

# The Environmental Long-Memory Space-Time Series Prediction

T. Di Battista<sup>1\*</sup> and G. Visini<sup>2</sup>

<sup>1</sup>Department of Quantitative Methods and Economic Theory, University of Pescara viale Pindaro, 42, Italy

<sup>2</sup>Department of Statistics, University of Messina via dei Verdi, 58, Italy

**ABSTRACT.** Temporal predictions by using long memory time series models have recently attracted much attention. In real surveys, such as those on environmental pollution, the observed data have long memory characteristics. In particular, the time series of pollution usually show “persistence” in the sense that their correlation functions decrease to zero at a much slower rate than the exponential rate implied by a short memory time series. In the literature, several contributions for the multistep prediction of univariate long memory time series have been proposed. However, in pollution studies, the phenomena are generally observed in several points of a study area so that a space-time series is available. In this context, we consider a multivariate extension of the univariate methodology in order to develop a multistep long memory space-time series prediction. For an easier evaluation of the procedure proposed, we focus our attention on bivariate long memory space-time series in ecological framework. The two proposed approaches of multistep prediction of the concentration of carbon monoxide in the Bergamo (Italy)-District are compared.

*Keywords:* AIC criterion, indirect inference, multistep prediction, VAR models, VARFI models

## 1. Introduction

Recently, the large collection of satellite data and environmental monitoring (Byun et al., 2003; Guo et al., 2003; Manobavan et al., 2003) have increased the need of new developments in the field of time and space-time-series. In the standard Box-Jenkins method, it is required to assume that the series are stationary. In nonstationary cases, the first difference will be well behaved, as long as there are no seasonal components (Box and Jenkins, 1994). In particular, it is hoped that difference  $\Delta x_t$ , of the series  $\{x_t\}$  ( $t = 1, 2, \dots$ ), will have rapidly decaying autocorrelations and with any trend behaviour, so that it can be well described by a stationary invertible ARMA model. In practice, many time-series processes are nonstationary as is the case with numerous environmental, biological or economic processes and therefore more general methods need to be developed (Taqqu et al., 2003). In these cases, we may have that the sample autocorrelation decays very slowly and the periodogram is completely dominated by the low frequency components (Giraitis et al., 2001). The first differences are better behaved as the sample autocorrelations decay more rapidly than for the original data. A look at the periodogram shows that most of the power is concentrated near zero frequency. Differencing the series a second time seems to go too far in removing the trend behaviour. The periodogram of  $\Delta^2 x_t$  shows almost no power near zero frequency, indicating that the low frequency components have been eliminated. Thus, data may have been over differenced. It seems that the “right” number of differences is between 1 and 2. The long memory models provide us with a way to

define such a fractional difference. We will say that a stationary time series  $\{x_t\}$  has long memory if there is a nonzero  $d \in (-0.5, 0.5)$  such that the spectral density obeys a power law, i.e.  $f(\lambda) \approx k\lambda^{-2d}$  as  $\lambda \rightarrow 0^+$ . Thus, as  $\lambda \rightarrow 0$ ,  $f(\lambda)$  tends either to  $\infty$  (if  $d > 0$ ) or to zero (if  $d < 0$ ). If  $d = 0$  we say that  $\{x_t\}$  has short memory. In this case,  $f(0)$  will be positive and finite. All stationary invertible ARMA processes are short memory. Another phenomenon indicating the inadequacy of ARMA models for some stationary time series and showing the need of long memory models is the failure of the central limit theorem. If the series  $\{x_t\}$  is *iid* for each  $t$  with finite variance, as it is well known, the sample mean  $\bar{x}$  is asymptotically normal and has a variance proportional to  $1/n$ . It is possible to prove (Fox et al., 1986) that the Central Limit Theorem will still be valid if the observations are not *iid* under the condition that the autocovariances decay rapidly (Giraitis et al., 1990). This is the case of the stationary invertible ARMA processes in which the covariance function decreases exponentially fast to zero. It has been found, however, that certain observed time series although apparently stationary, seem to violate the Central Limit Theorem in that the variance of  $\bar{x}$  seems to go to zero more slowly than  $1/n$ . Mandelbrot and Wallis (1969) found similar behaviour in a variety of geophysical time series including rainfall, earthquake frequencies and sunspot numbers. In this setting, a theoretical analysis shows that if the autocovariances decay to zero slowly enough, then the central limit theorem can indeed fail (Karanasos et al., 2004).

