APPLICATION OF ARMA AND GARCH MODELS ON TIME SERIES OF KOMERČNÍ BANKA STOCKS

Markéta Sedláková
University of South Bohemia, Faculty of Economics,
Department of Applied Mathematics and Informatics,
Studentská 13, 370 05 České Budějovice, Czech Republic
e-mail: marketa.dvorakova029@gmail.com

Abstract

One of the main goals of entities in the securities market is to buy stocks cheaply at the right time (buy undervalued stocks) and also to sell stocks expensively at the right time (sell overvalued stocks).

It is typical for undervalued stocks to be traded at a low price at some point due to their fundamental characteristics. In the future, undervalued stocks can be expected to rise in price, which can bring a capital gain to an investor who bought at a low price. Undervalued stocks are therefore recommended for purchase.

An important prerequisite is therefore the accurate timing of buy and sell signals. The problem, however, is that no one knows exactly when the right time will come, because the market price of stocks is affected by many factors that have an impact on fluctuations in the market value of stocks. For this reason, volatility modeling is coming to the forefront of the interests of many financial analysts and investors.

Keywords

ARMA model; GARCH model; Stocks; Volatility; Komerční banka.

Introduction

The so-called Mathematical modeling provides methods by which phenomena and activities that take place in a person's daily life can be studied. Thus, mathematical modeling makes it very easy to display complex questions or problems through mathematical equations or functions. One of the areas where mathematical models can be used very well is the area of finance, in which the financial market has a key position. Based on supply and demand, there is a movement of money and capital between different entities. The main platform for these financial transactions is stock exchanges, where the main attractions include stocks of large companies.

However, the problem with these transactions is a high degree of uncertainty, as stock volatility is largely unpredictable. Reasons include, for example, measures or regulations of governments in a given country, market expectations, financial or other crises, or the political situation in a given country. For this reason, prediction or modeling of stock volatility comes to the forefront of the interests of many investors, economists, speculators or financial analysts.

1   Research Subject

The aim of this article is to analyze the behavior of Komerční banka stocks within a given time interval using the application of methods and analyzes. This behavioral analysis will be

performed based on the residue distribution calculated from the GARCH model, which will be estimated from the data. The work will analyze both the expected changes that follow the normal distribution, but mainly unexpected changes – the so-called heavy tails.

The article uses as a data source the time series of Komerční banka stocks in the period from January 2018 to February 2019. For the analysis itself, the statistical software R was used as a tool.

2 Research Process

First of all, the source data must be treated so that there is no non-stationarity and inhomogeneity in the data. The so-called autocorrelation also occurs as an undesirable phenomenon in the case of time series. Failure to respect the autocorrelation of residues results in skewed estimates of unknown parameters, which also affects other characteristics. If autocorrelation occurs in the time series, then the residues are not independent. The time series therefore need to be cleaned so that the resulting p-values are not skewed. Therefore, time series will be logarithmized first, which will remove their multiplicative character (multiplicative changes of time series will be converted to additive changes). Then the method of the so-called 1st order difference (differentiation of source data) will be applied, which will remove the non-stationarity of this data. To verify that the given values are independent, the so-called autocorrelation function (independence verification) and partial autocorrelation function (independence verification when removing the influence of the third quantity) will be used. The autocorrelation function (ACF) is a suitable imaging tool for detecting visible patterns in data. The ACF value then indicates at different time intervals (Lag) whether there is any form of automatic correlation in the data.

Next, the GARCH model (with parameters p, q) will be used, which is a model for examining time series volatility. Using the GARCH model, heteroskedasticity will be removed from the data. At this point, the already adjusted data show the characteristics of white noise, which means that all influences are removed – so we obtain independent, equally distributed (same variance) of the random variable. In addition to the above assumptions (homogeneous, stationary series and non-correlation of logarithmic returns), the third assumption of volatility models, which is the normality of logarithmic returns, must be observed. Histogram display can be used to verify the normality of residues.

3 Literary Research and Formulas

3.1 Time Series Issues

The time series represents the so-called numerical variable, the values of which are largely dependent on the time in which these values were obtained. It is basically a sequence of chronologically arranged observations. The time points at which the data were obtained are usually equally distant from each other [5].

Description through descriptive statistics can provide a sufficient idea of the properties of time series as a single data unit but does not provide information about its time evolution [5].

Time series can be classified on the basis of various aspects [5]:

- according to the nature of the data, the values of which form a time series,
- interval time series - the data depends on the length of the interval that is monitored,
- instantaneous time series - data refer to a specific moment,
- according to the periodicity with which the data are monitored,
- time series of annual data,

- short-term time series,
- by type of data monitored,
- time series of absolute indicators,
- time series of derived characteristics – e.g. cumulative time series.

