Jitka Zichová

Some applications of time series models to financial data

Acta Universitatis Carolinae. Mathematica et Physica, Vol. 52 (2011), No. 1, 93--102

Persistent URL: http://dml.cz/dmlcz/143671

Some Applications of Time Series Models to Financial Data

JITKA ZICHOVÁ

Praha

Received May 17, 2010
Revised July 30, 2010

Data sets in economics and finance have often the form of time series. The article is devoted to an application of simple univariate and multivariate autoregressive models to a two-dimensional collection of exchange rates. Parameter estimates obtained using special methods constructed for non-negative time series are compared with the outputs of standard estimation procedures implemented in commonly used software products. Later on, the attention is paid to the predictive capability of our models.

1. Introduction

In economics and finance, we often observe time series, i.e. ordered sequences of records of a variable. Different models are available to describe such data and to predict their future development. Simple linear models such as the first order autoregression are frequently applied and implemented in commonly used statistical software packages.

Some special procedures for parameter estimation in non-negative autoregressive models were proposed in last decades [4,1,2,3]. Their small sample behaviour was investigated in simulation studies which confirmed satisfactory convergence properties. The aim of this article is to study the forecasting quality on real data sets and to compare selected univariate and multivariate models estimated using the mentioned approach with models analyzed by means of standard methods. Some series of exchange rates were used for this purpose.

Department of Probability and Statistics, Faculty of Mathematics and Physics, Charles University in Prague, Sokolovská 83, 186 75 Praha 8, Czech Republic

This research was supported by the project MSM0021620839.

Key words and phrases. Time series, AR models, parameter estimation

E-mail address: zichova@karlin.mff.cuni.cz

2. Autoregressive models

There are many books containing a detailed description of autoregressive models, e.g. [5,6]. Let us briefly recall basic definitions and notations. The autoregressive process of first order AR(1) satisfies the equation

$$X_t = bX_{t-1} + e_t, \tag{1}$$

where X_t is the value at time t and the error sequence $e_t, t = 1, 2, \ldots$ called white noise is created by uncorrelated random variables with constant mean and variance and with a distribution function F. The parameter $-1 < b < 1$ is usually estimated by the maximum likelihood method or by the least squares method.

Bell and Smith investigated in [4] the AR(1) with non-negative white noise and $F(d) - F(c) < 1$ for all $0 < c < d < \infty$. They proposed a simple estimate

$$b^* = \min\left(\frac{X_2}{X_1}, \frac{X_3}{X_2}, \ldots, \frac{X_n}{X_{n-1}}\right) \tag{2}$$

of the parameter $0 < b < 1$ based on an observed non-negative series X_t of length n. The white noise e_t is e.g. the exponentially distributed one with mean λ studied by Anděl in [1]. He showed that b^* can be obtained by maximizing the conditional likelihood function of the variables $X_2, X_3, \ldots, X_n$ having fixed $X_1 = x_1$ which is given by

$$L = \lambda^{-n+1} \exp\left(-\sum_{t=2}^{n} \frac{X_t - bX_{t-1}}{\lambda}\right) \tag{3}$$

under the conditions

$$X_t - bX_{t-1} \geq 0, \quad t = 2, 3, \ldots, n. \tag{4}$$

Later on, this result was generalized in [2] to the second order autoregression

$$X_t = b_1 X_{t-1} + b_2 X_{t-2} + e_t,$$

and in [3] to the multivariate AR(1) in which a vector time series is described using the model of type (1) with a matrix parameter.

3. Data sets

Results for two series of exchange rates, CZK/EUR and CZK/USD, are presented in the following text. Both data sets contain 86 monthly average exchange rates recorded from January 1999, when EURO was introduced, to February 2006 and published by Czech National Bank (web pages, June 2010: `http://www.cnb.cz/cnb/STAT.ARADY_PKG.SESTAVY_DZDROJE?p_zdrojid=KURZY&p_lang=EN`). The means are 32.884 for CZK/EUR and 31.502 for CZK/USD. The data were subjected to logarithmic transformation. In our case, the variance was reduced and the nonnegativity of the series was conserved. The mean of transformed CZK/EUR is 3.490, in the case

