ARIMA MODEL FOR PREDICTING THE DEVELOPMENT OF THE PRICE OF GOLD: EUROPEAN APPROACH

Ligita Gaspareniene[1,a,*] and Rita Remeikiene[2,b]

[1]Mykolas Romeris University, Ateities st. 20, LT-08303 Vilnius, Lithuania
[2]Mykolas Romeris University, Ateities st. 20, LT-08303 Vilnius, Lithuania
[a]ligitagasparenine@mruni.eu, [b]rita.remeikine@mruni.eu
Corresponding author

Abstract: Time series analysis has a long tradition in economics. The foundations of the current time series analysis, focusing on modelling of the development of one time series, were laid in the interwar period. In recent years, many new methods and approaches have emerged in the field of economic time series analysis. Also, the conditions for their practical application in our country have significantly improved in recent years, which is due to the relatively easy availability of a relatively wide range of types of quality software and the growing length of the analysed time series. In this research paper, we focused on the analysis of time series using the ARIMA model to predict the price of gold by the end of 2020. Those who did not believe that the gold price could strengthen significantly shortly before the outbreak of COVID-19 they are not enough to wonder. Its value returned to green numbers after the fall in the price of the commodity, which caused growing concerns after the outbreak of the world virus pandemic. It was the reason for the analysis of this particular commodity. The introduction contains the necessary information on the problem and basic knowledge of the theory of time series and their modelling, characterizes the Box - Jenkins methodology, and finally, the statistical software used. The method describes the procedure of data processing and the creation of the ARIMA model. The results section presents results and predictions in the last part of the discussion and conclusion; the results and interpretations of the results are summarized.

Keywords: prediction, gold, ARIMA modelling, time series

JEL Classification: C53, C22, C51

1. Introduction

The analysis of time series, including forecasts of their future behavior according to Cipra and Romera (1991), is one of the most critical areas for the development of current statistics because it successfully copes with the description of dynamic systems with which we often come into contact. The data that make up the time series arise as chronologically arranged observations, and the important thing for them is that they are arranged chronologically over time. (Wei, 2005; Lindner et al., 2018)

The methodology developed by Box and Jenkins (2015) was the most influential apparatus in the 1970s and 1980s. Mathematicians around the world have used it to model economic processes. This apparatus, also known as the Box - Jenkins methodology, is still used in time series modeling, although it is a relatively older methodology. Box - Jenkins methodology as

ARIMA model for predicting the development
of the price of gold: European approach
Authors: Ligita Gaspareniene, Rita Remeikiene

an essential element of the construction of time series, according to Martinez-Alvarezet et al. (2015), uses a residual component, which may consist of correlated (dependent) random variables. Thus, we can use these models to process time series with interdependent observations, which is even the focus of their procedures. (Buera and Lucas, 2018) According to Lu and AbouRizk (2009), the Box-Jenkins methodology creates the most accurate prediction models for each data set. (Mustafa and Yusof, 2011) These methods also provide a more systematic approach to building, analyzing, and forecasting time series models. In general, the Box-Jenkins methodology uses the latest observations as default values and then explains past forecast errors to set the future period correctly. (Brockwell and Davis, 2003) A Box-Jenkins method is a prognostic tool based on its ability to deal with complex situations, its adaptability in data processing dependent on time series data, its advanced mathematical and statistical procedures, its operation in risk analysis and uncertainty analysis, its ease of implementation. (Huarng et al., 2020)

According to Nelson (1998), most methods defining statistical relationships between variables require that forecast errors not be correlated. Real data often fail to meet this requirement. If the Durbin-Watson statistic shows that autocorrelation is present, the variance and regression analysis will be invalid and likely to be misleading. These databases can be analyzed by time series methods such as autoregressive integrated moving average models (ARIMA). Time series modeling using ARIMA models is a comprehensive methodology that allows for analysis without the need for time-independent error, which is necessary for most other statistical and modeling methods (Kilian and Taylor, 1991; Gertler and Gilchrist, 2018). ARIMA models are most often used in social sciences and marketing predictions. Besides, ARIMA models can be used as rescue methods if the data collected for linear or logistic regression modeling show autocorrelation. As Lamore (2016) stated, the ARIMA methodology emphasizes not the construction of a one-equation or multi-equation model, such as e.g., in regression analysis, but to analyze the own stochastic properties of economic time series according to the philosophy "let the data speak for themselves." In regression models, the dependent variable Y explained by several explanatory variables - regressors.

