Computational Statistics (1999) 14:443–467



# Forecasting time series by functional PCA. Discussion of several weighted approaches

Ana M. Aguilera[1], Francisco A. Ocaña[2] and Mariano J. Valderrama[2]

[1] Departamento de Estadística e I.O., Facultad de Ciencias, Universidad de Granada. Campus de Fuentenueva, s/n. 18071-Granada, Spain

[2] Departamento de Estadística e I.O., Facultad de Farmacia, Universidad de Granada. Campus de Cartuja, s/n. 18071-Granada, Spain.

**Summary**

In this paper a functional principal component model is applied to forecast a continuous time series that has been observed only at discrete time points not necessarily equally spaced. To take into account the natural order among the sample paths obtained after cutting the series into pieces, a weighted estimation of the principal components is proposed. In order to estimate the weighted functional principal component analysis, a cubic spline interpolation of the sample paths between their discrete observations is performed. Finally, an application with simulated data is developed where model fitting and forecasting results using different types of weightings on equally and unequally spaced data are given and discussed. The forecasting performance of the estimated functional principal component models is also compared with multivariate principal component regression models.

**Keywords:** Time series; Principal components; Orthogonal expansions; Weighted functional estimation; Interpolating splines.

# 1 INTRODUCTION

This work is motivated by our interest in modelling and forecasting continuous time series where the basic observational unit is a curve which is observed only at discrete time points. As long as most of methodologies in the fields of Statistics and Econometrics have been specially designed for predicting discrete time series, we propose a forecasting approach based on Functional Data Analysis that takes into account the continuous nature of data.

The objective of this paper is to adapt the functional Principal Component Prediction (PCP) model proposed in Aguilera *et al.* (1997) to forecast a continuous time series in a future interval from discrete-time observations in the past. This functional approach allows the forecasting of a time series from unequally spaced time values (i.e. such a situation arises when there are omitted or missing time series values), which is one of its main advantages if we take into account that the literature on this general class of problem, known as Discrete Autoregressive Integrated Moving Average (DARIMA), is actually sparse and complex. Moreover, PCP models do not impose restrictive hypothesis as stationarity on the stochastic process generating the data and allow not only to forecast a continuous time stochastic process in a future interval but also its reconstruction between the discretization time points in the past.

The basic tool to construct the PCP models is the principal component analysis of stochastic processes set up by Deville (1974) as a natural generalization of the classic PCA of a finite set of variables. As the observed data are curves, rather than the vectors of the standard multivariate analysis, this reduction dimension technique is usually known as Functional Principal Component Analysis (FPCA). An interesting perspective on the analysis of functional data can be seen in Ramsay and Dalzell (1991) and in the recent book by Ramsay and Silverman (1997). Other forecasting models for continuous time series that are based on FPCA are the Autoregressive Hilbertian (ARH) models introduced by Bosq (1991). Besse and Cardot (1994) have recently developed a spline approximation for ARH models and a nice application with climatological real data.

The PCP models are obtained as an extension of principal component regression of multiple responses (MPCR) to forecast an infinite set of responses (the process variables in the future) from an infinite set of predictors (the process variables in the past). These models are based on linear regression of the principal components associated to the process in the future against its principal components in the recent past.

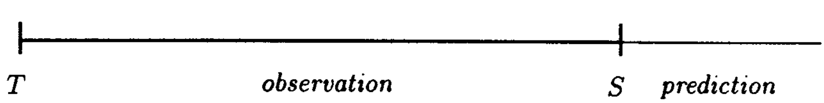
The natural estimators of the principal factors from a set of independent sample paths are the solutions to a second kind integral equation whose kernel is the sample covariance function (Deville 1974). The asymptotic theory on FPCA has been set up by Dauxois *et al.* (1982). In many real situations the data consist of a set of curves which can be observed only in a finite number of times. In this case multivariate PCA of the observed data could give erroneous results, specially with unequally spaced temporary observations (see Castro *et al.* 1986). Because of this the approximation of functional PCA from discrete-time observations of independent sample paths has been studied by many authors such as Deville (1974), Saporta (1985), Besse and Ramsay (1986), Besse (1991), and Aguilera *et al.* (1995), among others. A method for smoothed principal component analysis of functional data is discussed by Rice and Silverman (1991). A simpler alternative approach to FPCA that allows the principal components to be estimated with a single choice of the smoothing parameter is investigated in Silverman (1996), where this methodology is applied to the pinch force of human fingers data analyzed by means of FDA techniques in Ramsay *et al.* (1995).

In the application developed in this paper, the principal factors have been approximated by those of the natural cubic splines interpolating to the simulated sample curves between the observed data. Aguilera *et al.* (1996) have shown on simulated data that the FPCA of cubic spline interpolation of smooth curves corrects the failings of the multivariate PCA with the advantage of giving an optimum reconstruction of a continuous process not only at the discretization time knots but also on the whole interval of time.

In Section 2 we propose to cut the time series in order to obtain a sample of curves for estimating the PCP models. Then, the theoretical framework about PCP model formulation is briefly described. A weighted estimation of the sample covariance kernel that takes into account the natural order among sample curves is proposed in Section 3. Finally, in Section 4, several PCP approaches are used to model and forecast a simulated narrow-band process by using different types of weighting functions. The forecasting performance of the PCP models so obtained is finally compared with MPCR models.

# 2 FORMULATING PCP MODELS

In order to use a PCP model to forecast a time series $\{x(t) : t \geq T\}$ from the time S (S>T),




we propose to cut the series in periods of amplitude h (h>0) (see Figure 1) and define the following process by rescaling:

$$\{X_w(t) = x(wh + t) : t \in [T, T + h); \quad w = 0, 1, \ldots.\}. \tag{1}$$

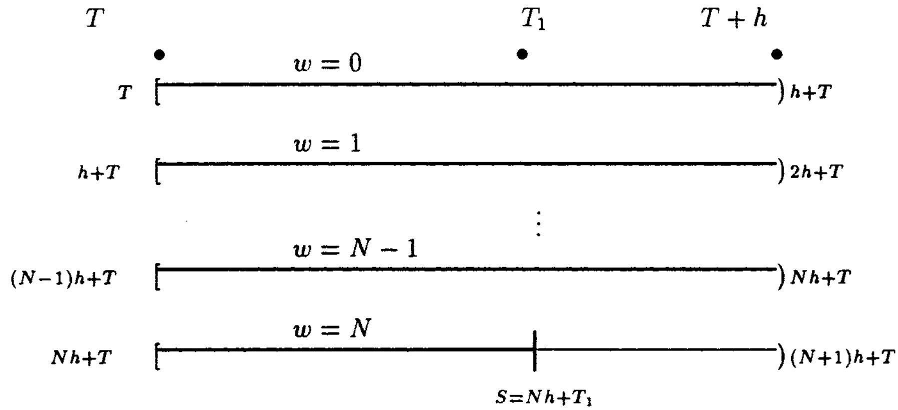



Fig. 1. Sample paths obtained after cutting the original series

Then, we use the PCP models introduced by Aguilera *et al.* (1997) to forecast the process $\{X(s)\}$ in a future interval $[T_2, T+h)$ in terms of its evolution in a past interval $[T, T_1]$ with $T_1 \leq T_2$. Finally, we propose this model for forecasting the original series $x(t)$ in the interval $[Nh+T_2, Nh+(T+h))$ for a large $N \in \mathbb{N}$.

In practice, the choice of the cut-off point is simple enough when there is a well defined seasonal period as in the case of many real time series. On the other hand, the selection of the past and future intervals is determined by the forecasting period that is previously fixed by the design.

Let us now consider a second order and quadratic mean continuous random process, $\{X(t) : t \in [T, T+h)\}$, whose sample functions have squares integrable over $[T, T+h)$.

The main feature of the PCP models is that a double functional PCA is performed, one in a past interval $[T, T_1]$ and the other in a future interval $[T_2, T+h)$ with $T < T_1 \leq T_2 < T+h$ $(h > 0)$.