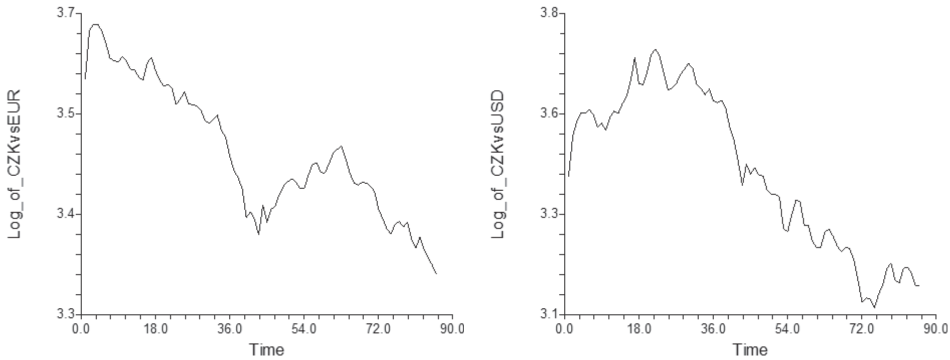

FIGURE 1. Log of monthly exchange rates CZK/EUR and CZK/USD

of transformed CZK/USD we get 3.433. Figure 1 shows the logarithmic transformations of our exchange rate series. The time values on the horizontal axis are numbers of months 1,2, ...

Univariate models were analyzed by means of the statistical software product NCSS calculating maximum likelihood parameter estimates (see [5]). The library Time Series Pack of the software system Mathematica (see [7]) was chosen as a useful tool for working with multivariate models. The Hannan-Rissanen procedure is used there to find an adequate model for a given data collection. Parameter estimates can be then constructed by conditional maximum likelihood method. The approach by Bell and Smith in the univariate AR(1), further denoted as BS method, and its generalization to multivariate AR(1) were applied and the results were compared with the outputs of NCSS and Mathematica.

The quality of forecasting an unknown future was checked in selected models. The series were shortened to comprise only observations from January 1999 to February 2005. In the estimated models based on the reduced series, forecasts were constructed for the time horizon from March 2005 to February 2006 using the shortened data. The comparison of these 12 forecasts with the observed values enables us to see the ability of each method with respect to forecasting the unknown future.

4. Univariate AR(1) model

The estimates of the parameter b in the model (1) for both mentioned data sets were calculated from the logarithmic series of length 86 using the procedures described above. Future forecasts were constructed from the shortened series being compared with the observed development of exchange rates.

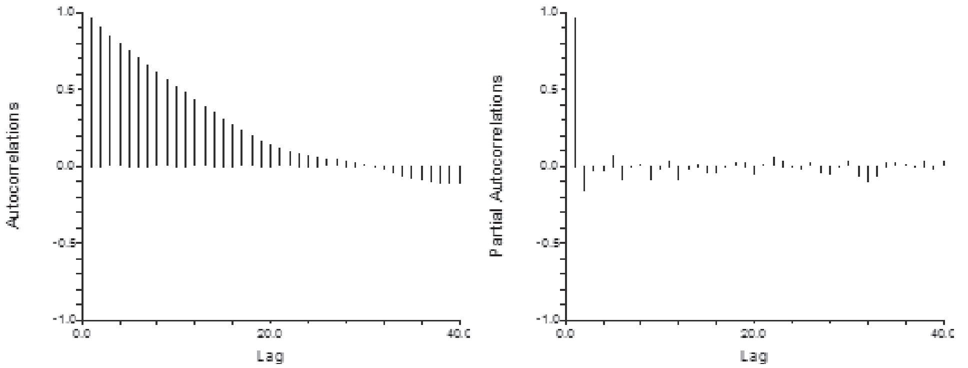

FIGURE 2. ACF and PACF of logarithmic monthly exchange rates CZK/EUR

BS method

The formula (2) gave the following estimates of the parameter $b : b^*_{EUR} = 0.9903$ for the series CZK/EUR and $b^*_{USD} = 0.9783$ for the series CZK/USD. The stationarity condition for the AR(1) process is $|b| < 1$ (see [6], p. 33) so that, in accordance with the graphs of the logarithmic series showing a decreasing trend, we get almost nonstationary estimated AR(1) models for both currencies.

