G. A. Medvedev. On Fitting the Autoregressive Investment Models to real Financial data, Transactions of the 26th International Congress of Actuaries, Birmingham, June 1998, Volume 7 “Investment”, pp. 187-211.

ON FITTING THE AUTOREGRESSIVE INVESTMENT MODELS TO REAL FINANCIAL DATA

G.A. Medvedev, Belarus

Summary

The successful investment policy is an integral part of successful activity of the insurance company. The return to the shareholders of the insurance company usually thought of as comprising the underwriting result and investment income. The investment income is very important even for an insurance company, which writes mainly a short tail business. For the successful activity the insurance company needs the appropriate investment policy as well as in good investment control.

For this purpose knowledge of the analysis of processes of a behavior of the various interest rates represents large interest. Recently many authors use the stochastic differential equations for description of processes of a development of a various sort of the interest rates. As it is known, solutions of such equations are the Markov processes and the observations of these processes in discrete instants will form time series circumscribed by autoregressive models of the first order. They are popular among those, who are interested in the analysis of financial data. In this connection it is useful to know about that as far as the models with real data will be precisely matched. In the present paper such problems with the special attention to the correspondence of correlation properties of real financial data to correlation properties of processes generated by autoregressive models are considered.

The 12 time series of the following financial data were exposed to a research UK Share Price, Dec 1918 - Jun 1995 (919 monthly values). UK Dividend Yield Rate for Shares, Dec 1918 - Jun 1995 (919 monthly values). UK Retail Prices Index, Jun 1900 - May 1995 (1145 monthly values). UK Wages Index, Jun 1920 - Apr 1995 (904 monthly values). Internal Rate of Yield on UK 2.5% Consoles, Jun 1900 - Jun 1995 (1146 monthly values). US Treasury Securities, Jan 1991 - Dec 1995 (for 1250 business days): Short-term debt instruments (the 3-month’s Bills); Medium-term debt instruments (the 3-year’s Notes); Long-term debt instruments (the 30-year’s Bonds). Currency Exchange Rates, Jan 1991 - Dec 1995 (for 1700 business days): Swiss Franc versus US $; German Mark versus US $; British Pound versus US $; Japanese Yen versus US $.

The analysis has revealed correlation properties of investigated time series of financial data and has shown, that real financial data have other correlation properties than their autoregressive models.

1. Introduction

1.1 The successful investment policy is an integral part of successful activity of an insurance company. The return to the shareholders of the insurance company is usually thought of as comprising the underwriting result and investment income. The investment income is very important even for an insurance company, which writes mainly a short tail business. The appropriate investment policy as well as good investment management are required for the successful activity of the insurance company. Knowledge of the analysis of processes of a behavior of the various interest rates is very interesting for this purpose. The various aspects of this problem have been considered in such widely known papers as Boyle (1980), Wilkie (1986, 1995), Tilley (1992), Vetzal (1992) etc., and also in common actuarial text-books (see for example, Hart, Buchanan & Howe (1996)).

1.2 It is interesting to know not only mathematical models of description of dynamics of the interest rates, but also and properties of these processes, which would allow to make in this or that measure the successful forecast of such dynamics. As it follows from the references on this problem by the most popular mathematical models of dynamics of the interest rates and derivative financial performances, depended on them, are the stochastic differential equations (for the analysis in continuous time) and Markov chains or autoregressive models (for the analysis in discrete time). The financial data are quoted in discrete time and are represented in the form of a time series. For an investigation of dynamic properties it seems more preferable to use the autoregressive models.

1.3 The temporal dependence between values of time series is very important for the purposes of prediction. The simplest forms of expression of such dependence are the correlation connections. The correlations are enough easily calculated by sample data. In the mean square theory of a prediction the optimum procedures of forecast are determined just through correlation properties of observable processes. A research of correlation properties of financial time series therefore is of interest, and also correlation properties of those mathematical models, which are taken for description of these financial time series.

1.4 The purpose of this paper is to present to the persons of the actuarial profession the features of correlation properties of the time series, which are generated by the autoregressive models, and to check up, how a time series of real financial data satisfy to these correlation properties. The 12 time series of the following financial data were exposed to a research

1) UK Share Price (more exactly the value of the all-share index without reinvestment of dividends), Dec 1918 - Jun 1995 (919 monthly values).

2) UK Dividend Yield Rate for Shares, Dec 1918 - Jun 1995 (919 monthly values).
3) UK Retail Prices Index (rate of inflation), Jun 1900 - May 1995 (1145 monthly values).
4) UK Wages Index (rate of wage inflation), Jun 1920 - Apr 1995 (904 monthly values).
5) Internal Rate of Yield on UK 2.5% Consoles, Jun 1900 - Jun 1995 (1146 monthly values).
6) US Treasury Securities, Jan 1991 - Dec 1995 (for 1250 business days):
    6.1) Short-term debt instruments (the 3-month's Bills);
    6.2) Medium-term debt instruments (the 3-year's Notes);
    6.3) Long-term debt instruments (the 30-year's Bonds).
7) Currency Exchange Rates, Jan 1991 - Dec 1995 (for 1700 business days):
    7.1) Swiss Franc (SFr) versus US $;
    7.2) German Mark (DM) versus US $;
    7.3) British Pound (BP) versus US $;
    7.4) Japanese Yen (Y) versus US $.

2. Autoregressive Models and Stochastic Differential Equations

2.1 An observed financial time series $X_1, X_2, \ldots , X_n$ can be thought of as a sample realization of a stochastic process. However, all real processes, explicating in a nature and community, vary in continuous time irrespective of the fact, which the mathematical models identify them. The processes circumscribing a conjuncture of the financial market are not elimination. Therefore it is necessary to set the more exact accordance between the mathematical models of the time series (process with discrete time) and the stochastic process in the continuous time. The most popular mathematical models of the stochastic processes in the continuous time are the stochastic differential equations that can be written as

$$dx = \mu(x,t)\, dt + \theta(x,t)\, dW(t)$$

where $\mu(x,t)$ and $\theta(x,t)$ are the continuously differentiable deterministic functions of its arguments; $W(t)$ is a standard Wiener process. In accordance with the Doob (1953) theorem in these conditions the stochastic process $x(t)$ is a Markov process. For our purposes it is sufficient to consider a more simple case of the linear stochastic differential equation when

$$dx = (a(t)\, x + b(t))\, dt + c(t)\, dW(t) \tag{1}$$

where $a(t)$, $b(t)$, and $c(t)$ are again the continuously differentiable deterministic functions. Note that the stochastic process $x(t)$ in these equations can be both a scalar and a vector process. Consider the more general case of vector process $x(t)$. In this case $a(t)$ and $c(t)$ are the matrix functions of the appropriate dimensions and $b(t)$ is the vector of the dimension of the vector $x(t)$. The solution of the equation for this case can be written in the form (more exactly, the solution is stochastically equivalent to following stochastic process)

$$x(t) = U(t,s)x(s) + \int_s^t U(t,v)b(v)dv + \xi(s,t) \tag{2}$$

where $x(s)$ is given initial vector of process at time s ; $U(t,s)$ is the fundamental matrix of the solution (FMS) of the homogeneous deterministic system

$$\frac{dx}{dt} = a(t)x(t)$$

with initial condition $x(s)$ at time s . $\xi(s,t)$ is the random vector with the zero mean and the correlation matrix

$$\int_s^t U(t,v)c(v)c^T(v)U^T(t,v)dv .$$

2.2 It is convenient to transform the solution to some different and more useful form. For it we introduce some new notation. Because the correlation matrix is the positive definite matrix then it is possible to introduce the other positive definite matrix $\sigma(s,t)$ by the relation

$$\sigma^2(s,t) = \int_s^t U(t,v)c(v)c^T(v)U^T(t,v)dv .$$

Let us designate

$$\mathrm{z}(t) = \int_u^t U(t,v)b(v)dv , \quad Z(t) = x(t) - \mathrm{z}(t) . \tag{3}$$

As the matrix $U(t,s)$ has a property of a factorization $U(t,s) = U(t,v)U(v,s)$ for any t, s, and v, it is possible to write for any u (concerning a parameter u it will be told below)

$$\int_s^t U(t,v)b(v)dv = \mathrm{z}(t) - U(t,s)\, z(s) .$$

Using the new notation the solution (2) of system (1) can be represented in the form

$$Z(t) = U(t,s)\, Z(s) + \sigma(s,t)\, V(t) \qquad (4)$$

where $V(t)$ is the normal random vector that has the following expectations

$$\mathbf{E}\{V(t)\} = 0,\ \mathbf{E}\{V(t)\, V^T(t)\} = \boldsymbol{I}\ ,\ \mathbf{E}\{V(t)\, V^T(s)\} = 0\ ,\ \text{for any } t \neq s\ ,$$

where $\boldsymbol{I}$ is the identity matrix.

The solution of the equation (2) in the form (4) is similar of the many-dimensional autoregression with the variable coefficients. Make this similarity more evident. For it consider a particular case of the constant matrix coefficients in the equation (2). That is $a(t)$, $b(t)$, and $c(t)$ are independent on t . In this case FMS is a matrix that depends only on the single argument, $U(t,s) = U(t-s)$ and $\sigma(s,t) = \sigma(t-s)$.