Deville (1974) defined the *ith* principal component associated to the process $\{X(t)\}$ in the interval $[T, T_1]$ as

$$\xi_i = \int_T^{T_1} (X(t) - \mu(t)) f_i(t) dt,$$

where $f_i$, called the *ith* principal factor, is the normalized eigenfunction corresponding to the *ith* largest eigenvalue $\lambda_i$ of the covariance kernel $C(t, s)$, and $\mu$ is the mean function of the process. That is, $f_i$ and $\lambda_i$ are the solution to the following second kind integral equation:

$$\int_T^{T_1} C(t,s)f(s)ds = \lambda_i f_i(t)\,, \qquad t \in [T, T_1]\,.$$

The principal components defined above have the same optimal properties that in the finite case. That is, $\xi_i$ is the normalized generalized linear combination of the process variables having maximum variance, $\lambda_i$, out of all generalized linear combinations which are uncorrelated with $\{\xi_j\}_{j=1}^{i-1}$. Thus, the variance explained by the *ith* principal component is $V_i^P = \lambda_i / V^P$, with $V^P = \sum_i \lambda_i$ being the total variance of the process in the past interval $[T, T_1]$.

Similarly, let $g_j$ and $\eta_j$ denote the principal factors and components associated to the process $\{X(s)\}$, respectively, in the future interval $[T_2, T+h)$. The total variance of the process in the future is given by $V^F = \sum_j \alpha_j$, where $\alpha_j$ denotes the variance of the principal component $\eta_j$. Therefore, the variance explained by the *jth* principal component $\eta_j$ is given by the ratio $V_j^F = \alpha_j / V^F$.

Then a PCP model for the process in the future (Aguilera *et al.* 1997) is given by

$$\tilde{X}^q(s) \;=\; \mu(s) + \sum_{j=1}^{q} \tilde{\eta}_j^{p_j}\, g_j(s)\,, \qquad s \in [T_2, T+h), \tag{2}$$

where $\tilde{\eta}_j^{p_j}$ is the least-squares linear estimator for the principal component $\eta_j$ against the first $p_j$ principal components in the past. That is,

$$\tilde{\eta}_j^{p_j} \;=\; \sum_{i=1}^{p_j} \frac{E[\eta_j \xi_i]}{\lambda_i}\, \xi_i \;=\; \sum_{i=1}^{p_j} \beta_i^j\, \xi_i. \tag{3}$$

Aguilera *et al.* (1997) have proved that the linear prediction given by the model (2) converges in quadratic mean to the least-squares linear estimator of $X(s)$ given the process variables $\{X(t) : t \in [T, T_1]\}$, for each $s \in [T_2, T+h)$. The PCP model defined by (2) will be denoted as PCP($q; p_1, \ldots, p_q$).

Finally, the total mean-square prediction error for the model (2) is

$$\begin{aligned} \epsilon^2 \;&=\; E\left[\int_{T_2}^{T+h} (X(s) \;-\; \tilde{X}^q(s))^2\, ds\right] \\ &=\; \sum_{j=1}^{q} \epsilon_j^2 + \sum_{l=q+1}^{\infty} \alpha_l \\ &=\; V^F - \sum_{j=1}^{q} \sum_{i=1}^{p_j} \alpha_j\, \rho^2(\eta_j, \xi_i)\,, \end{aligned} \tag{4}$$

where $\epsilon_j^2$ is the mean-square error for the linear model (3) given by

$$\epsilon_j^2 = E[(\eta_j - \tilde{\eta}_j^{p_j})^2] = \alpha_j \left(1 - \sum_{i=1}^{p_j} \rho^2(\eta_j, \xi_i)\right), \qquad (5)$$

and $\rho^2(\eta_j, \xi_i)$ is the square of the linear correlation coefficient between the principal components $\eta_j$ and $\xi_i$. Let us also observed that the quantity $\sum_{j=1}^{q} \sum_{i=1}^{p_j} \alpha_j \rho^2(\eta_j, \xi_i)$ represents the total variability of the linear predictor $\tilde{X}^q$.

In practice, it has been proved that the principal components with the largest variances in the past are not necessarily the best predictors (see Jackson 1991) because the components with the smallest variances could be highly correlated with the principal components in the future. In Aguilera *et al.* (1997) the principal components in the past are entered in each linear model (3) in order of magnitude of the squares of their correlations with the response principal component following a stepwise procedure. In the following section we will propose a new criterion for selecting PCP models that takes into account the principle of parsimony in model building and the proportion of future variance explained by a PCP model defined as

$$\frac{\epsilon^2}{V^F} = 1 - \sum_{j=1}^{q} \sum_{i=1}^{p_j} \frac{\alpha_j}{V^F} \rho^2(\eta_j, \xi_i). \qquad (6)$$

# 3 ESTIMATING WEIGHTED PCP MODELS

Given discrete values $\{x(wh + t_i) : t_i \in [T, T + h); i = 0, 1, \ldots, m; w = 0, \ldots, N-1\}$ and $\{x(Nh + t_i) : i = 0, \ldots, k\}$ $(k = 1, \ldots, m-2)$ of a time series $x(t)$, our objective is to adapt the defined PCP model to forecast the series in the interval $[Nh + t_k, Nh + t_m]$.

By cutting in pieces the observed series we have discrete time observations, $\{X_w(t_i) = x(wh + t_i) : t_i \in [T, T + h)\}$, from N sample paths, $\{X_w(t) : w = 0, \ldots, N-1\}$, of the continuous time process $\{X(t) : t \in [T, T + h)\}$ defined by (1). In this case the past and the future intervals to construct the PCP model are defined by $[T, t_k]$ and $[t_k, T + h)$, respectively.

The weighted estimation of the PCP models is summarized as follows:

1. Estimate the principal factors and components in the past and the future intervals that will be denoted by $(\hat{f}_i, \hat{\xi}_i)$ and $(\hat{g}_j, \hat{\eta}_j)$, respectively.

   In order to consider the chronological order among the sample curves, we propose to weight the sample paths giving higher importance to the recent observations. Then, the weighted estimators of the principal factors are, in each interval, the solutions to the second kind integral equation whose kernel is the following weighted estimation of the covariance function:

$$\hat{C}(t,s) = \frac{N}{(N-1)S_N} \sum_{w=0}^{N-1} P_w \left(X_w(t) - \bar{X}(t)\right)\left(X_w(s) - \bar{X}(s)\right), \quad (7)$$

where $\bar{X}$ is the weighted estimation of the mean defined as

$$\bar{X}(t) = \frac{1}{S_N} \sum_{w=0}^{N-1} P_w X_w(t),$$

and $P_w$ is the weight for the sample-path $w$ with $S_N = \sum_{w=0}^{N-1} P_w$.

On the other hand, as the sample paths are only observed in a finite set of time points in each interval, we propose to replace each of the original sample curves by its natural cubic spline interpolation between the observed data and to approximate the sample principal factors of the original process by means of the eigenfunctions of a weighted estimation of the interpolated process covariance function. The estimation of the sample principal factors of the cubic spline interpolated process in terms of the basis of cubic B-splines has been studied in detail in Aguilera *et al.* (1996).

2. Select jointly the optimum set $\{\eta_j : j \in J\}$ of future principal components to be introduced in the PCP model as response variables, and the optimum set $\{\xi_i : i \in I_j\}$ of past principal components to be considered in the linear regression model (3) as the best predictors for each response principal component $\eta_j$.

   Taking into account the principle of parsimony in model building, in this paper we propose a new criterion for selecting PCP models that is based on the proportion of error reduction (see Equation (6)) for each pair of principal components $(\hat{\eta}_j, \hat{\xi}_i)$ defined by

$$P(j,i) = \frac{\hat{\alpha}_j}{\hat{V}^F} r^2(\hat{\eta}_j, \hat{\xi}_i),$$

   where $\hat{V}^F = \sum_j \hat{\alpha}_j$ is the sample total variance of the process in the future and $r^2(\hat{\eta}_j, \hat{\xi}_i)$ are the squares of the sample linear correlations between the future and the past principal components. Let us observe that $P(j,i)$ is the proportion of future variance explained by a PCP model that considers only the principal components $\hat{\eta}_j$ and $\hat{\xi}_i$.