Analysis in NCSS

The autocorrelation (ACF) and partial autocorrelation (PACF) functions of the logarithmic series CZK/EUR are outlined in figure 2. The ACF and PACF of logarithmic exchange rates CZK/USD behave in the same way. Slowly decreasing ACF and PACF which cuts off after time lag 1 indicate an almost nonstationary AR(1) (see [6], p. 33). The maximum likelihood parameter estimates are similar to those calculated by the BS method: $\hat{b}_{EUR} = 0.9994$ and $\hat{b}_{USD} = 0.9993$.

Denote the forecasted series by $\hat{X}_t$. We have residuals $X_t - \hat{X}_t$, $t = 1, 2, \ldots, 86$ as the estimates of the white noise values. The average of the squared residuals and its square root RMS are measures of forecasting quality in the time horizon in which we observed the values of the time series. We obtained $RMS_{EUR} = 0.013$ and $RMS_{USD} = 0.029$ in the analysis of the logarithmic series. We can consider the residuals $X_t - \hat{X}_t$ to be uncorrelated in the case of CZK/EUR. It was confirmed by a significance test performed in NCSS and we can see it in the graph of the autocorrelation function of the residuals in figure 3. In the logarithmic series CZK/USD, only the first autocorrelation value 0.243 is higher than the lower bound of the significance test 0.216 so that the residuals seem to follow a MA(1) model (see [6], p. 47). We present graphs of the ACF of both series.

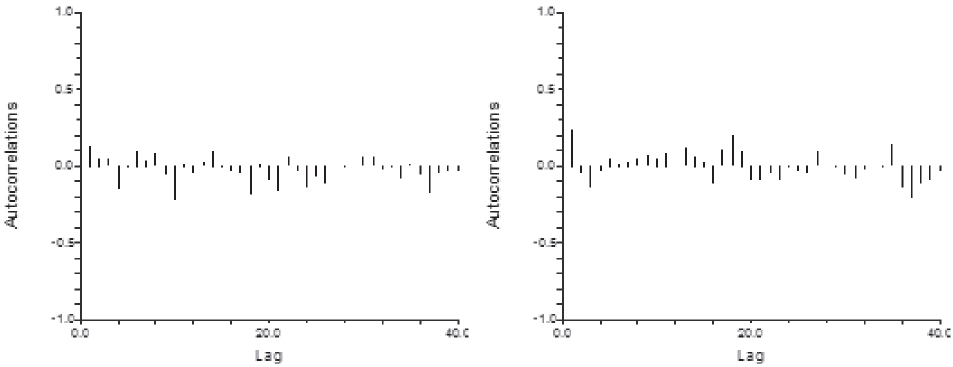

FIGURE 3. ACF of AR(1) residuals – CZK/EUR (left) and CZK/USD (right)

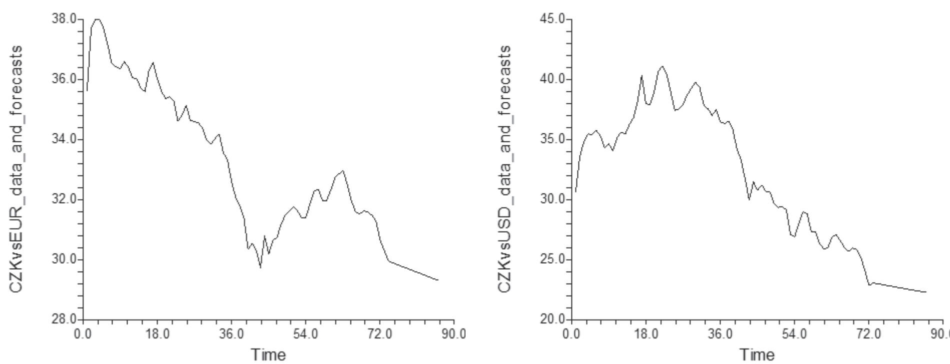

FIGURE 4. Data and forecasts in univariate AR(1) – CZK/EUR and CZK/USD

Since the estimates by the BS method and by NCSS are very similar, only the models estimated by NCSS were used to forecast the future behaviour of the series in one year from the reduced data. Figure 4 shows 74 observations and 12 forecasts for the original series CZK/EUR and CZK/USD. Figure 5 presents the observed series from March 2005 to February 2006 (line), the forecasts for the same time horizon based on the observations till February 2005 (triangles) and 95 percent prediction limits (circles). We can see from the values of RMS and from the pictures that the forecasting quality is very good in the case of CZK/EUR and satisfactory for CZK/USD even if the logarithmic transformation led to almost nonstationary AR(1) model.