In this paper, the applicability of the multivariate long memory time series in environmental framework has been considered. In particular, the article deals with environmental processes concerning the observation of the concentration of

---

\* Corresponding author: dibattis@dmqte.unich.it

carbon monoxide on two areas of the District of Bergamo. In other words, a spatial temporal process with long memory dependence is available. However, having examined only two spatial location, the spatial temporal process can be reduced to a bivariate temporal process. Thus, we can model the temporal component by using a Vector Autoregressive Fractionally Integrated model (VARFI) that allows us to generate multistep prediction of the process. In a univariate setting, Bhansali and Kokoszka (2001) proposed two different approaches, called type-I and type-II, to predict the long memory time series. In the former, the long memory time series is treated as a standard short memory time series and its multistep predictions are obtained by using the existing modelling approaches. The latter, by contrast, seeks to model the long memory stochastic characteristics of the observed time series by a fractional process such that its fractional difference follows a standard short memory process. In other words, the type I approach does not require the estimation of the long memory parameter while the type II requires this estimation process. In this paper, we propose a multivariate extension of both approaches in order to develop a multi-step long memory time series prediction in environmental surveys. The first one constructs the multi-step prediction of the bivariate long memory time series by a Vector Autoregressive model (VAR) with the order selected by the Akaike Information Criterion (AIC). For the univariate case, a justification of this approach can be found in Crato and Ray (1996). The second one seeks to model the long memory stochastic characteristics of the observed time series by a Vector Autoregressive Fractionally Integrated model (VARFI). The parameters have been estimated by the indirect estimation method as shown in Gourieroux et al. (1993).

The paper is structured as follows: Section 1 provides a brief discussion of the properties of multivariate autoregressive fractionally integrated moving average models (VARFIMA) and indirect estimation methods for estimating VARFIMA models. In section 2 the two forecasting approaches are developed. In section 3, the results of a simulation study designed to compare the two approaches to multi-step prediction are discussed. Finally in section 4 we find an application to environmental data.

## 2. Statistical Models

As said before, our approach is based on two methodologies, the VAR and VARFIMA approaches. In this section, we will focalise our attention on VARFIMA models. In fact, while VAR models are widely known and used in the literature (Hannan, 1970; Lütkepohl, 1993), we believe that a brief description of VARFIMA models to be useful because of their complexity.

The Vector Autoregressive Fractionally Integrated Moving Average model, denoted VARFIMA (p, **d**, q), can be formally expressed as

$$\Phi(L)\nabla^{\mathbf{d}}\mathbf{Y}_t = \Theta(L)\boldsymbol{\varepsilon}_t \quad (1)$$

where  $\mathbf{Y}_t$  is a  $k \times 1$  vector process,  $\nabla^{\mathbf{d}}$  is a  $k \times 1$  fractional differencing operator with  $\mathbf{d} = \{d_1, d_2, \dots, d_k\}$ ,  $\boldsymbol{\varepsilon}_t \sim iid N(0, \Sigma)$  where  $\Sigma$  is a  $k \times k$  matrix of variance-covariance,  $\Phi(L) = \Phi_0 - \sum \Phi_j L^j$  where  $\Phi_j$  is a  $k \times k$  matrix of the autoregressive parameters,  $\Theta(L) = \Theta_0 + \sum \Theta_j L^j$  where  $\Theta_j$  is a  $k \times k$  matrix of moving average parameters, and  $L$  is the lag operator (Baillie, 1996).

We shall call  $\mathbf{Y}_t$  the VARFIMA[ $p, (d_1, d_2, \dots, d_k), q$ ] process where  $p$  is the maximum order of the autoregressive component and  $q$  the maximum order of the moving average component. The univariate version with  $k = 1$  is referred to as the ARFIMA( $p, d_1, q$ ) process which has been extensively studied by Hosking (1981), Granger (1980), Granger and Jøgeaux (1980), and Sowell (1992).

The characteristics of the fractional vector time series  $\mathbf{Y}_t$  can be obtained from the univariate case as developed by Sowell (1989):

- (1)  $\mathbf{Y}_t$  is stationary if  $d_i < 1/2$  for  $i = 1, 2, \dots, k$ ;
- (2)  $\mathbf{Y}_t$  has an invertible moving average representation if  $d_i > -1/2$ ;
- (3) if the spectral density of  $\mathbf{Y}_t$  is denoted by  $f_y(\lambda)$  then as  $\lambda \rightarrow 0$

$$f_y(\lambda) = [k_{ij} \lambda^{-(d_i+d_j)}] \quad (2)$$

where each  $k_{ij}$  is a constant and it is independent from  $d_i$  and  $d_j$ ;

- (4) if the autocovariances of  $\mathbf{Y}_t$  are denoted by  $\gamma_y(s) = E[\mathbf{Y}_t \mathbf{Y}'_{t-s}]$  then as  $s \rightarrow \infty$ ,  $\gamma_y(s) \sim [h_{ij} s^{d_i+d_j-1}]$  where each  $h_{ij}$  is a constant and it is independent from  $d_i$  and  $d_j$ .

