COMPARISON OF ACCURACY OF SELECTED MODELS OF TIME SERIES AT DEVELOPMENT OF THE CZECH NATIONAL BANK ASSETS

Bohdan Linda, Jana Kubanová

University of Pardubice, Faculty of Economics and Administration, Institute of Mathematics

Abstract: *A lot of ways how to estimate the values of the time series for the future period are known. The basic approach starts from principle of linear regression methods. The Box-Jenkins methodology is very often used in financier when the time series are analysed. The paper deals with application of the bootstap principle in this methodology. The classical and resampling methods are compared at extrapolation of the values of the time series of assets of the Czech national bank.*

Key words: *extrapolation, parameters estimate, bootstrap method, autoregressive models, moving blocks overlapping and not overlapping methods*

Theoretical resources and methods

Extrapolation of the time series used to be very often task in the science and practical areas in economy, finance, insurance, industry and many other branches. The point of interest of this paper is to forecast the future behaviour and development of the indicator “assets of the Czech National Bank”.

The month data of this indicator were collected (from 31.1.2002 till 30.9.2007, n = 69). The first step was to purge the data from the calendar variations. The autoregression between terms X_{t-1} and X_t was verified by the Durbin-Watson test [3]. The test statistics value is 0,908. The basic condition for the simple linear regression is not fulfilled [7].

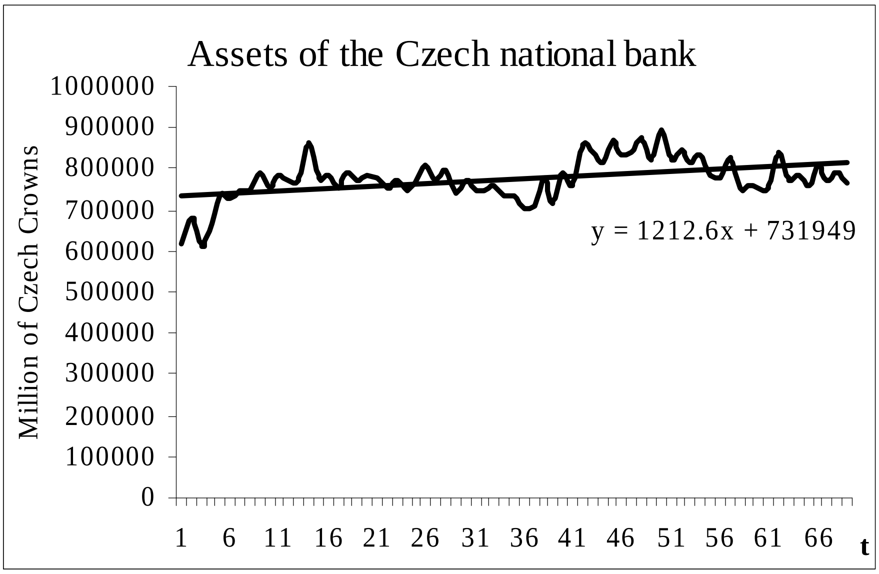

Fig.1 Assets of the Czech National Bank

We came out of the autoregressive models theory. Let $..., X_{-2}, X_{-1}, X_0, X_1, X_2, X_3, ...$ be a time series. When for their single elements holds true $E(X_t) = m$ and $\mathrm{cov}(X_t, X_{t+r}) = k(r)$ for every integer t, r then this time series is weakly stationary stochastic process. When it holds in addition

$$X_t = a\ X_{t-1} + X_t$$

for every t and where X_t are independent identically distributed random variables with mean 0 and standard deviation s , then we say, that the process is first-order autoregressive process AR(1). Random variables X_t are called residuals or the white noise.

a is an unknown real parameter and $|a| \leq 1$, [1].

