Nonlinear Analysis: Modelling and Control, 2002, Vol. 7, No. 2, 53–68

Functional Data Analysis of Payment Systems

A. Laukaitis[1], **A. Račkauskas**[2]

[1]Institute of Mathematics and Informatics, Akademijos 4, 2600 Vilnius, Lithuania
algirdas@vb.lt

[2]Vilnius University, Naugarduko 24, 2006 Vilnius, Lithuania
Institute of Mathematics and Informatics, Akademijos 4, 2600 Vilnius, Lithuania
alfredas.rackauskas@maf.vu.lt

Received: date 03.07.2002
Accepted: date 23.10.2002

Abstract. In this paper for a credit cards payment system as robust predictor of transactions number and transactions intensity is proposed by means of functional autoregressive model. Intraday economic time series are treated as random continuous functions projected onto low dimensional subspace. Both B-splines and Fourier bases are considered for data smoothing.

Keywords: high frequency data, functional AR model, functional data analysis, B-splines, Fourier bases, credit cards payment system.

AMS classifications: 62M10.

1 Introduction

Most econometric models used in financial institutions are adapted to work with highly summarized data. For example, liquidity riskcontrol (cash flow), profit risk analysis ect. are based on a daily summarized data. In reports to central institutions a usual time lag is either one month or quarter or one year. Similar time lag is used in macroeconometrics.

The objective of this paper is an empirical study of transactions intensity in ATM (auto tele machine) and POS (point of sale) networks using a high frequency data. Our study can be described as an exploratory data approach: Analysis $\Rightarrow$ Model $\Rightarrow$ Conclusions. Treating intraday time series as random functions in a space spanned by finite dimensional functional bases, we intensively explore methods of functional data analysis. Although the support of functions is taken to be a one day different time lags like a one week or a one month can be treated as well. The choice of a one day time lag in our case is trade of between a complication of functional space (number of bases functions) and a complication of discreet model (number of parameters involved). In our empirical analysis we use data from Vilnius Bank data warehouse*. Since we are looking into almost 50% of the whole Lithuanian market of transactions mad in electronic way, the data we are analyzing can present some interest for macroeconomic analysis of consumption processes as well.

The number of transactions at time t is interpreted as double stochastic Poisson process $N(t)$. Its conditional intensity $\Lambda(t)$:

$$EN(t)|F = \Lambda(t) = \int_0^t \lambda(s)\,ds$$

is a stochastic process. Assuming that in a small time interval the intensity Λ is constant, its efficient estimator is empirical mean. Combining this with a smoothing technique we obtain observations $\Lambda_i(t)$ and $\lambda_i(t), i = 1,\ldots,n$ of the stochastic processes $\Lambda(t)$ and $\lambda(t)$. Considering observations $\lambda_i(t)$ as random functions in a certain Hilbert function space, we fit the following functional autoregression model

$$\lambda_i(t) - \lambda_{i-7}(t) = \rho(\lambda_{i-1} - \lambda_{i-8})(t) + \varepsilon_i(t),$$

where ρ is a Fredholm type operator defined by a kernel function $k: \rho\lambda(t) = \int k(t,s)\lambda(s)\,ds$. The kernel k is estimated by the principal component approach. *Ex poste* prediction is used to check the model.

Due to simplicity of the model we can consider different type of functional of stochastic processes $\lambda(t)$ and $\Lambda(t)$ such as the global maximum $\max_t \lambda(t)$; the corresponding time of the global maximum $\arg\max_t \lambda_i(t)$; the square integrated density $\int \lambda^2(s)\,ds$ ect. Results related to these functionals will be presented elsewhere.

Recently there were several attempts to analyze economical data on transactions level. In recent paper Müller and Zumbach (2000) have presented a

* To avoid any particular financial details, all data used in this work have been linearly transformed

new time series operator technique to analyze inhomogeneous time series. Müller (2000) showed usefulness of this approach by calculating Value-at-Risk (VAR) when tick by tick data are concerned. Rydberg and Shephard (1999) proposed compound Poisson processes %to model trade-by-trade financial data. They %have used compound processes as the basis to model trading processes, with trading occurring at times determined by Cox process. Our work is related to Rydberg and Shephard (1999) since we model intra-day economic activity as double stochastic Poisson process. Rydberg and Shephard (1999) used Ornstein-Uhlenbeck process to model Poisson intensity parameter whereas we model Poisson intensity parameter as a random function, which is an element of a subspace, spanned by some finite functional bases.

