indeed, I am aware this is not the only possible procedure in the framework of non-linear models (Harvill and Ray, 2005, e.g.). Anyway, the results given in the previous reference for FCAR models do not seem to clearly support a particular method, so I use the most widespread one among practitioners. Moreover, other approaches, such as a direct multi-step approach, imply the identification of different models for different forecasting horizons and this could be an issue in the present context for at least two reasons: first, the computational burden is much heavier than the present approach; second, the choice of the state dependent variable becomes more questionable.

The parameters of the ARMA models as well as the bandwidth h (or the k parameter) were re-estimated at each period t. Moreover, a trimming was adopted as in Stock and Watson (1998), that is forecasts with exceptional values were excluded and replaced by a no-change forecast so as to simulate a human intervention on the automatic generated forecasts[1]. Quantitatively this was confined to 76 cases out of 2292 forecasts generated (3.3%); this concerned essentially four series which contain 60 out of the 76 cases considered.

Forecasting performance was evaluated with reference to some usual indicators. In particular, denoting with y_t the true observation of variable y at time t and with $\hat{y}_{st}$ the s-step ahead forecast for variable y at time t, and with $1, \ldots, \tau$ the interval of evaluation, I calculated the following measures:

- mean error (ME): $\frac{1}{\tau}\sum_{t=1}^{\tau}(y_t - \hat{y}_{ts})$;
- mean absolute error (MAE): $\frac{1}{\tau}\sum_{t=1}^{\tau}|y_t - \hat{y}_{ts}|$;
- root mean squareserror (RMSE): $\sqrt{\frac{1}{\tau}\sum_{t=1}^{\tau}(y_t - \hat{y}_{ts})^2}$;
- median error (MedE): $Med\left\{y_t - \hat{y}_{ts}\right\}_{t=1,\ldots,\tau}$;
- median absolute error (MedAE): $Med\left\{|y_t - \hat{y}_{ts}|\right\}_{t=1,\ldots,\tau}$.

In addition to the previous measures, which mainly address the question of how close is the forecast value to the realised one, I use a further evaluation criterion, which is given by the performance of a given model in correctly predicting the direction of change in the time series to be forecast. In fact, it could well be the case that a forecasting model is very good at forecasting a variable producing small errors, while being inaccurate at forecasting the sign of its change (and vice versa). Indeed, in some contexts, a correct sign forecast could be a valuable asset in evaluating the prediction ability. Having said that, I use also the following criterion:

- fraction of corrected directional forecasts: $\frac{1}{\tau}\sum_{t=1}^{\tau}\mathbb{I}_{(y_t - y_{t-1})(\hat{y}_{ts} - y_{t-1})=1}$.

The Diebold-Mariano test was used to assess the significance of differences in forecasting accuracy among the various models. In

[1] Forecasts which produced a change exceeding the maximum observed in the past of the series were excluded.

particular the variant proposed by Harvey et al. (1998) was used. Let us denote with e_{it} the forecasting errors stemming from model i at time t, then when comparing τ forecasts stemming from two competing models i and j the Diebold-Mariano statistics is:

$$DM = \frac{\tau^{-1}\sum_{t=1}^{\tau}[g(e_{it}) - g(e_{jt})]}{\sqrt{\tau^{-1}2\pi \hat{f_d(0)}}} \tag{11}$$

where $f_d(0)$ is a consistent estimate of the spectral density of $\tau^{-1}\sum_{t=1}^{\tau}[g(e_{it}) - g(e_{jt})]$ at frequency 0. The Diebold-Mariano statistics should be confronted with a standard normal distribution. In this paper I consider the function $g(\cdot) = |\cdot|$. In particular I used the variant of the test proposed by Harvey et al. (1998) where the forecasting horizon s is also taken into account:

$$DM^* = \left[\frac{\tau + 1 - 2s + \tau^{-1}s(s-1)}{\tau}\right]^{1/2} DM. \tag{12}$$

The authors propose to compare such a statistic with the Student t distribution with $\tau - 1$ degrees of freedom.