   This criterion for choosing the *best* PCP model follow three basic steps:

   (a) Select the maximum number $q$ of principal components $\hat{\eta}_j$ to be introduced as response variables in the PCP model. We will used the simplest rule which consists in choosing a cut-off (somewhere between 80 and 99%) and retaining the first $q$ principal components in the future whose percentage of cumulative variance is greater than or equal to this cut-off. Other criterions for deciding how many components should be used in principal component reconstructions can be seen in Jackson (1991).

(b) Estimate the linear correlations $r_{j,i} = r(\hat{\eta}_j, \hat{\xi}_i)$ between each of the first $q$ principal components in the future and each of the principal components in the past and test their statistical significance.

(c) Select those pairs of future-past principal components $(\eta_j, \xi_i)$ with significative linear correlation and add them to the PCP model in order of magnitud of their proportions of error reduction until the relative proportion of future variance (RPV) explained by the PCP model is as close to one as possible (greater or equal to a cut-off previuously fixed as, for example, 0.9).

The relative proportion of future variance explained by a PCP model is defined by the ratio between its explained proportion of future variance and the maximum proportion of future variance explained by a PCP model (MPV) that is given by the sum of proportions of error reduction of those pairs of future-past principal components with significative linear correlation.

3. Estimate the selected PCP$(r, p_1, \ldots, p_r)$ model as

$$\tilde{X}^r(s) = \bar{X}(s) + \sum_{j \in J} \tilde{\hat{\eta}}_j^{p_j} \hat{g}_j(s), \quad s \in [t_{k+1}, t_m], \quad \mathrm{Card}(J) = r, \quad (r \leq q), \tag{8}$$

where, for each principal component $\hat{\eta}_j$ $(j = 1, \ldots, r)$, the parameters of the linear regression model

$$\tilde{\hat{\eta}}_j^{p_j} = \sum_{i \in I_j} \hat{\beta}_i^j \hat{\xi}_i, \quad I_j \subset \mathbb{N}, \quad \mathrm{Card}(I_j) = p_j, \tag{9}$$

are estimated by weighted least-squares linear regression.

4. Finally, estimate the mean-square prediction error by

$$\hat{\epsilon}^2(s) = \frac{N}{(N-1)S_N} \sum_{w=0}^{N-1} P_w \left( IX_w(s) - \tilde{X}^q(s) \right)^2, \quad s \in [t_{k+1}, t_m], \tag{10}$$

where $IX_w(s)$ is the cubic spline interpolating to the sample path $w$ between the observed knots in the future interval.

Moreover, the sample total mean-square prediction error is given by

$$\hat{\epsilon}^2 = \hat{V}^F - \sum_{j \in J} \sum_{i \in I_j} \hat{\lambda}_i (\hat{\beta}_i^j)^2 = \hat{V}^F - \sum_{j \in J} \hat{\alpha}_j \sum_{i \in I_j} r^2(\hat{\eta}_j, \hat{\xi}_i), \tag{11}$$

where $\hat{V}^F = \sum_j \hat{\alpha}_j$ is the sample total variance of the process in the future, $\hat{\lambda}_i$ are the eigenvalues of the weigthed covariance kernel of the cubic spline interpolation of the process in the past, and $r^2(\hat{\eta}_j, \hat{\xi}_i)$ are the squares of the sample linear correlations between the future and the past principal components.

*Choosing the weights*

In order to choose an optimal weighting function we propose to estimate the maximum proportion of future variance explained (maximum proportion of error reduction ) by a PCP model and to select a weight scheme maximizing this quantity with the minimum number of parameters.

*Computational remarks*

To perform the computations involved by PCP model estimation, we have developed the statistical system SMCP$^2$ (see Figure 2) which consists of a set of libraries, coded in Turbo Pascal by using Object Oriented Programming, and two executable programs. On the one hand, PCAP program provides different numerical methods for approximating the sample principal factors and components using discrete-time observations from continuous curves. On the other hand, REGRECOM program performs regression models for the principal components and all the computations related to PCP and MPCR models. The SMCP$^2$ program provides accurate estimation of PCP models in short time for large data set so that a new PCP model for predicting a different time period is easily computed. Lectors interested in this program can get in touch with the authors.

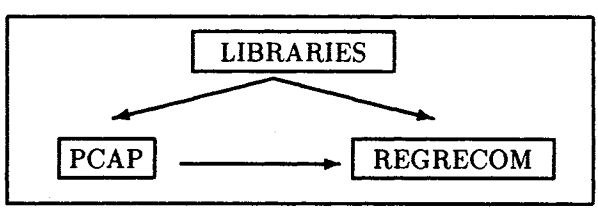



Fig. 2. SMCP$^2$ statistical software

# 4 FORECASTING SIMULATED DATA

In this section, we are now going to consider different kinds of weightings to forecast a simulated narrow-band process.

## 4.1 Simulated data description

Let $I$ be a real interval defined by

$$I \stackrel{def}{=} [0, m\,h) = \bigcup_{i=0}^{m-1} I_i \,,$$

where the $m$ subintervals $I_i$ of amplitud $h$ are a partition of $I$ given by

$$I_i \stackrel{def}{=} [i\,h,\,(i+1)\,h)\,, \qquad \forall i = 0, 1, \ldots, m-1\,.$$

Let us consider the stochastic process $\{O(t) : t \in I\}$ defined on a probability space $(\Omega, \mathcal{A}, P)$ and the interval $I$ as follows:

$$O(t) = \sum_{i=0}^{m-1} \left\{ \sum_{j=0}^{P-1} A_j\, \nu_{j,i} \cos\left( 2\,\pi\, j\, \frac{(t-ih)}{h} + \Theta_j \right) \right\} \chi_i(t)\,,$$

where

$\chi_i$ is the indicator function on the interval $I_i$,

$\{A_j\}_{j=0}^{P-1}$ and $\{\Theta_j\}_{j=0}^{P-1}$ are two sequences of independent real random variables both defined on the probability space $(\Omega, \mathcal{A}, P)$,

$\nu_{j,i}$ is the weight of the *j-th* frecuency on the interval $I_i$,

$P$ is the length of the spectrum.

In order to forecast the process $O$ by using the weighted PCP models described in this paper, we have simulated a realization $o(t)$ of this process in the interval $[0, 250)$ from a partition of $m = 50$ intervals of amplitud $h = 5$, and considered in each interval $I_i$ a number of $P = 30$ frecuencies with weights $v_{j,i}$ defined as

$$\nu_{j,i} = \begin{cases} 0.9^{i-j} & \text{if } j \leq i \\ 0 & \text{if } i < j \end{cases} .$$

Let us observe that in each interval $I_i$ we only consider the frecuencies $j \leq i$ with $\nu_{i,i} = 1$ and $\nu_{j,i} \simeq 0$ if $j << i$. Moreover, the real values $\{A_j\}_{j=0}^{P-1}$ and $\{\Theta_j\}_{j=0}^{P-1}$ corresponding to this realization have been randomly generated by the following probability distributions:

$$\{A_j\}_{j=0}^{P-1} \rightsquigarrow \mathcal{U}[10, 20] \quad \text{and} \quad \{\Theta_j\}_{j=0}^{P-1} \rightsquigarrow \mathcal{U}[0, 2\pi].$$