To satisfy the above mention condition, the data had to be transformed to obtain the stationary series. We can see in the picture 2 (left panel) that the difference of the series is stationary, the regression line y = -203,42x + 9554,8 was calculated by the least square method and both its parameters were tested. The null hypothesis was not rejected for both parameters; it means that parameters of the mentioned regression line can be equal to zero. The differences can have the normal distribution as signifies the histogram in the figure 2, right panel. This hypothesis was verified by the χ^2 test

After these steps we tried to use following methods to estimate the future development of the time series:

- simple linear regression model (although it is not recommended),
- classical AR(1) model,
- bias reduced estimate of the parameter α based on model oriented bootstrap method [signification AR(1)bias],
- bias reduced estimate of the parameter α based on bootstrap not-overlapping moving blocks, length of the block 4 and 7 [significations MBnol (4), MBnol (7)],
- bias reduced estimate of the parameter α based on bootstrap overlapping moving blocks, length of the block 4 and 7, [significations MBol (4), MBol (7)].

All these methods are described in [5], [6].

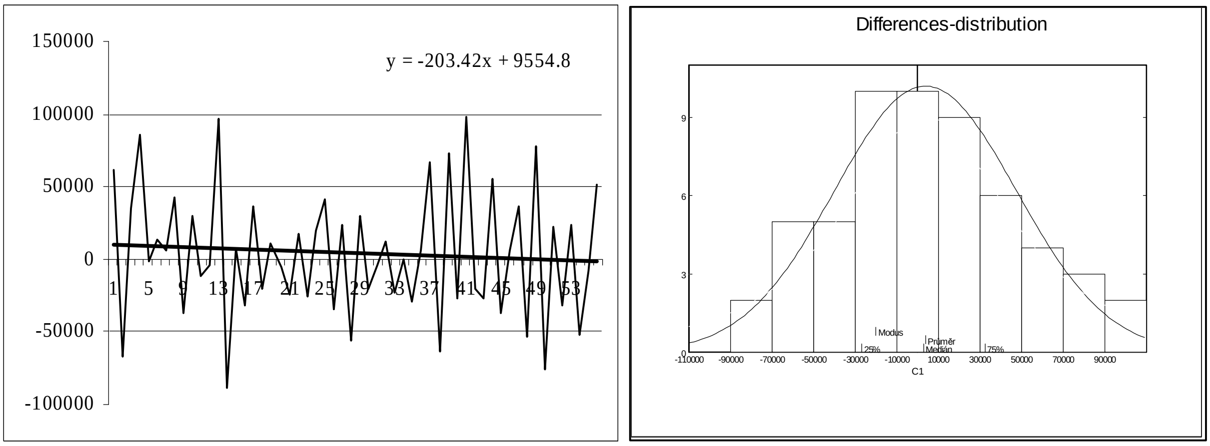

Fig.2 Differences (left) and their distribution (right)

The verification of the accuracy of the estimate

To decide which of the above mentioned methods provide the best results and which should be used, we considered 12 periods shorter time series. We used all these methods to estimate the values of assets of the Czech national bank during 12 time periods (1 year) and compared the obtained results with real values. We can see in the figure 3 that the greatest differences between the real and estimated value were obtained when the simple linear regression model was used. This result was expected, in the beginning of this paper is written that the assumption of this model were not fulfilled. We can state in a similar way that the first-order autoregressive model doesn't provide the correct results as well. The best results (it means the results that are most similar to real values) were obtained by the help of the bootstrap methods, first of all methods of moving blocks with length 4 of the block [2].