2.3 Let us assume, that the values of process that is determined by equation (2) only in the discrete instants from some subset are accessible to observation. Let this subset is $\{t_i \mid i = 1, 2, ... ;\ t_{i+1} - t_i = h$ for any $i\}$. And let us put $Z(t_k) = Z_k$, $V(t_k) = V_k$. Then (4) can be written in the form

$$Z_k = U(h)\, Z_{k-1} + \sigma(h)\, V_k \qquad (5)$$

that is the many-dimensional (vector) autoregression. Let us note, that as the matrix $U(h)$ is nondegenerate and $\mathbf{E}\{V_k\} = 0$ for any k, from (5) follows, that $\mathbf{E}\{Z_k\} = 0$. Returning to initial notation $x(t) = z(t) + Z(t)$ and setting $x(t_k) = X_k$, we see that this is equivalent to equality

$$\mathbf{E}\{Z_k\} = \mathbf{E}\{X_k\} - z(t_k) = 0 \quad \text{or} \quad \mathbf{E}\{X_k\} = z(t_k)$$

If the matrix

$$\int_0^\infty U(t) U^T(t)\, dt$$

exists then such stochastic process is stationary. In this case it is relevant to consider a sense of the function $z(t)$ in (3). In a common sense the integral $z(t)$ under the lower limit of integration $u = -\infty$ gives the expectation of the stationary process in the form

$$z(t) = \int_{-\infty}^{t} U(t-v)\, b\, dv = \int_0^\infty U(s)\, b\, ds = \mu$$

In alternative case the lower limit of integration u can be considered as the initial instant of development of process then the function $z(t)$ can be considered as expectation of process of an establishment to a stationary average value, that can be a trend of an stochastic process.

Thus for a stationary process the relation (5) can be rewritten by the initial notation as

$$X_k = \mu + U(h)(X_{k-1} - \mu) + \sigma(h) V_k \tag{6}$$

When the dimension of model n is equal to 1, all above-stated relations become scalar as well as all terms of the equations of a autoregressive models (5) and (6).

2.4 As an example we shall consider widely known the Wilkie (1986) investment model that describes the stochastic behavior of the inflation process. On this model the U. K. Retail Prices Index $Q(t)$, based on annual (h = 1) data from the period 1919 to 1982, will form the time series for the force $I(t) = \ln(Q(t)/Q(t\text{-}1))$ of inflation over the year (t-1,t)

$$I(t) = 0.05 + 0.6\times(I(t-1) - 0.05) + 0.05\times V(t) \tag{7}$$

where $V(t)$ is a series of independent, identically distributed unit normal variables, i.e. $\mathbf{E}\{V(t)\} = 0$, $\mathbf{E}\{V^2(t)\} = 1$, and $\mathbf{E}\{V(t)V(s)\} = 0$ $t \neq s$. For this case

$$U(h) = e^{-ah} = 0.6,\ z(t) = \mu = b/a = 0.05\ ,\ \sigma(h) = c\sqrt{\left(1-e^{-2ah}\right)/2a} = 0.05,\ h = 1.$$

Hence the stochastic differential equation (1) that corresponds to this autoregression takes the form

$$dx = (-0.511\,x + 0.0255)\,dt + 0.0632\,dW(t)$$

2.5 Other the Wilkie (1995) investment model is two-dimensional. It is an autoregressive model in that the force of the wage inflation $J(t)$ over year (t – 1, t) is added to the force of the inflation $I(t)$. Then for $\tilde{I}(t) = I(t) - \mu_I$ and $\tilde{J}(t) = J(t) - \mu_J$ the autoregressive model had been obtained by Wilkie in the form

$$\begin{pmatrix} \tilde{I}(t) \\ \tilde{J}(t) \end{pmatrix} = \begin{pmatrix} 0.1817 & 0.5927 \\ 0.1724 & 0.5618 \end{pmatrix} \begin{pmatrix} \tilde{I}(t-1) \\ \tilde{J}(t-1) \end{pmatrix} + \begin{pmatrix} 0.0382 & 0.0142 \\ 0.0142 & 0.0303 \end{pmatrix} \begin{pmatrix} V_1(t) \\ V_2(t) \end{pmatrix} \tag{8}$$

where μ_I = 0.0359 and μ_J = 0.0509, and $V_1(t)$, $V_2(t)$ are independent, identically distributed unit normal variables, i.e. $\mathbf{E}\{V_k(t)\} = 0$, $\mathbf{E}\{V_k{}^2(t)\} = 1$ for k = 1,2 and $\mathbf{E}\{V_1(t)V_2(t)\} = 0$ for any t . In this case the fundamental matrix is

$$U(t) = \begin{pmatrix} 0.2445 \times e^{\lambda t} + 9.5173 \times e^{\mu t} & 0.7970 \times \left(e^{\lambda t} - e^{\mu t}\right) \\ 0.2318 \times \left(e^{\lambda t} - e^{\mu t}\right) & 0.2445 \times e^{\mu t} + 9.5173 \times e^{\lambda t} \end{pmatrix}$$

where $\lambda = -0.2962$, $\mu = -20.6248$.

The stochastic differential equation (1) that corresponds to the two-dimensional autoregression (8) takes the form

$$\begin{pmatrix} d\tilde{I} \\ d\tilde{J} \end{pmatrix} = \left(a \begin{pmatrix} \tilde{I}(t) \\ \tilde{J}(t) \end{pmatrix} + b \right) dt + c \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix},$$

where

$$a = \begin{pmatrix} -3.3812 & 3.2547 \\ 0.9465 & -17.5398 \end{pmatrix}, \quad b = \begin{pmatrix} 0.05335 \\ 0.00381 \end{pmatrix}, \quad c = \begin{pmatrix} 0.01495 & 0.02038 \\ 0.00185 & 0.00252 \end{pmatrix}.$$

2.6 Thus, we see, that the use of observations of a process $x(t)$, that is a solution of the stochastic differential equation in discrete instants , results in the autoregression model of the first order. According to the theorem of the Doob the time series generated by such model make up a Markov process. The other theorem of the Doob (1953) states, that the scalar stationary stochastic process is Markov one if and only if its correlation function is exponential. It means that for the time series that are generated by the stochastic differential equations should have a property

$$C(u) = Cov\left\{X_k X_{k+u}\right\} / \sqrt{Var\left\{X_k\right\} Var\left\{X_{k+u}\right\}} = \exp\{-\rho u\} \quad , \ u \geq 0 . \tag{9}$$

where the positive parameter ρ is determined in appropriate way. In a consequence we shall take advantage of these properties for the analysis of a goodness of fit of financial time series to models of an autoregression.

3. Some notes about the stationary and correlation properties of autoregressive processes

3.1 More often authors deal with stationary models of financial time series when the factors of a model are constant. This model generates the stationary (of second order) time series X_k , that has properties

$$\mathbf{E}\{X_k^2\} < \infty , \ \mathbf{E}\{X_k\} = m , \ \mathbf{E}\{X_k X_{k+u}\} = f(|u|)$$

for all k, u = 0, ±1, ±2, ... Note that the stationarity properties are always determined for infinity index set and are used for theoretical analysis while the time series observed in practice are given for finite index set. Therefore for practical use the stationarity properties should be modified. In practice the situation becomes more complicated as the systematic and seasonal components can be added to a stationary time series. Usually the first step in the analysis of real time series is to plot the data. This provides to make the classical decomposition of time series on three components (see for example Brockwell & Davis (1987)):

$$m_k + s_k + X_k$$

a slowly changing function (trend component) m_k , a seasonal component s_k , and a random component X_k . The trend component can include both a systematic deterministic component and a slowly changing random component of stationary time series. If a observation time period is less than a seasonal cycle then seasonal component s_k can be included too into the trend m_k . Thus even if the systematic and seasonal components are absent then a trend component m_k can be determined as a slowly changing random component. In this case decomposition is rather conditional. Later under an analysis of real financial data we take such decomposition. The random component X_k is usually considered as a realization of the stationary process and often described as an time series generated by the autoregressive model.

3.2 Most often used model is an autoregressive-moving average (ARMA) model. ARMA(p,q) - the autoregresive-moving average model of degree (p,q) generate a stationary time series X_k by following difference equation

$$X_k = \sum_{i=1}^{p} a_i X_{k-i} + \sum_{j=0}^{q} b_j V_{k-j} \ , \ k = 0, \pm 1, \pm 2, ... \qquad (10)$$

where a_i and b_j are constant and V_k is a standard white noise process, $\mathbf{E}\{V_k\} = 0$, $\mathbf{E}\{V_k^2\} = 1$, $\mathbf{E}\{V_k V_{k+u}\} = 0$, $u \neq 0$. Note that if some time series Y_k is a linear transformation of another time series X_k that is an ARMA process then Y_k is an ARMA process too. For example if $Y_k = \sum_{t=0}^{r} c_t X_{k-t}$ where X_k is an ARMA(p,q) process generated by (10) then the time series Y_k is the ARMA($p,q+r$) process

$$Y_k = \sum_{i=1}^{p} a_i Y_{k-i} + \sum_{j=0}^{q+r} d_j V_{k-j} \ , \ k = 0, \pm 1, \pm 2, ... \qquad (11)$$

where the coefficients d_j are easy determined by b_j and c_j . It is important for example to analyze of the yield to maturity. The yield to maturity is a measure of the average rate of return that will be earned on a bond if it is bought now and held until maturity. If the underlying short interest rate X_k is generated by the ARMA(1,0) (i.e.

AR(1)) process as often it is assumed then the yield to maturity Y_k for a default-free discount bond maturing in time T is generated by the ARMA(1,T) process.