Time series are the result of observations made at discrete time points. Some of them are then discreet in their nature (as an example, time series of total production of a certain agricultural crop for individual years), others need to be "discretized" first. Thus, time series can be created by discretization of values of a continuously changing quantity (e.g. a series of values of amplitude of a signal at given time points), accumulation of values of monitored quantity for a given time period (daily precipitation totals in meteorology) or by averaging values of considered quantity in given time interval (average daily temperatures) [4].

If there is a choice, then it is recommended to choose a compromise solution. The high density of observation time points allows the characteristics of the time series to be well captured, but calculation difficulties can occur. The choice of equidistant intervals between adjacent observations should be a matter of course. As part of the analysis of economic time series, problems associated with the calendar may occur (different lengths of calendar months, different number of working days per month, moving holidays). In such cases, a so-called "standard month" of 30 days or a standard number of working days per month is usually introduced, or the observed data are accumulated. The length of a time series is defined as the total number of observations in the time series, not as the time span between the first and last observations [4].

3.2 Time Series Autocorrelation

A key assumption underlying the linear regression model (LRM) commonly used in applied econometric studies is a sufficient limitation of a phenomenon called autocorrelation [6].

An important feature of time series is their (potential) serial correlation. Therefore, a thorough analysis and visualization of these correlations is needed. The autocorrelation between two random variables X_t and $X\_(t+k)$ can be described as follows [2]:

$$Cor(X_{t+k}, X_t) = \frac{Cov(X_{t+k}, X_t)}{\sqrt{Var(X_{t+k})}Var(X_t)} \tag{1}$$

Since moments are required for stationary data to be constant over time, autocorrelation can be written for these values as a function of delay [2]:

$$\rho(k) = Cor(X_{t+k}, X_t) \tag{2}$$

The most common autocorrelation test in the regression model is the borderline Durbin - Watson test, which is used to test the independence of residues in the normal regression model. The test finds application when the data are obtained sequentially, and the values of the dependent variable form a time series [3].

Durbin - Watson test is calculated as [2]:

$$\widetilde{D} = \frac{\sum_{t=2}^{n}(r_t - r_{t-1})^2}{\sum_{t=1}^{n} r_t^2} \tag{3}$$

3.3   Model AR-1

The autoregressive model of a time series is based on the assumption that any value in a time series depends on the previous value of that series. The autoregressive model AR (p) of order p can be defined as follows [1]:

$$y_t = b_1 y_{t-1} + b_2 y_{t-2} + \cdots + b_p y_{t-p} + \epsilon_t \tag{4}$$

Where
$b_1, b_2, \ldots, b_p$ are the coefficients within the autoregressive process,
ϵ_t is the so-called white noise (current value), and
y_t is the new series value calculated based on the previous values.

3.4   MA Moving Sum Model

The process MA (q) – (Moving Average) - of order q can be written as follows [1]:

$$y_t = \epsilon_t + w_1 \epsilon_{t-1} + w_2 \epsilon_{t-2} + \cdots + w_q \epsilon_{t-q} \tag{5}$$

Where
w are the parameters of the model,
ϵ_t is white noise.

3.5   ARMA Model

By combining the already mentioned processes AR (p) and MA (q), a mixed process in the form of ARMA (p, q) can be obtained [1].

The condition of stationarity of the ARMA process coincides with the condition of stationarity of the AR process (p) and the condition of process invertibility is the same as the condition of process invertibility MA (q). The mean value of the ARMA process is also zero (as in the previous AR and MA processes) and its autocorrelation function satisfies a similar system of difference equations as in the AR process. Several alternative options can be selected for writing the ARMA process in the form of a difference equation, a linear process or in an invertible form. For each AR model of order p, an equivalent MA model with a sufficient number q of the interference element can be found. Economic or business time series can be modeled using a relatively small number of p and q elements within the AR, MA or ARMA model. The aim is to find or determine the smallest number of p and q elements needed for satisfactory time series prediction [1].

