SHS Web of Conferences **39**, 01007 (2017) DOI: 10.1051/shsconf/20173901007
IES2017

Prediction of stock price developments using the Box-Jenkins method

Bořivoj Groda[1], *Jaromír* Vrbka[1*]

[1]Institute of Technology and Business, School of Expertness and Valuation, Okružní 517/10, 37001 České Budějovice, Czech Republic

Abstract: Stock prices develop in a non-linear way. Naturally, the stock price prediction is one of the most important issues at stock markets. Therefore, a variety of methods and technologies is devoted to the prediction of these prices. The present article predicts the future development of the stock price of ČEZ, a. s., on the Prague Stock Exchange using the ARIMA method - the Box-Jenkins method. The analysis employs the final price of the last trading day in a given month, from February 2012 to September 2017. The data come from the Prague Stock Exchange database. Statistica software is used for processing the data, namely advanced time series prediction methods, the ARIMA tool, and autocorrelation functions. First, the current stock development of ČEZ, a.s., was graphically evaluated, and this was followed by a stock price prediction for the next 60 days in which the shares would be traded. Lastly, the prediction residues were analysed. It was confirmed that the calculation was done correctly, but with little accuracy. The conclusion is an assertion that the Box-Jenkins method is not a suitable tool for prediction.
Key words: Box-Jenkins, ARIMA, prediction, shares

1 Introduction

Successful forecasting of future share prices development can bring a considerable profit for investors. The effective market hypothesis indicates that share prices reflect all information currently available, and that the changes in prices which are not based on newly discovered information are unpredictable [1]. However, some experts disagree with this statement, and therefore there are a number of methods and technologies associated with this concept, which enable to gain information on future development of share prices [2].

The basic characteristics of share prices are sensitivity, stationarity and asymmetric volatility. Investors naturally want to be able to forecast the share price movement, despite of the fact that it is a stochastic process. The development of share prices is a non-linear and dynamic process. There are a number of macroeconomic, industry and company factors which may have influence on share prices. These include for example global indices of share prices, overall economic activity, exchange rate, interest rate etc. [3]. Share prices

* Corresponding author: vrbka@mail.vstecb.cz

forecasting is thus very useful and attracts interest of researchers and investors, who carry out subjective investment judgements on the basis of objective technical indicators [4]. Expected share price is one of the important issues of stock market research. Accurate prediction of share prices, which is the basis for deciding on financial investment, is probably the biggest challenge for capital investment [5]. As referred to above, currently there are a number of applications and new technologies for share prices forecasting, such as neural networks, the ARIMA method etc. [6]. In this contribution, the ARIMA method (Box-Jenkins) is used for analysing and share prices forecasting.

ARIMA, i.e. a model of Autoregressive integrated moving average, is one of the most widely used models in time series forecasting. At present, the ARIMA model is often used for its unique predictive capability also for share prices development forecasting [7]. Frances [8] favours this statement claiming that the ARIMA models are the most frequently used models for predicting time series that can be stationary transformation. ARIMA can therefore enable identification of time series characteristics as well as predicting their behaviour in the future. Lags of differenced series appearing in the forecasting equation are called Autoregressive terms, while lags of the forecast errors are called moving average terms. As a general rule, a time series which needs to be differenced to be made stationary is said to be an integrated version of a stationary series [9]. Random variable which is a time series is considered stationary if its statistic characteristics are constant over time [10]. The models are used mainly for short-term forecasts, when there are no data for explanatory variables, or in case the model has poor predictive power [11].

Forecasting future stock market values on the basis of both past and current data series is thus one of the most required financial applications. The objective of the contribution is to forecast the future development of ČEZ share prices on the Prague Stock Market using Box-Jenkins method.

2 Data and methods

ČEZ, a.s. is one of the most important business entities in the Czech Republic. Its operations, structure, vision and business activities are characterized as follows [12]: "ČEZ group is an integrated energy cluster operating in a number of Central and South-Eastern countries and Turkey, with headquarters in the Czech Republic. Its core business consists of production, distribution, trading and sale in the field of electricity and heat, trading and sale of natural gas and coal mining. ČEZ group currently employs nearly 27,000 employees.