In contrast, in ARIMA methods, the dependent variable Y at time t is explained by the values of the same Y in past tense moments and at the same time by error terms in current or past moments. Unlike regression models and models of simultaneous equations, which are based on economic theory, ARIMA models are not directly dependent on the method. Technical dependencies are expressed indirectly through the observed values in past time points.Auto-Regressive Integrated Moving Average (ARIMA) models are one of the most important time series models used in financial market forecasting over the past three decades but not very often used to forecast gold prices. (Abdullah, 2012)

2. Methodology

There are many forecast models used in forecasting gold price. The models which has been used recently are Box-Jenkins models (Miswan et al., 2013; Khashei et al., 2008), back propagation neural network (Yuan, 2012; Zhou et al., 2012; Parisi et al., 2008), system dynamics model (Tharmmaphornphilas et al., 2012), varying-coefficient regression model (Zhang et al., 2011), data mining methods (Mustaffa and Yusof, 2011; Bozanic et al., 2018), jump-and-dip diffusion (Shafiee and Topal, 2010), artificial intelligence models, multiple linear regression models (Ismail et al., 2009; Kramer et al., 2016) or can be the hybrid of the above mentioned models (Khashei et al., 2009, 2008; Hadavandi et al., 2010; Asadi et al., 2012, Larmore, 2016 Bazilier and Girard, 2020).

In the current unstable economic environment, it is not possible to model time series by classical methods. For this reason, new techniques and procedures have been developed to forecast future economic indicators' future values, which quickly adapt to changes in the development of time series. We are talking about the models of the Box - Jenkins methodology, which are the subject of this paper. (Cinnirella and Streb, 2017)

The Box-Jenkins methodology is especially popular among forecasters, who demand better results than the simultaneous equations of models based on economic theory provide. Box-Jenkins models are general enough to allow for non-stationarity and coping with seasonality. (Gibson et al., 2019) However, as they are free from the influence of economic theory, they are not intended to test the economic hypothesis and do not provide estimates of crucial elasticity parameters. Box - Jenkins methodology takes the residual component as the basis for the construction of time series models. It can not only process time series with mutually correlated observations, but even the focus of their procedures lies in the investigation of such dependencies, i.e., in the so-called correlation analysis. (Ramcharan, 2020)

Depending on how many times the difference is necessary to make the series stationary, the time series integration order is determined. Therefore, in the case of non-stationarity, ARIMA models are used, characterized not only by the parameters *(p, q)* but also by the parameter *d* determining the order of integration.

There is a possibility that after the transformation of the integrated process by the difference of the *d*-th series, the resulting process shows such autocorrelations and partial autocorrelations that they can be expressed in the form of a stationary and invertible ARMA model *(p, q)*. Then the original integrated process represented in the way:

$$\emptyset_p(B)(\pm B)^d y_t = \theta_q(B)a_t \tag{1}$$

We call the autoregressive integrated process of moving averages of the series *p, d, q,* and denote them as ARIMA *(p, d, q).*

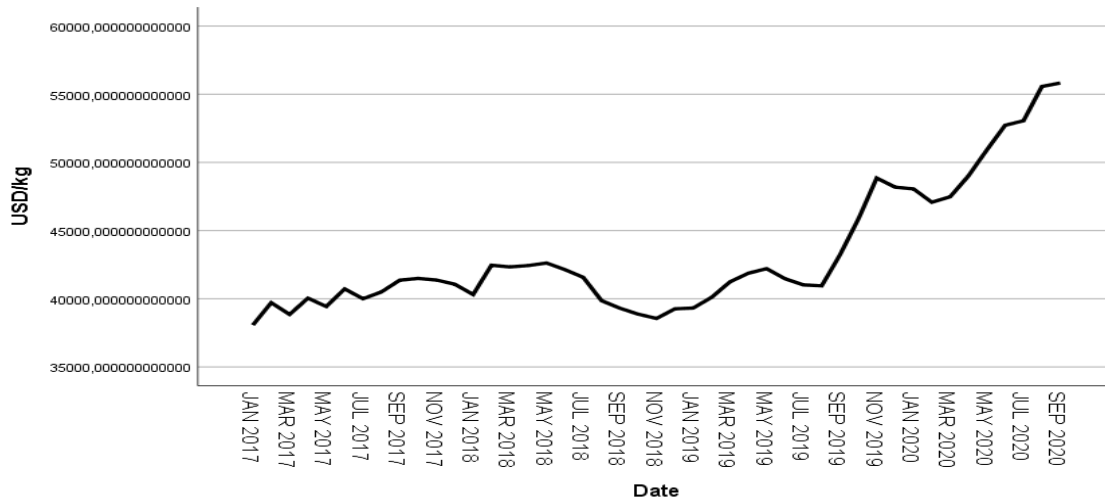A case is an integrated process I (d) written mostly in a simple form.