3.3 There are some methods to calculate the correlation function of the ARMA(p,q) processes (see for example Brockwell & Davis (1987)). We will represent here the most simple cases that are necessary for our purposes. For the AR(1) (i.e. ARMA(1,0)) process

$$X_k = aX_{k-1} + \sigma V_k$$

the correlation function is

$$C(u) = E\{X_k X_{k+u}\} / \sqrt{E\left\{X_k^2\right\} E\left\{X_{k+u}^2\right\}} = a^u , \ u > 0 .$$

For the AR(2) process

$$X_k = a_1 X_{k-1} + a_2 X_{k-2} + \sigma V_k$$

the correlation function is

$$C(u) = \frac{\left(g_2^2 - 1\right) g_1^{1-u} - \left(g_1^2 - 1\right) g_2^{1-u}}{\left(g_2^2 - 1\right) g_1 - \left(g_1^2 - 1\right) g_2} \ , u > 0 . \qquad (12)$$

where g_1 and g_2 are the different roots of the equation $1 - a_1 g - a_2 g^2 = 0$ (for stationary time series these roots are $|g_k| > 1$, $k = 1, 2$). If these roots are complex conjugate, $g_1 = fe^{-i\varphi}$ and $g_2 = fe^{i\varphi}$, then this form is transforming to

$$C(u) = f^{-u} \frac{\sin(u\varphi + \psi)}{\sin\psi} \quad , \qquad \tan\psi = \frac{f^2 + 1}{f^2 - 1} \tan\varphi . \qquad (13)$$

Thus for all time series, that are generated by the AR(1) models with positive coefficient a, the correlation functions are a positive monotone decrease functions. This follows also from the properties of stochastic differential equations that generate the Markov processes (see 2.6 above). Conversely if the correlation function of some time series takes the negative values then such time series can not be generated by the AR(1) model. It means too that the processes, which are observed by these time series, cannot be generated by the stochastic differential equations.

4. Real Financial Data Analysis

4.1 The analysis of real financial data was made uniformly for all time series (with small exceptions) and was carried out under the following design:

- The plot was constructed for full volume of series.
- The slowly changing component (the trend) m_k was selected. In the capacity of this component the approximation of time series by the method of least squares in the form of a power polynomial of the sixth order got out in all necessary cases.

$$m_k = b_0 + b_1 k + b_2 k^2 + b_3 k^4 + b_5 k^5 + b_6 k^6 , \quad k = 0, 1, 2, \ldots , N - 1 ,$$

where N is the series length (sample size).

- As some authors (see e.g. Black & Karasinski(1991), Tilley (1992)) offer instead of initial time series to consider a series composed from logarithms of data therefore such series were investigated in the considered analysis too.
- The random component of time series X_k was determined as a residual between an initial series and slowly changing component (trend) m_k . Thus obtained time series (initial and residual) were considered as process of an autoregression which was subject to the further analysis.
- The factors of an autoregression of models of a type AR(1) and AR(2) were determined for initial time series and series of residuals on a method of least squares.
- The correlation functions for time series were calculated on base of theoretical relations (see above 3.3) for the autoregressive models and the sample correlation functions. Because of the coefficients of trend functions were calculated by sample data the trend can be considered as a random function too. Therefore the correlation functions of trends were also calculated. Then were made a comparison of all these correlation functions.
- The results are submitted as the graphs of time series, trends and correlation functions. The factors of autoregressive models and trends are submitted in the form of tables.

4.2 ***UK Share Price*** (more exactly the value of the all-share index without reinvestment of dividends), Dec 1918 - Jun 1995 (919 monthly values). About sources of these data see Appendix F of Wilkie (1995). The general view of data is submitted on Figure 1. This Figure shows the logarithms of the relations the share price index (*SPI*) of month k to the share price index of month $k - 1$ where k is a serial number of month in a sequence of months from December 1918 till June 1995 . From figure it is visible, that the trend of transformed share price indexes is practically absent, therefore it was not investigated for these data. This time series had been considered very detailed by Wilkie (1995) therefore we bring only the autoregressive factors of the AR(1) (a = 0.1193) and AR(2) (a_1 = 0.1273, a_2 = –0.0689) models for this time series and the theoretical and sample correlation functions (on Figure 2).

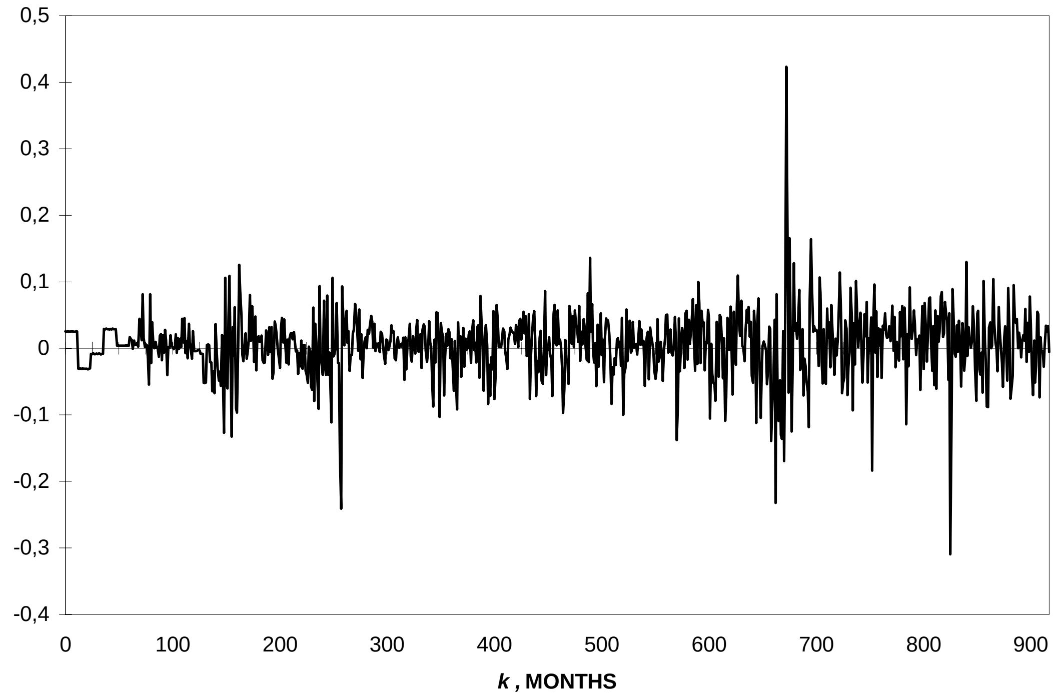

Figure 1. The data about the Share Price Indices (*SRI*) that are transformed by relation $\ln\left(SRI(k)/SRI(k-1)\right)$ where k is the serial number of month of the sequence from December 1918 (k = 0) to June 1995 (k = 918).

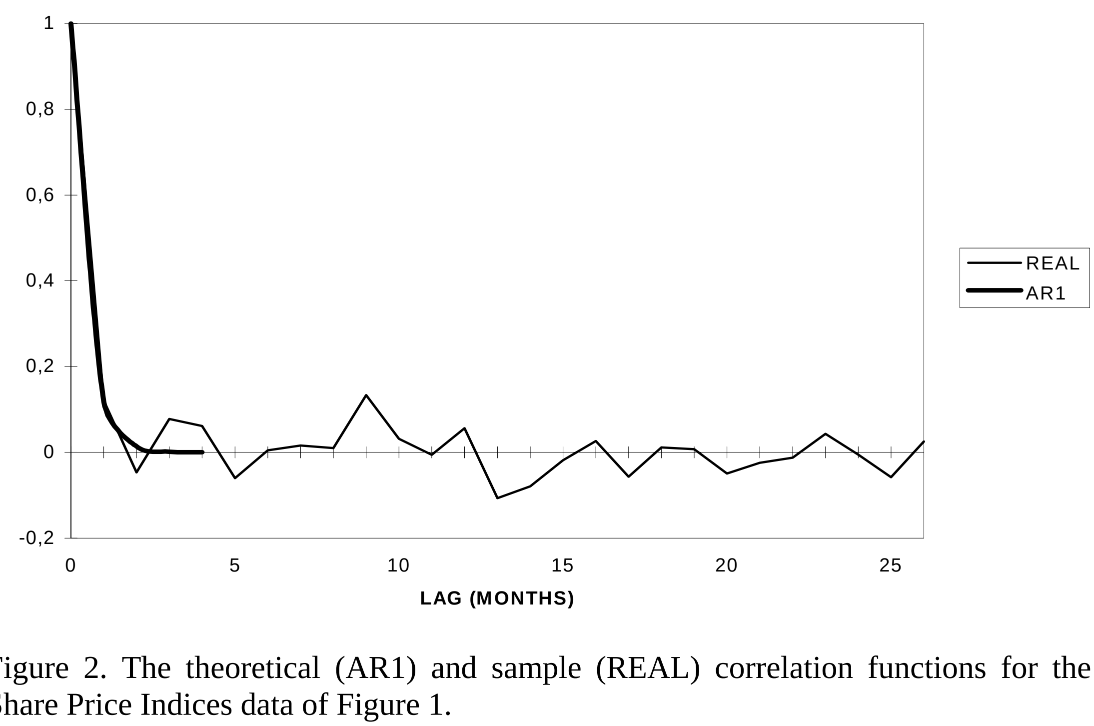

Figure 2. The theoretical (AR1) and sample (REAL) correlation functions for the Share Price Indices data of Figure 1.

4.3 ***UK Dividend Yield Rate for Shares***, Dec 1918 - Jun 1995 (919 monthly values). About sources of these data see Appendix F of Wilkie (1995). Note that this time series had been considered also detailed by Wilkie (1995). The general view of data is submitted on Figure 3. This figure shows the logarithms of the relations the dividend yield rate of month k to the dividend yield rate of month $k-1$ where k is a serial number of month in a sequence of months from December 1918 till June 1995 . On this figure the slowly changing component (trend) of these data is shown too.