The most significant shareholder of the parent company ČEZ, a.s., is the Czech Republic with a holding in the capital of the company (as of June 14, 2017) of nearly 70%. The shares of ČEZ, a. s., are traded on Prague and Warsaw Share Exchange, where they are a part of the PX and WIG-CEE share exchange indices.

ČEZ mission is to provide safe, reliable and positive energy to customers as well as the whole company. Its aim is to bring innovations to meet energy needs and thus contribute to a better quality of life. Its strategy reflects the fundamental transformation of European energy market. ČEZ wants to operate its energy assets in the most efficient way and to adapt to the growing share of decentralized and non-emission production. Another priority is to offer a wide range of products and services for the customers, along with selling electricity and gas. The third priority is to invest actively in prospective energy assets with a focus on the Central European region and in supporting modern technologies at an early stage of their development.

In the Czech Republic, the ČEZ group are active in coal mining, generation and distribution of electricity and heat, trading in electricity and other commodities, sell electricity, heat and natural gas to end-customers and provide other services. The

production portfolio consists of nuclear, coal, gas, water, photovoltaic, wind and biogas sources."

Data on share prices between February 2012 and September 2017 are available. Specifically, the data on the final price at the end of the last trading day in a month will be used. In total, 69 records will be available for analysis. The data come from the Prague Stock Exchange database.

For the data processing, DELL Statistica V12 will be used. Advanced methods of Time series/ ion will be applied. Subsequently, the tools of ARIMA and autocorrelation function will be selected.

As a target variable, share prices will be selected. Next, we will specify 20 variable backups. We will assume seasonality of data at the level of each month of the year. The parameter of seasonal variation will be set to 12. Parameter p (Autoregressive constant) as well as parameter q (moving average) are set at level 1. The number of movements will be 15, the p-value for sharpening is 0.05.

First, we will review graphically the current development of ČEZ share prices. Next we will focus on forecasting the value of ČEZ shares. The development of prices will be predicted for the following 60 days on which the shares will be traded. Finally, residual stock will be analysed.

3 Results

ČEZ share prices development in the monitored period can be seen in Figure 1.

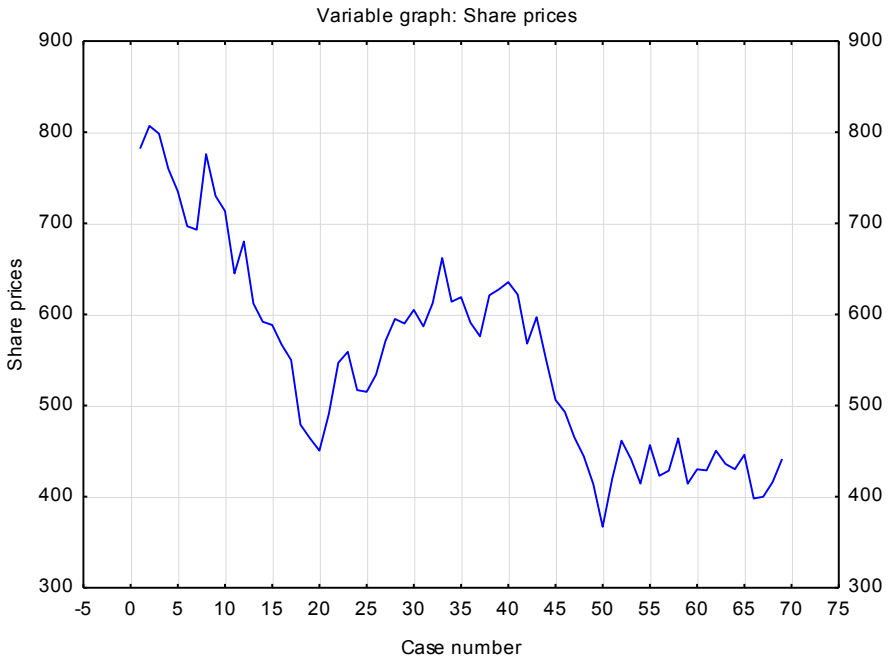

Figure 1 ČEZ share prices between February 2012 and September 2017

ČEZ share prices reached the peak at the beginning of the monitored period. At that time, the price was above 800 CZK per share. Subsequently, the price fell below 700 CZK per share. Then the share price grew to about 780 CZK and immediately decreased to 450 CZK per share. Approximately in the middle of the monitored period, share price increased

to more than 650 CZK. In February 2016, the price was between 400 (or more precisely, 399.9) and 450 (or more precisely, 455.5) CZK per share.