$$(1 - B)^d y_t = a_t \tag{2}$$

3. Results

This paper aimed to analyze the time series of gold development and then determine the price forecast until the end of 2020. We worked with data from January 2017 every month to May 2020. Our task was to use the ARIMA model to predict price developments until the end of 2020.

Figure 1 shows the development of the price in USD per kg of gold, and we can state that it does not have a seasonal pattern and, at the same time, does not appear to be stationary.
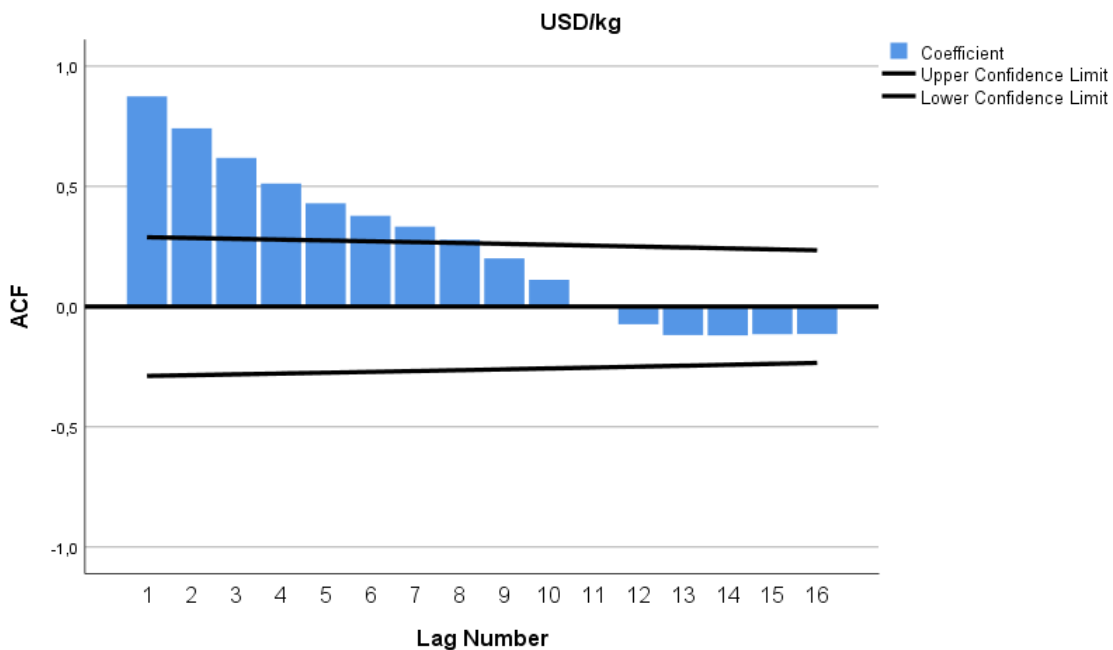
ARIMA model for predicting the development
of the price of gold: European approach
Authors: Ligita Gaspareniene, Rita Remeikiene