The indirect estimator proposed by Martin and Wilkins (1999), in order to estimate the parameters of both ARFIMA and VARFIMA models, is considered. This approach is based on the Indirect Inference criterion proposed by Gourieroux et al. (1993).

An Indirect approach is motivated by the fact that econometric models often lead to complex formulations, which may even be such that it is impossible or extremely difficult to efficiently estimate the parameters of interest because of the intractability of the likelihood function. In this way, the method circumvents the problems associated with the complicated likelihood function arising from VARFIMA models, while it generates consistent and asymptotically normal parameter estimates under fairly general conditions (Martin and Wilkins, 1999).

To make it easier we consider the bivariate Vector Autoregressive Fractionally Integrated models (VARFI) without the moving average component. Implementation of the indirect estimator involves generating a simulated process using the VARFI model for a given set of parameter values. The real data and the simulated data are then used to estimate an auxiliary model (in this case we refer to a VAR model). The two sets of parameter values obtained for the indirect model using the two data sets are then calibrated by choosing

the parameters of the VARFI model.

The main steps of this method are: (1) to choose an auxiliary model and an estimation method of it. In this context we use a VAR(p) model with the order p selected by the Akaike Criterion (AIC), estimated by OLS method; and (2) to simulate a VARFI model for a given set of parameter values. The simulation method that we use in this framework is known as the truncated auto regression method.

Given the model in (1), with

$$\nabla^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} L^j$$

where  $\Gamma(\cdot)$  is the Gamma function, the simulation model can be written as

$$\Phi(L) \left( \sum_{j=0}^l \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} L^j \right) Y_t = \delta + \Theta(L)\omega_t \quad (3)$$

where  $\omega_t \sim iid(0, \mathbf{I})$  and  $l$  is the truncation parameter for the infinite order differencing filter.

To begin the algorithm, it is necessary to choose  $l + p$  starting values for simulating  $Y_t$ . We set these values equal to zero. This approach may be inaccurate if the long memory property of the model takes a long time to forget these values. To overcome this start-up problem we simulate the model for  $T + \tau$  observations and truncate the first  $\tau$  observations.

To compute the indirect estimates of a VARFI(p, d, 0) model with a VAR auxiliary model and the truncated simulator, the algorithm proceeds as follows:

- 1) draw a set of random numbers  $\varpi_t$  from a  $N(0, \mathbf{I})$  distribution;
- 2) choose an initial set of parameter estimates for the VARFI(1,  $\mathbf{d}$ ) model,  $\Psi^{(0)} = \{\mathbf{d}^{(0)}; \Phi_i^{(0)}, i=1, 2, \dots, p\}$  and simulate it;  $d_1$  and  $d_2$  are chosen in the interval of stationary (0 - 0.5) whilst the autoregressive coefficients matrix is selected to assure the process to be stationary; in other words, we impose that the eigenvalues of  $\Phi(L)$  are less than 1 in module;
- 3) simulate the VARFI model; these observations represent the actual data  $Y_t^{ACT}$ ;
- 4) estimate the VAR(p) auxiliary model with the lag length p set to 2;  $\Phi^{ACT}$  is computed;
- 5) draw a second set of random numbers different from the first one, from an  $N(0, \mathbf{I})$  distribution;
- 6) simulate a VARFI(1,  $\mathbf{d}$ ) process which represents now the simulation data  $Y_t^{SIM}$ ;
- 7) estimate a VAR(2) model obtaining  $\Phi_h^{SIM}$  for the  $h$ th simulation path;
- 8) repeat steps 6 and 7, for  $h = 1, 2, \dots, H$  times by using the same set of random numbers each time from step 5; and
- 9) calibrate the parameter vector  $\Psi^{(0)}$  to minimize the

following function:

$$\hat{\Psi} = \underset{\Psi}{\text{Min}} \sum_{h=1}^{niter} (\Phi^{ACT} - \Phi_h^{SIM})^2 \quad (4)$$

### 3. Forecasts for VARFI Models: Two Approaches

In real applications the main purpose of the analysis is to forecast the future values of the time series based on the data collected to the present. As known, the best linear predictors are those that minimize the mean squared error. In this section we compare two distinct approaches for constructing the linear least squares forecasts of unknown future values of long-range dependent process. As mentioned before, the first approach is essentially based on the application of the VAR model. In other words, the vector long memory time series is treated on pair with short memory time series without recognizing its special structure.

However the method requires the choice of a model to approximate the autocorrelation structure of the process. We propose to fit the long memory process by a VAR model with the order selected by the Akaike Criterion (AIC) as a good approximation of a VARFI model. As known, the order of the VAR model can be very high in the presence of a long memory component. In this context, the first step of the analysis is the estimation of the model parameters by the ordinary least squares, which are then used for constructing the multistep forecasts at all lead times.