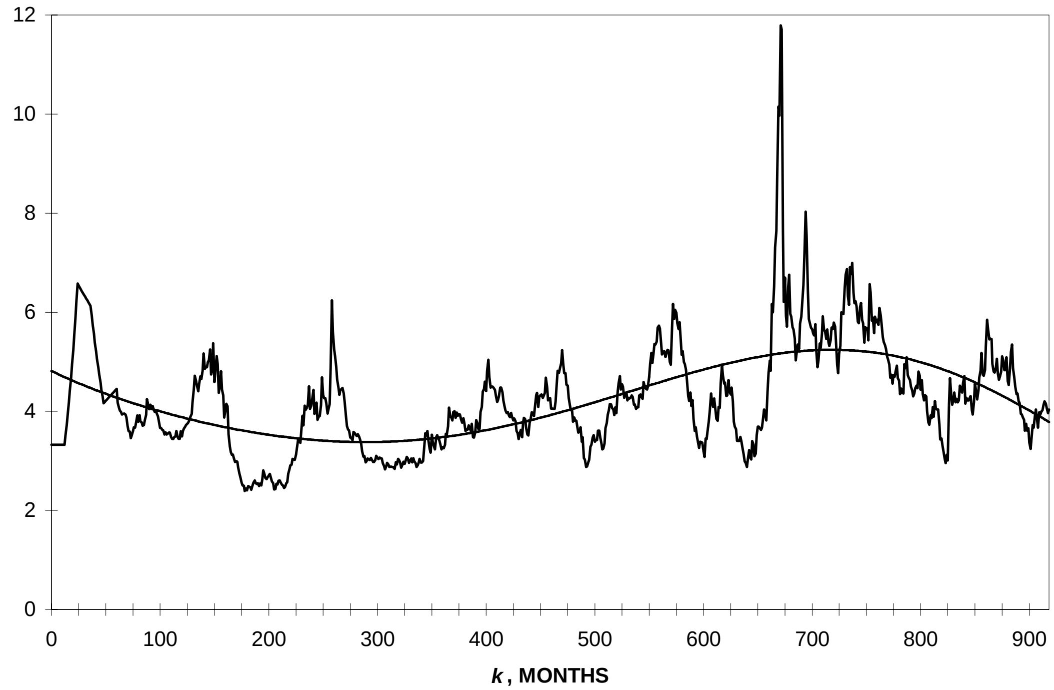

Figure 3. The data about the Dividend Yield Rate for Shares (*DYR*) that are transformed by relation $\ln\left(DYR(k+1)/DYR(k)\right)$ where k is the serial number of month of the sequence from December 1918 (k = 0) to June 1995 (k = 918).

Figure 4 shows some correlation functions for data of Figure 3. Among them the sample correlation function of the transformed dividend yield rate for shares (briefly called YIELD, note that this function repeats the graph on Figure 4.3 from Wilkie (1995)), the sample correlation function of the logarithms of the transformed dividend yield rate for shares (briefly called LN(YIELD)), the correlation function of the trend (TREND) and the theoretical correlation function of AR(1) model for data of Figure 3. For these data the AR(1) model has the factor a = 0.9982 and the AR(2) model has the factors a_1 = 1.1500, a_2 = – 0.1521. Figure 5 shows the correlation functions for the residuals that are obtained by subtraction the trend from the sample data: the sample correlation function (REAL) and the theoretical correlation functions for AR(1) model (a = 0.9581) and AR(2) model (a_1 = 1.1262, a_2 = – 0.1753).

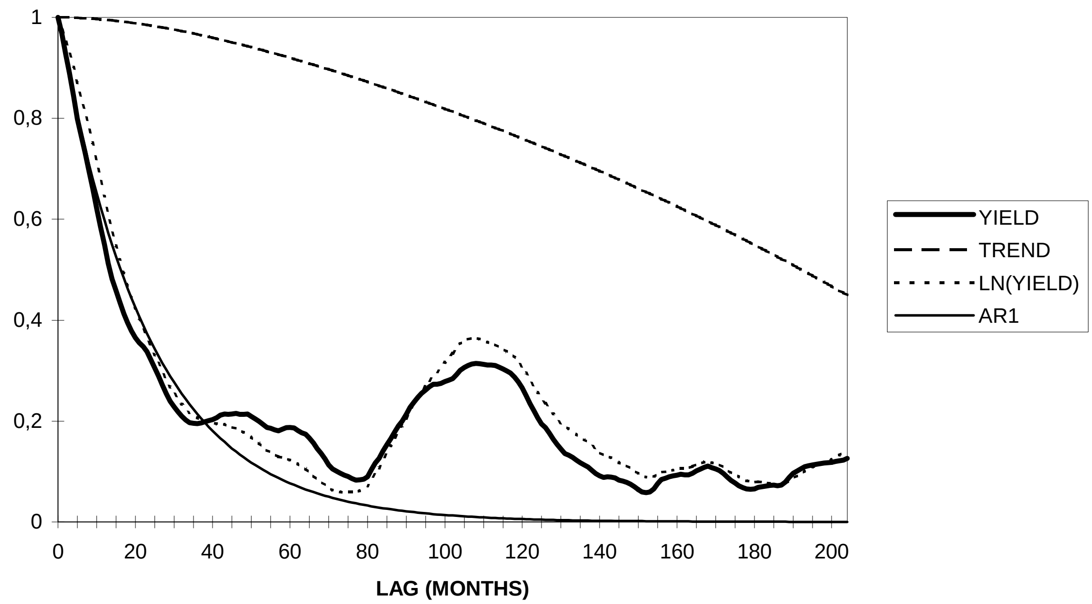

Figure 4. Correlation functions for data of Figure 3.

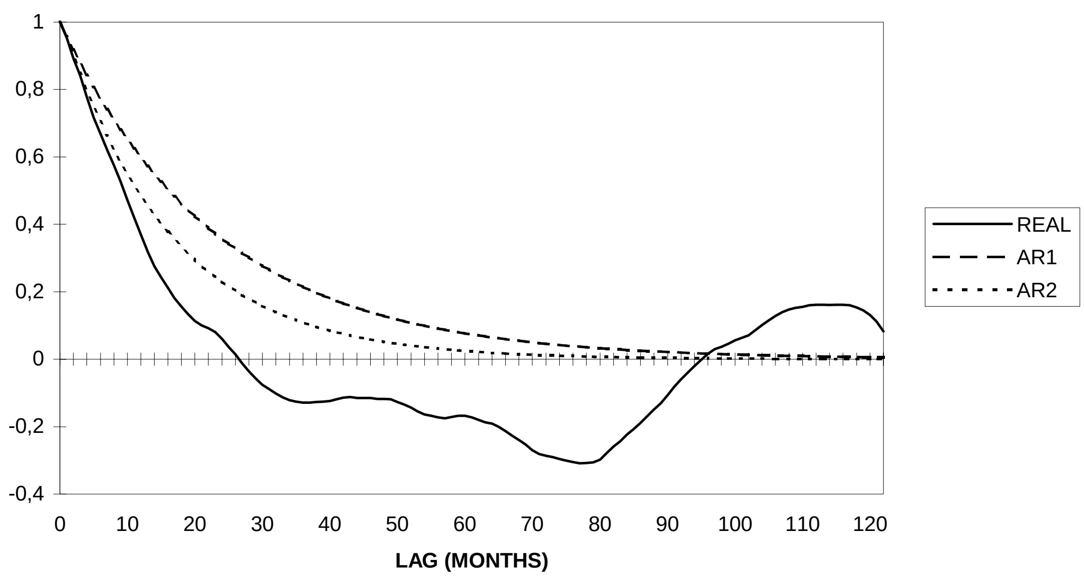

Figure 5. Correlation functions for the residuals of data of Figure 3.

From these figures it is possible to note that the logarithm of data of Figure 3 has practically the same sample correlation function what has also data. Next the correlation function of the residuals of data after subtraction of trend can take the rather significant negative values though the correlation functions of the AR(1) and AR(2) models are always nonnegative.

4.4 ***UK Retail Prices Index*** (rate of inflation), Jun 1900 - May 1995 (1145 monthly values). About sources of these data too see Appendix F of Wilkie (1995). Note that this time series had been considered also very detailed by Wilkie (1986, 1995). The general view of data is not submitted here because of it had been represented on Figure 2.2 in Wilkie (1995). Here we represent on Figure 6 only the correlation functions for this time series: sample function (REAL, this function was more detailed represented on Figures 2.3 - 2.5 in Wilkie (1995)) and two theoretical functions for AR(1) and AR(2) models with the autoregressive factors a = 0.4612 and a_1 = 0.3334, a_2 = 0.2770 respectively .

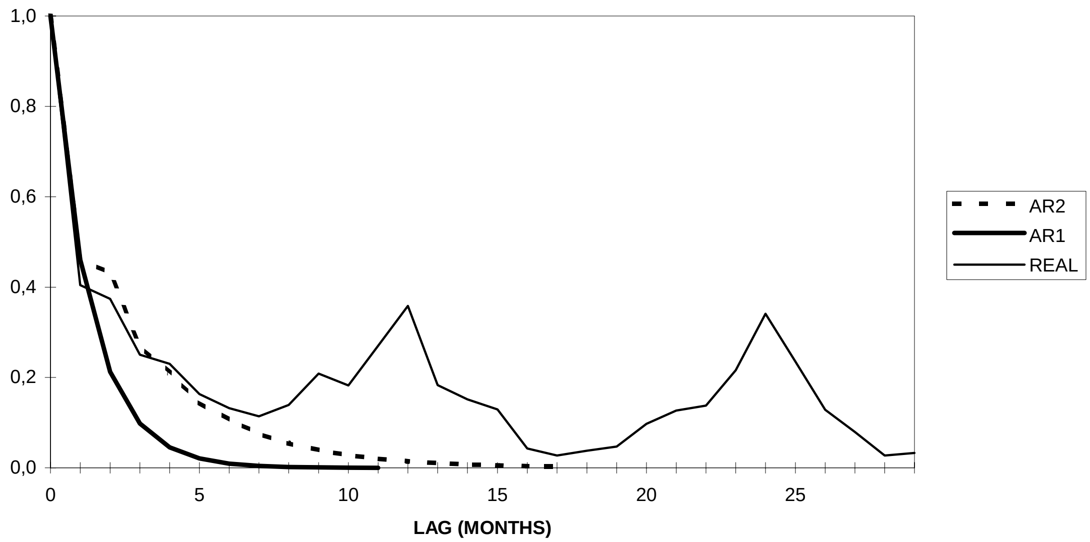

Figure 6. Correlation functions of monthly inflation.