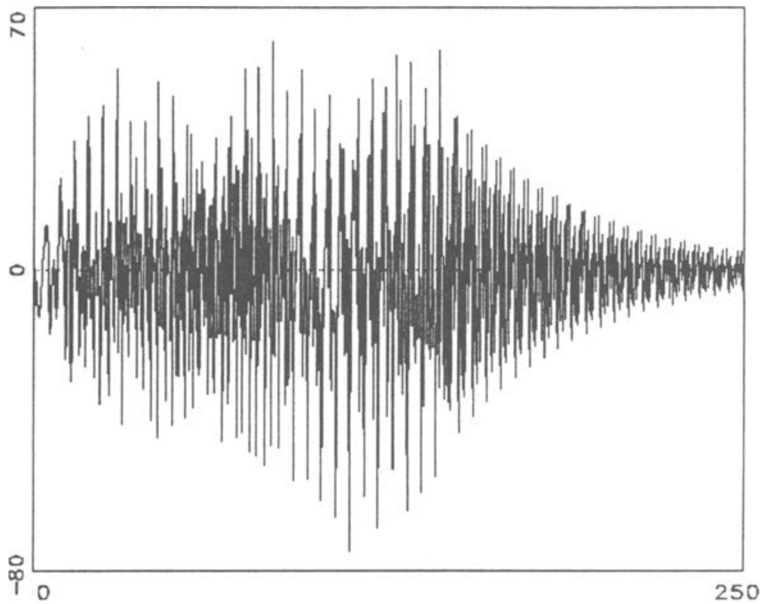



Fig. 3. Simulated time series plot

The result of this simulation is the continuous time series drawn in Figure 3. Taking into account the special definition of this process, we have cut the simulated time series, denoted by $o(t)$, in periods of amplitud $h = 5$, obtaining by rescaling fifty sample paths of a new process $X$ defined as

$$\{X_w(t) = o(5w + t) : t \in [0, 5);\ w = 0, \ldots, 49\}.$$

The first forty-five sample paths have been used for fitting a weighted PCP model for predicting the process $X$ in the future interval $[3, 5)$ from its evolution in the past interval $[0, 3]$ and the final five sample paths (last five periods of amplitud five) will be used to measure the accuracy of the forecasts provided by the adjusted weighted PCP models. Let us observed that this weighted PCP model gives the forecasting of the last five sample-paths $\{X_w : w = 45, 46, 47, 48, 49\}$ in the interval $[3, 5)$ from their discrete values simulated on the interval $[0, 3]$ and, as a consequence, the forecasting of the original realization $o(t)$ in the intervals $[5w + 3, 5(w + 1))$, for each $w = 45, \ldots, 49$.

In order to simulate discrete values of the process $O$ in the interval $[0, 250)$ we have considered in this paper two diferent schemes for the distribution of the discretization time-knots in each interval of amplitud $h = 5$.

## 4.2 Building weighted PCP and MPCR models for equally spaced data

The realization $o(t)$ has been evaluated at one-thousand equally spaced time points in the interval $[0, 250)$. That is, these time knots are given in each interval $I_i$ by $t_k^i = 0.25k + 5i\,(k = 0, \ldots, 19; i = 0, \ldots, 49)$. As a consequence the sample paths of the process $X$ defined by rescaling are evaluated at twenty equally spaced points in the interval $[0, 5)$. In order to estimate a PCP model we have used, for each of the forty-five first sample paths of X, thirteen equally spaced observations in the past interval $[0, 3]$ and eight in the future interval $[3, 5)$, where the observation at time 3 is included in both past and future intervals. That is, we have forty-five replicates of a thirteen dimensional vector as the past and forty five replicates of a eight dimensional vector as the future.

On the one hand, the sample principal factors have been estimated, in each period, by using multivariate PCA of the discrete data matrix. On the other hand, uniform weighting ($P_w = 1$ $(w = 0, \ldots, 44)$); linear weights ($P_w = w$) and exponential weights ($P_w = \exp w$), have been used to estimate the covariance function of the cubic spline interpolated forty-five sample paths.

The percentages of variance explained by the principal components associated to these PCAs appear in Table 1. Let us observe that the variances explained by the principal components in the future are very different when using exponential weighting (only the first principal component explains more than a 99% of the total variability). With the other three principal component approaches, the first six principal components in the future explain more than a 96% of the total variance. This implies that we are going to construct the PCP model with no more than these six principal components as response variables.

In order to use the criterion proposed in the previous section for selecting the *best* PCP model for each type of weights, we have ordered the pairs of future-past principal components with significative linear correlation in terms of their proportions of error reduction. The results appear in Tables 2, 3, 4 and 5 for the multivariate PCA, uniform, linear and exponential weighted FPCA, respectively. The explanation of columnns in this tables is as follows: APV represents the accumulated percentage of error reduction (percentaje of future variance explained by a PCP model with the pairs of principal components entered until each row); RPV is the relative accumulated percentage of error reduction defined as the ratio between the corresponding APV and the MPV (maximum percentage of future variance explained by a PCP model with all pairs of future-past principal components with significative linear correlation); RMSE is the relative total mean squared prediction error obtained as the ratio between the total mean squared prediction error $\hat{\epsilon}^2$ and the number of knots in the future interval in the multivariate case or the amplitud of the future interval in the functional case.

Table 1: Total variances and percentage of variance explained by the principal components for multivariate, unweighted and weighted PCA in the two periods $[0,3]$ and $[3,5)$ in the case of equally spaced data.

[0,3)

| | | | Weighted FPCA | |
|---|---|---|---|---|
| p.c. | Multivariate | Unweighted | Linear | Exponential |
| 1 | 39.8216 % | 44.1334% | 42.7903% | 99.9969% |
| 2 | 24.5861 % | 26.1528% | 33.5323% | 2.3E-3% |
| 3 | 12.7267 % | 9.1084% | 7.8960% | 4.0E-4% |
| 4 | 6.5469 % | 6.7669% | 4.9249% | 3.1E-4% |
| 5 | 4.4427 % | 3.8634% | 2.8333% | 7.0E-5% |
| 6 | 3.2068 % | 2.8222% | 2.4065% | 2.0E-5% |
| 7 | 2.3491 % | 2.0743% | 1.6895% | 1.0E-5% |
| 8 | 2.1582 % | 1.6071% | 1.2753% | 0.0000% |
| 9 | 1.5001 % | 1.1429% | 0.9744% | 0.0000% |
| 10 | 1.1350 % | 0.9784% | 0.8368% | 0.0000% |
| 11 | 0.6356 % | 0.6742% | 0.4047% | 0.0000% |
| 12 | 0.5209 % | 0.4482% | 0.2773% | 0.0000% |
| 13 | 0.3702 % | 0.2278% | 0.1588% | 0.0000% |
| $\hat{V}$ | 2880.4225 | 593.2246 | 488.1441 | 1.1668 |

[3,5)

| | | | Weighted FPCA | |
|---|---|---|---|---|
| p.c. | Multivariate | Unweighted | Linear | Exponential |
| 1 | 32.7616% | 37.5866% | 48.7817% | 99.9967% |
| 2 | 21.9684% | 23.4455% | 20.0786% | 2.7E-3% |
| 3 | 18.3058% | 15.3254% | 12.4656% | 4.2E-4% |
| 4 | 13.5720% | 12.8143% | 10.4475% | 9.0E-5% |
| 5 | 6.5253% | 4.5259% | 3.5305% | 4.0E-5% |
| 6 | 3.5111% | 3.8347% | 2.4242% | 1.0E-5% |
| 7 | 2.2542% | 1.8755% | 1.7106% | 0.0000% |
| 8 | 1.1017% | 0.5923% | 0.5614% | 0.0000% |
| $\hat{V}$ | 1697.3036 | 312.8011 | 242.3424 | 0.6428 |

Table 2: Pairs of future-past principal components ordered in terms of their percentage of error reduction for the multivariate estimation of PCA in the case of equally spaced data.