A comprehensive description of methods for exploring and estimating functional data can be found in the recent book by Ramsay and Silverman (1997). We found very useful the code of MATLAB (S-PLUS is also available) where examples and functions of functional data analysis are presented. We refer to Bosq (1991, 1999), Pumo (1998), Besse, Cardot and Stephenson (2000), Cardot, Ferraty and Sarda (1999), Besse, Cardot, and Ferraty (1997) for different models of functional data.

Our paper is organized in the following way. A data description and some preliminary analysis are presented in section 2. In section 3 we describe exploration and identification technique of functional time series under consideration. In section 4 we estimate and check the model. Some technical details are given in the appendix.

2 Data Description and Preliminary Analysis

We tackle the problem of mathematical prediction of dynamics of the intensity of transactions made in Vilnius Bank ATM and POS networks which is the largest bank of Lithuania. Transactions made in ATM and POS networks between 03/11/2000 and 10/02/2001 were extracted from the data warehouse.

To explain our work let us first describe a similar standard vector task, which can be solved with the existing tools. To this aim first consider a problem of predicting daily number of transactions for one day ahead. Using daily summarized data (fig. 1) and standard SAS tools, the best model fitting such data is $ARIMA\ (0,1,1)\ (0,1,1)_s$.

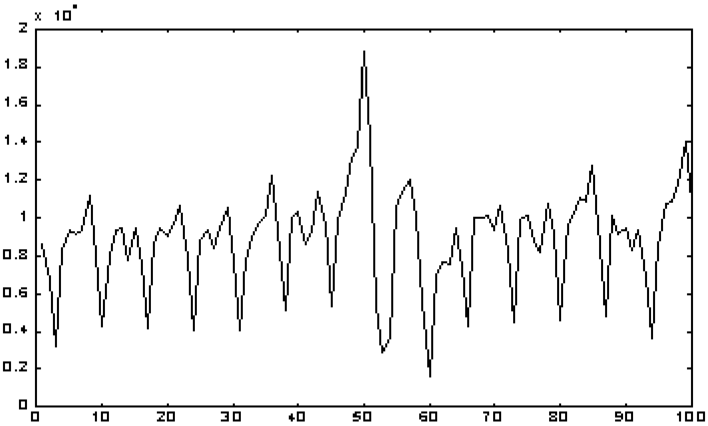

Fig. 1. Number of transactions over 100 days in ATM network

Now let us split each day into two equal parts, doubling the number of observations. So observations with 24 hours time lag now are replaced by observations separated by 12 hours time interval. Fig. 2 displays the data.

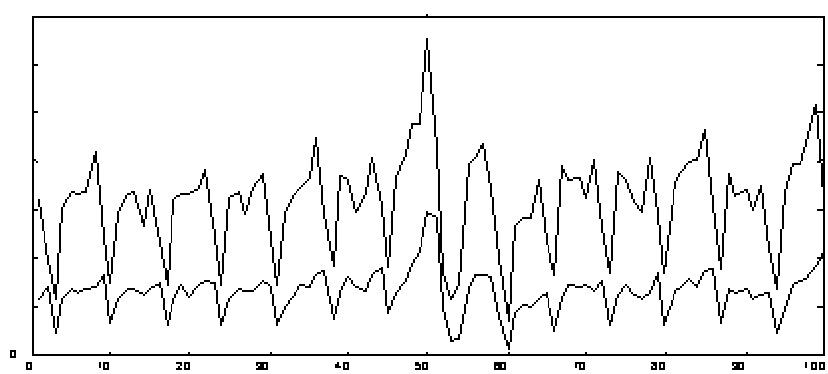

Fig. 2. Bivariate time series

The upper curve corresponds to the number of transactions measured daily between 12 and 24 hours whereas the lower curve represents the number of transactions between 0 and 12 hours. Number of days remains 100. Considering the data as bivariate time series we fitted *VAR*(7) model. Removing Christmas effect the model simplified to *VAR*(1). Next let us repeat the previous step and again double the number of observations. Now we are observing a state of ATM network after every 6 hours. So we localize some intra-day economic activity. However, after several such steps we failed to find an appropriate model for the obtained in this way data using standard time series tools. The STATESPACE procedure, which is designed for automatical selection of the best state space model, was used but in our case the maximum likelihood estimates failed to converge.