The first approach, for the univariate case, is justified by Crato and Ray (1996) who demonstrated, by a simulation study, that only a high number of observations and a strong persistence motivate the use of models that take into account the presence of long memory parameters. Starting from the last consideration, in this paper we tried to extend their results to the multivariate case.

On the other hand, the second approach proposed is implemented by explicitly recognising that the observed process has long memory and by postulating that its  $d$ th fractional difference,  $0 < d < 0.5$ , follows a short memory model. This approach requires the choice of the stochastic model to being fitted to the observed time series. In this context, we propose the VARFI model to perform the analysis.

### 4. Simulation Results

The relative predictive performance of the two approaches has been compared on simulated time series. The following five VARFI models were simulated by using the truncated autoregression method (Martin and Wilkins, 1999) with the choice of the initial set of parameters as shown in point 2 of section 2 (see Table 1).

Bivariate time series of length  $N = 600$  from each of these five models were generated with the innovation process  $\epsilon_t$  simulated as a sequence of  $MNV(0, \mathbf{I})$ . For each simulated bivariate long memory time series, the multistep forecasts  $H = 15$  steps.

**Table 1.** Initial Sets of Parameters for Simulated VARFI(1, *d*) Models

VARFI(1, <i>d</i> ) Iteration	Parameters Estimation $\Psi^{(0)} = \{d^{(0)}; \Phi_i^{(0)}, i = 1, 2, \dots, p\}$
1	$\Psi^{(0)} = \{(0.1, 0.2); (-0.296, 1.119; 0.653, 0.282)\}$
2	$\Psi^{(0)} = \{(0.2, 0.2); (0.487, 0.278; 0.259, 1.080)\}$
3	$\Psi^{(0)} = \{(0.3, 0.3); (0.547, 0.568; 0.568, -0.018)\}$
4	$\Psi^{(0)} = \{(0.3, 0.4); (0.912, -0.082; 0.051, -0.711)\}$
5	$\Psi^{(0)} = \{(0.4, 0.4); (0.981, -0.176; 0.474, -0.247)\}$

The final objective of the simulation study has been to compare the relative behaviour of the two approaches for multistep prediction of bivariate long memory time series.

We use the criterion of the proportionate change in the cumulative simulated mean squared errors of prediction up to *H* step ahead ( $H = 1, 2, \dots, 15$ ) for making the comparisons discussed above. A definition of how this measure was computed is given in generic terms for two different approaches of multistep prediction, called Approach A and B respectively (Bhansali and Kokoszka, 2002):

$$PCSMSE(H) = \frac{\{CSMSE_A(H) - CSMSE_B(H)\}}{CSMSE_B(H)} \quad (5)$$

where  $CSMSE_A(H) = \sum SMSE_A(h)$  denotes the cumulative simulated mean squared error of *H*-step prediction for the approach A and

$$SMSE_A(h) = (Niter)^{-1} \sum_{j=1}^{Niter} \left\{ \hat{y}_{jNA}(h) - y_{j,N+h} \right\}^2$$

denotes the simulated *h*-step mean squared error of prediction for Approach A;  $CSMSE_B(H)$  and  $SMSE_B(h)$  are determined through replacing  $\hat{y}_{jNA}(h)$  by corresponding forecast  $\hat{y}_{jNB}(h)$  given by the Approach B. *Niter* represents the number of iterations.

$PCSMSE(H)$  measures the change in the cumulative simulated mean squared error of prediction when using Approach A in preference to the Approach B as a proportion of the cumulative simulated mean squared error of prediction of the Approach B. If for a given value of *H* it is positive, then this signals that the approach B has an advantage over approach A in that the latter has a larger cumulative simulated MSE of prediction up to *H* steps ahead than the former method and conversely a negative value of  $PCSMSE(H)$ .

The simulation results do not unequivocally point to the superiority of one approach over the other. These two approaches give quite similar results for all the simulated mod-

els. For the models 1, 2, 3 and 5 the cumulative simulated mean squared error of prediction for the second approach is larger than the first approach, even if not at all lead times, while the reverse is true for model 4.

In summary, the simulation results do not provide evidence for preferring either the first or the second approach. In any case, for moderate values of *d* ( $d < 0.3$ ), the first approach can represent a valid alternative to the second one. Conversely, for larger values of *d* ( $d \geq 0.3$ ) there is evidence to support the use of the second approach because the order of the VAR model becomes quite high (generally 3 or 4) involving a growth of the number of parameters estimated. This later doesn't respect the parsimonious criterion.

### 5. Application to Environmental Data

The proposed approaches were applied to a space-time series with long memory observed just on two spatial points. The choice of only two stations is motivated to better evaluate the performance of our proposal.