Figure 1: Development of the price of gold over time

Stationarity or non-stationarity can be seen in Figure 2 and is determined based on autocorrelations and partial autocorrelations.
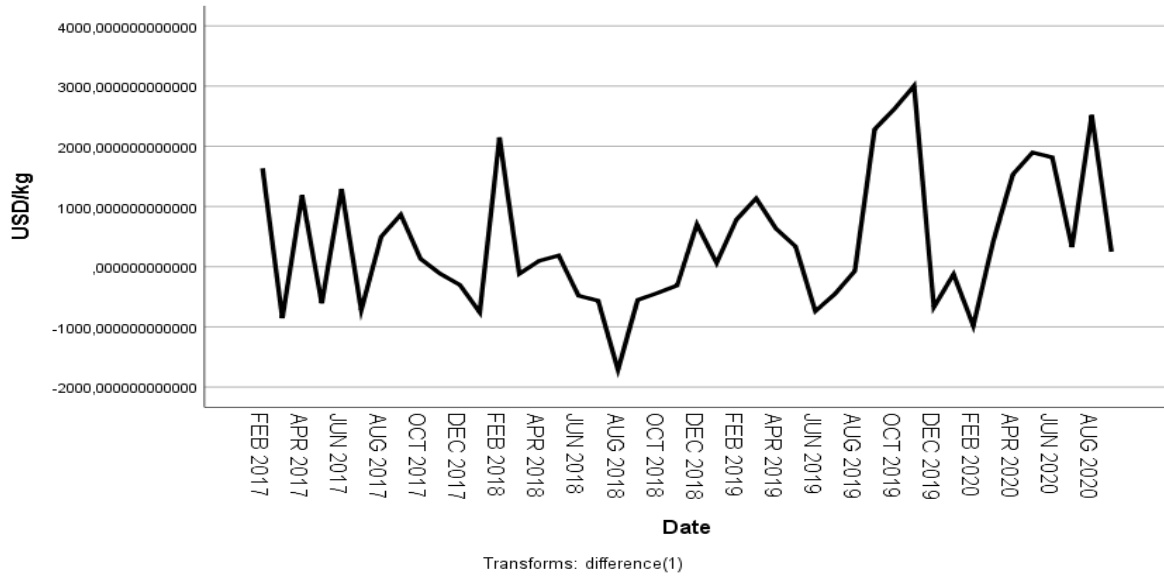
Figure 2: Non-stationariry