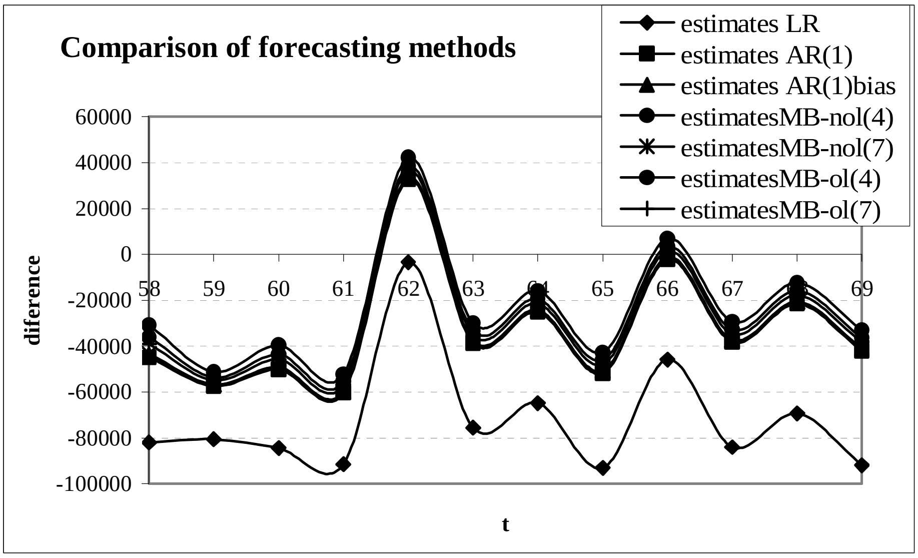

Fig.3 Differences between the real and estimated value

The real data of the assets of the Czech national bank are in the third column and estimated values are from fourth till tenth column. The bold printer numbers mean the best correspondence in the time period.

Tab.1 Estimates of the values of the time series

day	t	real data purged	estimates LR	estimates AR(1)	estimates AR(1)bias	estimates MBnol(4)	estimates MBnol(7)	estimates MBol(4)	estimates MBol(7)
31.10.2006	58	**753679**	835590	798632	797306	**784459**	797051	789812	792871
30.11.2006	59	**757251**	837705	814695	814114	**808478**	814002	810826	812168
31.12.2006	60	**755461**	839820	805678	804679	**794996**	804487	799030	801336
31.1.2007	61	**750394**	841935	810740	809975	**802563**	809828	805652	807416
28.2.2007	62	**840683**	**844050**	807899	807002	798315	806830	801935	804003
31.3.2007	63	**770724**	846165	809493	808671	**800700**	808513	804021	805919
30.4.2007	64	**783494**	848280	808598	807734	**799361**	807568	802850	804844
31.5.2007	65	**757329**	850396	809101	808260	**800113**	808098	803508	805447
30.6.2007	66	**806626**	852511	808819	807965	799691	**807801**	803139	805109
31.7.2007	67	**770638**	854626	808977	808131	**799928**	807968	803346	805299
31.8.2007	68	**787449**	856741	808888	808038	**799795**	807874	803229	805192
30.9.2007	69	**766907**	858856	808938	808090	**799869**	807927	803295	805252

The other questions that must be answered in connection with these estimates are the accuracy of the estimates and in connection with the bootstrap method the number of bootstrap replications that must be done [4].

The accuracy of the estimate is usually evaluated by its bias and standard error. Whereas the distribution function is unknown, it is problematical to state these values. The bootstrap method is one of very elegant way of solution of this problem. We can see the smallest values of bias when bootstrap method was used to estimate the parameter α in the model AR(1). Similarly the small values of bias provide the moving block methods not overlapping, when the length of the block equal 7 was selected. The greatest values are observed when the method of moving blocks-not overlapping, the length of the block 4, was applied. But not only bias provides the exhausting information about the quality of the estimate, we have to consider the standard error as well. The values of the standard error are presented in the table 3. We can state that these values are very similar; all methods provide the analogical quality of exactness.