4   Practical Research

The models will be applied to the time series of Komerční banka stocks. This is a series with the length of 285 observations. This is the daily development of stocks with the exception of weekends and holidays from January 2018 to February 2019. The development of stocks in this period is shown in Figure 1.
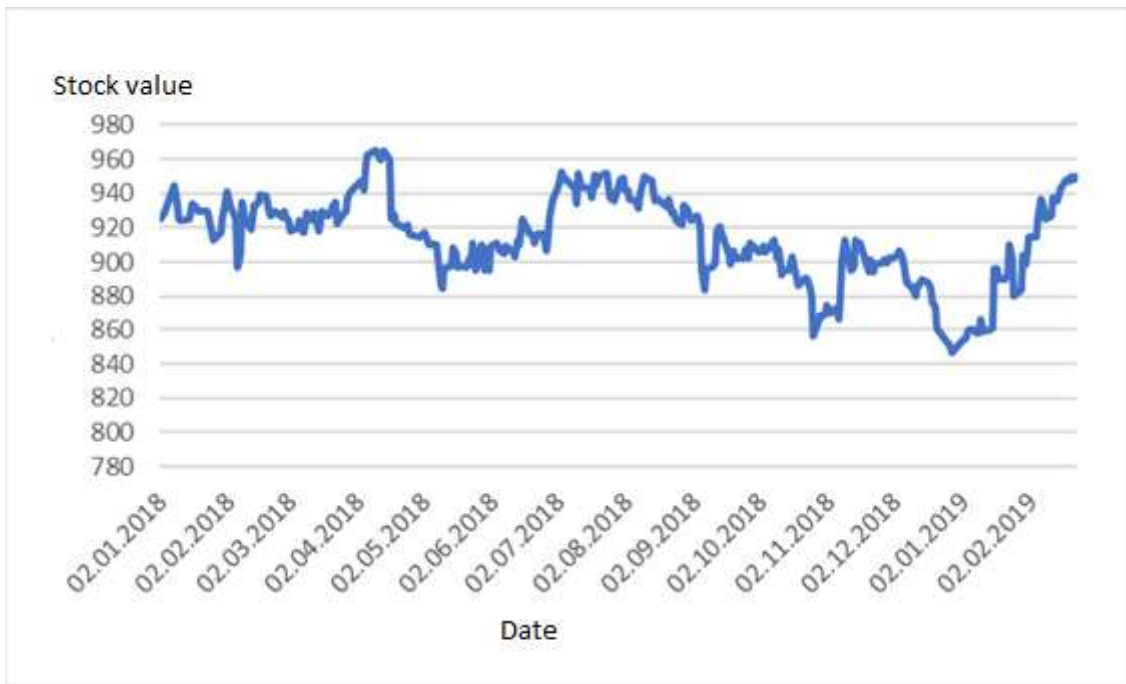
**Fig. 1:** *Daily development of Komerční banka stocks*

At first glance, it is clear, that the series is inhomogeneous and non-stationary. For verification, the so-called autocorrelation function is used by default, which can be seen in Figure 2, and the partial autocorrelation function shown in Figure 3. But first, the time series must be inserted into the software R and commands for autocorrelation (verification of stationarity) and partial autocorrelation function (verification of the influence of the third quantity) must be used.
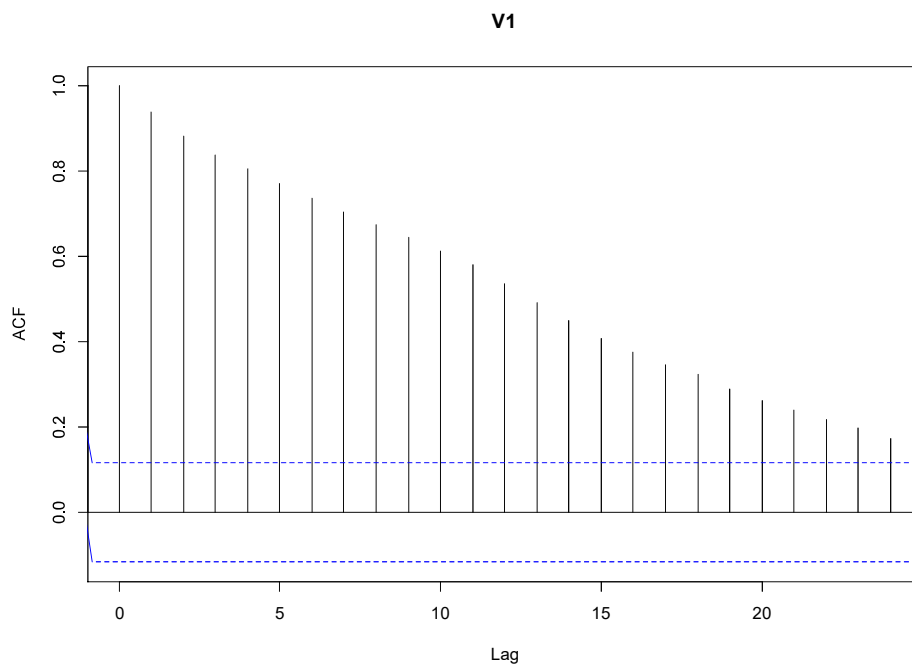
Fig. 2: *Autocorrelation function (ACF) of Komerční banka stocks*

The autocorrelation function (ACF) is a suitable imaging tool for detecting visible patterns in data. The ACF value at different time intervals (Lag) indicates whether there is any form of automatic correlation in the data. In Figure 2 you can see clearly visible patterns between the data. In other words, a clear violation of the presumption of independence can be noticed.