The figure shows there is no regular movement of share prices. A question therefore arises, whether the Box-Jenkins method, which is based on seasonal cycles of time series, is the right tool for forecasting ČEZ share prices.

This could be better seen from the graph in Figure 2.

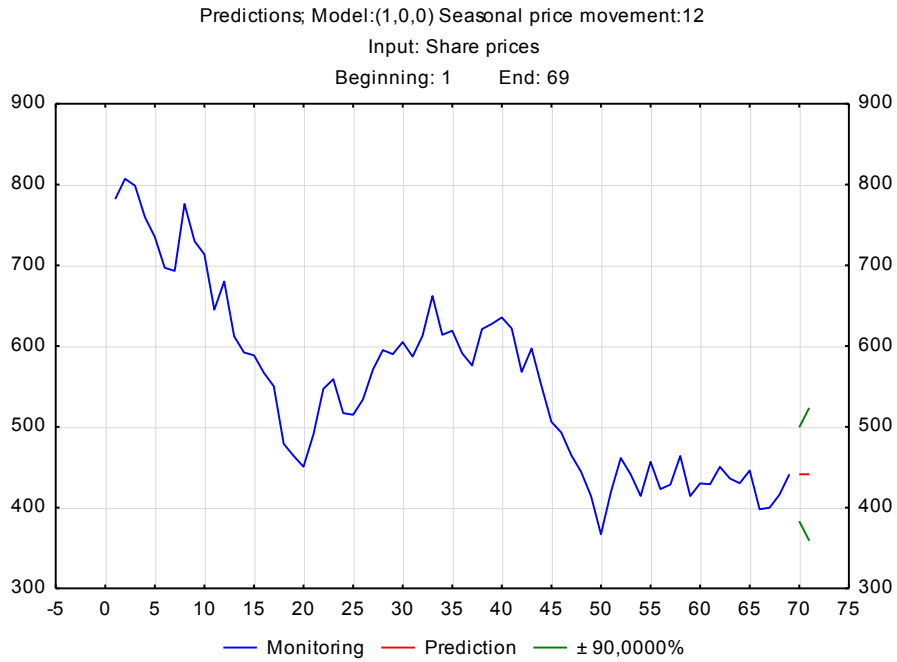

Figure 2. Prediction of future development of ČEZ share prices

The graph in Figure 2 illustrates the development of share prices in the monitored period. In addition, it also shows forecasts for the following two months of trading on the Prague Stock Exchange.

The actual forecast values are listed in Table 1

Table 1 Forecasting values of the future ČEZ share prices development.

Case number	**Forecasts; Model:(1,0,0) Seasonal movement:12 (Development of share prices - ČEZ) Input: Share price Beginning : 1 End : 69**			
	Forecast	Lower 90.0000%	High end 90.0000%	Standard error
70	441.3000	383.1765	499.4235	34.84797
71	441.3001	359.1010	523.4991	49.28247

The figure 2 as well as the table 1, shows that the forecast for both following months is almost the same. The values differ by 0.0001 CZK. However, the Box-Jenkins method indicates the pessimistic and optimistic variant, or lower-end and high-end limit of the forecast interval. The limit is, however, set at the level of +- 10 %. Therefore, this leads to situation, when in October 2017, the share price is assumed to be in the interval of ⟨383.1765; 499.4235⟩. The difference between the high-end and lower-end limit is 116 CZK. Similarly, in November 2017, the share price is supposed to be in the interval of ⟨359.1010; 523.4991⟩. In this month, the share price is thus predicted with an inaccuracy of more than 164 CZK. In November, we are thus working on the interval that is close to half of the lower predicted price. The question is whether it is possible to consider such inaccuracy adequate and whether it is possible to work with it in practice. We will therefore try to analyse the forecast residuals (that is, the difference between predicted and real price of ČEZ shares) and find out whether better results can be achieved using the Box-Jenkins method.