In what follows I try to summarize the main results of the forecasting exercise. First of all, the results presented are relative to the FCAR model with fixed bandwidth and to the ARMA one. Actually the FCAR model with a variable bandwidth (k-NN estimator) resulted always in a poorer forecasting performance than the one with fixed bandwidth. Moreover the naive (random walk) forecast results are also discarded because they are almost always significantly outperformed by all the other methods[2].

Table 3 presents the aggregate results concerning forecasts with horizon from 1 to 6 step-ahead; in particular the percentage of cases where FCAR model outperforms ARMA model are shown. Considering 31 time series with 12 set of forecasts for each time series, we have a total of 186 possible comparisons for horizons 1 to 6 and 186 for horizons 7 to 12. The results are broken down by the degree of non-linearity of the series, considering separately the series for which only 0 or 1 tests rejected linearity (at 95% confidence level), series for which this was true for 2 or 3 tests, and finally series were almost all tests (4–5) rejected linearity.

[2]Obviously, the complete results are available from the author.

Table 3: Forecasting performance at 1 to 6 step-ahead: percentage of cases the FCAR model performs better than the ARMA model model. Breakdown by non-linearity score.

# series	non-linearity[a]	MAE	RMSE	MedAE	directional error
12	0–1	47.2%	34.7%	48.6%	74.6%
12	2–3	48.6%	50.0%	51.4%	94.2%
7	4–5	23.8%	14.3%	35.7%	82.9%
31	all	42.5%	36.0%	46.8%	84.6%

[a] "non-linearity" stands for degree of non-linearity as measured by the number of tests which rejected linearity. So, first row refer to the 12 series which were mainly judged linear because only at most 1 test refused the linearity hypothesis, etc.

What emerges is that a percentage between 36% and 47% of the cases considered, depending upon the criterion chosen, see an improvement in the forecasting performance with the use of FCAR model. The criterion chosen influences the results, with the RMSE favouring more the ARMA model: this is likely to be associated to the presence of a few large errors in some of the forecasts produced with the FCAR model, and is consistent with the better results obtained by the latter comparing the median absolute error.

A very good result for the non-linear model comes out checking the directional error, which almost invariably picks up the FCAR model as the best performing model.

For all the criteria considered there is an improvement in the forecasting performance of the FCAR model passing from the more "linear" series (first row of table 3) to the intermediate ones (second row). On the other hand, there is a drop in the performance with the more non-linear series

Passing to the longest forecasting horizons (7 to 12 step-ahead) the performance of the non-linear model appears to be even better on average, as shown in table 4: in over 44% of cases the FCAR model outperforms the ARMA one, considering the MAE and RMSE criteria, while this percentage rises to 60% when MedAE is considered. On the other hand, there is a deterioration of the performance for the directional error: nevertheless, according to this criterion still more than 60% of the series are best forecast with the FCAR model, irrespective of their linearity.

The significance of the difference in forecasting performance is

Table 4: Forecasting performance at 7 to 12 step-ahead: percentage of cases the FCAR model performs better than the ARMA model model. Breakdown by non-linearity score.

# series	non-linearity[a]	MAE	RMSE	MedAE	directional error
12	0–1	51.4%	55.6%	62.5%	62.2%
12	2–3	50.0%	43.1%	68.1%	61.4%
7	4–5	31.0%	28.6%	42.9%	60.6%
31	all	46.2%	44.6%	60.2%	61.5%

[a]see footnote in table 3

carried out by means of the DM test (12), and a summary of the results is presented in tables 5 and 6.

The results do not show a clear cut pattern, nor with reference to the linearity of the series, neither with reference to the forecasting horizon; actually, when the forecasting horizons 1–6 are considered in half the series (16 out of 31) there are no significant differences among ARMA and FCAR forecasts, while for 7–12 horizons this number rises to 21. Therefore, the informative content of the test appears not to be very high in this context, and it surely deserves some deeper analysis.