As a next example consider the economic activity in ATM and POS networks at a particular time of a day. Namely, consider the number of transactions $\Lambda(t)$ made in a unit of time (in our case it equals to ≈ 1.4 min.). Fig. 3 displays the data.

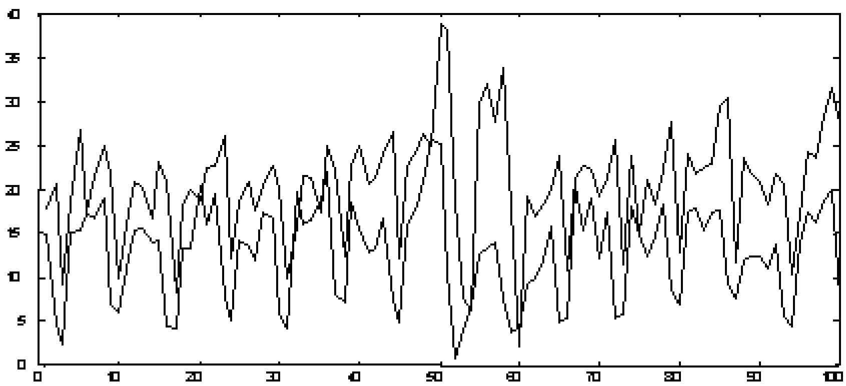

Fig. 3. Number of transactions at a particular time

Lower curve represents $\Lambda(t)$ at 19:20 hour whereas upper curve at 12:15 hour. The *ARIMA* (0,1,1) $(0,1,1)_s$ model was chosen automatically by SAS tools when analysing each series separately. Modelling both time series as a vector we obtained *VAR*(1) after censoring and *VAR*(7) without censoring. However such empirical analysis becomes more complicated when considering the number of transactions at more than two time instances.

These examples motivate the functional data approach and give an idea for a functional data model.

Now let us describe the data under investigation more precisely. One day is normalized to be the interval [0,1]. As was mentioned above, the number of transactions we interpret as double stochastic Poisson process $(N^q(t), t \in [0,1])$, where q = 1 corresponds to ATM network whereas q = 2 corresponds to POS network. The corresponding conditional intensities $\Lambda^q(t)$, $t \in [0,1]$ given by

$$\Lambda^q(t) = \int_0^t \lambda^q(s)\,ds = EN^q(t)F$$

are stochastic processes for which we have observations

$$\{\Lambda_i^q(t_j),\quad i = 1,\ldots,n,\quad j = 0,1,\ldots,1024\},\quad q \in \{1,2\}.$$

Here n denotes the number of days which we are analysing, and it equals to either 100 or 86 after censoring (removing two weeks of Christmas effect). In what follows $t_j = j\,/1024$, $j = 0, 1,\ldots,1024$. As one can see, one day (more precisely twenty-four hours) is divided into 1024 equal time intervals. Motivation for choosing this number is that first, we want to construct an empirical process by counting the number of transactions during a time interval approximately equal to one minute, and second, the number of intervals is reasonable to be equal to a power of two.

Using the observations of the intensity functions $\Lambda^q(t)$, we build the observations for its derivative $\lambda^q(t)$. Namely, we consider

$$\{\lambda_i^q(t_j),\quad i = 1,\ldots,n,\quad j = 1,\ldots,1024\},$$

where

$$\lambda^q(t_j) = \Lambda(t_j) - \Lambda(t_{j-1}),\quad j = 1,\ldots,1024\,.$$

Fig. 4 displays the data of one randomly chosen day. In the left side we see the number of transactions in ATM and POS networks whereas in the right we see the corresponding differentiated processes.

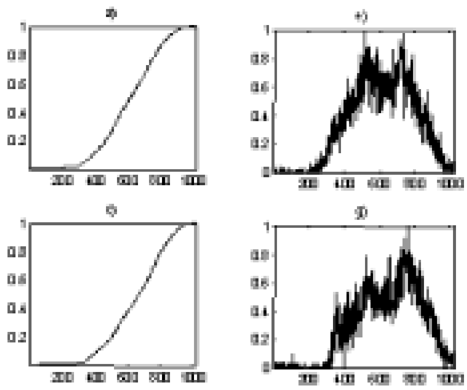

Fig. 4. The numbers of transactions and the intensities of transactions

3 Exploration and Identification

In what follows $\left(x_i(t_j)\right)$ will denote either of the four sets of observations: $\Lambda_i^q(t_j),\ q=1,2,\ \ \lambda_i^q(t_j),\ q=1,2$.