Figure 3 shows graphical illustration of residuals.

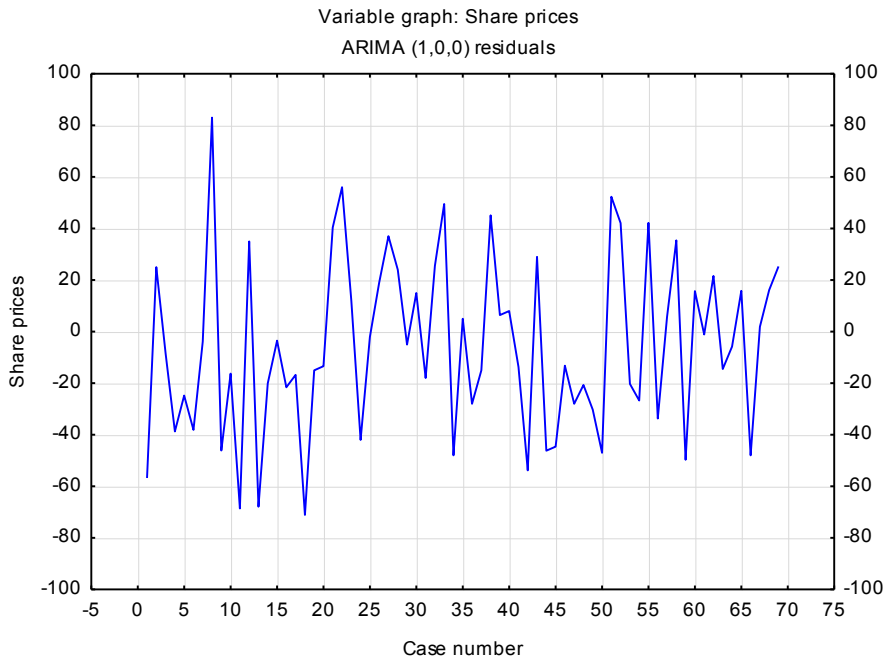

Figure 3. Residuals of share prices forecasts

The shape of the curve connecting residuals shows that the differences of predicted and real share prices development are rather big and the model suggested is not optimal. Two types of results could be achieved. We could either obtain a relatively precise model (with a minimal differences within the interval of the high-end and low-end limit), whose accuracy would be relatively low, or a model with a wider range between lower-end and high-end limit. However, share prices will most likely move within the predicted interval. In our case, the second alternative has been chosen, that is, wider range of interval with a high probability that we will move within the given interval.

Figure 4 illustrates this situation.

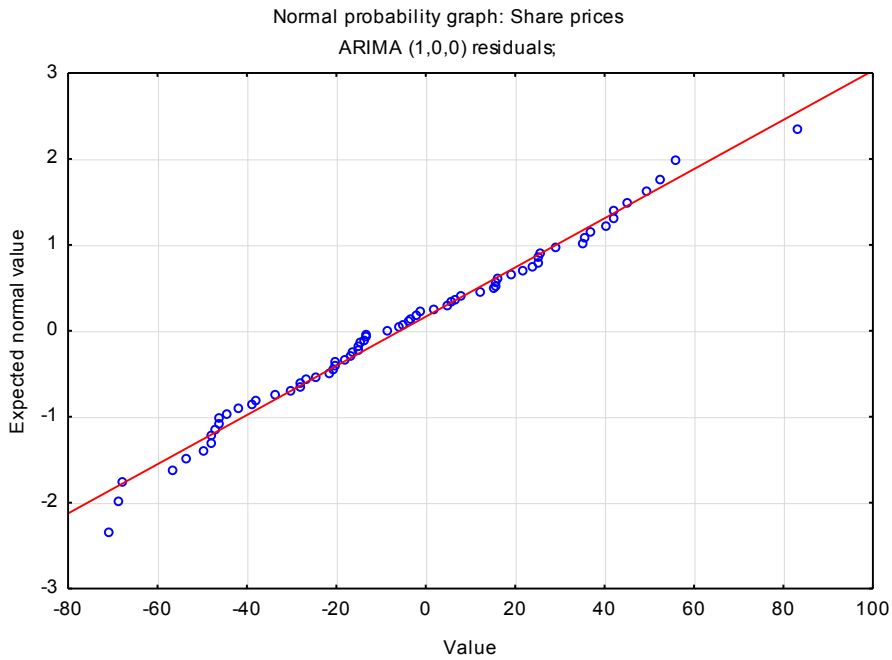

Figure 4. Normal probability of share prices

The figure shows the residuals fitted with the normal probability curve. Figure 5 shows histogram of residuals.