Table 5: Test of forecasting accuracy at 1–6 step ahead. Percentage of cases the test is significant at 95%. Breakdown by non-linearity score.

# series	non-linearity[a]	FCAR better than ARMA	ARMA better than FCAR
12	0–1	2.8%	9.7%
12	2–3	20.8%	8.3%
7	4–5	4.8%	11.1%
31	all	10.2%	9.5%

[a]see footnote in table 3.

5 CONCLUDING REMARKS

The purpose of this paper is to evaluate the forecasting ability for real, mainly economic, time series of a non-linear non-parametric model

Table 6: Test of forecasting accuracy at 7–12 step ahead. Percentage of cases the test is significant at 95%. Breakdown by non-linearity score.

# series	non-linearity[a]	FCAR better than ARMA	ARMA better than FCAR
12	0–1	11.1%	1.4%
12	2–3	9.7%	11.1%
7	4–5	2.4%	19.0%
31	all	8.6%	9.1%

[a]see footnote in table 3.

(functional coefficient autoregressive) with that of a classical linear, parametric one (autoregressive, moving average). The comparison was carried out by trying to be as close as possible to a real exercise: models were identified using just a sub-sample of the available observations, while the remaining were used to generate forecasts.

The comparison was carried out over a variety of evaluation criteria. The results are encouraging, in the sense that the forecasting performance of the FCAR model is superior to that of the ARMA one in a non-negligible number of cases. A somewhat bad new is represented by the fact that, while the main motivation for using a FCAR model lies in the non-linear nature of the series at hand, nevertheless, there is no clear connection between the results of some linearity tests and the forecasting improvement obtainable from the FCAR model.

Further research is planned to shed some more light with regard to different aspects. Among others: the link between linearity test diagnostic and the forecasting performance of non-linear models; the use of a linear combination of lagged values of the variable of interest as state variable; the comparison of prediction intervals instead of just the point ones.

REFERENCES

Mehmet Balcilar. *`mFilter`: Miscellaneous time series filters*, 2007. URL `http://www.mbalcilar.net/mFilter`, `http://www.r-project.org`. R package version 0.1-3.

Peter J. Brockwell and Richard A. Davis. *Time Series: Theory and Methods*. Springer, New York, 1991.

Zongw Cai, Jianqing Fan, and Qiwei Yao. Functional-coefficient regression models for nonlinear time series. *Journal of the American Statistical Association*, 95(451):941–956, Sep 2000.

Rong Chen and Lon-Mu Liu. Functional coefficient autoregressive models: Estimation and tests of hypotheses. *Journal of Time Series Analysis*, 22(2):151–173, 2001.

Rong Chen and Ruey S. Tsay. Functional-coefficient autoregressive models. *Journal of the American Statistical Association*, 88(421): 298–308, Mar 1993.

J. Fan and I. Gijbels. *Local Polynomial Modeling and Its Applications*. Chapman and Hall, London, 1996.

Jianqing Fan and Qiwei Yao. *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer, New York, 2003.

D. I. Harvey, S. J. Leybourne, and P. Newbold. Tests for forecast encompassing. *Journal of Business and Economic Statistics*, 16: 254–259, 1998.

Jane L. Harvill and Bonnnie K. Ray. A note on multi-step forecasting with functional coefficient autoregressive models. *International Journal of Forecasting*, 21:717–727, 2005.

M. J. Hinich. testing for gaussianity and linearity of a stationary time-series. *Journal of Time Series Analysis*, 3:169–176, 1982.

Rob J. Hyndman. *`fma`: Data sets from "Forecasting: methods and applications" by Makridakis, Wheelwright & Hyndman (1998)*, 2007a. URL `http://www.robhyndman.info/Rlibrary/forecast/`. R package version 1.09.