3.1 Choice of a functional subspace. Our next step is to convert a raw functtional data $\left(x_i(t_j)\right)$ into a true functional form $\left(y_i(t), t\in[0,1]\right)$. To this aim we use smoothing by function bases. Hence

$$y_i(t)=\sum_{j=1}^{m} c_{ij}B_j(t),$$

where $\left(B_j(t)\right)$ are either cubic *B*-splines or Fourier bases. By choosing the bases we are making hypothesis that our functions of interest can be approximated either by polynomials (in the case of *B*-splines) or by a fixed number of frequencies (in the case of Fourier bases). One of the key issues at this step is the number of coefficients *m* one has to choose. Our motivation for *m* = 32 coefficients we have used is that firstly 1/32 of a day represents 45 minutes and secondly $32=\sqrt{1024}$. One can check that in many cases square root of the number of observations for the dimension gives optimal in a sense approximation.

Coefficients $C=(c_{ij})$ were found by the least square method:

$$C=\left(B'B\right)^{-1}B'X,$$

where $B=\left(B_j(t_l),\ j=1,\ldots,32,\ l=1,\ldots,1024\right)$ is the bases matrix and $X=\left(x_i(t_l),\ i=1,\ldots,32,\ l=1,\ldots,1024\right)$ is our data matrix. Fig. 5 displays

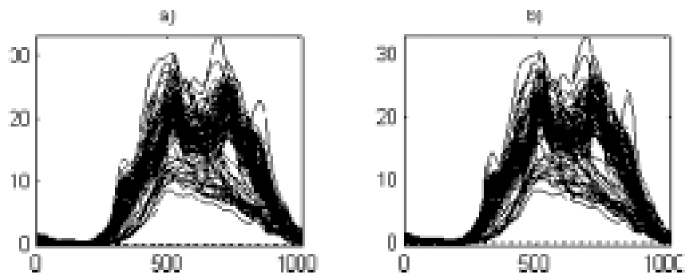

Fig. 5. Smoothed data

functional observations of ATM network: a) intensity observations smoothed with *B*-splines, b) intensity observations smoothed with Fourier bases. As we can see, there is no difference on the first glance between our data presented in *B*-splines or Fourier bases. Similar situation occurs for POS network. So in the rest of the paper we will consider functional data $\left(y_i(t),\ t\in[0,1],\ i=1,\dots,n\right)$ obtained smoothing by *B*-splines only.

3.2 Principal component analysis. Covariance and cross-covariance analysis of the functional time series $\left(y_i(t),\ t\in[0,1],\ i=1,\dots,n\right)$ suggest for differentiation of the functional time series at lag 7. Differentiating pointwise we obtain

$$z_i(t) = y_i(t) - y_{i-7}(t), \quad t\in[0,1], \quad i=1,\dots,n, \tag{1}$$

Now we consider functional time series $\left(z_i(t)\right)$ defined by (1), as observations of a stationary sequence in the space $L_2(0,1)$. The estimate of the covariance function

$$\hat{h}_z(s,t) = \frac{1}{n}\sum_{k=1}^{n} z_k(s) z_k(t), \quad s,t\in[0,1]$$

defines the operator $\hat{h}_z : L_2(0,1)\to L_2(0,1)$ by

$$\hat{h}_z u(t) = \int_0^1 \hat{h}_z(s,t) u(s)\,ds, \quad t\in[0,1].$$

This operator is self-adjoint, nonnegative and admits the spectral decomposition

$$\hat{h}_z = \sum_j \lambda_j \phi_j \otimes \phi_j, \tag{2}$$

where (ϕ_j) is the orthogonal basis in the space $L_2(0,1)$ consisting of eigenfunctions of the operator $\hat{h}_z$ and (λ_j) denotes the corresponding eigenvalues assuming that $\lambda_1 > \lambda_2 > \dots$. The first four eigenfunctions are shown in fig. 6.