**Series akcie**
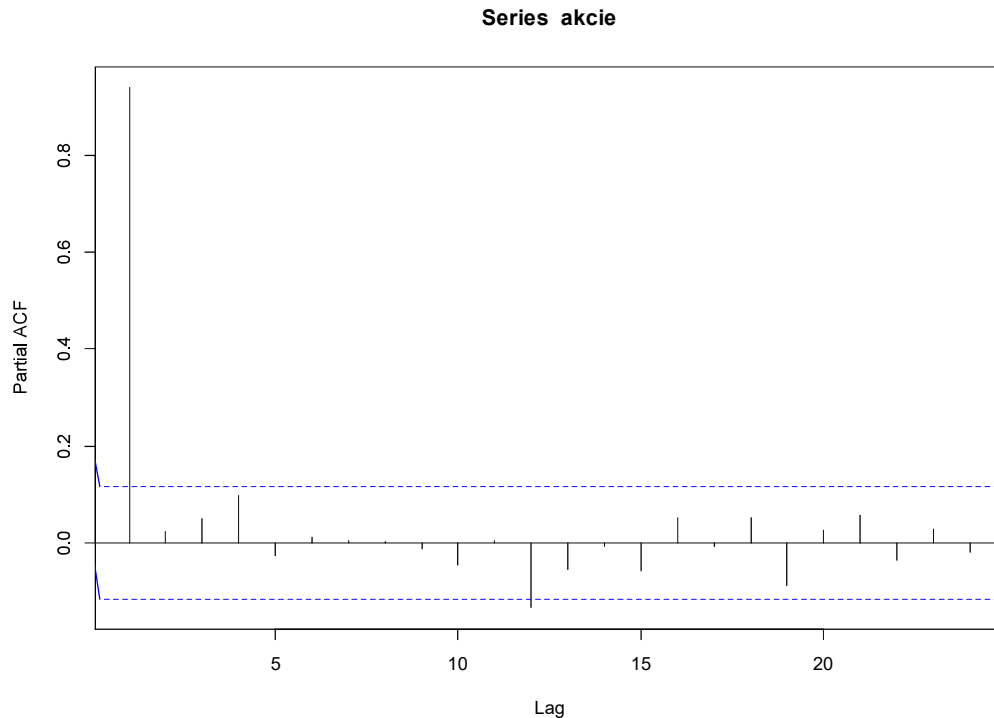
Source: Own
Fig. 3: *Partial autocorrelation function (PACF) of Komerční banka stocks*

The autocorrelation function gradually decreases and the first PACF value is close to 1, which indicates that the series is not stationary.

In order to apply the models, a number of KB stocks need to be transformed. Thus, it will be necessary to logarithm the data, which will remove their inhomogeneity, and subsequently it will be necessary to make a difference, which will remove their non-stationarity.