The sample correlation functions of the Retail Prices Index (rate of inflation) has an explicitly expressed seasonal component with a cycle 12 months. For creation of an autoregressive model in this case it is necessary or to eliminate a seasonal component from initial time series (see 3.1) or to build a model of an autoregression of greater than 12th order.

4.5 ***UK Wages Index*** (rate of wage inflation), Jun 1920 - Apr 1995 (904 monthly values). About sources of these data too see Appendix F of Wilkie (1995). This time series had been considered also by Wilkie (1995). The general view of data is not submitted here. Here we represent on Figure 7 only the correlation functions for this time series: sample function (REAL) and two theoretical functions for AR(1) and AR(2) models with the autoregressive factors a = 0.3060 and a_1 = 0.2511, a_2 = 0.1735 respectively. The remarks about form of the sample correlation function in this case are the same as for previous one. The sample correlation functions of the Wage Index (rate of wage inflation) has an explicitly expressed seasonal component with a cycle 12 months. Therefore problem of a construction of an autoregressive model in this case is same as well as in previous.

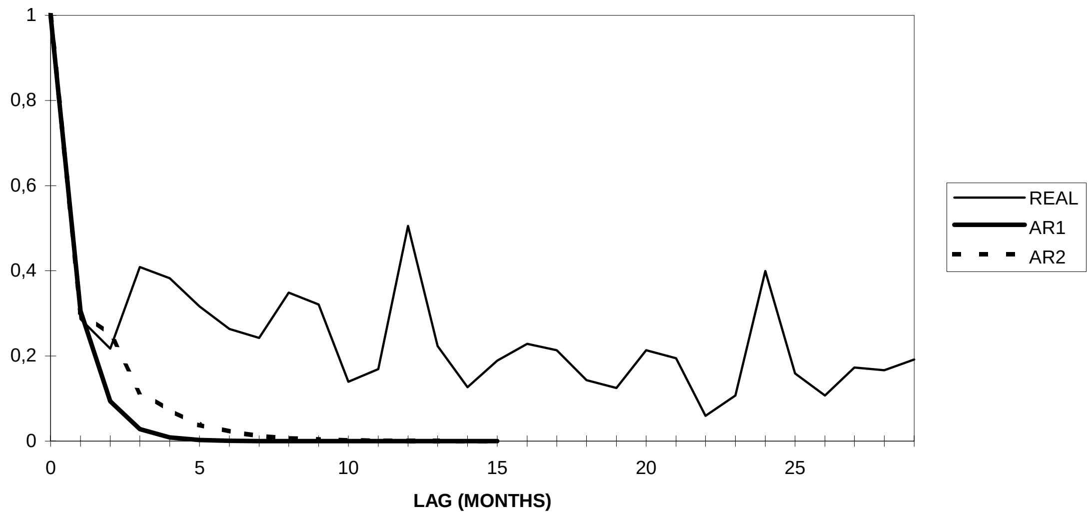

Figure 7. Correlation functions of monthly rate of wage inflation.

4.6 ***Internal Rate of Yield on UK 2.5% Consols***, Jun 1900 - Jun 1995 (1146 monthly values). About sources of these data too see Appendix F of Wilkie (1995). The general view of data is submitted on Figure 8. This figure shows the Internal Rate of Yield on UK 2.5% Consols as function of a serial number of month in a sequence of months from June 1900 till June 1995 . On this figure the slowly changing component (trend) of these data is shown too.

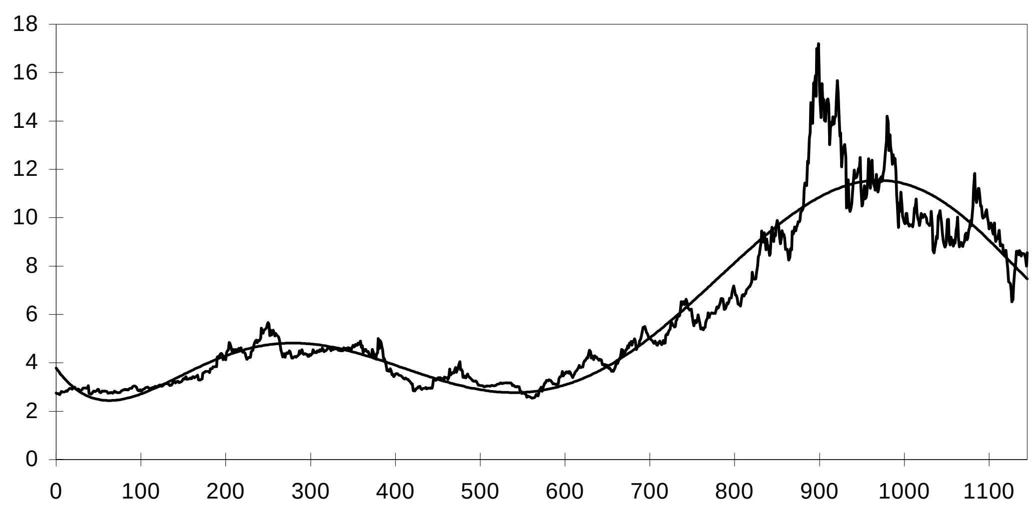

Figure 8. Consols yield and its trend as function of serial number k of month from June 1900 (k = 0) till June 1995 (k = 1145).

Figure 9 shows some correlation functions for data of Figure 8. Among them the sample correlation function of the yield on consols (briefly called YIELD), the sample correlation function of the logarithms of the yield on consols (briefly called LN(YIELD)), the correlation function of the trend (TREND) and the theoretical correlation functions of the AR(1) and AR(2) models for data of Figure 8. For these data the AR(1) model has the factor a = 0.9998 and the AR(2) model has the factors a_1 = 0.9568, a_2 = 0.0431. Figure 10 shows the correlation functions for the residuals that are obtained by subtraction the trend from the sample data: the sample correlation function (RES) and the theoretical correlation functions for AR(1) model (a = 0.9581) and AR(2) model (a_1 = 1.1262, a_2 = – 0.1753).

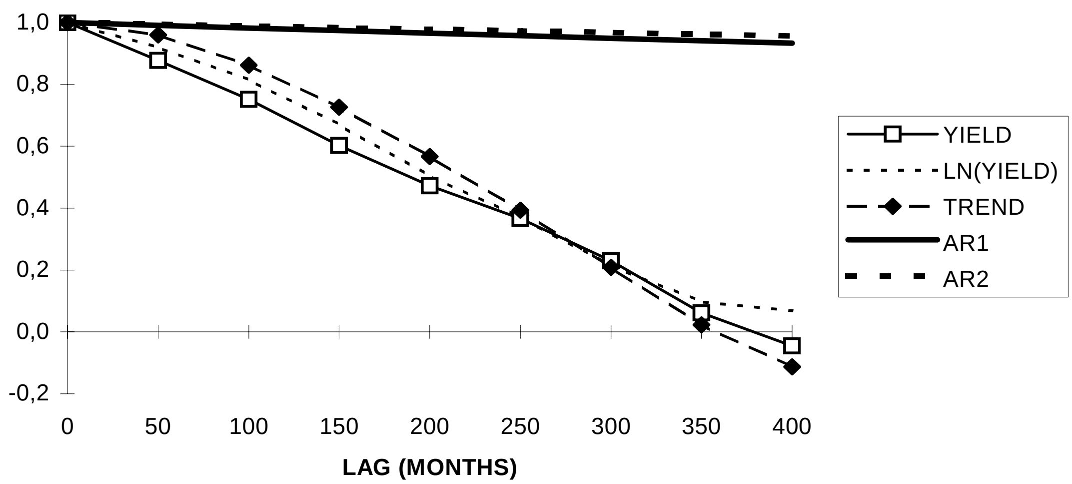

Figure 9. Correlation functions of consols yield.

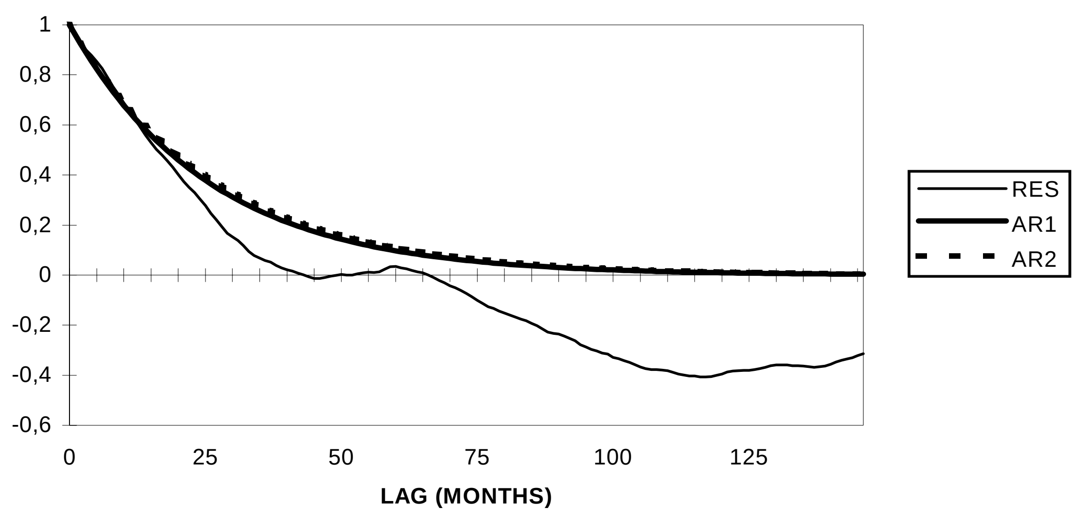

Figure 10. Correlation functions for the residuals of data of Figure 8.