Furthermore, it is necessary to eliminate the non-stationarity trend by taking the change from the first period to the next - first difference (Figure 3). After this step, the graph looks to be stationary and therefore we continue to other graphs.
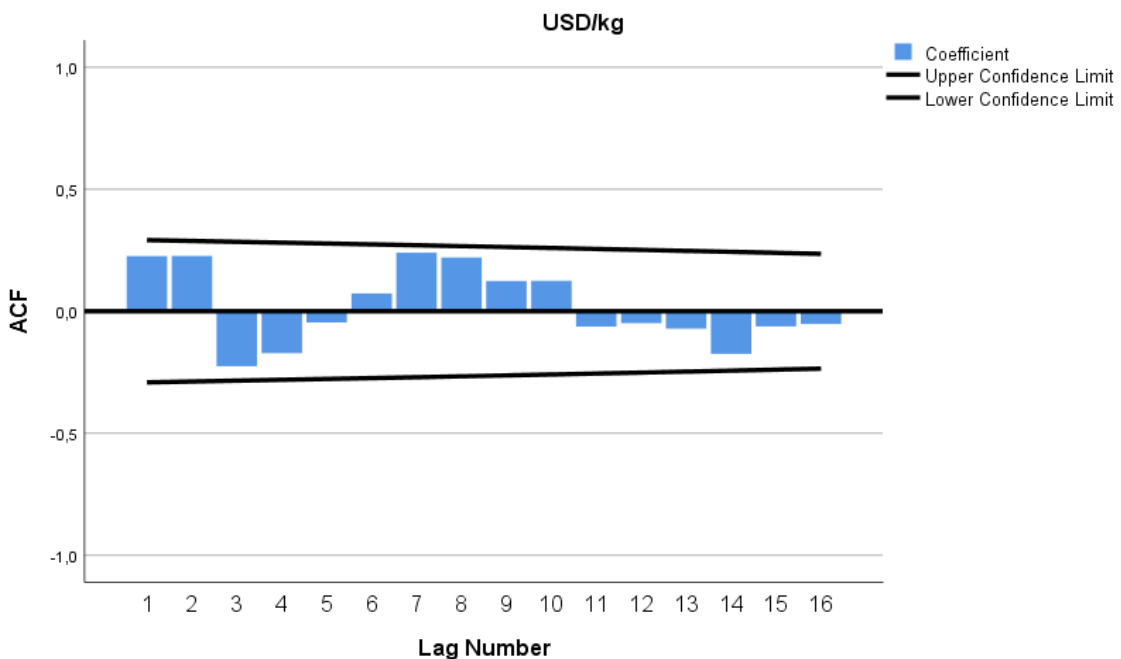
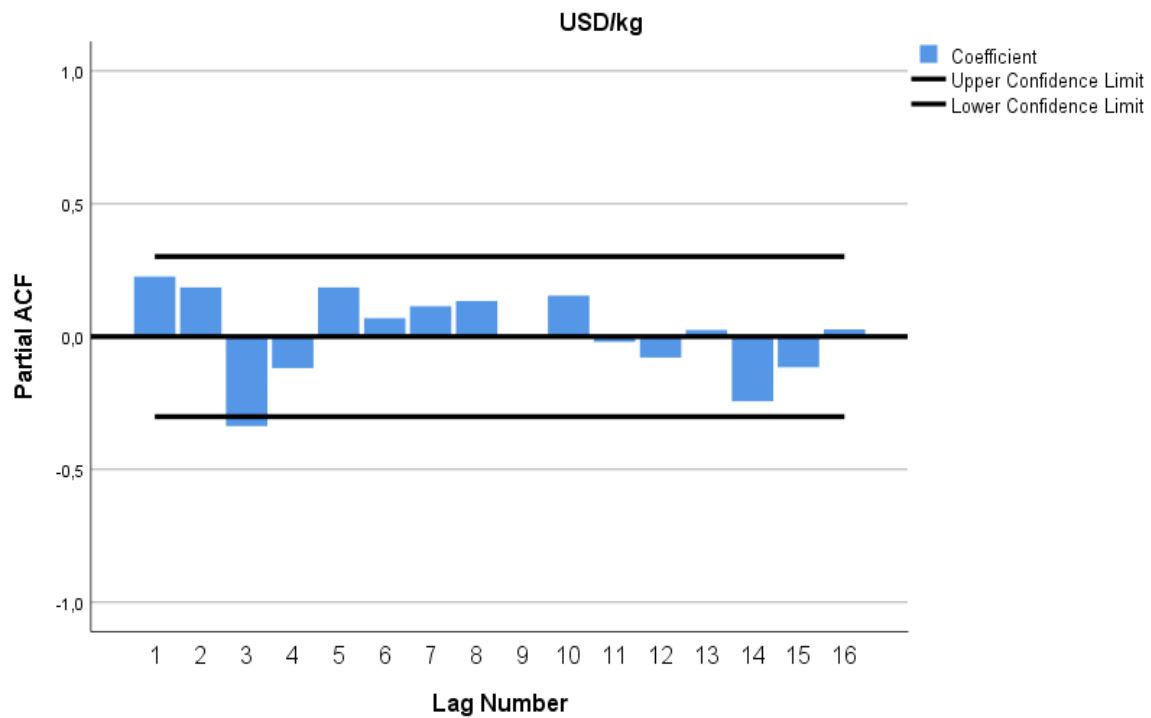*Figure 3: Graph of the first difference*

Source: Own output

The non-stationarity of the time series is also confirmed by the shape of ACF and PACF.The ACF and PACF sample, in fact, together with the time series plot of the residues in Figure 4 and 5, indicate that the residues are stationary and can be modeled using the ARIMA model (0,1,1).

Figure 4: ACF

Source: Own output

ARIMA model for predicting the development
of the price of gold: European approach
Authors: Ligita Gaspareniene, Rita Remeikiene