After using commands for logarithm the data and making a difference, it is now possible to identify the model by estimating the autocorrelation and partial autocorrelation functions, which are shown in Figures 4 and 5. It can be deduced that this is indeed a stationary series, because ACF values no longer gradually decrease but move in a certain interval.
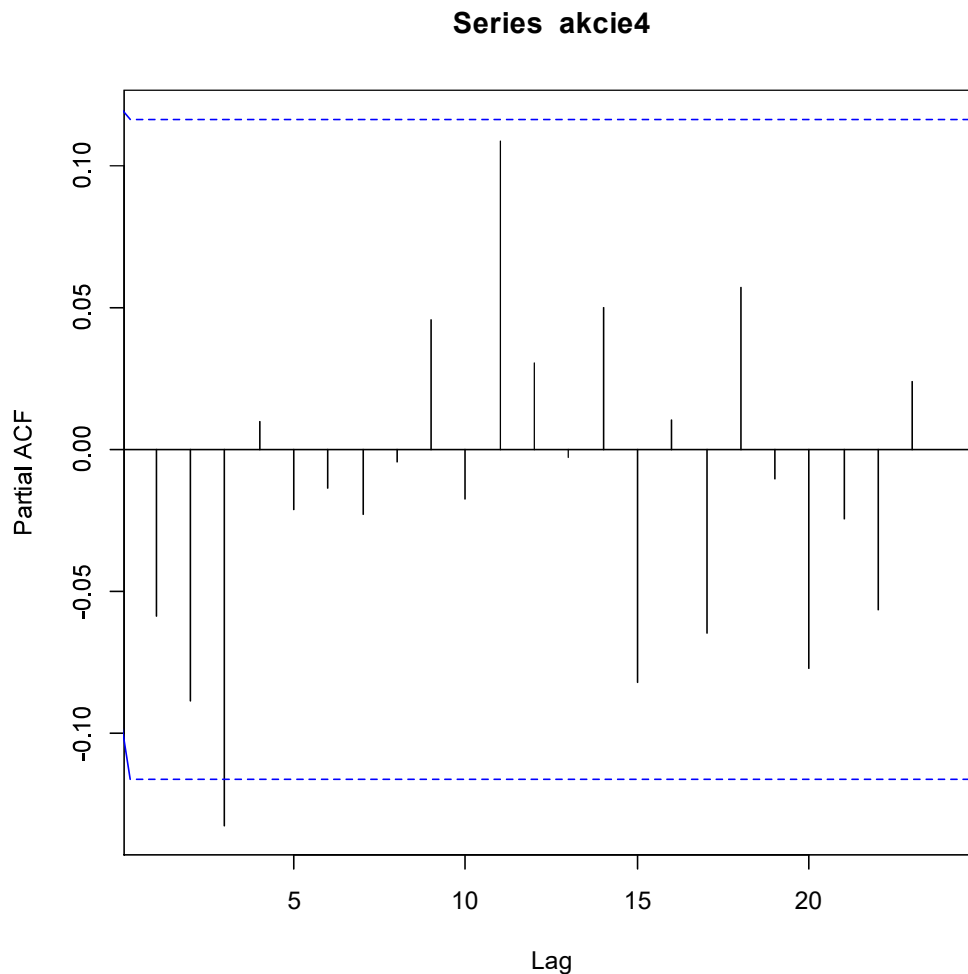
Series akcie4
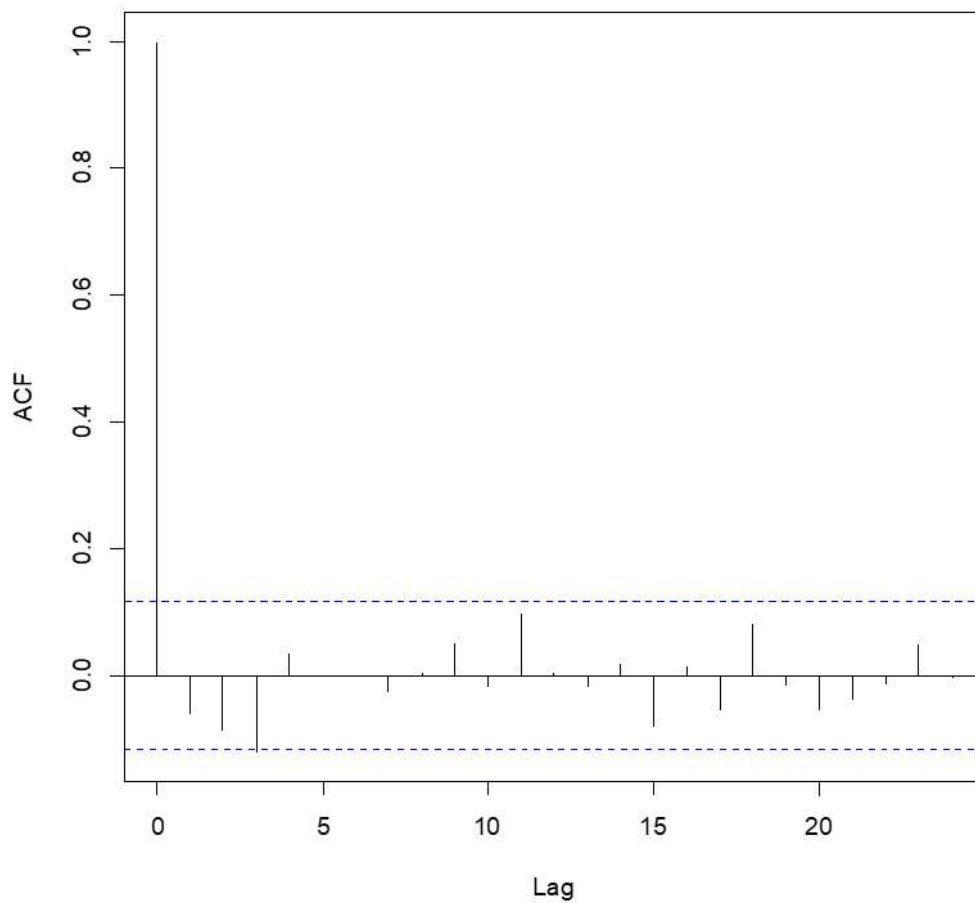
Source: Own
Fig. 4: *ACF stationary series of Komerční banka stocks*

The dashed line in the correlogram determines the truncation points (confidence interval). Values between zero and this limit are considered insignificant, i.e. zero. Thus, it is clear, that the series does not contain any MA process because the ACF values are zero.