| Multivariate PCA | | | | | | |
|---|---|---|---|---|---|---|
| Pair | f.p.c. | p.p.c. | APV | RPV | RMSE | t-value |
| 1 | 1 | 2 | 20.218 | 29.323 | 13.0103 | 8.325 |
| 2 | 2 | 1 | 34.461 | 49.981 | 11.7919 | 8.904 |
| 3 | 1 | 4 | 44.527 | 64.580 | 10.8486 | 7.248 |
| 4 | 3 | 3 | 48.578 | 70.455 | 10.4450 | 2.463 |
| 5 | 4 | 7 | 52.287 | 75.834 | 10.0613 | 4.021 |
| 6 | 3 | 4 | 55.655 | 80.719 | 9.6996 | 3.114 |
| 7 | 4 | 3 | 58.770 | 85.237 | 9.3528 | 3.579 |
| 8 | 3 | 8 | 61.727 | 89.526 | 9.0111 | 2.586 |
| 9 | 3 | 1 | 63.689 | 92.371 | 8.7771 | 2.272 |
| 10 | 5 | 10 | 64.702 | 93.841 | 8.6538 | 2.812 |
| 11 | 6 | 11 | 65.666 | 95.239 | 8.5348 | 2.730 |
| 12 | 5 | 5 | 66.627 | 96.632 | 8.4146 | 2.725 |
| 13 | 5 | 13 | 67.372 | 97.713 | 8.3201 | 3.403 |
| 14 | 6 | 8 | 68.067 | 98.721 | 8.2310 | 2.264 |
| 15 | 6 | 9 | 68.579 | 99.463 | 8.1648 | 2.708 |
| 16 | 6 | 4 | 68.949 | 100 | 8.1166 | 2.252 |

Table 3: Pairs of future-past principal components ordered in terms of their percentage of error reduction for the unweighted estimation of FPCA in the case of equally spaced data.

| Uniform FPCA | | | | | | |
|---|---|---|---|---|---|---|
| Pair | f.p.c. | p.p.c. | APV | RPV | RMSE | t-value |
| 1 | 1 | 2 | 22.433 | 32.976 | 11.7748 | 7.978 |
| 2 | 2 | 1 | 38.525 | 56.632 | 10.4824 | 9.701 |
| 3 | 1 | 4 | 45.124 | 66.332 | 9.9039 | 3.026 |
| 4 | 3 | 3 | 50.285 | 73.919 | 9.4267 | 4.673 |
| 5 | 4 | 7 | 54.617 | 80.287 | 9.0066 | 4.686 |
| 6 | 3 | 4 | 58.462 | 85.939 | 8.6166 | 3.795 |
| 7 | 4 | 3 | 61.334 | 90.160 | 8.3134 | 3.524 |
| 8 | 3 | 8 | 62.781 | 92.288 | 8.1564 | 2.118 |
| 9 | 3 | 1 | 64.161 | 94.316 | 8.0038 | 2.062 |
| 10 | 5 | 10 | 65.136 | 95.750 | 7.8941 | 3.438 |
| 11 | 6 | 11 | 65.843 | 96.789 | 7.8137 | 3.117 |
| 12 | 5 | 5 | 66.420 | 97.637 | 7.7474 | 2.507 |
| 13 | 5 | 13 | 66.910 | 98.358 | 7.6907 | 2.285 |
| 14 | 6 | 8 | 67.323 | 98.964 | 7.6425 | 2.278 |
| 15 | 6 | 9 | 67.689 | 99.503 | 7.5996 | 2.131 |
| 16 | 6 | 4 | 68.027 | 100 | 7.5597 | 2.039 |

Table 4: Pairs of future-past principal components ordered in terms of their percentage of error reduction for the linear weighted estimation of FPCA in the case of equally spaced data.

| Linear FPCA | | | | | | |
|---|---|---|---|---|---|---|
| Pair | f.p.c. | p.p.c. | APV | RPV | RMSE | t-value |
| 1 | 1 | 2 | 35.455 | 45.591 | 9.4543 | 10.696 |
| 2 | 2 | 1 | 51.097 | 65.705 | 8.2293 | 12.313 |
| 3 | 1 | 3 | 59.028 | 75.903 | 7.5325 | 2.889 |
| 4 | 4 | 5 | 63.135 | 81.185 | 7.1450 | 5.278 |
| 5 | 3 | 4 | 66.564 | 85.595 | 6.8046 | 4.039 |
| 6 | 3 | 3 | 69.549 | 89.433 | 6.4937 | 3.679 |
| 7 | 4 | 7 | 71.087 | 91.410 | 6.3277 | 2.724 |
| 8 | 4 | 3 | 72.405 | 93.105 | 6.1818 | 2.491 |
| 9 | 4 | 4 | 73.586 | 94.624 | 6.0480 | 2.341 |
| 10 | 3 | 2 | 74.739 | 96.107 | 5.9145 | 2.094 |
| 11 | 5 | 10 | 75.878 | 97.571 | 5.7797 | 4.524 |
| 12 | 6 | 9 | 76.573 | 98.465 | 5.6957 | 4.160 |
| 13 | 6 | 11 | 77.055 | 99.085 | 5.6368 | 3.267 |
| 14 | 5 | 13 | 77.424 | 99.559 | 5.5914 | 2.239 |
| 15 | 5 | 3 | 77.767 | 100 | 5.5488 | 2.150 |

Table 5: Pairs of future-past principal components ordered in terms of their percentage of error reduction for the exponential weighted estimation of FPCA in the case of equally spaced data.

| Exponential FPCA | | | | | | |
|---|---|---|---|---|---|---|
| Pair | f.p.c. | p.p.c. | APV | RPV | RMSE | t-value |
| 1 | 1 | 1 | 99.996 | 99.997 | 0.0039 | 2308.154 |
| 2 | 2 | 2 | 99.998 | 100 | 0.0024 | 24.479 |

From Tables 2, 3, 4 and 5 we deduce that exponential is the optimum weighting function in this case because a PCP(1;1) model with only a parameter explains more than a 99.99% percentage of the future variance of this process. However, the MPV explained by a PCP model is very similar, about a 68%, for the multivariate and the unweighted approaches meanwhile the MPV explained by the linear weighted PCP models increases until 77.8%. On the other hand, in order to explain a 65% of the future variance of this process, a PCP models needs eleven parameters with the multivariate approach, ten with the unweighted functional estimation and only five (the half) with the linear weighting function.

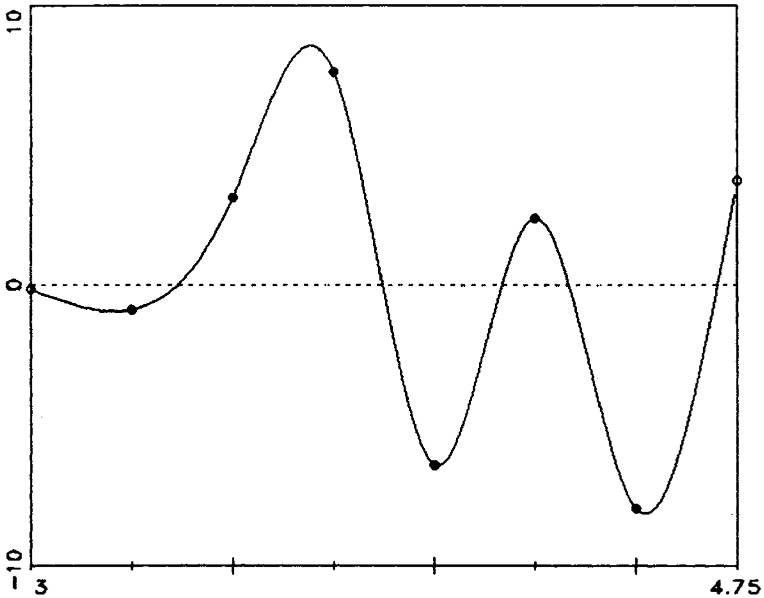



Fig. 4. Simulated series $o(t)$ (dotted line) and its forecasting (solid line) in the period [228,230) (sample path $X_{45}$ in the interval [3,5)) by using the exponential PCP(1,1) model in the case of equally spaced data