Tab.2 Bias of the estimates of the parameter a **, R is number of bootstrap replications**

BIAS	estimates	estimates	estimates	estimates	estimates
R	AR(1)bias	MBnol(4)	MBnol(7)	MBol(4)	MBol(7)
200	0.026	0.279	0.032	0.173	0.115
300	0.025	0.278	0.032	0.173	0.113
400	0.025	0.278	0.032	0.173	0.112
500	0.025	0.278	0.031	0.173	0.113
600	0.025	0.278	0.031	0.173	0.113
700	0.026	0.278	0.031	0.173	0.113
800	0.026	0.278	0.031	0.173	0.113
900	0.026	0.278	0.031	0.173	0.113
1000	0.026	0.278	0.031	0.173	0.113

Tab.3 Standard error of the estimates of the parameter a

ST.ERROR	estimates	estimates	estimates	estimates	estimates
R	AR(1)bias	MBnol(4)	MBnol(7)	MBol(4)	MBol(7)
200	0.113	0.106	0.076	0.112	0.112
300	0.111	0.105	0.075	0.111	0.112
400	0.111	0.105	0.076	0.111	0.112
500	0.111	0.105	0.076	0.111	0.111
600	0.111	0.105	0.075	0.111	0.111
700	0.111	0.104	0.075	0.111	0.111
800	0.111	0.104	0.075	0.111	0.110
900	0.111	0.104	0.075	0.111	0.111
1000	0.111	0.104	0.075	0.111	0.110

Now we consider the development of bias and standard error of the estimate when 200 till 1000 bootstrap replications were made and method of moving blocks-not overlapping, the length of the block is 4, applied. Any convergence after approximate 300 bootstrap replications is obvious. The values of bias converge to 0,278, the values of the standard error to 0,104 when this method was applied. The number of replications that must be done used to be common question that was solved in the past. We can append that in present development of the computational technique it is not problem to provide some thousands of simulations. 1000 replications was done, but this number seemed to be quite restful, the values started to stabilize far earlier. The running of bias and standard error of the estimate was very similar when different methods of estimate were used.

The prognostication of the time series

The method that provided the "best results" was applied in the second step and the prognoses of the future development were calculated. We selected the methods of the moving blocks not overlapping with the length of the step 4. We could see that 300, respectively 600 bootstrap replications should be enough, that the values of bias and standard error of the estimate of the parameter α don't change any more.

The average value of bias (when 20 series of 1000 bootstrap replications were made and method of moving blocks-not overlapping used) was 0,278 when we made the experiment and tried to find the best method. The value 0,287 of bias is calculated when the total time series was used. The standard error calculated for the entire time series is lower (0,096) then the standard error for the experimental time series (0,104).

Table 4. The bias and standard error of the estimates (method of moving blocks-not-overlapping, *m* = length of the blocks)

R	*m*	200	300	400	500	600	700	800	900	1000
bias	4	0.289	0.288	0.288	0.287	0.287	0.287	0.287	0.287	0.287
St.Error	4	0.097	0.096	0.096	0.096	0.096	0.096	0.095	0.096	0.096