Data concern the daily maximum value of carbon monoxide concentration on two stations of the Bergamo District, which are named Nembro and Seriate. The measurement units are micrograms of Carbon per cubic meter. The raw data were provided by the Environmental Agency (ARPA) of Lombardy of Region Italy. The time-series are composed by 1500 observations covering the period of 1998 to 2001. As we can better see in the comments, data are suitable in order to be analyzed by a long time series approach.

In fact, preliminary analysis shows that the series cannot be considered a realization of a temporal stationary process. To remove heteroschedasticity, a logarithm transformation has been applied. As shown in Figure 1, the variability structure of the two series exhibits a strong seasonal component. As the width of the seasonal component is of order 365 days, we stabilized the data by the difference of order 365 of the logarithm data.

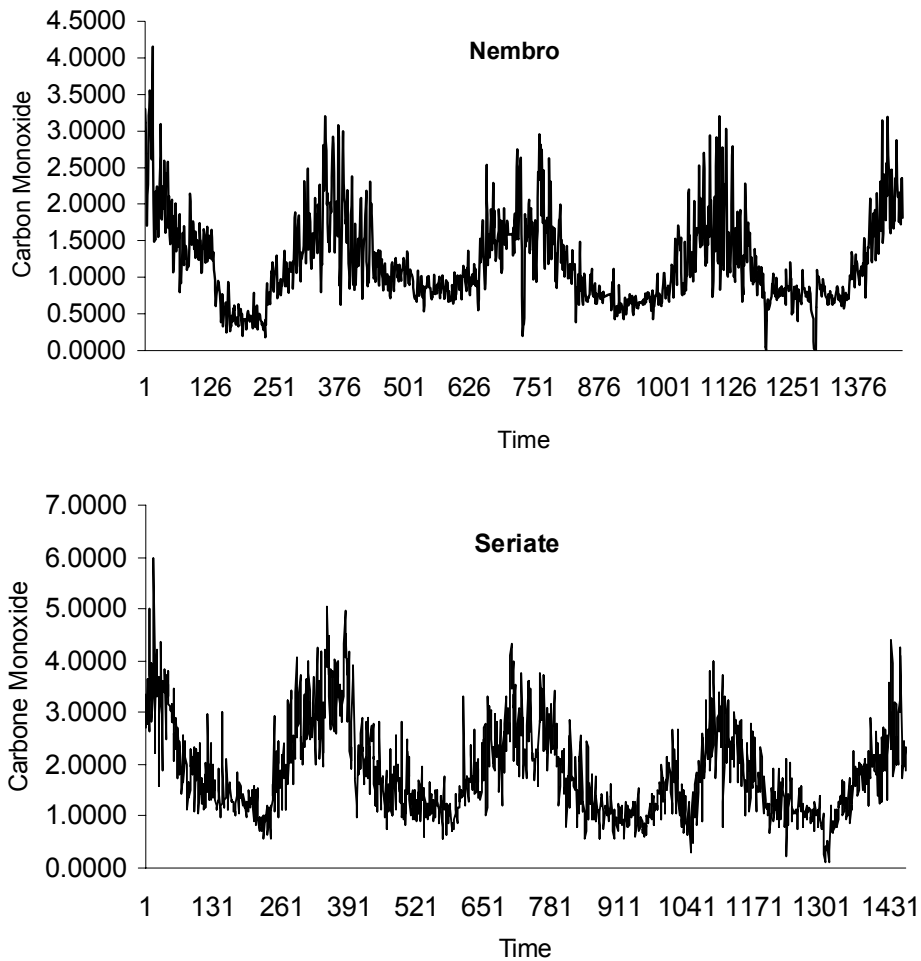
Further analysis, on stationary data, shows the presence of long memory components as confirmed by the autocorrelation function which decreases to zero at a very slow rate. This is confirmed through the Rescaled-range test (R/S), the Lobato-Robinson Test (Lob-Rob) and the Rescaled variance test (V/S) for the null hypothesis of no long memory time series (Beran, 1994). Table 2 shows test results at a significance value of 5%. All tests are generally significant which means that both the time series are generated by a long memory process.

Thus, the described stationary series can be used both for the application of the indirect estimation method and for the application of the proposed multistep prediction method.