Therefore, the selected exponential PCP(1;1) model is given by

$$\tilde{X}^1(s) = \bar{X}(s) + \tilde{\hat{\eta}}_1^{\,1}\,\hat{g}_1(s), \qquad s \in [3, 4.75],$$

where the weighted least-squares linear regression model for $\tilde{\hat{\eta}}_1^{\,1}$ is given by

$$\tilde{\hat{\eta}}_1^{\,1} = -0.7422\,\hat{\xi}_1.$$

In order to evaluate the forecasting performance of the estimated PCP(1;1) model, we have computed different types of mean square errors (MSE) with respect to the last five sample paths and the predictions at the discretization knots. The results appear in Tables 6 and 7. As an example, the forecasting for the first of the five sample paths used for evaluating the forecasting performance of this model have been drawn in Figure 4. Finally, Figure 5 represents the curve of mean squared predicction error defined as

$$MSE(s) = \sqrt{\frac{1}{5}\sum_{w=45}^{49}(X_w(s) - \tilde{X}^1_w(s))^2}\,, \qquad s \in [3, 5). \tag{12}$$

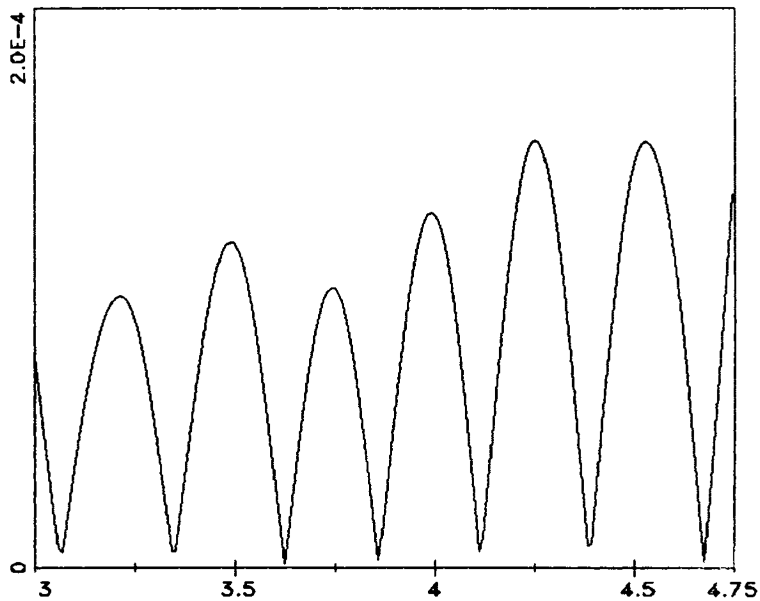



Fig. 5. Mean squared prediction error given by exponential PCP(1,1) model in the case of equally spaced data

Table 6: Mean squared error given by the exponential PCP(1;1) model for each of the last five sample paths in the case of equally spaced simulated data.

$$MSE(w) = \sqrt{\tfrac{1}{8}\sum_{k=12}^{19}(X_w(0.25k) - \tilde{X}_w^1(0.25k))^2}$$

| w | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $MSE(w)$ | 5.8E-5 | 9.1E-5 | 1.2E-4 | 1.4E-4 | 1.7E-4 |

Table 7: Mean squared error given by the exponential PCP(1;1) model for each of the equally spaced discretization time points.

$$MSE(t_k^0) = \sqrt{\tfrac{1}{5}\sum_{w=45}^{49}(X_w(t_k^0) - \tilde{X}_w^1(t_k^0))^2}$$

| $t_k^0$ | 3.00 | 3.25 | 3.50 | 3.75 | 4.00 | 4.25 | 4.50 | 4.75 |
|---|---|---|---|---|---|---|---|---|
| $MSE(t_k^0)$ | 7.5E-5 | 8.8E-5 | 1.2E-4 | 1.0E-4 | 1.3E-4 | 1.5E-4 | 1.5E-4 | 1.5E-4 |

## 4.3 Building weighted PCP and MPCR models for unequally spaced data

In this section, the realization $o(t)$ has been evaluated at one-thousand unequally spaced time points in the interval $[0, 250)$. In order to obtain twenty unequally spaced knots in the interval [0,5) including the knots 0 and 3, we have randomly generated eighteen values of a uniform distribution on $[0, 5)$. Once the simulated time points have been ordered, and denoted by $t_k$ $(k = 0, \ldots, 19)$, with $t_0 = 0$ and $t_{14} = 3$, the rest of time points are generated from these in each interval $I_i$ as follows: $t_k^i = t_k + 5i$ $(i = 0, \ldots, 49)$. As a consequence the sample paths of the associated process $X$ are evaluated at twenty unequally spaced points in the interval $[0, 5)$. In order to estimate a PCP model we have used, for each of the forty-five first sample paths of X, forteen unequally spaced observations in the past interval $[0, 3]$ and seven in the future interval $[3, 5)$.

The sample principal factors have been again estimated, in each period, by using multivariate PCA of the discrete data matrix, and uniform, linear and exponential weights for estimating the covariance function of the cubic splines interpolating to the forty-five sample paths.

The percentages of variance explained by the principal components associated to these PCAs appear in Table 8. The variances explained by the principal components in the future are again very different when using exponential weighting (only the first principal component explains more than a 99% of the total variability). On the other hand, the first five principal components in the future explain more than a 97% of the total variance when using the other three principal component approaches. Therefore, the PCP model is going to be constructed with no more than these five principal components as response variables.

In order to select the *best* PCP model for each type of weights, the pairs of future-past principal components with significative linear correlation have been ordered again in terms of their proportions of error reduction. The results appear in Tables 9, 10, 11 and 12 for the multivariate PCA, uniform, linear and exponential weighted FPCA, respectively. The explanation of columnns in this tables is the same that in the previous case of equally spaced data. The exponential approach is again the optimum because a PCP(1;1) model explains more than a 99.9% of the future variability. Let us observe that the differences among the MPVs for the multivariate, uniform and linear approaches is not so big. However the number of parameters in a PCP model that explains the same percentage of future variance is much smaller for the uniform and linear weighted FPCA. For example, a PCP model explaining at least a 65% of the future variance have 9, 5 and 2 parameters for the multivariate, uniform and linear approaches, respectively. As it was expected the results given by multivariate PCA are really diferent to the ones given by unweighted functional PCA when the data are unequally spaced.

Table 8: Total variances and percentage of variance explained by the principal components for multivariate, unweighted and weighted PCA in the two periods $[0, 3]$ and $[3, 5)$ in the case of unequally spaced data.

[0,3)

| | | | Weighted FPCA | |
|---|---|---|---|---|
| p.c. | Multivariate | Unweighted | Linear | Exponential |
| 1 | 35.6757 % | 64.9781% | 71.6266 % | 99.9957 % |
| 2 | 32.2936 % | 29.3650% | 24.2170 % | 3.9E-3 % |
| 3 | 10.9124 % | 1.4674% | 1.1482 % | 2.8E-4 % |
| 4 | 6.0895 % | 1.2284% | 0.9888 % | 9.0E-5 % |
| 5 | 3.7700 % | 0.9060% | 0.6892 % | 4.0E-5 % |
| 6 | 3.1080 % | 0.7816% | 0.5318 % | 0.0000 % |
| 7 | 2.7741 % | 0.4059% | 0.3040 % | 0.000% |
| 8 | 1.5638 % | 0.3448% | 0.2234 % | 0.000% |
| 9 | 1.4138 % | 0.2245% | 0.1145 % | 0.000% |
| 10 | 1.1373 % | 0.1401% | 0.0935 % | 0.000% |
| 11 | 0.7496 % | 0.0804% | 0.0349 % | 0.000% |
| 12 | 0.3202 % | 0.0566% | 0.0215 % | 0.000% |
| 13 | 0.1903 % | 0.0197% | 0.0057 % | 0.000% |
| 14 | 0.0017 % | 0.0015% | 0.0010 % | 0.000% |
| $\hat{V}$ | 2988.7725 | 3326.3798 | 3164.7995 | 4.3613 |