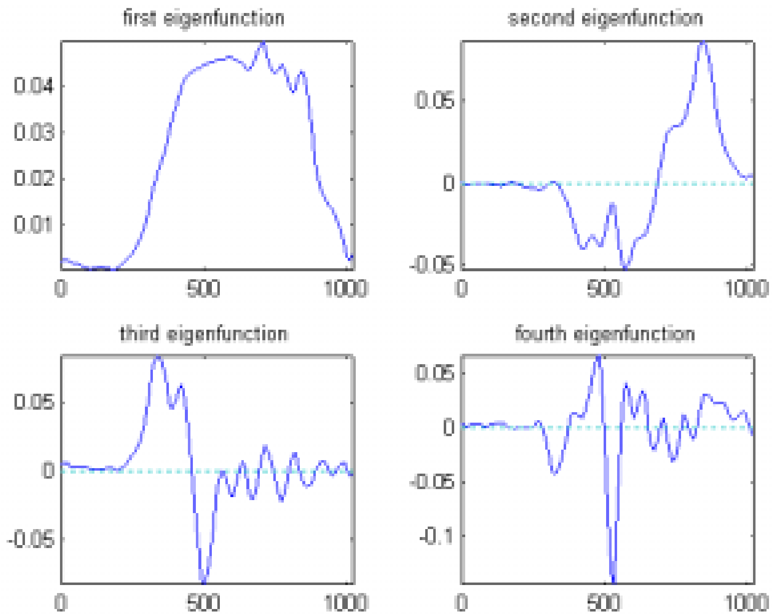

Fig. 6. First four eigenfunctions

When eigenfunctions and corresponding eigenvalues are found one can rerepresent observations in the new bases, and thus obtaining

$$\hat{z}_k(t) = \sum_{j=1}^{M} \lambda_{jk}\phi_j(t), \quad j = 1, \ldots, n,$$

where

$$\lambda_{jk} = \int_0^1 z_k(t)\phi_j(t)\,dt, \quad k = 1, \ldots, n, \quad j = 1, \ldots, M .$$

The crucial point here is the choise of the dimension *M*. We used several arguments. In the left side of the fig. 7 one see the functional time series $(z_i(t))$ projected onto the first three eigenfunctions of the covariance operator whereas in the right side of the fig. 7, projected onto the next three eigen-functions. By rotating three dimensional representation of functional time series we find that the first eigenfunction clearly separates Sundays, the second separates Saturdays and the third can be used to separate Fridays. By visual rotation of the next three eigenfunctions it is difficult to find more information. So this technique allows us to judge about a dimension supporting the main information.

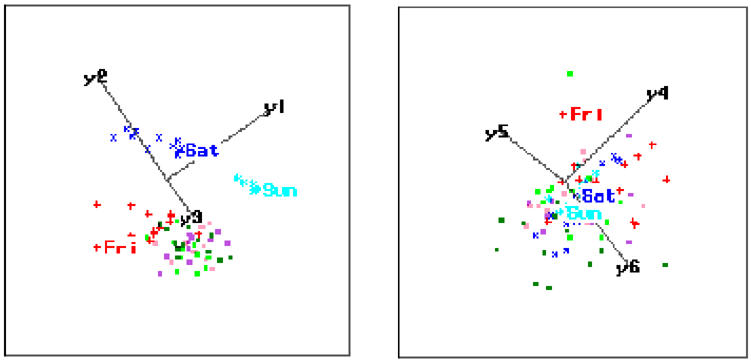

Fig. 7. Functional observations scored by first three (corresp. next three) eigenfunctions

Another technique that can be used to estimate a dimension is based on the white noise test of functional data representation in each eigenfunction successively. One starts with $M = 1$ and tests the following null hypothesis: the time series $\{\lambda_{11}, \lambda_{12}, \ldots, \lambda_{1n}\}$ follows a white noise. Next one tests the null hypothesis for the time series $\{\lambda_{21}, \lambda_{22}, \ldots, \lambda_{2n}\}$ ect. The procedure stop when coefficients starts to behave like a white noise. We have used SAS SPECTRA procedure. Both the Fisher's Kappa statistic and the Kolmogorov-Smirnov statistic rejected the null hypothesis for the first four eigenfunctions.

Finally, we used the *ex poste* prediction to find the best in a sense dimension for the functional data $(\hat{z}_k(t))$ (see the next section below). In all approaches the dimension $M = 4$ was found to be satisfactory, and has been used for the final model.