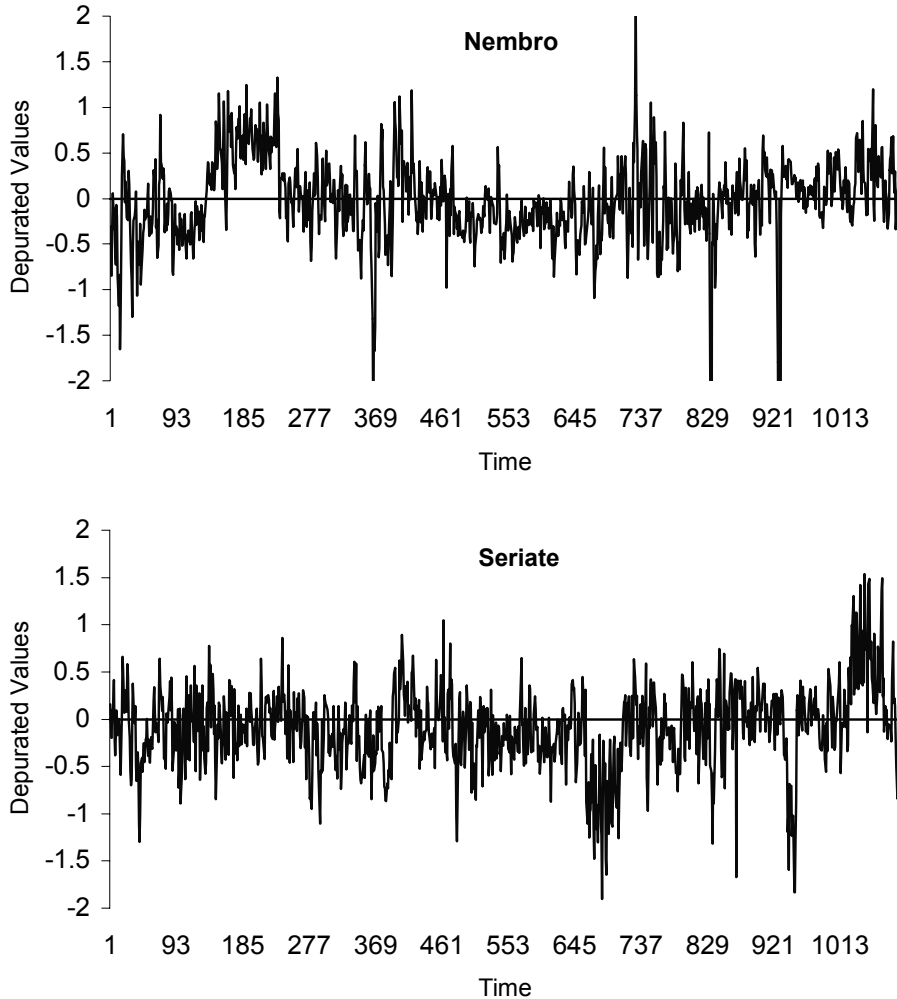
In particular we performed the OLS estimation on the transformed data of a VAR model with order  $p = 5$  selected by the Akaike Criterion (AIC) and the indirect estimation method of the parameters of a VARFI (1, **d**) model (see Table 1).

**Table 2.** Results of Long Memory Test

Truncation lag for correlogram		0	4	8	10	12	26	50	Critical value 5%
V/S statistic	First series	1.091	0.384	0.087	0.129	0.192	0.121	0.0685	0.1869
	Second series	1.106	0.276	0.198	0.177	0.161	0.104	0.074	
R/S statistic	First series	135.651	58.65	96.45	48.623	51.892	31.125	41.562	1.747
	Second series	127.527	63.69	53.947	51.053	48.684	39.098	32.907	
Truncation frequency periodogram		10	15	20	25	30	35	40	Critical value 5%
Lobato-robinson	First series	-0.183	-1.21	-1.625	-1.992	-2.521	-4.589	-4.989	-1.96
	Second series	-0.222	-1.01	-1.845	-2.668	-3.469	-4.441	-5.326	



**Figure 1.** Concentration of carbon monoxide in Nembro and Seriate for 1998 to 2001.



**Figure 2.** Data deputed by the non-stationary in variance and the seasonal component.

From each of the two models we generated a multi-step prediction for  $H = 15$  periods ahead by the two proposed approaches of the multi-step prediction. In order to test the goodness of temporal forecasts, the last 15 observations for each series have been excluded from the analysis. Figure 3 compares observed data and predicted data generated by VAR(5) model and VARFI(1,  $\mathbf{d}$ ) model estimated on the data.

Finally, we compared the behaviour of the two approaches by the *PCSMSE* index (5) introduced in section 3. This index results to be positive for the time series of Nembro. That means that the VARFI approach has an advantage over the VAR approach, while it is negative for the time series of Seriate except for the central steps of prediction (see Figure 4).

This confirms the simulation results in the sense that the VAR model represents a good fit of the correlation structure of a long memory process generated by a VARFI (1,  $\mathbf{d}$ ) model even if the lag order necessary to produce forecasts similar to them generated by a VARFI (1,  $\mathbf{d}$ ) model can be very high if

the fractional component is strong and generally greater than 0.3. In this case, the estimated VAR model is not parsimonious as the high dimension of the parameter vectors that must be estimated.

On the other hand, considering that the VARFI model is computationally more intensive than a VAR model, even with a high order, we can conclude that the univariate results obtained by Bhansali and Kokoszka (2001) can be extended to the multivariate analysis.