Remark: Slowly decreasing ACF of logarithmic series indicates the possibility of using the transformation of first differences. Having done this, we ascertained that all values of both ACF and PACF of the differenced logarithmic CZK/EUR can be considered equal to zero which corresponds to the parameter estimate close to one. The differences represent a sequence of uncorrelated random variables. The differences of

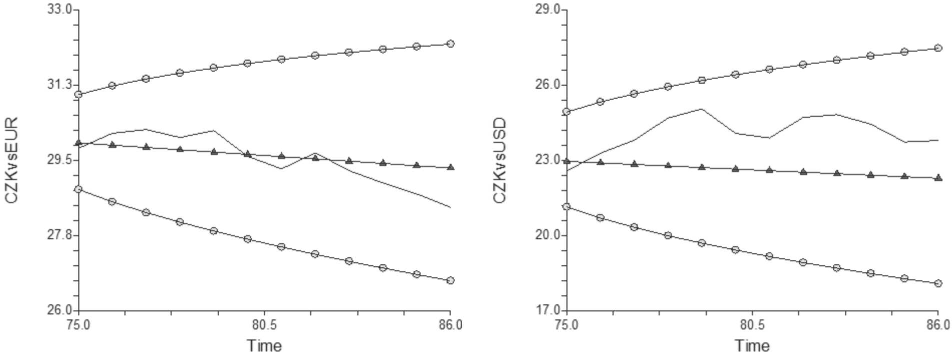

FIGURE 5. Forecasts and actual development – CZK/EUR and CZK/USD

the observations of the original series have the same property. In the original and logarithmically transformed series CZK/USD, ACF and PACF of first differences have only the value at time lag one slightly exceeding the lower bound of the significance test.

5. Multivariate AR(1) model

The multivariate AR(1) model is defined by the equation

$$\boldsymbol{X}_t = \boldsymbol{U}\boldsymbol{X}_{t-1} + \boldsymbol{e}_t.$$

We suppose the parameter matrix $\boldsymbol{U}$ with elements $u_{ij} \geq 0$, $i, j = 1, 2, \ldots, d$ and all eigenvalues inside the unit circle. As our data are two-dimensional, $d = 2$. The theory of multivariate autoregressions, their properties and stationarity conditions can be found in [6].

Let the white noise vectors $\boldsymbol{e}_t$, $t = 1, 2, \ldots$ be non-negative, independent and identically distributed having a continuous distribution with finite second moments. Under some technical assumptions concerning the distribution function of the white noise (see [3]), a simple generalization of the formula (2) follows to the estimates of the elements of $\boldsymbol{U}$ based on a realization $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ of the process $\boldsymbol{X}_t$ in the form

$$u^{+}_{ij} = \min_{2 \leq t \leq n} \frac{X_{t,i}}{X_{t-1,j}} \tag{5}$$

which was introduced and studied in [3]. The estimate (2) in the univariate AR(1) can be obtained by solving the maximization problem (3), (4) based on the exponential white noise assumption. Simulation studies showed that its convergence to the true parameter value is sufficiently fast even in the cases of other white noise distributions. Therefore Anděl in [3] proposed a construction of the estimate $\boldsymbol{U}^*$ of $\boldsymbol{U}$ by solving the maximization of

$$u_{11}\frac{1}{n}\sum_{t=1}^{n}X_{t,1}+u_{12}\frac{1}{n}\sum_{t=1}^{n}X_{t,2},$$
$$u_{21}\frac{1}{n}\sum_{t=1}^{n}X_{t,1}+u_{22}\frac{1}{n}\sum_{t=1}^{n}X_{t,2} \qquad (6)$$

under the conditions

$$X_{t1}-u_{11}X_{t-1,1}-u_{12}X_{t-1,2}\geq 0,$$
$$X_{t2}-u_{21}X_{t-1,1}-u_{22}X_{t-1,2}\geq 0,$$
$$u_{ij}\geq 0,\ i,j=1,2. \qquad (7)$$