4 Estimating and Using FAR(1) Model

For the functional time series $(z_i(t),\ t \in [0,1],\ i = 1, \ldots, n)$ we consider the following functional auto-regressive model of the first order (FAR(1)): for i=1, …, n

$$z_i(t) = \int_0^1 \beta(s,t)\, z_{i-1}(s)\,ds + \varepsilon_i(t) \quad \text{for each} \quad t \in [0,1],$$

where we assume that

$$\int_0^1 \int_0^1 \beta^2(s,t)\,ds\,dt < \infty \ .$$

The error terms $(\varepsilon_i(t), t \in [0,1])$ are assumed to be independent random-functions such that $E\varepsilon_i(t) = 0$ for each $t \in [0,1]$; $E\int_0^1 \varepsilon_i^2(t)\,dt = \sigma^2 < \infty$ and

$$E\int_0^1 \varepsilon_i(t) z_j(t)\,dt = 0 \ \text{for } j < i.$$

To estimate the function $\beta(s,t)$, $s,t \in [0,1]$ we proceed as follows. Since (ϕ_j) the eigenfunctions of the covariance operator constitute the orthonormal bases in the spase $L_2(0, 1)$ we can represent

$$\beta(s,t) = \sum_{j,k=1}^{\infty} d_{jk}\phi_j(s)\,\phi_k(t), \quad s,t \in [0,1].$$

Thus, our estimator of the function β is given by

$$\widehat{\beta}(s,t) = \sum_{u,v=1}^{M} \widehat{d}_{uv}\phi_u(s)\,\phi_v(t), \quad s,t \in [0,1],$$

where

$$\widehat{d}_{uv} = \frac{(n-1)^{-1}\sum_{i=1}^{n-1}\langle z_{i-1}, \phi_v\rangle\langle z_i, \phi_u\rangle}{n^{-1}\sum_{i=1}^{n}\langle z_i, \phi_v\rangle\langle z_i, \phi_u\rangle} \tag{3}$$

is a consistent estimator of the coefficient d_{uv} (argumentation is given in the appendix). Fig. 8 shows the function $\widehat{\beta}$. The choise of M is very important. It is clear however that M should be equal to the number of principal components used for functional data representation. As already discussed above, we take $M = 4$.

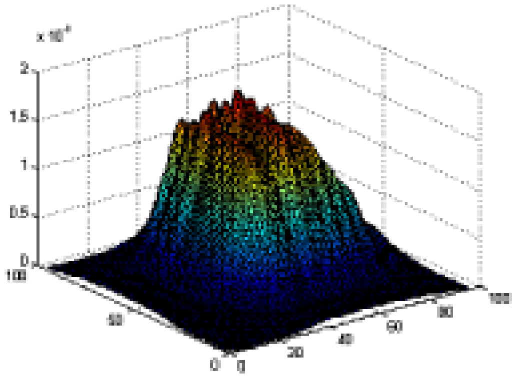

Fig. 8. Estimated kernel β

Finally let us discuss the prediction capability of the model. Consider the prediction of the function $y_{n+1}(t)$, $t \in [0,1]$. According to the model, for each $t \in [0,1]$

$$\hat{y}_{n+1}(t) = y_{n-6}(t) + \int_0^1 \hat{\beta}(s,t)[y_n(s) - y_{n-7}(s)]ds .$$

Fig. 9 displays the actual and the forecasted process for nine days. The processes are evaluated every 15 minutes (96 point in x axis).

Different types of forecasting errors can be considered. The most widely accepted measures of performance of the estimator $\hat{y}_{n+1}$ is its mean integrated square error given by

$$MISE = E \int_0^1 (y_{n+1}(t) - \hat{y}_{n+1}(t))^2 dt ,$$

and the mean uniform error $\Delta_\infty = E \sup_t |y_{n+1}(t) - \hat{y}_{n+1}(t)|$.

Both these quantities we evaluated by choosing ten days that where not used in the parameter estimation (table 1). It is clear that the predicted values strongly depends on the number of principal components used in the model. The table 1 shows this dependence via MISE and Δ_∞ errors.