Figure 5: PAFC

Source: Own output

The model we have compiled is therefore also satisfactory on the basis of this test. The suitability of using the first difference is also confirmed by figure 6, in which it can be seen that after using the first difference, no autocorrelation or partial autocorrelation coefficient is statistically significant at the significance level of 0.05. It is necessary to check the residuals on graph 6 and see if there is white noise or there is random independent errors no correlation which both graphs in graph 6 meet because none of the residuals is significantly different from zero.
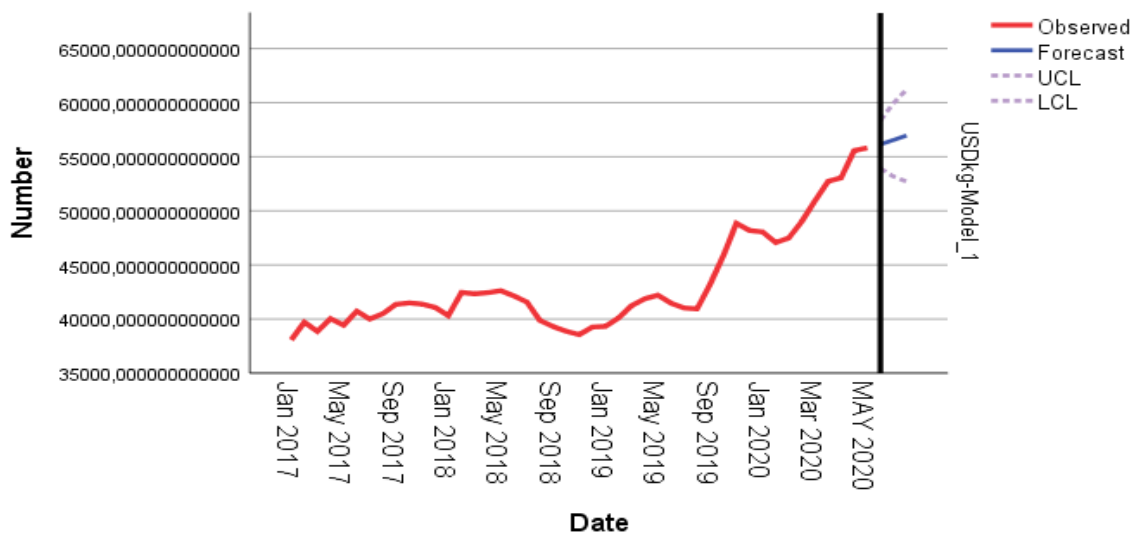
Ekonomicko-manazerske spektrum
2020, Volume 14, Issue 1, pp. 87-96

Figure 6: Residual ACF and residual PACF
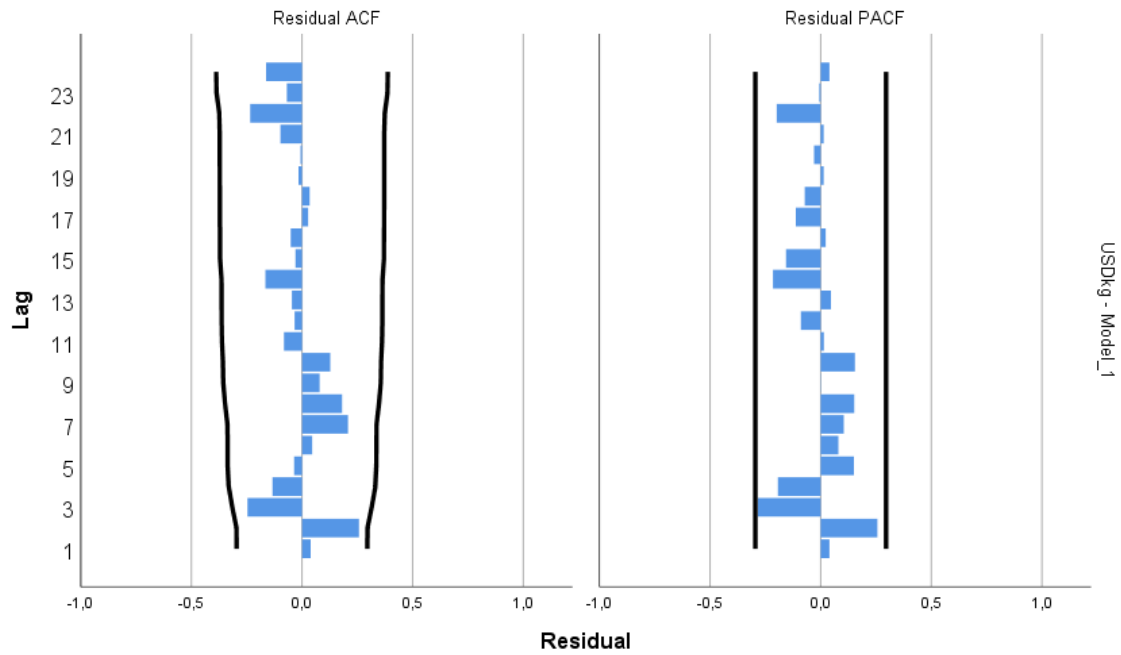
Source: Own output

The model can be used for the prediction, which is shown in Figure 7. There is a blue line showing the development and at the same time the prediction intervals are indicated.