Rob J. Hyndman. *`forecast`: Forecasting functions for time series*, 2007b. URL `http://www.robhyndman.info/Rlibrary/forecast/`. R package version 1.09.

D. M. Keenan. A tukey nonadditivity-type test for time-series nonlinearity. *Biometrika*, 72:39–44, 1985.

T.H. Lee, H. White, and C.W.J. Granger. Testing for neglected nonlinearity in time series models. *Journal of Econometrics*, 56: 269–290, 1993.

James H. Stock and Mark W. Watson. A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. NBER Working Papers 6607, National Bureau of Economic Research, Inc, Jun 1998.

R Development Core Team. *`R`: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2007. URL `http://www.R-project.org`. ISBN 3-900051-07-0.

T. Teräsvirta, C.F. Lin, and C.W.J. Granger. Power of the neural network linearity test. *Journal of Time Series Analysis*, 14:209–220, 1993.

Timo Teräsvirta. Forecasting economic variables with nonlinear models. In Graham Elliot, Clive W.J. Granger, and Allan Timmermann, editors, *Handbook of Economic Forecasting*, volume I, chapter 8, pages 413–57. Elsevier, 2006.

Adrian Trapletti and Kurt Hornik. *`tseries`: Time Series Analysis and Computational Finance*, 2006. URL `http://CRAN.R-project.org/`. R package version 0.10-4.

Rolf Tschernig and Lijian Yang. Nonparametric lag selection for time series. *Journal of Time Series Analysis*, 21(4):457–487, 2000.

A DATA - DESCRIPTION

Table 7: Series Description

Name	description
airpass	The classic Box & Jenkins airline data. Monthly totals of international airline passengers (1949–1956).
bricksq	Australian quarterly clay brick production: 1956–1994.
condmilk	Manufacturer's Stocks of evaporated and sweetened condensed milk.
copper	Yearly copper prices, 1800–1997 (in constant 1997 dollars).
dj	Dow-Jones index on 251 trading days ending 26 Aug 1994.
elec	Australian monthly electricity production: Jan 1956 – Aug 1995.
hsales	Monthly sales of new one-family houses sold in the USA since 1973.
huron	Level of Lake Huron in feet (reduced by 570 feet): 1875–1972.
ibmclose	Daily closing IBM stock price.
labour	Number of persons in the civilian labour force in Australia each month (Feb 1978 - Aug 1995).
lynx	Annual number of lynx trapped in McKenzie river district of northwest Canada: 1821–1934.
milk	Average monthly milk production per cow over 14 years.
roomnights	Total room nights occupied at hotel, motel and guest house in Victoria, Australia: Jan 1980 - June 1995.
takings	Total monthly takings from accommodation at hotel, motel and guest house in Victoria, Australia: Jan 1980 - June 1995.
motion	Monthly employment figures for the motion picture industry (SIC Code 78): Jan 1955 – Dec 1970.
oilprice	Oil prices in constant 1997 dollars: 1870–1997.
vehicles	US Sales of motor vehicles and parts:Jan 1971 - Dec 1991.
pigs	Monthly total number of pigs slaughtered in Victoria, Australia (Jan 1980 – Aug 1995).

Table 7: Series Description

Name	description
ukdeaths	Monthly total deaths and serious injuries on UK roads: Jan 1975 – Dec 1984. In February 1983, new legislation came into force requiring seat belts to be worn.
uselec	Monthly total generation of electricity by the U.S. electric industry (Jan 1985 - Oct 1996.
wagesuk	Real daily wages in pound, England: 1260–1994.
writing	Industry sales for printing and writing paper (in thousands of French francs): Jan 1963 – Dec 1972.
itipi	Italian industrial production index (wda): Jan 1981 Feb 2007
rs	discount rate on 91-Day treasury bills quarterly, 1954 – 1987
rl	yield on long-term treasury bonds rl quarterly, 1954 – 1987
M1	seasonally adjusted real U.S. money M1 quarterly, 1954 – 1987
GNP	seasonally adjusted real U.S. GNP in 1982 Dollars quarterly, 1954 – 1987
co2	Atmospheric concentrations of CO2 are expressed in parts per million (ppm) and reported in the preliminary 1997 SIO manometric mole fraction scale. The values for February, March and April of 1964 were missing and have been obtained by interpolating linearly between the values for January and May of 1964.
gas	Australian monthly gas production: 1956–1995.
unemp	Quarterly US unemployment series from 1959.1 to 2000.4.
usgdp	Quarterly real US gdp from 1947.1 to 2006.2