Series akcie4

Fig. 5: *PACF stationary series of Komerční banka stocks*

The partial autocorrelation function helps determine the order of the AR process. However, the monitored series also does not include this process, because all PACF values are insignificant. Correlations could not be demonstrated in this time series, so the whole series is considered white noise. The GARCH model will therefore be used for modeling.

### 4.1 Estimation of Volatility Models on the Time Series of Komerční Banka Stocks

Now the volatility model will be applied to the same time series of KB stocks. The GARCH model (*p, q*) will be used. However, before the GARCH model is applied, the basic assumptions for modeling the volatility of a given series must be verified. A significance level of 0.05 is considered for all tests below.

The given time series must again be transformed into a stationary series by means of logarithmization and subsequently difference. Figures 4 and 5 are proof of the stationary series. Furthermore, the non-correlation of logarithmic returns is determined, for example using the ACF and PACF functions. As already shown above, the given time series does not contain a correlation of random variables.

The third basic assumption of volatility models is the normality of logarithmic returns. The Jarque-Ber test can be used to verify normality. For JB test, the null hypothesis is followed, for which the normality of the distribution of logarithmic returns is assumed, as well as the

alternative hypothesis, which states that logarithmic returns do not have a normal distribution. The *p*-value is very low (2.2e-16 in Figure 7), which is less than the significance level of 0.05, so the normality of logarithmic returns must be rejected.

A histogram was also used as confirmation. The fact that the logarithmic yields do not have a normal distribution is shown in Figure 6, which shows a histogram of the actual distribution of the logarithms of the series yields. Logarithmic returns have a sharper distribution, which is typical for financial series.

Histogram of akcie4
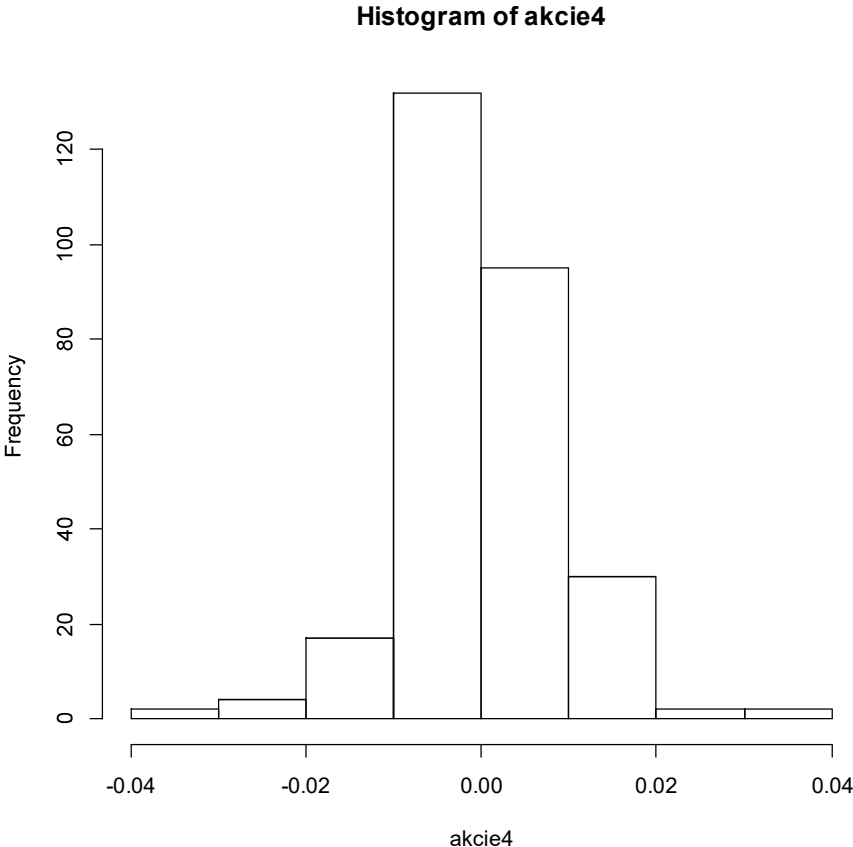
Source: Own
Fig. 6: *Histogram*

Although the Jarque-Ber test did not show the existence of a normal distribution in the data, this time series can be used in the volatility model. The GARCH (1, 1) model will be applied to the time series, which is the most used model for examining time series volatility. The results from the R studio are shown below in Figure 7.

```
> summary(arch.y)

Call:
garch(x = y, order = c(1, 1))

Model:
GARCH(1,1)

Residuals:
    Min      1Q  Median      3Q     Max
-4.2741 -0.5311 -0.0111  0.5278  4.6664

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)
a0 3.564e-05  2.527e-05    1.411    0.158
a1 9.140e-02  6.998e-02    1.306    0.192
b1 4.876e-01  3.488e-01    1.398    0.162

Diagnostic Tests:
        Jarque Bera Test

data:  Residuals
X-squared = 118.47, df = 2, p-value < 2.2e-16


        Box-Ljung test

data:  Squared.Residuals
X-squared = 0.47471, df = 1, p-value = 0.4908
```