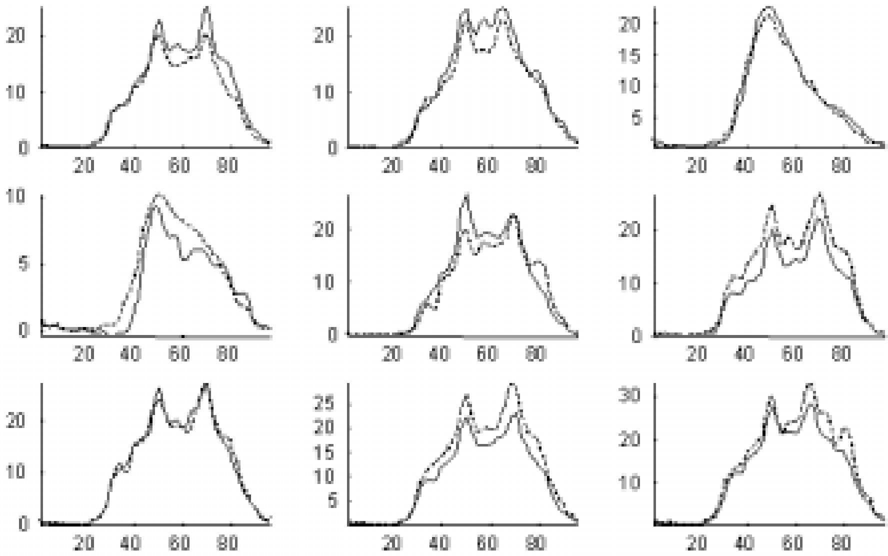

Fig. 9. Predicted values

Error\M	1	2	3	4	5	6	7	8
Δ_∞	5.50	5.02	4.99	4.95	5.05	4.97	5.04	5.07
MISE	5.59	4.74	4.73	4.54	4.72	4.76	4.76	4.81

Table 1. Prediction errors

Appendix

Let H denote the Hilbert space $L_2(0, 1)$ with the norm $\|x\|^2 = \int_0^1 x^2(t)\,dt$ and with the inner product $\langle x, y\rangle = \int_0^1 x(t)\,y(t)\,dt$. A sequence $\{\xi_i = \xi_i(t),\ t \in [0,1],\ i \geq 1\}$ of random functions with values in H is said to follow a Functional Auto-Regressive process of the first order (FAR(1)) if for each $i \in N$ $E\xi_i = 0,\ E\|\xi_i\|^2 < \infty$ and

$$\xi_i = \rho\,\xi_{i-1} + \varepsilon_i\,. \qquad (4)$$

Error terms $(\varepsilon_i(t),\ t \in [0,1])$ are assumed to be independent random elements with values in H such that $E\varepsilon_i = 0, E\|\varepsilon_i\|^2 = \sigma^2 < \infty$ and $E\langle \varepsilon_i, \xi_j \rangle = 0$ for $j < i$. The operator $\rho : H \to H$ is linear compact and satisfies $\|\rho\| < 1$. Covariance operator Γ_0 of $\{\xi_i\}$ is symmetric positive operator defined by $\Gamma_0 = E\xi_i \otimes \xi_i$ so that $\langle \Gamma_0 x, z\rangle = E\langle \xi_i, x\rangle\langle \xi_i, z\rangle$ for $x, z \in H$. A lag one cross covariance operator is defined by $\Gamma_1 = E\xi_{i-1} \otimes \xi_i$ so that $\langle \Gamma_1 x, z\rangle = E\langle y_{i-1}, x\rangle\langle \xi_i, z\rangle$ for $x, z \in H$. Let $\beta : [0,1]^2 \to R$: satisfies

$$\int_0^1 \int_0^1 \beta^2(s,t)\, ds\, dt < \infty .$$

Consider the operator $\rho : H \to H$ defined by the kernel β:

$$\rho x(s) = \int_0^1 \beta(s,t)\, x(t) dt, \text{ for } s \in [0,1].$$

Consider any orthogonal basis (ϕ_i) on the space H. Then one can represent

$$\beta(s,t) = \sum_{j,k=1}^{\infty} d_{jk} \phi_j(s) \phi_k(t), \quad s, t \in [0,1].$$

Hence

$$\rho x(s) = \int_0^1 \sum_{j,k=1}^{\infty} d_{jk} \phi_j(s) \phi_k(t) x(t)\, dt = \sum_{j,k=1}^{\infty} d_{jk} \langle x, \phi_k \rangle \phi_j(s), \quad s, t \in [0,1].$$

For the model 4 we have

$$\xi_i = \sum_{j,k=1}^{\infty} d_{jk} \langle \xi_{i-1}, \phi_k \rangle \phi_j + \varepsilon_i, \quad i = 1, \ldots, N .$$