The remarks about results represented on Figures 9 and 10 are same as well as in a ratio of results on Figures 4 and 5. The theoretical correlation functions decrease essentially slower than the sample correlation functions. And the sample correlation function can take a negative values.

4.7 ***US Treasury Securities***, Jan 1991 - Dec 1995 (for 1250 business days). Among them Short-term debt instruments (the 3-month's Bills), Medium-term debt instruments (the 3-year's Notes), Long-term debt instruments (the 30-year's Bonds). These data had been taken from yield curve rates that are updated by the Internet (from address gopher://una.hh.lib.umich.edu:70/00/ebb/monetary/yc/ycurve.tre). The general view of data is represented on Figure 11. This figure shows the US Treasury yield rates as functions of a serial number of business day in a sequence of business days from 2 January 1991 till 5 January 1996 . On this figure the slowly changing components (trends) of these data are shown too.

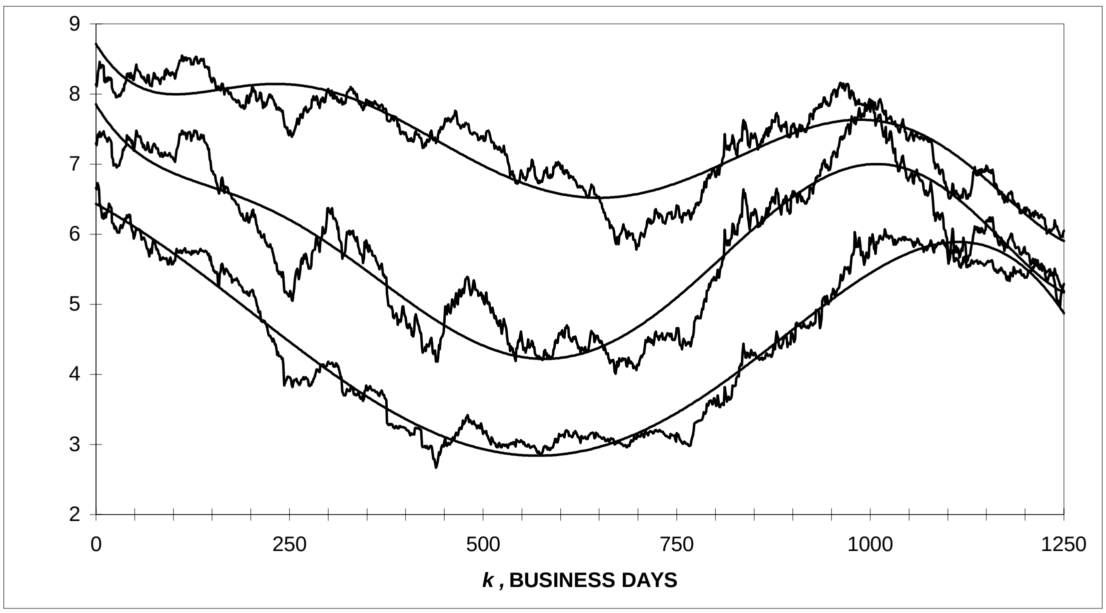

Figure 11. US Tresury yield rates (%) as the functions of serial number k of business day from 2 January 1991 (k = 0) to 5 January 1996 (k = 1250) for the 3-month's Bills (lower curve), the 3-year's Notes (curve in a middle), and the 30-year's Bonds (upper curve). The smooth curves are the relevant polynomial trends.

The character of correlation functions for all variants of the yield rates is identical therefore for economy of a place we bring here only two figures of correlation functions. Figure 12 shows some correlation functions for data of Figure 11. Among them the sample correlation function of the yield rates for 30-year's Bonds (briefly called RATE), the sample correlation function of the logarithms of these yield rates (briefly called LN(RATE)), the correlation function of the trend (TREND) and the theoretical correlation functions of the AR(1) model for data of Figure 11 (upper curve). For these data the AR(1) model has the factor a = 0.9998. Figure 13 shows the correlation functions for the residuals that are obtained by subtraction the trend from the sample data of the yield rates of the 3-month's Bills (lower curve on Figure 11): the sample correlation function (REAL) and the theoretical correlation functions for AR(1) model (a = 0.97898) and AR(2) model (a_1 = 1.0739, a_2 = – 0.09696).

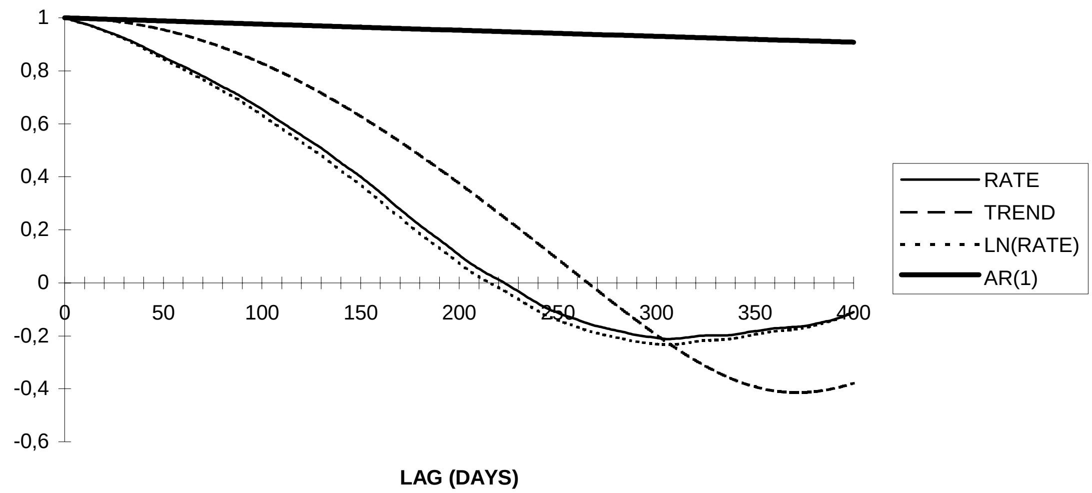

Figure 12. Correlation functions for the yield rates for 30-year's Bonds.

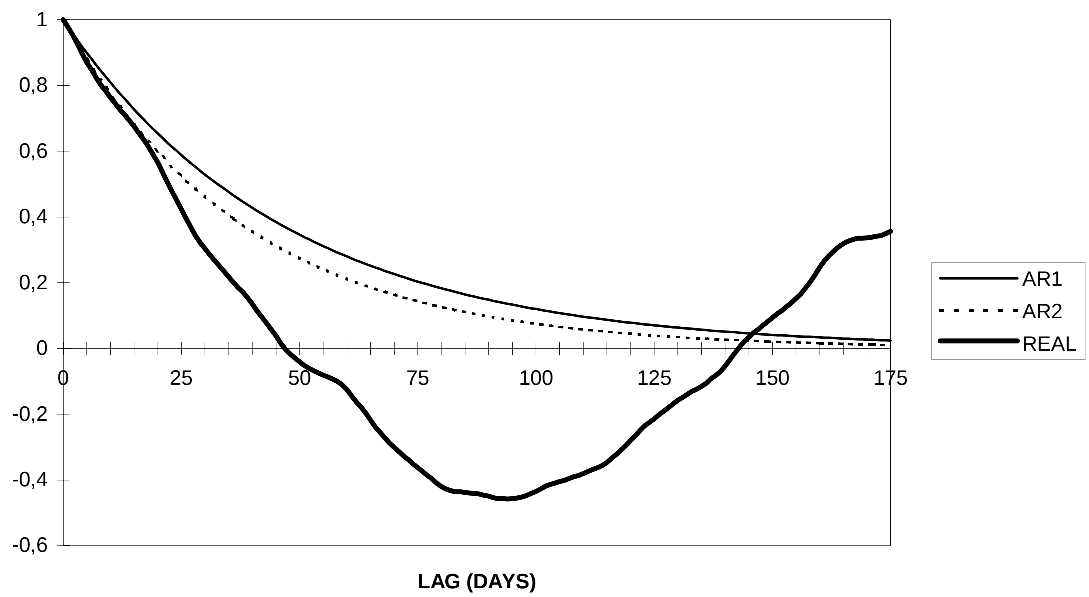

Figure 13. Correlation functions for the residuals of data for the 3-month's Bills (lower curve of Figure 11).

4.8 ***Currency Exchange Rates***, Jan 1991 - Dec 1995 (about 1700 business days). Among them British Pound (BP) versus US \$, German Mark (DM) versus US \$, Japanese Yen (Y) versus US \$, and Swiss Franc (SFr) versus US \$. The time series of these financial data had been received from the Trade System «Dow Jones Telerate».