## 6. Concluding Remarks

In this paper we have stressed the opportunity to implement the long time series in the environmental framework. Long-range dependence has become a key aspect of time series modelling in a wide variety of areas including hydrology, physics, econometrics and many others fields. Thus, estimations of long-memory models have been considered by a large number of authors.

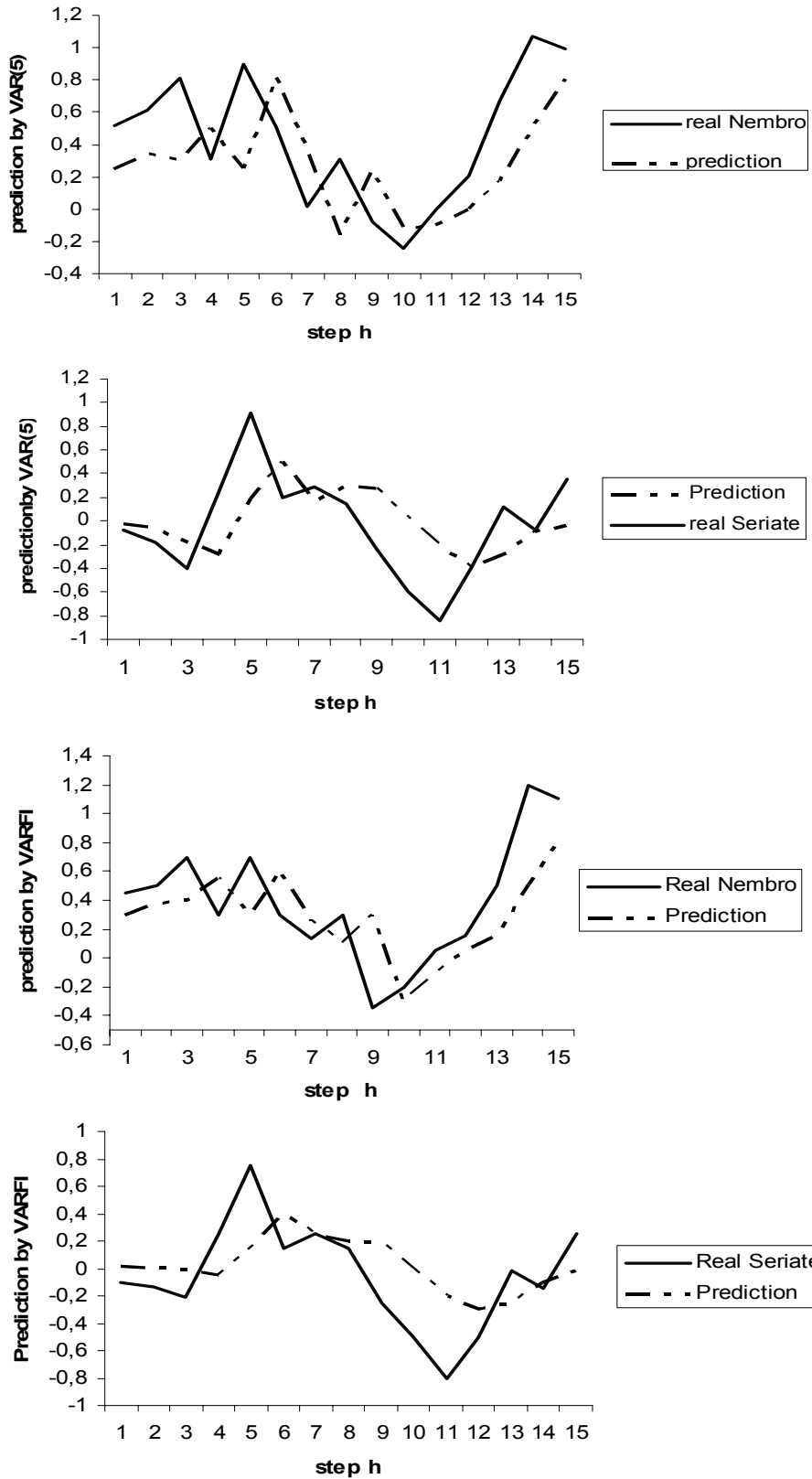
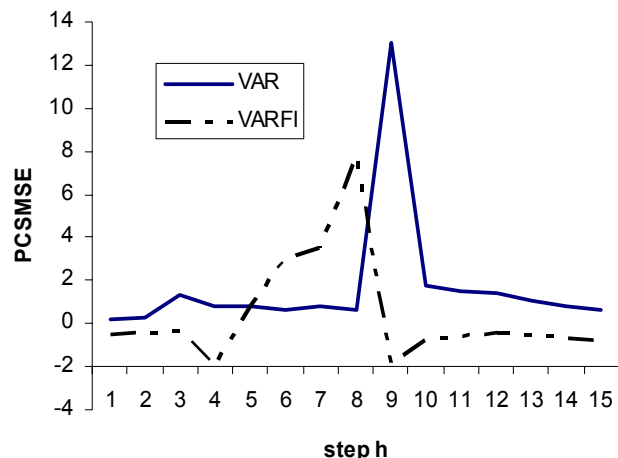


Figure 3. Multistep prediction for H = 15 steps ahead from a VAR(5) model and a VARFI model.