“Multiplying” the last equation by ϕ_v one obtains

$$\langle \xi_i, \phi_v \rangle = \sum_{j,k=1}^{\infty} d_{jk} \langle \xi_{i-1}, \phi_k \rangle \langle \phi_j, \phi_v \rangle + \langle \varepsilon_i, \phi_v \rangle = \sum_{k=1}^{\infty} d_{vk} \langle \xi_{i-1}, \phi_k \rangle + \langle \varepsilon_i, \phi_v \rangle .$$

Multiplying by $\langle \xi_{i-1}, \phi_u \rangle$ and taking expectation we get

$$E\langle \xi_i, \phi_v \rangle \langle \xi_{i-1}, \phi_u \rangle = \sum_{k=1}^{\infty} d_{vk} E\langle \xi_{i-1}, \phi_k \rangle \langle \xi_{i-1}, \phi_u \rangle + E\langle \varepsilon_i, \phi_v \rangle \langle \xi_{i-1}, \phi_u \rangle .$$

Since

$$E\langle \xi_i, \phi_v \rangle \langle \xi_{i-1}, \phi_u \rangle = 0 ,$$

$$E\langle \xi_{i-1}, \phi_k \rangle \langle \xi_{i-1}, \phi_u \rangle = \langle \Gamma_0 \phi_k, \phi_u \rangle = \lambda_k \langle \phi_k, \phi_u \rangle = \lambda_k \delta_{ku} ,$$

$$E\langle \xi_i, \phi_v \rangle \langle \xi_{i-1}, \phi_u \rangle = \langle \Gamma_1 \phi_v, \phi_u \rangle$$

we have

$$\langle \Gamma_1 \phi_v, \phi_u \rangle = \lambda_u d_{vu} .$$

Hence

$$d_{vu} = \frac{\langle \Gamma_1 \phi_v, \phi_u \rangle}{\lambda_u} , \qquad (5)$$

provided $\lambda_u \neq 0$.

A consistent estimator of the coefficient d_{uv} is obtained by

$$\widehat{d}_{uv} = \frac{(n-1)^{-1} \sum_{i=1}^{n-1} \langle y_{i-1,} \phi_v \rangle \langle y_i, \phi_u \rangle}{n^{-1} \sum_{i=1}^{n} \langle y_{i,} \phi_v \rangle \langle y_i, \phi_u \rangle} . \qquad (6).$$

References

1. Besse P., Cardot H., Stephenson D. Autoregressive forecasting of some functional climatic variations, *Scandinavian Journal of Statistics*, **4**, p. 673-688, 2000
2. Besse P., Cardot H., Ferraty F. Simultaneous nonparametric regressions of unbalanced longitudinal data, *Computational Statistics and Data Analysis*, **24**, p. 255–270, 1997
3. Bosq D. Modelization, non-parametric estimation and prediction for continuous time processes, In: *Nonparametric Functional Estimation and Related Topics,* Roussas G. (Ed.), NATO ASI series, p. 509–529, 1991

4. Cardot H. Nonparametric Estimation of Smoothed Principal Components Analysis of Sampled Noisy Functions, *Journal of Nonparametric Statistics*, **12**, p. 503–538, 2000

5. Cardot H., Ferraty F., Sarda P. Functional Linear Model, *Statistics and Probability Letters*, **45**, p. 11–22, 1999

6. Engle R.F., Russell J.E. Forecasting Transaction Rates: The Autoregressive Conditional Duration Model, *UCSD Discussion Paper*, 1995

7. Engle R.F. The Econometrics of Ultra High Frequency Data, *Econometrica*, **68**, p. 1–22, 2000

8. Müller U.A. Volatility Computed by Time Series Operators at High Frequency, *UAM.1999-09-01*, 2000

9. Müller U.A. Zumbach G., Operators on Inhomogeneous Time Series, *GOZ 1998-11-01*, 2000

10. Rydberg T.H., Shephard N. A modelling framework for the prices and times of trades made at the New York stock exchange, *Nuffield College, Oxford, Working paper series*, 1999

11. Ramsay J., Silverman B. *Functional Data Analysis*, Springer-Verlag, 1997

12. Pumo B. Prediction of Continuous Time Processes by *H*-valued Autoregressive process, *Statistical Inference for Stochastic Processes*, p. 297-309, 1998