This approach was derived assuming exponential distribution of the independent components of the bivariate white noise $\boldsymbol{e}_t$ and maximizing the conditional likelihood function as in univariate AR(1). Nevertheless, the estimation procedures (5) and (6), (7) are not equivalent in the multivariate model. Both lead to estimates which are strongly consistent, this means they have good convergence properties in large samples, but small sample convergence is faster in the case of the estimates obtained by solving (6), (7) which we call A1 method in the following text. It was applied to the two-dimensional series containing logarithmic transformations of the monthly averages of exchange rates CZK/EUR and CZK/USD. The estimation algorithm implemented in Time Series Pack of the software system Mathematica was also used to estimate the parameter matrix and the received results were compared.

A1 method

Having solved the optimization problem in the logarithmic series of length 86 we got this estimate of the parameter matrix $\boldsymbol{U}$:

$$\boldsymbol{U}^{*}=\begin{pmatrix}0.9903 & 0\\ 0 & 0.9783\end{pmatrix}.$$

We can see immediately that AR(1) models were estimated for both components of the vector process (CZK/EUR, CZK/USD) with the same parameter values as in the application of the BS method to the univariate series.

Analysis in Time Series Pack

First order autoregression was recommended as an appropriate model for our data by the Hannan-Rissanen optimization algorithm ([7], p. 95). Conditional maximum likelihood estimate of the parameter matrix calculated from the same series as above is

$$\hat{\boldsymbol{U}}=\begin{pmatrix}1.0097 & -0.0107\\ 0.0238 & 0.9750\end{pmatrix}.$$

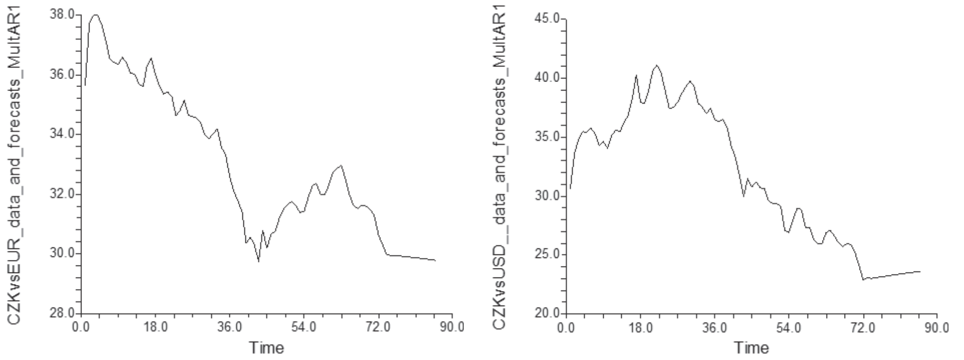

Figure 6. Data and forecasts in multivariate AR(1) – CZK/EUR and CZK/USD

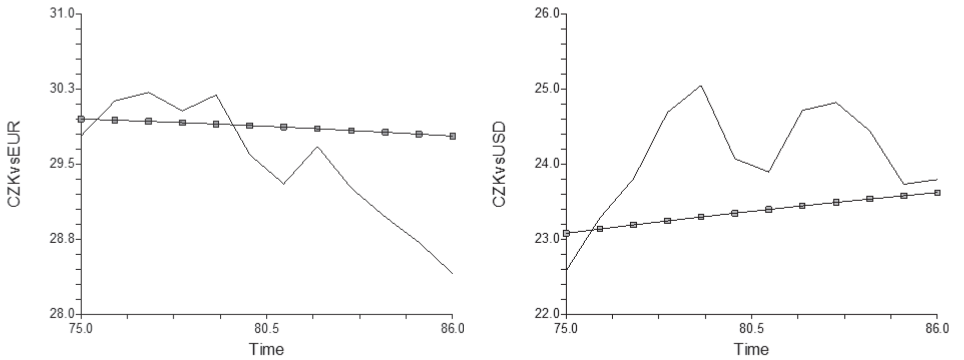

Figure 7. Forecasts and actual development – CZK/EUR and CZK/USD

The eigenvalues of the above matrix are 0.9992 and 0.9855, which means that the stationarity condition of the model is satisfied (see [6], pp. 339–340). The residuals of the model showed a satisfactory correlation structure.