The general view of data is represented on Figures 14 - 17. This figures show the Currency Exchange Rates as functions of a serial number of business day in a sequence of business days from January 1991 till December 1995 . On this figures the slowly changing components (trends) of these data are shown too.

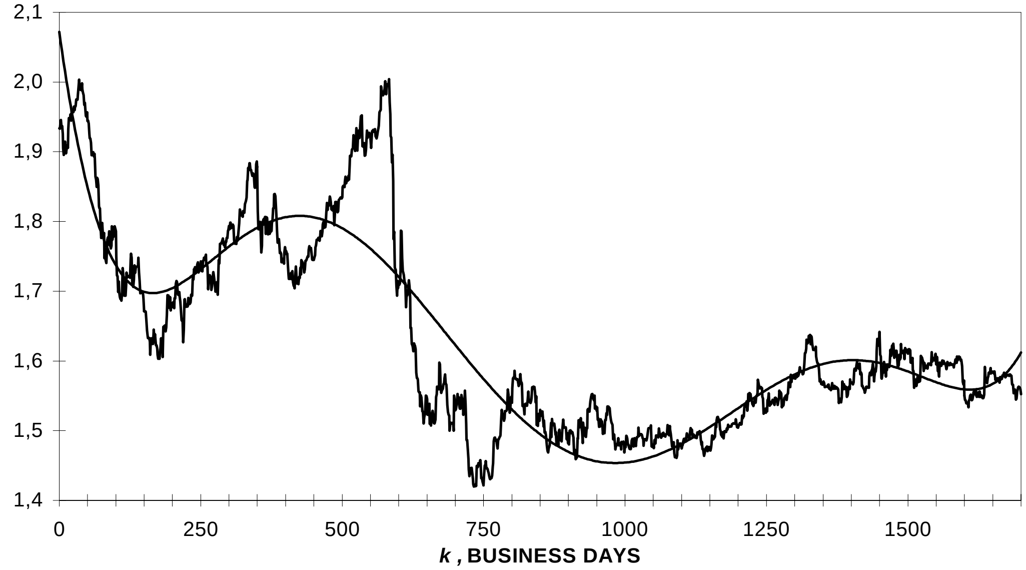

Figure 14. British Pound versus US $ as the functions of serial number k of business day from 1 January 1991 (k = 0) to 31 December 1995 (k = 1728).

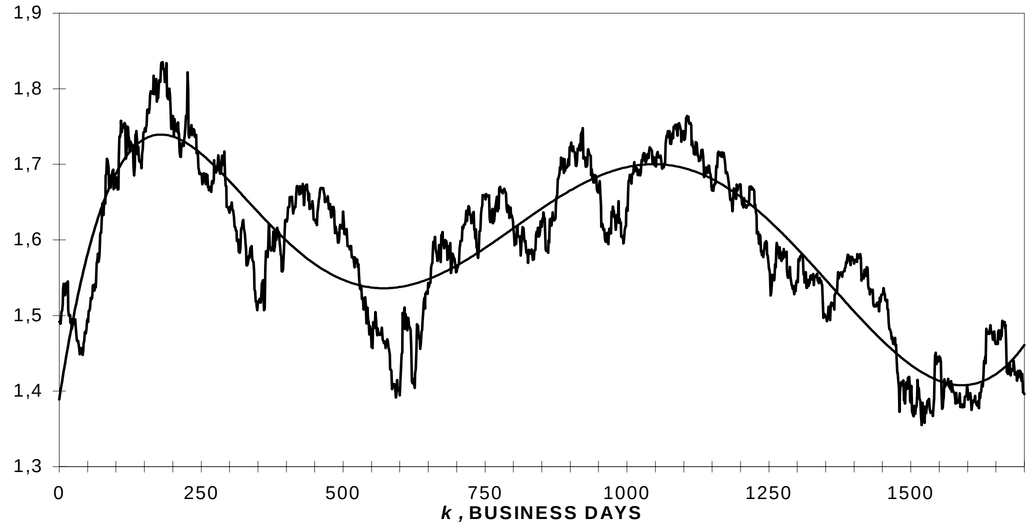

Figure 15. German Mark versus US $ as the functions of serial number k of business day from 1 January 1991 (k = 0) to 31 December 1995 (k = 1763).

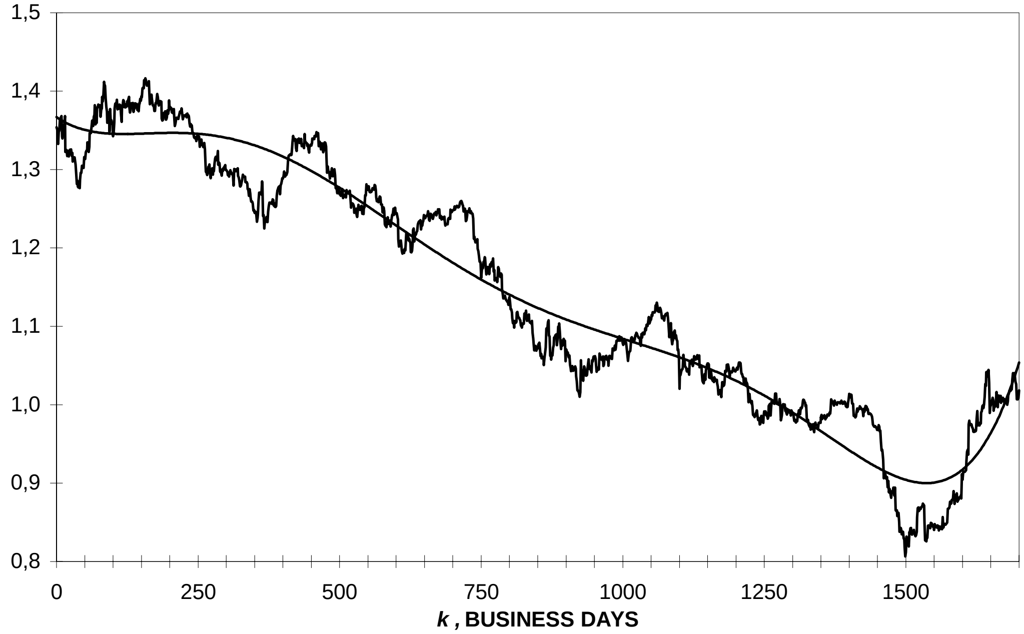

Figure 16. Japanese Yen (divided by 100) versus US \$ as the functions of serial number k of business day from 1 Jan 1991 (k = 0) to 29 Dec 1995 (k = 1740).

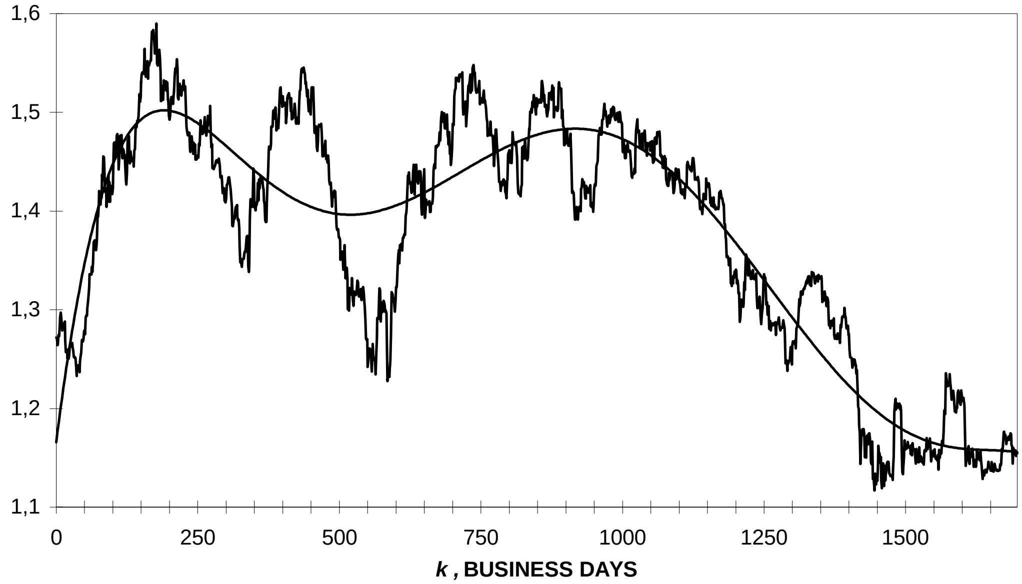

Figure 17. Swiss Franc versus US \$ as the functions of serial number k of business day from 1 January 1991 (k = 0) to 29 December 1995 (k = 1697).

Figure 18 for example shows some correlation functions for data of Figure 17. Among them the sample correlation function of the yield rates for exchange rates of SFr v. US \$ (briefly called RATE), the sample correlation function of the logarithms of these exchange rates (briefly called LN(RATE)), the correlation function of the trend (TREND) and the theoretical correlation functions of the AR(1) model for sample data of Figure 17. For these data the AR(1) model has the factor a = 0.9999 (see the Table below). For other data these correlation functions are similar. Figure 19 shows only sample correlation functions for all four exchange rates. The data on factors of the AR(1) and AR(2) models of the exchange rates are shown in the Table together with factors of the AR(1) and AR(2) models of an autoregression for deviations of the exchange rates from their trends.