Figure 7: Graphical representation of the prediction the development of gold prices in 3 months

Source: Own output

93   *ISSN 1337-0839 (print) / 2585-7258(online)*

ARIMA model for predicting the development
of the price of gold: European approach
Authors: Ligita Gaspareniene, Rita Remeikiene

Table 1: Numerical expression of the forecast

Forecast				
Model		Jun 2020	July 2020	Aug 2020
	Forecast	56150.919946471620	56557.505132934566	56964.090319397510
USD/kg-Model_1	UCL	58373.118487915370	59951.388581915530	61218.349758235156
	LCL	53928.721405027870	53163.621683953600	52709.830880559864

[*]For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.

Source: Own output

Table 1 shows the plotting of forecasted value with the help of the model values for three months. At the same time, the development of the upper control limit and lower control limit is shown here

4. Discussion

We have provided an interesting example of using the ARIMA model to reproduce and explain time series. We had in mind the explanation of the assumptions and support of general models. We believe that this material will serve as a useful case study for teachers and students.

Our goal is to obtain new results in the field of modeling and forecasting the price of gold, and at the same time it is interesting to interpret the results from the ARIMA model.

At the same time, it is essential to point out the variability of the exchange rate and the fact that any change in the environment can be reflected in the development. While coronavirus hurts some financial markets, it offers room for unprecedented growth in others. The best example is the gold exchange. The price of the precious metal has been rising almost continuously throughout the pandemic.

At the beginning of the second half of April, he even managed to achieve the best result since 2012. Fulfilling the forecasts for further growth would bring new highs. More optimistic predictions even speak of reaching the highest values since the beginning of the century. It will suffice if uncertainty persists in the financial markets, and gold can quickly rewrite historical tables.

5. Conclusion

In this research paper, we focused on analysing the performance of the price of gold from traded values over the last 4 years using the ARIMA model (0, 1, 1), which helps us predict future gold values because it provides a model that meets the criteria for Compliance. At the same time, it is necessary to take into account certain limitations when predicting data using this technique, as it is only used for short-term detection of small discrepancies in the data. If there is a sudden change in government policies or economic instability, or another change that is difficult to capture, the model becomes ineffective to predict. The forecast of the gold price using the ARIMA model was made with the basic assumption that it is a perfectly linear pattern, and therefore it is possible to implement nonlinear forecasting techniques using soft calculation techniques with the term less white noise (Guha and Bandyopadhyay, 2016).

References

Abdullah, L. (2012). ARIMA model for gold bullion coin selling prices forecasting. *International Journal of Advances in Applied Sciences*, *1*(4), 153-158.

Asadi, S., Tavakoli, A. and Hejazi, S. R. (2012). A new hybrid for improvement of auto-regressive integrated moving average models applying particle swarm optimization. *Expert Systems with Applications*, *39*, pp. 5332-5337.

Bazilier, R. and Girard, V. (2020). The gold digger and the machine. Evidence on the distributive effect of the artisanal and industrial gold rushes in Burkina Faso. *Journal of Development Economics, 143.*

Box, G. E., Jenkins, G. M., Reinsel, G. C. and Ljung, G. M. (2015). Time series analysis: forecasting and control. *John Wiley & Sons.*

Bozanic, Z., Roulstone, D. T. and Van Buskirk, A. (2018). Management earnings forecasts and other forward-looking statements. *Journal of Accounting & Economics, 65*(1), pp. 1-20.

Brockwell, P. J. and Davis, R. A. (2003). Introduction to Time Series and Forecasting. Springer: New York, NY.

Buera, F. J. and Lucas, R. E. Jr. (2018). Idea flows and economic growth. *Annual Review of Economics*, *10*, pp. 315-345.

Cipra, T. and Romera, R. (1991). Robust Kalman filter and its application in time series analysis. *Kybernetika*, *27*(6), pp. 481-494.

Cinnirella, F. and Streb, J. (2017). The role of human capital and innovation in economic development: Evidence from post-Malthusian Prussia. *Journal of Economic Growth, 22*(2), pp. 193-227.

Guha, B. and Bandyopadhyay, G. (2016). Gold price forecasting using ARIMA model. *Journal of Advanced Management Science*, *4*(2).

Huarng, K.-H., Rey-Marti, A., Guaita-Martinez, J.-M. (2020). Knowledge, business, and innovation: Economies and sustainability of future growth. *Journal of Business Research, 112*, pp. 236-239.