Both estimation procedures gave parameter matrices with nonsignificant elements outside the diagonal and eigenvalues close to one, this means almost nonstationary models corresponding with the results in the univariate AR(1).

The quality of future forecasting was checked in the model estimated by Time Series Pack. The series were shortened in order to have 74 observations and the forecasts for the remaining 12 values were calculated. We show the observed 74 values followed by 12 predicted values in both series in figure 6.

Figure 7 presents the detail of the development of the exchange rates from March 2005 to February 2006 (line) and the predicted values for this time period (squares).

Finally, let us compare the quality of forecasting an unknown future in the univariate and multivariate AR(1) model. The following table shows the values of root

Table 1. RMS in univariate and multivariate AR(1)

Data	Univariate AR(1)	Multivariate AR(1)
CZK/EUR	0.430	0.630
CZK/USD	1.649	0.949

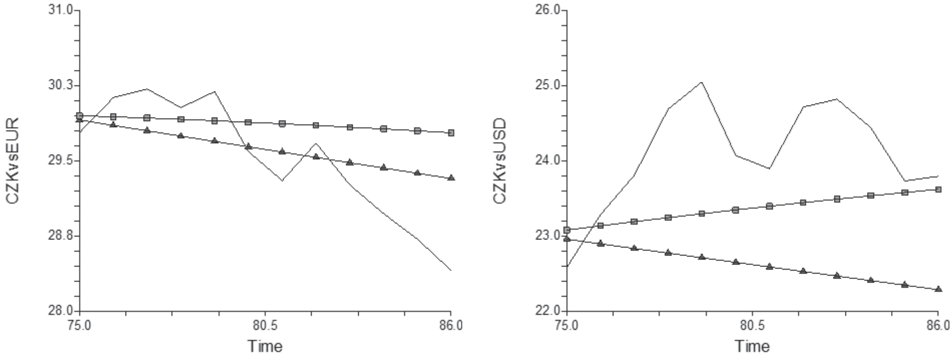

Figure 8. Forecasts and actual development – CZK/EUR and CZK/USD

mean square (RMS) of residuals calculated as differences of observed and forecasted values of the original data from March 2005 to February 2006.

Figure 8 shows the observed data from March 2005 to February 2006 (line), the forecasts are represented by triangles in univariate AR(1) and by squares in multivariate AR(1).

We can see from the values of RMS and from the pictures that the quality of future forecasts in both univariate and multivariate model is similar for the series CZK/EUR. In the case of CZK/USD, the multivariate AR(1) model seems to have better capability of predicting the unknown future. It can be noticed that the univariate model predicted for CZK/USD the globally decreasing trend whereas the multivariate one predicted the possibility of local increasing phases which corresponds to the actual development of the series from March 2005.

References

[1] Anděl, J.: *On AR(1) processes with exponential white noise.* Commun. Statist. Theory Methods **17** (1988), 1481–1495.

[2] Anděl, J.: *Non-negative autoregressive processes.* J. Time Ser. Anal. **10** (1989), 1–11.

[3] Anděl, J.: *Nonnegative multivariate AR(1) processes.* Kybernetika **28** (3) (1992), 213–226.

[4] Bell, C. B., Smith, E. P.: *Inference for non-negative autoregressive schemes.* Commun. Statist. Theory Methods **15** (1986), 2267–2293.

[5] Box, G. E. P., Jenkins, G. M.: Time Series Analysis. Forecasting and Control. Holden Day, San Francisco (1970).

[6] Wei, W. W. S.: Time Series Analysis. Univariate and Multivariate Methods. Addison-Wesley, New York (1990).

[7] Time Series Pack. Reference and User's Guide. Mathematica Applications Library. Wolfram Research, Champaign, 1995.

[8] `http://www.cnb.cz.` Web pages of the Czech National Bank, section Statistics, ARAD – time series system, Data Table KURZY.