B DATA MAIN FEATURES

Table 8: Results of unit root tests [3]

series	ADF test	PP test	KPSS test
airpass	*	***	**
bricksq	***	***	
condmilk		**	
copper		***	***
dj			***
elec	***	***	**
hsales	*	***	
huron		**	***
ibmclose			***
labour	**	***	
lynx	***	***	
milk		**	**
roomnights	**	***	
takings		***	*
motion	**	**	***
oilprice	*	*	**
vehicles	**	***	
pigs		***	
ukdeaths		***	*
uselec	*	***	
wagesuk			***
writing	*	***	
itipi	***	***	
rs	**		***
rl			***
M1			***
GNP			***
co2	***	***	***
gas	**	***	***
unemp			***

[3] The null hypothesis of the ADF and PP test is the presence of a unit root, while null hypothesis for KPSS test is stationarity. Tests marked with '*' are significant at 10%, those marked with '**' are significant at 5%, the ones with '***' at 1%.

Table 8: Results of unit root tests [3]

series	ADF test	PP test	KPSS test
usgdp			***

Table 9: Series main features

serie	transformation	difference	frequency	length	number of non-linearity tests signif. at 5%
airpass	log	first+seasonal	12	144	1
bricksq	none	seasonal	4	155	2
condmilk	log	seasonal	12	120	2
copper	square-root	first	1	198	1
dj	log	first	365	292	2
elec	square-root	seasonal	12	476	2
hsales	square-root	seasonal	12	275	1
huron	none	first	1	98	0
ibmclose	none	first	365	369	4
labour	square-root	seasonal	12	211	4
lynx	log	none	1	114	0
milk	none	first+seasonal	12	168	2
roomnights	log	seasonal	12	186	2
takings	square-root	seasonal	12	186	2
motion	log	seasonal	12	192	0
oilprice	log	first	1	128	1
vehicles	square-root	seasonal	12	252	3
pigs	none	seasonal	12	188	3
ukdeaths	log	seasonal	12	120	3
uselec	log	seasonal	12	142	4
wagesuk	log	first	1	735	4
writing	none	seasonal	12	120	2
itipi	none	seasonal	12	314	0
rs	square-root	first	4	136	1
rl	log	first	4	136	1
M1	log	first	4	136	4
GNP	square-root	first	4	136	0
co2	log	seasonal	12	468	4
gas	log	seasonal	12	476	5
unemp	log	first	4	168	2
usgdp	log	first	4	238	1

C MODELS

Table 10: Models

series	ARMA order (p,q)	FACR lags selected	FCAR state variable lag
airpass	(0,1)	12	6
bricksq	(4,4)	1	4
condmilk	(1,0)	1 12	11
copper	(1,2)	2 3	2
dj	(0,0)	2 7	6
elec	(1,1)	1 12	12
hsales	(1,0)	1 11 12 13	2
huron	(0,0)	2	3
ibmclose	(0,1)	1 2 10	7
labour	(1,1)	1 2 7 12	1
lynx	(2,3)	1 2 3	3
milk	(2,2)	12 13	12
roomnights	(1,2)	1 11	7
takings	(2,0)	1 12 13	1
motion	(1,0)	1 10 11	6
oilprice	(2,2)	3	2
vehicles	(1,0)	1 2 4 12	12
pigs	(4,1)	1 3 12	7
ukdeaths	(0,0)+intercept	5 12	12
uselec	(1,0)	1 6	12
wagesuk	(0,3)+intercept	1 2	2
writing	(0,0)	4 12	9
itipi	(2,2)	1 2 9 12	3
rs	(1,2)	2	2
rl	(0,1)	1 2	4
M1	(1,0)	1	4
GNP	(1,0)	1 2	1
co2	(2,1)	1 4 12	11
gas	(4,4)	1 12 13	1
unemp	(2,0)	1	2
usgdp	(1,0)	1	3