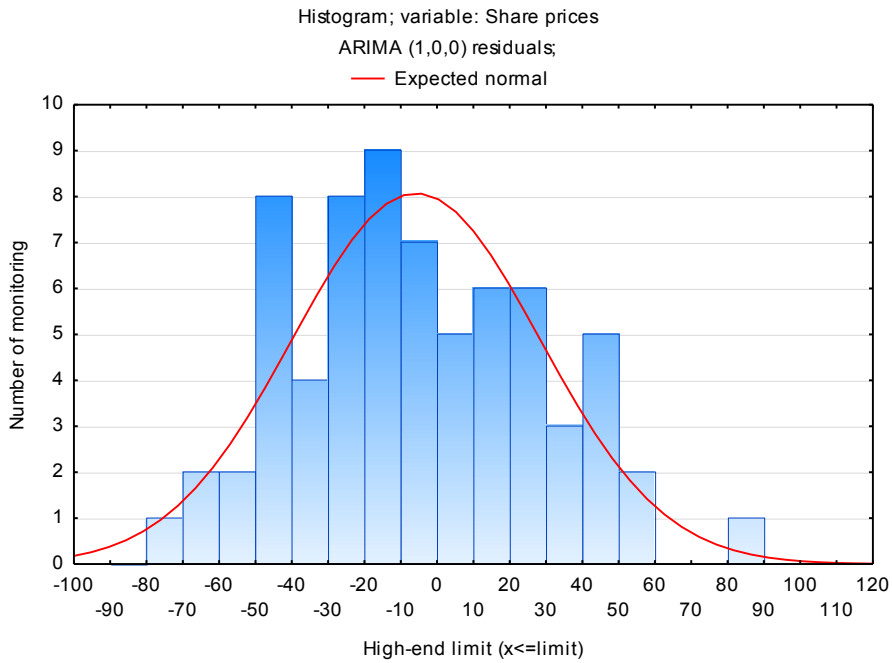

Figure 5. Histogram of residuals

The figure shows normal residuals distribution (replicating the Gaussian curve). However, the peak of the curve is at – 10 CZ. The Gaussian curve is shifted to the left, to negative values. Most probably, the sum of the residuals will be negative. This is, however, not a major problem. The most important thing is the overall number of residuals and their relative high values. This, in fact, confirms the accuracy of a poorly precise result.

Figure 6 shows autocorrelation functions of residuals.