[3,5)

| | | | Weighted FPCA | |
|---|---|---|---|---|
| p.c. | Multivariate | Unweighted | Linear | Exponential |
| 1 | 47.3558% | 70.4186% | 79.4791% | 99.9899 % |
| 2 | 22.5199% | 15.9749% | 9.3377% | 7.9E-3 % |
| 3 | 12.8540% | 6.7802% | 6.0246% | 2.1E-3 % |
| 4 | 9.6431% | 3.2263% | 2.3479% | 1.3E-4 % |
| 5 | 5.4272% | 2.3270% | 1.9874% | 6.0E-5 % |
| 6 | 1.6679% | 1.0658% | 0.6558% | 1.0E-5 % |
| 7 | 0.5321% | 0.2072% | 0.1675% | 0.0000 % |
| $\hat{V}$ | 2227.8403 | 616.8356 | 505.0300 | 0.2829 |

Table 9: Pairs of future-past principal components ordered in terms of their percentage of error reduction for the multivariate PCA in the case of unequally spaced data.

| Multivariate PCA | | | | | | |
|---|---|---|---|---|---|---|
| Pair | f.p.c. | p.p.c. | APV | RPV | RMSE | t-value |
| 1 | 1 | 1 | 30.852 | 42.338 | 14.8349 | 8.966 |
| 2 | 2 | 2 | 41.336 | 56.726 | 13.6640 | 6.120 |
| 3 | 1 | 3 | 49.362 | 67.740 | 12.6950 | 2.962 |
| 4 | 2 | 3 | 53.636 | 73.605 | 12.1474 | 3.174 |
| 5 | 2 | 1 | 57.823 | 79.351 | 11.5859 | 3.134 |
| 6 | 3 | 14 | 60.321 | 82.779 | 11.2376 | 3.220 |
| 7 | 4 | 2 | 62.507 | 85.779 | 10.9236 | 3.551 |
| 8 | 4 | 4 | 64.637 | 88.702 | 10.6088 | 3.491 |
| 9 | 3 | 6 | 66.220 | 90.874 | 10.3686 | 2.458 |
| 10 | 5 | 7 | 67.672 | 92.866 | 10.1434 | 3.962 |
| 11 | 4 | 3 | 68.897 | 94.548 | 9.9493 | 2.502 |
| 12 | 4 | 5 | 70.018 | 96.086 | 9.7684 | 2.378 |
| 13 | 4 | 7 | 70.889 | 97.282 | 9.6254 | 2.067 |
| 14 | 5 | 8 | 71.677 | 98.363 | 9.4943 | 2.702 |
| 15 | 5 | 3 | 72.389 | 99.340 | 9.3741 | 2.549 |
| 16 | 5 | 5 | 72.870 | 100 | 9.2922 | 2.044 |

Table 10: Pairs of future-past principal components ordered in terms of their percentage of error reduction for the non-weighted estimation of FPCA in the case of unequally spaced data.

| Uniform FPCA | | | | | | |
|---|---|---|---|---|---|---|
| Pair | f.p.c. | p.p.c. | APV | RPV | RMSE | t-value |
| 1 | 1 | 1 | 44.618 | 62.013 | 17.1067 | 8.623 |
| 2 | 1 | 4 | 54.886 | 76.285 | 15.4396 | 2.709 |
| 3 | 2 | 2 | 59.727 | 83.013 | 14.5878 | 4.324 |
| 4 | 2 | 5 | 63.651 | 88.467 | 13.8589 | 3.742 |
| 5 | 2 | 6 | 66.009 | 91.744 | 13.4019 | 2.729 |
| 6 | 3 | 3 | 67.978 | 94.481 | 13.0078 | 4.196 |
| 7 | 3 | 14 | 69.487 | 96.579 | 12.6976 | 3.509 |
| 8 | 2 | 9 | 70.900 | 98.541 | 12.4003 | 2.042 |
| 9 | 4 | 5 | 71.442 | 99.296 | 12.2841 | 2.949 |
| 10 | 4 | 2 | 71.949 | 100 | 12.1747 | 2.830 |

Table 11: Pairs of future-past principal components ordered in terms of their percentage of error reduction for the linear weighted estimation of FPCA in the case of unequally spaced data.

| Linear FPCA | | | | | | |
|---|---|---|---|---|---|---|
| Pair | f.p.c. | p.p.c. | APV | RPV | RMSE | t-value |
| 1 | 1 | 1 | 57.133 | 74.894 | 13.6181 | 10.485 |
| 2 | 1 | 4 | 66.283 | 86.889 | 12.0775 | 2.365 |
| 3 | 3 | 2 | 69.285 | 90.823 | 11.5274 | 6.534 |
| 4 | 2 | 6 | 70.912 | 92.955 | 11.2180 | 3.012 |
| 5 | 2 | 8 | 72.320 | 94.801 | 10.9432 | 2.763 |
| 6 | 2 | 9 | 73.344 | 96.143 | 10.7388 | 2.301 |
| 7 | 3 | 14 | 74.246 | 97.326 | 10.5556 | 2.752 |
| 8 | 2 | 3 | 75.045 | 98.373 | 10.3905 | 2.006 |
| 9 | 4 | 5 | 75.754 | 99.303 | 10.2418 | 4.314 |
| 10 | 4 | 11 | 76.022 | 99.654 | 10.1851 | 2.353 |
| 11 | 4 | 3 | 76.286 | 100 | 10.1289 | 2.333 |

Table 12: Pairs of future-past principal components ordered in terms of their percentage of error reduction for the exponential weighted estimation of FPCA in the case of unequally spaced data.

| Exponential FPCA | | | | | | |
|---|---|---|---|---|---|---|
| Pair | f.p.c. | p.p.c. | APV | RPV | RMSE | t-value |
| 1 | 1 | 1 | 99.988 | 99.999 | 0.0053 | 1608.175 |
| 2 | 2 | 2 | 99.989 | 100 | 0.0051 | 2.688 |

The selected exponential PCP(1;1) model is given by

$$\tilde{X}^1(s) = \bar{X}(s) + \tilde{\hat{\eta}}_1^{\,1}\,\hat{g}_1(s), \qquad s \in [3, 4.16736],$$

where the weighted least-squares linear regression model for $\tilde{\hat{\eta}}_1^{\,1}$ is given by

$$\tilde{\hat{\eta}}_1^{\,1} = -0.2547\,\hat{\xi}_1.$$

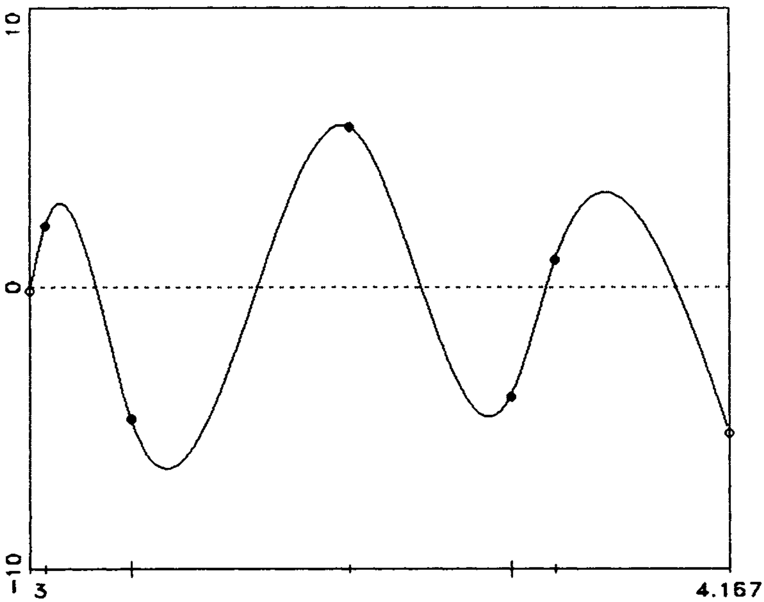



Fig. 6. Simulated series $o(t)$ (dotted line) and its forecasting (solid line) in the period [228,230) (sample path $X_{45}$ in the interval [3,5)) by using the exponential PCP(1,1) model in the case of unequally spaced data

In order to evaluate the forecasting performance of the estimated PCP(1;1) model, we have computed again different types of mean square errors (MSE) with respect to the last five sample paths and the predictions at the discretization knots. The results appear in Tables 13 and 14. The forecasting for the first of the five sample paths used for evaluating the forecasting performance of this model have been drawn in Figure 6. Finally, Figure 7 represents the curve of mean squared predicction error $MSE(s)$ given as in Equation (12).