Working Papers available:		
n. 40/04	R. BASILE S. DE NARDIS	Non linearità e dinamica della dimensione d'impresa in Italia
n. 41/04	G. BRUNO E. OTRANTO	Dating the Italian Business Cycle: a Comparison of Procedures
n. 42/04	C. PAPPALARDO G. PIRAS	Vector-auto-regression Approach to Forecast Italian Imports
n. 43/04	R. DE SANTIS	Has Trade Structure Any Importance in the Transmission of Currency Shocks? An Empirical Application for Central and Eastern European Acceding Countries to EU
n. 44/04	L. DE BENEDICTIS C. VICARELLI	Trade Potentials in Gravity Panel Data Models
n. 47/05	M. MALGARINI P. MARGANI B.M. MARTELLI	Re-engineering the ISAE manufacturing survey
n. 48/05	R. BASILE A. GIUNTA	Things change. Foreign market penetration and firms' behaviour in industrial districts: an empirical analysis
n. 49/05	C. CICCONI	Building smooth indicators nearly free of end-of-sample revisions
n. 50/05	T. CESARONI M. MALGARINI G. ROCCHETTI	L'inchiesta ISAE sugli investimenti delle imprese manifatturiere ed estrattive: aspetti metodologici e risultati
n. 51/05	G. ARBIA G. PIRAS	Convergence in per-capita GDP across European regions using panel data models extended to spatial autocorrelation effects
n. 52/05	L. DE BENEDICTIS R. DE SANTIS C. VICARELLI	Hub-and-Spoke or else? Free trade agreements in the "enlarged" European Union
n. 53/05	R. BASILE M. COSTANTINI S. DESTEFANIS	Unit root and cointegration tests for cross-sectionally correlated panels. Estimating regional production functions
n. 54/05	C. DE LUCIA M. MEACCI	Does job security matter for consumption? An analysis on Italian microdata

Working Papers available:		
n. 55/05	G. ARBIA R. BASILE G. PIRAS	Using Spatial Panel Data in Modelling Regional Growth and Convergence
n. 56/05	E. D'ELIA	Using the results of qualitative surveys in quantitative analysis
n. 57/05	D. ANTONUCCI A. GIRARDI	Structural changes and deviations from the PPP within the Euro Area
n. 58/05	M. MALGARINI P. MARGANI	Psychology, consumer sentiment and household expenditures: a disaggregated analysis
n. 59/05	P. MARGANI R. RICCIUTI	Equivalenza Ricardiana in economia aperta: un'analisi dinamica su dati panel
n. 60/05	M. BOSCHI A. GIRARDI	Euro Area inflation: long-run determinants and short-run dynamics
n. 61/05	M. BOVI	Book-Tax Gap. An Income Horse Race
n. 62/06	M. BOVI	The Cyclical Behavior of Shadow and Regular Employment
n. 63/06	G. BRUNO C. LUPI C. PAPPALARDO G. PIRAS	The cross-country effects of EU holidays on domestic GDP's
n. 64/06	M. COZZOLINO F. DI NICOLA M. RAITANO	Il futuro dei fondi pensione: opportunità e scelte sulla destinazione del TFR
n. 65/06	S. LEPROUX M. MALGARINI	Clima di fiducia e spesa delle famiglie in Italia: un'analisi disaggregata secondo il reddito degli intervistati
n. 66/06	M. BOVI	Consumers Sentiment and Cognitive Macroeconometrics Paradoxes and Explanations
n. 67/06	G. ROCCHETTI	Modelli di business nel mercato del software e partecipazione delle imprese italiane al fenomeno open source
n. 68/06	L. CROSILLA	La stagionalità delle inchieste ISAE su imprese e consumatori: aspetti metodologici ed evidenza empirica