Hadavandi, E., Ghanbari, A. and Abbasian-Naghneh, S. (2010). Developing a Time Series Model Based On Particle Swarm Optimization for Gold Price Forecasting. 3rd International Conference on Business Intelligence and Financial Engineering (BIFE 2010), *IEEE computer society*, pp. 337-340.

Gertler, M. and Gilchrist, S. (2018). What happened: Financial factors in the Great Recession. *Journal of Economic Perspectives, 32*(3), pp. 3-30.

Gibson, J., McKenzie, D., Rohorua, H. et al. (2019). The long-term impact of international migration on economic decision-making: Evidence from a migration lottery and lab-in-the-field experiments. *Journal of Development Economics, 138*, pp. 99-115.

Khashei, M., Hejazi, S. R. and Bijari, M. (2008). A new hybrid artificial neural networks and fuzzy regression model for time series forecasting. *Fuzzy Sets and Systems*, *159*, pp. 769-786.

Kilian, L. and Taylor, M.P. (1991). Why Is It So Difficult to Beat Random Walk Forecast of Exchange Rate. *Journal of International Economics*, *60*, pp. 85-107.

Kramer, S., Maas, V. and van Rinsum, M. (2016). Relative performance information, rank ordering and employee performance: A research note. *Management Accounting Research, 33*, pp. 16-24.

Larmore, E. (2016). Methodologies for forecast modeling for small areas with limited data availability and unique tax structures. (Doctoral dissertation).

Lindner, T., Klein, F. and Schmidt, S. (2018). The effect of internationalization on firm capital structure: A meta-analysis and exploration of institutional contingencies. *International Business Review, 27*(6), pp. 1238-1249.

Lu, Y. and AbouRizk, S. M. (2009). Automated box–jenkins forecasting modelling. *Automation in Construction*, *18*(5), pp. 547-558.

Martinez-Alvarez, F., Troncoso, A., Asencio-Cortes, G. and Riquelme, J. C. (2015). A survey on data mining techniques applied to electricity-related time series forecasting. *Energies*, *8*(11), pp. 13162-13193.

Miswan, N. H, Ping, P. Y. and Ahmad, M. H. (2013). On parameter estimation for Malaysian gold prices modelling and forecasting. *International Journal of Mathematical Analysis*, *7* (21-24), pp. 1059-1068.

Mustafa, Z. and Yusof, Y. (2011). Optimizing LSSVM using ABC for non-volatile financial prediction. *Australian Journal of Basic and Applied Sciences*, *5*(11), pp. 549-556.

ARIMA model for predicting the development
of the price of gold: European approach
Authors: Ligita Gaspareniene, Rita Remeikiene

Nelson, B. K. (1998). Time series analysis using autoregressive integrated moving average (ARIMA) models. *Academic emergency medicine*, 5(7), pp. 739-744.

Parisi, A., Parisi, F. and Diaz, D. (2008). Forecasting gold price changes: Rolling and recursive neural network models. *Journal of Multination Financial Management*, *18*, pp. 477-487.

Ramcharan, R. (2020). Banks´balance sheets and liquidation values: Evidence from real estate collateral. *Review of Financial Studies, 33*(2), pp. 504-535.

Shafiee, S. and Topal, E. (2010). An overview of global gold market and gold price forecasting. *Resources Policy*, *35*, pp. 178-189.

Tharmmaphornphilas, W., Lohasiriwat, H. and Vannasetta, P. (2012). Gold price modeling using system dynamics. *Engineering Journal*, *16*(5), pp. 57-67.

Yuan, G. (2012). Study on gold price forecasting technique based on neural network optimized by GA with projection pursuit algorithm. *Journal of Convergence Information Technology*, *7*(18), pp. 558-565.

Zhang, Q., Ma, J. and Wang, Y. (2011). Study on forecasting of gold price based on varying-coefficient regression model. *Key Engineering Materials*, *467*(469), pp. 1398-1403.

Zhou, S., Lai, K. K. and Yen, J. (2012). A dynamic meta-learning rate-based model for gold market forecasting. *Expert System with Applications*, *39*, pp. 6168-6173.

Wei, W. W. S. (2005). Time Series Analysis: Univariate and Multivariate Methods, 2nd Edition. *Addison Wesley: New York*, NY.