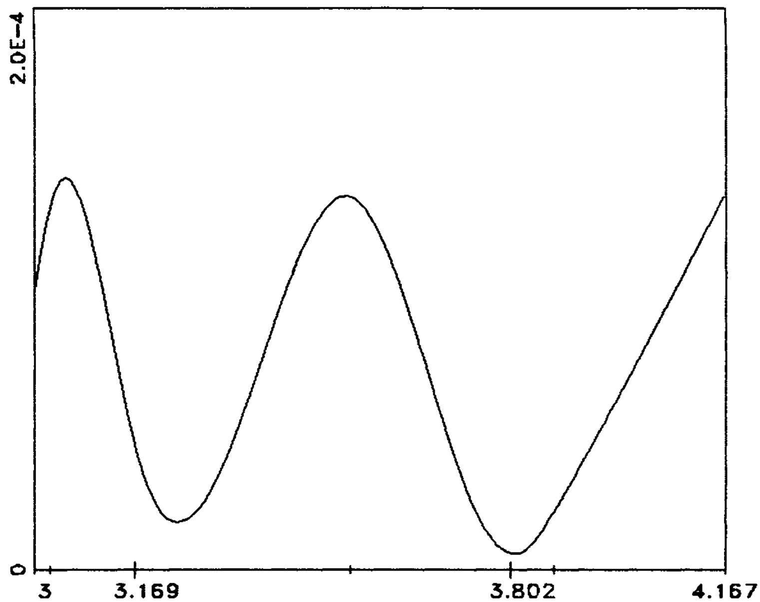



Fig. 7. Mean squared prediction error given by exponential PCP(1,1) model in the case of unequally spaced data

Table 13: Mean squared error given by the exponential PCP(1;1) model for each of the last five sample paths in the case of unequally spaced data.

| $MSE(w) = \sqrt{\frac{1}{7}\sum_{k=13}^{19}(X_w(t_k) - \tilde{X}_w^1(t_k))^2}$ | | | | | |
|---|---|---|---|---|---|
| $w$ | 45 | 46 | 47 | 48 | 49 |
| $MSE(w)$ | 4.7E-5 | 7.1E-5 | 9.4E-5 | 1.1E-4 | 1.3E-4 |

Table 14: Mean squared error given by the exponential PCP(1;1) model for each of the unequally spaced discretization time points.

| $MSE(t_k) = \sqrt{\frac{1}{5}\sum_{w=45}^{49}(X_w(t_k) - \tilde{X}_w^1(t_k))^2}$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $t_k$ | 3 | 3.0254 | 3.1687 | 3.5324 | 3.8025 | 3.8768 | 4.1674 |
| $MSE(t_k)$ | 9.9E-5 | 1.3E-4 | 4.4E-5 | 1.3E-4 | 6.1E-6 | 2.1E-5 | 1.3E-4 |

## 4.4 Discussion and concluding remarks

In this paper we have proposed a weighted estimation of PCP models to take into account the natural order among the sample paths obtained when a time series is cut into pieces. These PCP models have been originally developed to deal with a sample of independent realizations of a continuous time stochastic process and adapted for predicting univariate continuous time series when discrete values are not necessarily equally spaced.

From the analysis of simulated data in the last section we deduce that the exponential weighting drastically reduces the RMSE and provides the most parsimonious model with the highest proportion of variance explained in the future (MPV). In addition, we have tried to model without success the simulated time series by using the classic Box-Jenkins methodology.

The main advantage of the PCP models is that they predict and smooth the process between the observed knots. Moreover, Castro *et al.* (1986) and Aguilera *et al.* (1995) have shown (for quadrature approximation and orthogonal projection, respectively) that multivariate PCA is not robust and gives dreadful results when the sampling points are unequally spaced. In fact, from the previous application with simulated data, we can deduce that the results given by MPCR are similar to those given by the unweighted functional approach when the data are equally spaced. However, the unweighted functional approach improves in the case of unequally spaced data by reducing the MSE and the number of parameters in the model.

Summarizing, we consider exponential weighting to be suitable for adjusting a PCP model to this case because takes into account the change in the evolution of the simulated process in the last periods of amplitud five giving them higher weights.

## Acknowledgements

This research was supported in part by Project PB96-1436, Dirección General de Enseñanza Superior, Ministerio de Educación y Cultura, Spain.

## References

Aguilera, A.M., Gutiérrez, R., Ocaña, F.A. and Valderrama, M.J. (1995), Computational approaches to estimation in the principal component analysis of a stochastic process, *Appl. Stoch. Models Data Anal.*, **11**, 279-299.

Aguilera, A.M., Gutiérrez, R. and Valderrama, M.J. (1996), Approximation of estimators in the PCA of a stochastic process using B-splines, *Comm. Statist. (Simulation and Computation)*, **25**, 671-691.

Aguilera, A. M., Ocaña, F. A. y Valderrama, M. J. (1997), An approximated principal component prediction model for continuous time stochastic processes, *Appl. Stoch. Models Data Anal.*, **13**, 61-72.

Besse, P. and Ramsay, J.O. (1986), Principal component analysis of sample functions, *Psychometrika*, **51**, 285-311.

Besse, P. (1991), Approximation spline de l'ánalyse en composantes principales d'une variable aléatoire hilbertienne, *Annales de la Faculté des Sciences de Toulouse*, **12**, 329-346.

Besse, P. and Cardot, H. (1994), Approximation spline de la prévision d'un processus autorǵressif hilbertien d'ordre 1, Publication du Laboratoire de Statistique et Probabilits de Toulouse, **17**.

Bosq, D. (1991), Modelization, non-parametric estimation and prediction for continuous time processes, In *Nonparametric functional estimation and related topics*, Ed. G. Roussas. NATO, ASI Series, 509-529.

Castro, P.E., Lawton, W.H. and Sylvestre, E.A. (1986), Principal modes of variation for processes with continuous sample curves, *Technometrics*, **28**, 329-337.

Dauxois, J., Pousse, A. and Romain, Y. (1982), Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference, *J. Multivar. Anal.*, **12**, 136-154.

Deville, J.C. (1974), Méthodes statistiques et numériques de l'analyse harmonique, *Annales de l'INSEE*, **15**, 3-101.

Deville, J.C. (1978), Analyse et prevision des series chronologiques multiples non stationnaires, *Statist. Anal. Données*, **3**, 19-29.

Jackson, J.E. (1991), *A User's Guide to Principal Components*, Wiley, New York.

Ramsay, J.O. and Dalzell, C.J. (1991), Some tools for functional data analysis (with discussion), *J. R. Statist. Soc. B*, **53**, 539-572.

Ramsay, J.O., Wang, X. and Flanagan, R. (1995), A functional data analysis of the pinch force of human fingers, Appl. Statist., **44**, 17-30.

Rice, J.O. and Silverman, B.W. (1991), Estimating the mean and covariance structure nonparametrically when the data are curves, *J. R. Statist. Soc. B*, **53**, 233-243.

Ramsay, J.O. and Silverman, B.W. (1997), *Functional Data Analysis*, Springer-Verlag, New York.

Saporta, G. (1985), Data analysis for numerical and categorical individual time series, *Appl. Stoch. Models and Data Anal.*, **1**, 109-119.

Silverman, B.W. (1996), Smoothed functional principal components analysis by choice of norm, *Ann. Statist.*, **24**, 1-24.