Working Papers available:		
n. 69/06	C. OLDANI	Money demand & futures
n. 70/06	R. BASILE S. DE NARDIS A. GIRARDI	Pricing to market of italian exporting firms
n. 71/06	B.M. MARTELLI G. ROCCHETTII	The ISAE Market Services Survey: Methodological Upgrading, Survey Reliability
n. 72/06	M. FIORAMANTI	Predicting sovereign debt crises using artificial neural networks: a comparative approach
n. 73/06	S. ZECCHINI M. VENTURA	Public Credit Guarantees and SME Finance
n. 74/06	G. FERRI M. VENTURA	Macchinari del *made in Italy* e dinamiche dei distretti industriali
n. 75/07	R. BASILE	Intra-distribution dynamics of regional per-capita income in Europe: evidence from alternative conditional density estimators
n. 76/07	M. BOVI	National Accounts, Fiscal Rules and Fiscal Policy Mind the Hidden Gaps
n. 77/07	L. CROSILLA S. LEPROUX	Leading indicators on construction and retail trade sectors based on ISAE survey data
n. 78/07	R. CERQUETI M. COSTANTINI	Non parametric Fractional Cointegration Analysis
n. 79/07	R. DE SANTIS C. VICARELLI	The “deeper” and the “wider” EU strategies of trade integration
n. 80/07	S. de NARDIS R. DE SANTIS C. VICARELLI	The Euro’s Effects on Trade in a Dynamic Setting
n. 81/07	M. BOVI R. DELL’ANNO	The Changing Nature of the OECD Shadow Economy
n. 82/07	C. DE LUCIA	Did the FED Inflate a Housing Price Bubble? A Cointegration Analysis between the 1980s and the 1990s
n. 83/07	T. CESARONI	Inspecting the cyclical properties of the Italian Manufacturing Business survey data

Working Papers available:

n. 84/07	M. MALGARINI	Inventories and business cycle volatility: an analysis based on ISAE survey data
n. 85/07	D. MARCHESI	The Rule Incentives that Rule Civil Justice
n. 86/07	M. COSTANTINI S. de NARDIS	Estimates of Structural Changes in the Wage Equation: Some Evidence for Italy
n. 87/07	R. BASILE M. MANTUANO	La concentrazione geografica dell'industria in Italia: 1971-2001
n. 88/07	S. de NARDIS R. DE SANTIS C. VICARELLI	The single currency's effects on Eurozone sectoral trade: winners and losers?
n. 89/07	B.M. MARTELLI G. ROCCHETTII	Cyclical features of the ISAE business services series
n. 90/08	M. MALGARINI	Quantitative inflation perceptions and expectations of Italian Consumers
n. 91/08	P. L. SCANDIZZO M. VENTURA	Contingent valuation of natural resources: a case study for Sicily
n. 92/08	F. FULLONE B.M. MARTELLI	Re-thinking the ISAE Consumer Survey Processing Procedure
n. 93/08	M. BOVI P. CLAEYS	Treasury v dodgers. A tale of fiscal consolidation and tax evasion
n. 94/08	R. DI BIASE	Aliquote di imposta sul lavoro dipendente: analisi per figure tipo e con dati campionari
n. 95/08	M. BOVI	The "Psycho-analysis" of Common People's Forecast Errors. Evidence from European Consumer Surveys
n. 96/08	F. BUSATO A. GIRARDI A. ARGENTIERO	Technology and non-technology shocks in a two-sector econ
n. 97/08	A. GIRARDI	The Informational Content of Trades on the EuroMTS Platform