Source: Own
Fig. 7: *GARCH function (1,1)*

A more important parameter in the model is the parameter b1, so the resulting model GARCH (1, 1) has the form:

$$\rho_t{}^2 = 0{,}00003564 + 0{,}0914e_{t-1}{}^2 + 0{,}487\rho_{t-1}{}^2 \tag{6}$$

However, the *Pr* value is greater than the significance level of 0.05. For this reason, it is white noise. It can therefore be stated that the market values of Komerční banka stocks in the period under review are not affected by the impact of new information but are also not affected by past variance values. However, if this were not the case, based on the parameters of this function, it can be stated that the impact of new information ($e_{t-1}{}^2$) is not as important within this time series as past values of variance ($0{,}4876\rho_{t-1}{}^2$) and the modeling of market values of stocks should be based on past values, as this is the most significant factor influencing this time series.

The GARCH (0, 1) model can also be applied to the time series, which is also a frequently used model in the study of time series volatility. The results from the R studio are shown in Figure 8.

```
> summary(arch.y)

Call:
garch(x = y, order = c(0, 1))

Model:
GARCH(0,1)

Residuals:
     Min      1Q   Median       3Q      Max
-4.02586 -0.52718 -0.01033  0.52449  4.37558

Coefficient(s):
    Estimate  Std. Error  t value  Pr(>|t|)
a0 8.248e-05   5.314e-06   15.522    <2e-16 ***
a1 2.118e-02   6.142e-02    0.345      0.73
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Diagnostic Tests:
        Jarque Bera Test

data:  Residuals
X-squared = 89.696, df = 2, p-value < 2.2e-16


        Box-Ljung test

data:  Squared.Residuals
X-squared = 0.026868, df = 1, p-value = 0.8698
```

Source: Own
Fig. 8: *GARCH function (0,1)*

A more important parameter in the model is the parameter *a1*, the resulting model GARCH (0,1) has the form:

$$\rho_t{}^2 = 0{,}00008{,}248 + 0{,}02118 e_{t-1}{}^2 \tag{7}$$

The value of the indicator *Pr* is less than the significance level 0.05 only in the case of parameter *a0*. Therefore, it can be stated again that the impact of the new information is not significant in the model, therefore it is a white noise and the market values of Komerční banka stocks in the monitored period are not affected by the new information.

Conclusion

The aim of the article was to apply the ARMA and GARCH model to the time series of Komerční banka stocks using the analysis in the R software and to find out the behavior of these stocks. The most commonly used form of autocorrelation is the first-order AR (1) autoregression, which was also used.

Based on the GARCH (1, 1) and GARCH (0, 1) models, it was found that the examined time series of market values of Komerční banka stocks showed signs of white noise. The market values of Komerční banka stocks in the period under review are therefore not affected by the impact of new information but are also not affected by past variance values.

The issue of time series of stocks is a very specific topic, because what may apply in one market may not apply in another market, or even what applies to the market values of stocks of one company may not show the same results for another company.

The resulting values are also affected by the observed period, when in a certain period the results may be completely different from other periods. If the time series of market values of Komerční banka stocks did not show white noise values, it could be said that the impact of new information is not as important within the examined time series as past values of variance and stock value modeling should be based on past values, as it is the most important factor influencing the market value of stocks.

Literature

[1] DANEL, R.: *Predikce časové řady pomocí autoregresního modelu*. [online]. Czech Republic, Technická univerzita Ostrava, 2004. Available from WWW: http://homel.vsb.cz/~dan11/publikace/Danel_Autoregresni_model_predikce_casovych_rad.pdf

[2] DETTLING, M.: *Applied Time Series Analysis*. [online]. Switzerland, Zurich: University of Applied Sciences, 2018. Available from WWW: https://ethz.ch/content/dam/ethz/special-interest/math/statistics/sfs/Education/Advanced%20Studies%20in%20Applied%20Statistics/course-material-1719/Zeitreihen/ATSA_Script.pdf

[3] DUFOUR, J.-M.; DAGENAIS, M. G.: Durbin – Watson Tests for Serial Correlation in Regressions with Missing Observations. *Journal of Econometrics*. 1985, Vol. 27, Issue 3, pp. 371–381. DOI: 10.1016/0304-4076(85)90012-0