**Figure 4.** Representation of PCSMSE measure.

Unfortunately, in this setting, the recent literature proposed efficient solutions only for univariate time series, leaving unsolved the multivariate case. To solve this shortcoming we have proposed a multivariate extension of the univariate methodology in order to develop a multi-step long memory space-time series prediction. In particular, two different approaches have been considered. Simulations have been performed in order to assess the predictive performance of the two approaches.

Results show a good performance in both cases even if evidence is given in order to choose the method suitable according to the value of differences  $d$  used for nonstationary series. Furthermore, in order to show the applicability of the long memory models on environmental processes, the two proposed approaches have been applied in order to make predictions on the concentration of carbon monoxide observed in the Bergamo District (Italy).

## References

Baillie, R.T. (1996). Long memory processes and fractional integration in econometrics. *J. Econometrics*, 73, 5-59.

Beran, J. (1994). *Statistics for Long-memory Processes*, CRC Press, Florida, USA.

Bhansali, R.J. and Kokoszka, P.S. (2001). Prediction of long-memory time series: An overview. *Estadística*, 53, 41-96.

Bhansali, R.J. and Kokoszka, P.S. (2002). Computation of the forecast coefficients for multistep prediction of long-range dependent time series. *Int. J. Forecast.*, 19, 181-206.

Byun, D.W., Kim, S.T., Cheng, F.Y., Kim, S.B., Cuclis, A. and Moon, N.K. (2003). Information infrastructure for air quality modeling and analysis: Application of the Houston-Galveston ozone non-attainment area. *J. Environ. Inf.*, 2(2), 38-57.

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994). *Time Series Analysis, Forecasting and Control*, 3rd Edition, Englewood Cliffs, Prentice-Hall, NJ, USA.

Crato, N. and Ray, B.K. (1996). Model selection and forecasting of long-range dependent processes: Results of a simulation study. *J. Forecast.*, 15, 107-125.

Fox, R. and Taqqu, M.S. (1986). Large-sample proprieties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Stat.*, 14, 517-532.

Giraitis, L., Hidalgo, J. and Robinson, P.M. (2001). Gaussian estimation of parametric spectral density with unknown pole. *Ann. Stat.*, 86(1), 87-104.

Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate. *Prob. Theory Related Fields*, 29(4) 987-1023.

Gourieroux, C., Monfort, A. and Renault, E. (1993). Indirect inference. *J. Appl. Econometrics*, 8, 85-118.

Granger, C.W.J. (1980). Long memory relationships and the aggregation of dynamic models. *J. Econometrics*, 14, 227-238.

Granger, C.W.J. and Joyeux, R. (1980). An introduction to long memory time series models and fractional differencing. *J. Time Ser. Anal.*, 1, 15-29.

Guo, H.C., Lio, L. and Huang, G.H. (2003). A stochastic water quality forecasting system for the Yiluo river. *J. Environ. Inf.*, 1(2), 18-32.

Hannan, E.J. (1970). *Multiple Time Series*, John Wiley & Sons, New York, USA.

Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*, 68, 165-176.

Karanasos, M., Psaradakis, Z. and Sola, M. (2004). On the autocorrelation properties of long memory GARCH processes. *J. Time Ser. Anal.*, 25(2), 265-281.

Lütkepohl, H. (1993). *Introduction to Multiple Time series Analysis*, Springer-Verlag, Berlin, Germany.

Mandelbrot, B.B. and Wallis, J.R. (1969). Some long-run properties of geophysical records, *Water Resour. Res.*, 5, 422-437.

Manobavan, M., Lucas, N.S., Boyd, D.S. and Petford, N. (2003). The sensitivity and response of terrestrial South American vegetation to interannual climatic variability induced by the ENSO. *J. Environ. Inf.*, 2(2), 1-10.

Martin, V.L. and Wilkins, N.P. (1999). Indirect estimation of AR-FIMA and VARFIMA models. *J. Econometrics*, 93, 149-175.

Sowell, F. (1990). The fractional Unit Root Distribution. *Econometrica*, 58(2), 495-505.

Sowell, F.B. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series analysis. *J. Forecast.*, 13, 109-131.

Taqqu, M.S., Oppenheim, G. and Doukhan, P. (2003). *Theory and Applications of Long Range Dependence*, Birkhauser, Boston, USA.