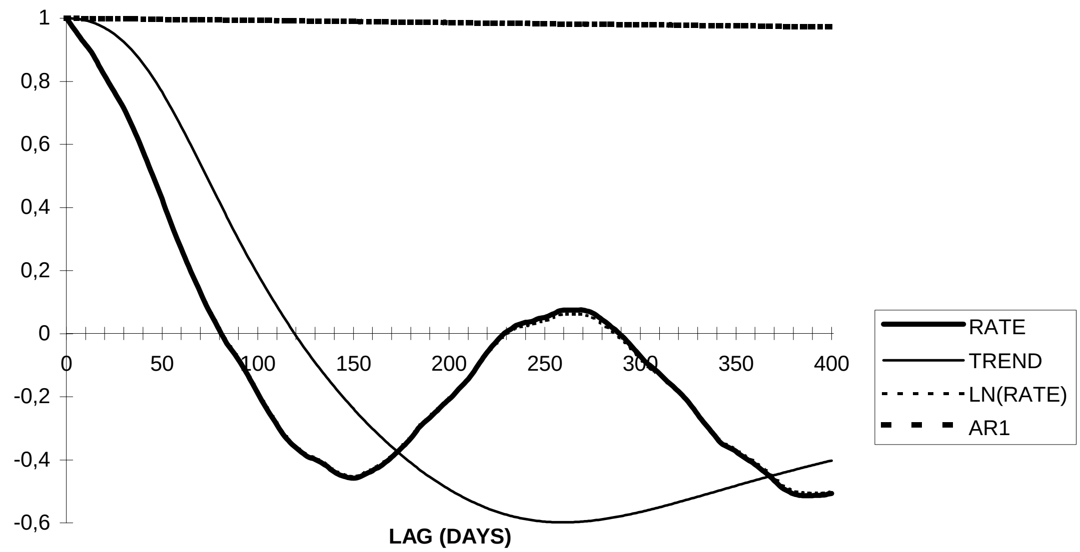

Figure 18. Correlation functions for exchange rate of SFr v. US $.

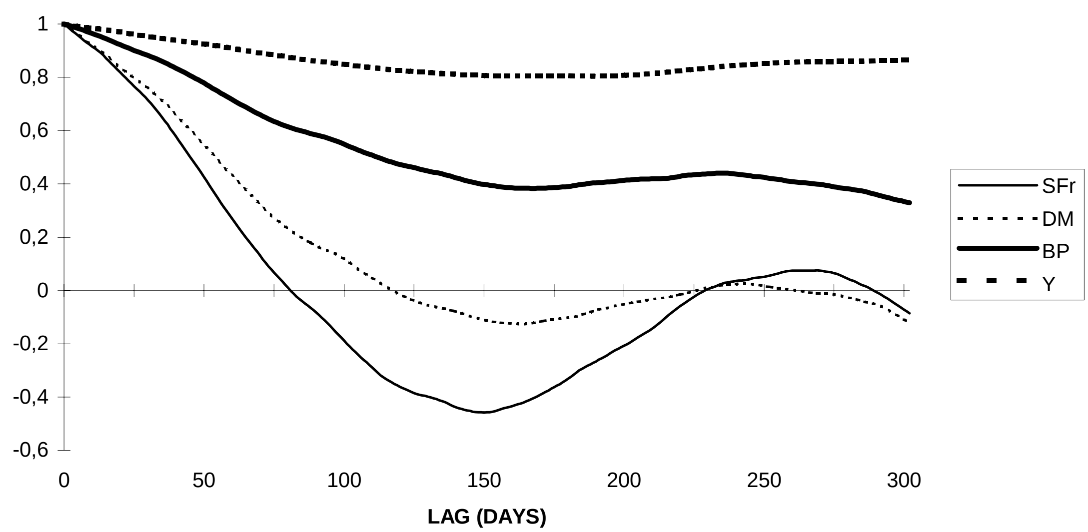

Figure 19. Sample correlation functions for exchange rates.

Note that sample correlation functions is usually significant more dynamical than correlation functions appropriate to their AR models .

Table of coefficients of autoregressive models

	SFr	DM	BP	Y
Rates				
AR(1), *a*	0,9999	0,9999	0,9998	0,9998
AR(2), *a*1	0,9967	0,9760	0,9837	0,9384
AR(2), *a*2	0,0032	0,0239	0,0161	0,0615
Residuals				
AR(1), *a*	0,9801	0,9775	0,9857	0,9847
AR(2),*a*1	0,9929	0,9709	0,9845	0,9343
AR(2),*a*2	-0,0131	0,0065	0,0011	0,0512

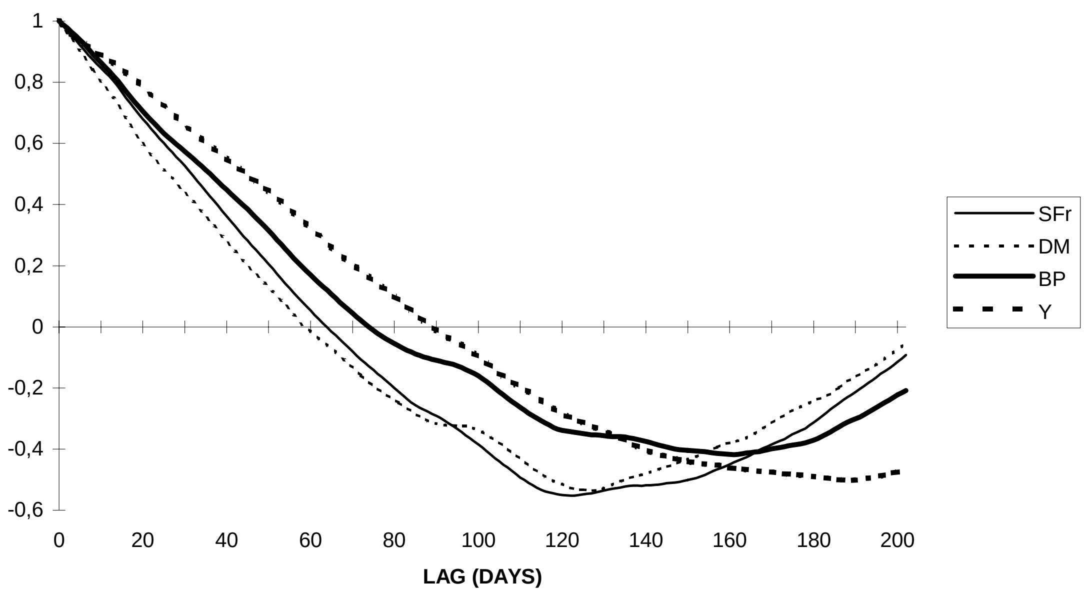

Figure 20. Sample correlation functions for deviations of exchange rates from their trends.

As well as in the previous cases the correlation functions appropriate to all autoregressions with factors from an above Table take only positive values monotonically decreasing to zero with increase of a lag. At the same time sample correlation functions both for exchange rates and for their deviations from a trend with increase of a lag confidently enter into a negative range of values (except the cases for BP and Y on Figure 19).

5. Conclusions

Let us summarize conclusions, which were practically already made above.

- The stochastic differential equations as the models of processes of the interest rates or their derivatives generate the Markov processes and bring to the autoregressive models of the first order. Let us note that the univariate stochastic differential equations can not bring to the AR(p) models (p > 1) or the ARMA(p,q) models (q > 0).
- The autoregressive models of the first order generate the Markov sequences of data with exponential function of correlation taking only positive values.
- The sample correlation functions of the considered time series of financial data have as rule a character of damping oscillations about a zero level taking both positive and negative values.
- The correlation functions (13) for autoregressive models of the second order with the complex roots of a polynomial of an autoregression could have such character. However for the AR(2) models of the considered financial series these roots were found real and the correlation functions, appropriate to them, were also monotonically decreasing positive functions. Thus the autoregressive models of the real financial data allow only the positive correlation for any lags while the sample correlation for the same data can be both positive and negative.
- The correlation dependence between values of the time series is very important for purposes of the forecast. In the mean square theory of a prediction the optimum procedures of forecast are determined extremely through correlation properties of treated data. Therefore the construction of the mathematical models of financial data that reflect correctly the correlation properties of these data is very desirable.
- Though the AR models are most appropriate for description of time series of financial data nevertheless the further researches are necessary for reaching their best goodness of fit to real data.

Author intend to develop a relevant stochastic recurrent models of financial data and to use them for the forecast problems.

Acknowledgements

The author is very grateful to Professor David Wilkie from Watson Wyatt UK, which kindly provided the data used in 4.2 - 4.6 of this paper, and to Professor Sam Cox from Georgia State University USA, which has offered to the author to investigate stochastic interest rate processes and has learned to get the financial data from Internet.

References

Black F. & Karasinski (1991). Bond and Bond Option Pricing when short rates are lognormal. Financial Analysts Journal, July - August, 52 -59.

Boyle P.P. (1980). Recent Models of the Term Structure of Interest Rates with Actuarial Applications. Transactions of the 21st International Congress of Actuaries, topics 4, 95 - 104.

Brockwell P.J. & Davis R.A. (1987). Time Series: Theory and Methods. Springer-Verlag, New York.

Doob J.L. (1953). Stochastic Processes. John Wiley & Sons, New York.

Hart D.G., Buchanan R.A. & Howe B.A. (1996). The Actuarial Practice of General Insurance. Institute of Actuaries of Australia, Sydney.

Tilley J.A. (1992). An Actuarial Layman's Guide to Building Stochastic Interest Rate Generators. Transactions of the Society of Actuaries, vol. 44, 509 - 538.

Vetzal K.R. (1994). A Survey of Stochastic Continuous Time Models of the Term Structure of Interest Rates. Insurance: Mathematics and Economics, vol. 14, 139-161.

Wilkie, A.D. (1986). A Stochastic Investment Model for Actuarial Use. Transactions of the Faculty of Actuaries, vol. 39, 341 - 373.

Wilkie, A.D. (1995). More on a Stochastic Asset Model for Actuarial Use. British Actuarial Journal, vol. 1, part V, 777 - 946.