[4] KŘIVÝ, I.: *Analýza časových řad*. [online]. Czech Republic, Ostrava, 2012. Available from WWW: https://web.osu.cz/~Bujok/files/ancas.pdf

[5] LITSCHMANNOVÁ, M.: *Úvod do analýzy časových řad*. [online]. Czech Republic, Technická univerzita Ostrava, 2010. Available from WWW: https://homel.vsb.cz/~lit40/ZS/Prezentace/ZS_12_CasRady.pdf

[6] McGUIRK, A.; SPANOS, A.: The Linear Regression Model with Autocorrelated Errors: Just Say No to Error Autocorrelation. In: *Annual Meeting of the American Agricultural Economics Association*. [online]. USA, Virginia Tech, 2002. Available from WWW: https://core.ac.uk/download/pdf/6407474.pdf

Ing. Markéta Sedláková

APLIKACE ARMA A GARCH MODELŮ NA ČASOVÉ ŘADĚ AKCIÍ KOMERČNÍ BANKY

Jedním z hlavních cílů subjektů na trhu cenných papírů je ve správný okamžik akcie levně nakoupit (nakoupit podhodnocené akcie) a rovněž ve správný okamžik akcie draze prodat (prodat nadhodnocené akcie).

Pro podhodnocené akcie je typické, že se v určitém okamžiku, vzhledem ke svým fundamentálním charakteristikám, obchodují za nízký kurz. Do budoucna lze u podhodnocených akcií předpokládat růst jejich ceny, což může investorovi, který nakoupil při nízkém kurzu, přinést kapitálový zisk. Podhodnocené akcie jsou tedy doporučovány k nákupu.

Důležitým předpokladem je tedy přesné načasování nákupních a prodejních signálů. Problém však je, že nikdo přesně neví, kdy nastane ten správný okamžik, protože tržní cena akcií je ovlivněna mnoha faktory, které mají dopad na kolísání tržních hodnot akcií. Z tohoto důvodu se modelování volatility dostává do popředí zájmu mnoha finančních analytiků a investorů.

ANWENDUNG VON ARMA UND GARCH MODELLEN AUF DIE ZEITREIHE DER AKTIEN VON KOMERČNÍ BANKA

Eines der Hauptziele von Unternehmen auf dem Wertpapiermarkt ist es, Aktien zum richtigen Zeitpunkt billig zu kaufen (unterbewertete Aktien zu kaufen) und Aktien zum richtigen Zeitpunkt teuer zu verkaufen (überbewertete Aktien zu verkaufen).

Es ist typisch, dass unterbewertete Aktien aufgrund ihrer fundamentalen Eigenschaften irgendwann zu einem niedrigen Preis gehandelt werden. In Zukunft ist mit einem Preisanstieg bei unterbewerteten Aktien zu rechnen, was einem Anleger, der zu einem niedrigen Preis gekauft hat, einen Kapitalgewinn bringen kann. Unterbewertete Aktien werden daher zum Kauf empfohlen.

Eine wichtige Voraussetzung ist daher das genaue Timing der Kauf- und Verkaufssignale. Das Problem ist jedoch, dass niemand genau weiß, wann der richtige Zeitpunkt kommt, da der Marktpreis von Aktien von vielen Faktoren beeinflusst wird, die sich auf Schwankungen des Marktwerts von Aktien auswirken. Aus diesem Grund tritt die Volatilitätsmodellierung in den Vordergrund der Interessen vieler Finanzanalysten und Investoren.

ZASTOSOWANIE MODELI ARMA I GARCH NA SZEREGU CZASOWYM AKCJI BANKU KOMERČNÍ BANKA

Jednym z głównych celów podmiotów działających na rynku papierów wartościowych jest tani zakup akcji we właściwym czasie (zakup akcji niedowartościowanych), a także sprzedaż akcji we właściwym czasie (sprzedaż akcji przewartościowanych).

Niedowartościowane akcje charakteryzują się tym, że w pewnym momencie obrót nimi ze względu na ich fundamentalne cechy odbywa się po niskim kursie. W przypadku akcji niedowartościowanych można się w przyszłości spodziewać wzrostu ich ceny, co może inwestorowi, który kupił po niskiej cenie, zapewnić zysk kapitałowy. Zakup niedowartościowanych akcji jest więc zalecany.

Ważną przesłanką jest zatem dobre trafienie terminu zleceń kupna i sprzedaży. Problem polega jednak na tym, że nikt nie wie dokładnie, kiedy nadejdzie ten właściwy moment, ponieważ na cenę rynkową akcji wpływa wiele czynników, które mają wpływ na wahania wartości rynkowej akcji. Dlatego też modelowanie wahań kursów leży w centrum zainteresowania wielu analityków finansowych i inwestorów.