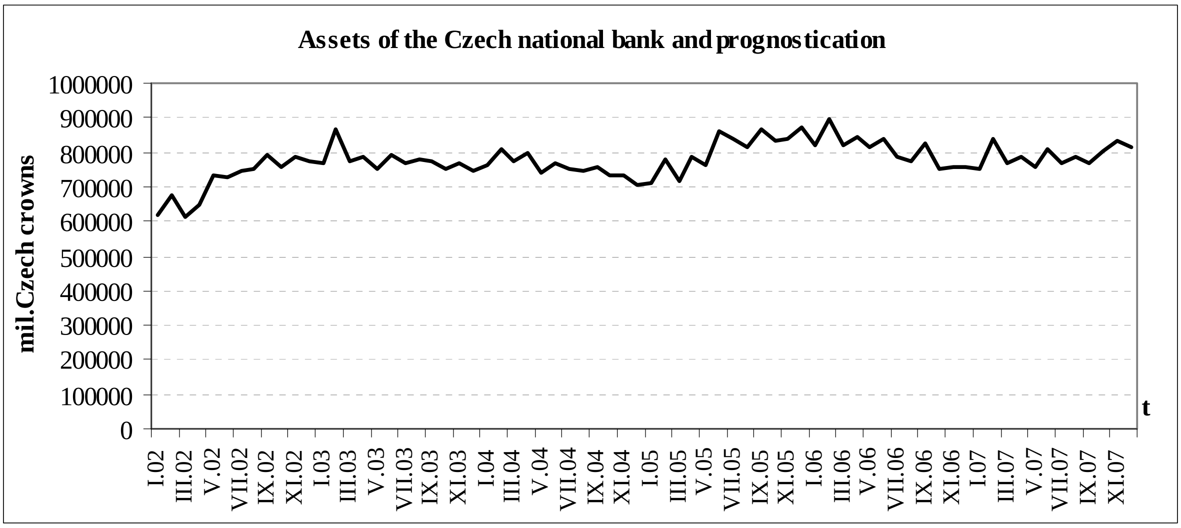

Fig 4. Assets of the Czech National Bank and prognosis for 3 consecutive months

The figure 4 presents the development of the assets of the Czech national bank from 31.1.2002 till 30.9.2007 and prognosis for 3 consecutive months, e.g. from 31.10.2007 till 31.12.2007. The concrete values of the prognosis are stated in the table 5.

Table 5. Prognosis of assets of the Czech national bank

date	31.10.2007	30.11.2007	31.12.2007
prognosis of assets	820652.6	819602.7	831635.4

Conclusion

The prognostications based on the Box-Jenkins methodology are dependent above all upon the quality of the parameters estimates of the used models. We usually aren't able to determine the exact probability distribution of these estimators or bias and standard error of them at least. The bootstrap methods were developed just for estimation of these characteristics. It is necessary to remember that the bootstrap methods are not absolutely exact but they provide the practicable solution in such cases when real situation require using of the complicated model. The mathematical complexity of the model needn't to be in relation with accuracy of the bootstrap analysis. Bootstrap methods are required to be a compromise solution for the cases when application of exact methods is impossible or too complicated.

Literature:

[1] ARLT, J. Moderní metody modelování ekonomických časových řad, Praha, Grada Publishing, 1999, 303 p. ISBN 80-7169-593-4

[2] Efron, B.,Tibshirany, R.: An Introduction to the Bootstrap. Chapman &Hall, New York, 1993.

[3] KANDEROVÁ M., Úradníček, V. Štatistika a pravděpodobnost pre ekonómov, 2.časť. Bratislava. OZ Financ 2007. 187 p. ISBN 978-80-969535-6-1

[4] Linda, B.: Resampling Methods Accuracy in Point Estimates. In: Sborník příspěvků 15. škola statistiky EKOMSTAT´2001 – Aplikace štatistických metód v hospodárskej praxi. Trenčianske Teplice 2001, s.64 – 67, ISBN 80-88946-10-7

[5] Linda B., Kubanová J. Přesnost odhadu v autoregresních modelech. In: Forum Statisticum Slovacum 4/2007. SŠDS Bratislava 2007. s.81-88. ISSN 1336-7420

[6] Linda B., Kubanová J.,SEMERÁK P. Přesnost odhadu v autoregresních modelech při použití metody klouzavých bloků. In: Forum Statisticum Slovacum 4/2007. SŠDS Bratislava 2007. s.89-93. ISSN 1336-7420
[7] Stankovičová I., Vojtková M. Viacrozmerné štatistické metódy s aplikáciami. Bratislava. IURA EDITION 2007, p. 261. ISBN 978-80-8078-152-1

Contact addresses:

doc. RNDr. Bohdan Linda, CSc.
doc. PaedDr. Jana Kubanová, CSc.
Studentská 95
532 10 Pardubice
Email: bohdan.linda@upce.cz , jana.kubanova@upce.cz
tel. 46 603 6020, 46 603 6046