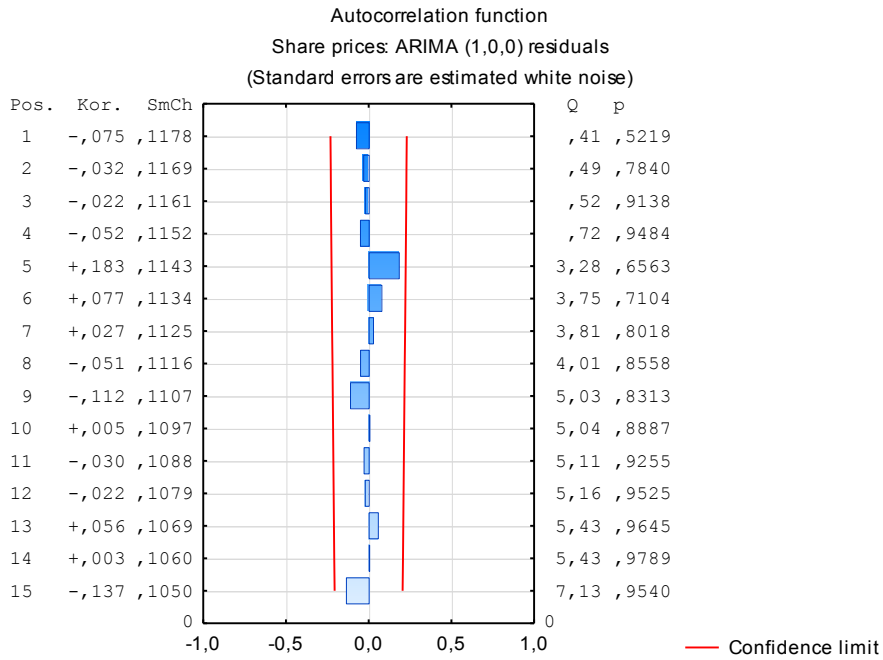

Figure 6. Autocorrelation function of residuals

Figure 6 shows a wide range of possible occurrence of the correct value. The figure clearly shows standard errors, correlations and especially confidence limits marked in red. It may therefore be concluded, that the result has been calculated correctly and that in October and November 2017, the ČEZ share price will be within the calculated interval.

4 Conclusion

The objective of the contribution is to forecast future development of ČEZ share prices on the Prague Stock Market using the Box-Jenkins method.

First of all, the data series was analysed. Next, forecasts of future ČEZ share prices development were calculated. The low-end and high-end limits of share prices forecasts for October and November were identified. However, the results indicate that there is a significant difference between low-end and high-end limits. Subsequently, the forecast residuals were analysed. Mainly due to autocorrelation functions of residuals and thus also elimination of white noise it was confirmed, that the calculation was carried out correctly, but with low accuracy. The question thus arises, whether the Box-Jenkins method is the right forecasting tool. It would therefore be convenient to validate the results by means of other method (with the same data set).

The objective of the contribution was achieved.

References

1. D. Walterová, M. Vochozka, Š. Bendová, The Subprime Mortgage and Bank Crisis in the Financial Market. *Littera Scripta,* **2**(1), 95-108 (2009)
2. X. Li, H. Zhao, K. Zheng, S. Sun, A DFS Model for Forecasting Stock Price. *Proceedings of the 2nd Workshop on Advanced Research and Technology in Industry Applications*, Dalian, China, **81**, 1678-1683 (2016)
3. Y. Ch. Wang, T. N. Nguyen, Forecasting Stock Prices for an Emerging Market: A Case of Vietnam. *Proceedings of the 2nd International Conference on Finance and Economics*, Ho Chi Minh City, Vietnam, 190-200 (2015)
4. Ch. Wu, P. Luo, Y. Li, K. Chen, Stock Price Forecasting: Hybrid Model of Artificial Intelligent Methods. *Engineering Economics,* **26**(1), (2015)
5. M. F. Anaghi, Y. Norouzi, A Model for Stock Price Forecasting Based on ARMA Systems. *Proceedings of the 2nd International Conference on Advances in Computational Tools for Engineering Applications,* Zouk Mosbeh, Lebanon, 265-268 (2012)
6. Y. S. Lu, J.L. Zhang, Forecasting stock price by SVMs regression. *Proceedings of the 11th International Conference on Artificial Intelligence: Methodology, Systems, and Applications*, Varna, Bulgaria, **3192**, 295-303 (2004)
7. P. F. Pai, Ch. S. Lin, A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega – International Journal of Management Science*, **33**(6), 497-505 (2005)
8. P. H. B. F Franses, The Econometric Modelling of Financial Time Series. *International Journal of Forecasting*, **16**(3), 426-427 (2000)
9. A. B. Sánchez, C. Ordóñez, F. S. Lasheras, F. De Cos Juez and J. Roca-Pardiñas, Forecasting SO 2 Pollution Incidents by means of Elman Artificial Neural Networks and ARIMA Models. *Abstract and Applied Analysis*, (2013)
10. J. Junttila, Structural breaks, ARIMA model and Finnish inflation forecasts. *International Journal of Forecasting*, **17**(2), 203-230 (2001)
11. G. Mélard, J. M. Pasteels, Automatic ARIMA modeling including interventions, using time series expert software. *International Journal of Forecasting*, **16**(4), 497-508 (2000)
12. ČEZ, 2017, *ČEZ, a.s. – about company* [online], Available at: https://www.cez.cz/cs/o-spolecnosti/cez/profil-spolecnosti.html (2017)