RESEARCH ARTICLE **(Open Access)**

Analysis and projections of wheat production in Bosnia and Herzegovina using ARIMA modeling

NEMANJA JALIĆ[1]*, ALEKSANDAR OSTOJIĆ[1], ŽELJKO VAŠKO[1]

[1]Department of agricultural economics and rural development, University of Banja Luka, Faculty of Agriculture, Bulevar Vojvode Petra Bojovića 1A, Bosnia and Herzegovina

*Nemanja Jalić; e-mail: nemanja.jalic@agro.unibl.org,

Abstract:

Wheat is one of the most important field crops in Bosnia and Herzegovina (B&H) and on second place in terms of sown areas and the volume of production. It is also one of the most widespread cereals in the world, and according to the total sown areas, wheat is in the first place. Wheat is especially important for nutrition of the population and is often used as a substitute for corn as fodder. Despite the large volume of production, B&H has a low degree of self-sufficiency in wheat. This paper analyzes wheat production in B&H based on the time series of data from 1998 to 2020, and the forecast a wheat production in the domestic market for a five-year period applaying ARIMA modeling. In the analyzed period, wheat production had a average volume about 256.000 tons a year and decreased at an average rate of (R_G-0.26%). However, the coefficient of variation shows the variability of production in the observed period (C_V = 20.77%). The subject of the research is the projection of wheat production in B&H using the EViews 10 program, where ARIMA models were applied for forecast purposes. The aim of the research is to forecast the wheat production for a period of five years (2021-2025) based on the Box-Jenkins methodology. The data required for forecasting were taken from the FAOSTAT database. Based on the selected ARIMA (1, 1, 10) model, the trend of decreasing in wheat production in B&H is expected and in the next five years. It can be concluded that in the forecasted period, production will fall by in average of 3.1% each year. Variability of the forecasted production trend may be influenced by extreme manifestations of climatic factors, as well as unpredictable price movements in the regional and world markets, as well as government policy measures in regard to grain production.

Keywords: Wheat production, Bosnia and Herzegovina, Forecasting, ARIMA

1. Introduction

Wheat is,after the corn, the most important crop in Bosnia and Herzegovina. Wheat is primarily used for human consumption (ground into flour and processed into bakery products), but a certain quantity is also used for feeding livestock, primarily due to more favorable trends in the price of wheat compared to the price of corn. World wheat production has been increasing in the last 10-15 years, and the actual production largely depends on weather conditions in certain years, with the significant oscillations. The greatest producer of wheat in the world is EU with about 20% of total world production, followed by China, India, Russian Federation, USA, etc. Wheat production is competitive and largely export - oriented. Therefore, wheat production and prices in the EU and other exporters affect the price in B&H. Wheat sown area in the Bosnia and Herzegovina is increasing. The increase in sown area under wheat, according to statistical data, was about 3% at the average annual rate in the period 2014 - 2019. So, the data indicates that domestic farmers are interested in wheat production. A characteristic of this primary agricultural product is that developed countries have small variations in yields and production, while in other countries the variations are often significant. The needs for wheat for human and livestock nutrition on the market cause an increase in production with the growth of subsidies for the production of this crop. Bosnia and Herzegovina is traditionally a net importer of wheat, so

*Corresponding author: Nemanja Jalic; E-mail: nemanja.jalic@agro.unibl.org
(Accepted for publication 04.09.2021)
***ISSN**: 2218-2020,* © *Agricultural University of Tirana*

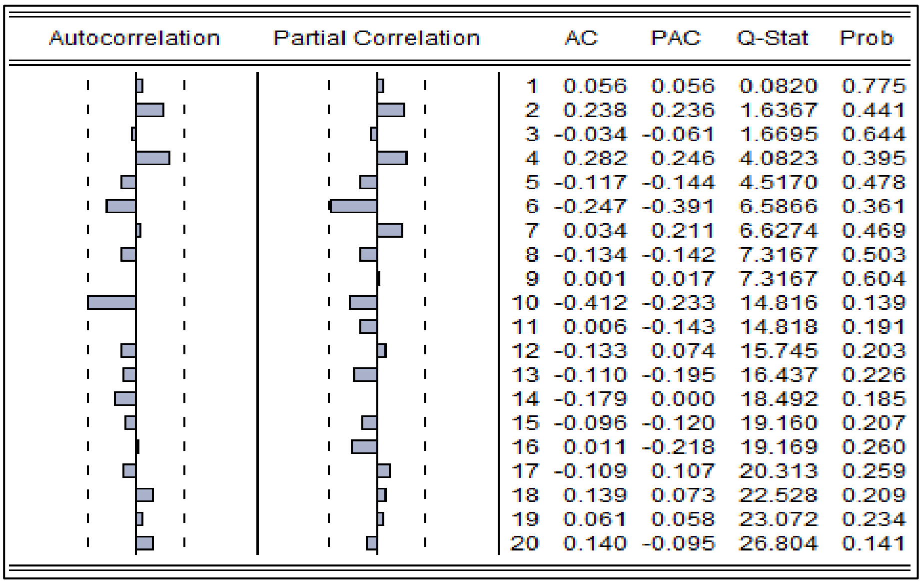
even today between 300 and 400 thousand tons of wheat are imported annually (Ostojic et al., 2020). Mainly high quality wheat is imported for the needs of the bakery and other food industry. Regardless of that, wheat is also exported from B&H. On average, about 43.000 tons per year are exported. Due to unsatisfied demand for wheat by own production, the shortage must be compensated by import.

By analyzing time series of wheat production in Bosnia and Herzegovina, we tried to describe and explain the observed time series, but also to predict its future movement. More researchers used the regression modelling to forecast what production (Karim et al., 2005; Prabakaran et al., 2013; Karim Ahmadzai and Eliw, 2019; Kovačević, 2016; Hamulczuk and Klimkowski, 2012). Their results confirmed that different models can be applied for forecastingin different areas. Thus, we can set fundamental goals in the analysis of time series data: first to describe the past time series, then to find and explainbest model for analysis and the last to predict the feature trends. The assumption that the methods of descriptive and analytical statistics are adequate for forecasting production and economic phenomena was also confirmed by the forecasting results obtained using ARIMA model by Nedeljkovic, 2019. The main objective in fitting the ARIMA model is to identify the stochastic process of the time series and predict the future values accurately (Sunil and Sindhu, 2020). Wheat production forecast is done by using ARIMA models which are developed in 1970 by Box and Jenkins (Box and Jenkins, 1970). This methodology is a powerful method of determination of mathematical models of different stochastic variables. In essence, it is a methodology of systematic multi-phase modeling for the identification and evaluation of models with the aim of finding the best adjusted data time series and making forecasts (Ilic et al., 2016).

2. Material and Methods

The research was conducted through the collection of secondary data (FAOSTAT yearly data of Bosnia and Herzegovina's production of wheat in tons for the period from 1998. to 2020.) and the instrument of collection was the method of the so-called field research. The Autoregressive Integrated Moving Average (ARIMA) model is used in this paper. ARIMA Models has an Autoregressive Model AR(p) as the first part, and Moving Average Model MA(q) a second part, and the third part I(d) represents the differences required by the time-series in order to be stationary (Alfaki and Masif, 2015). All the process of ARIMA modelling was done in the program EViews10.

Building ARIMA Model and Forecasting: In order to apply certain techniques for identifying and ARIMA model for the data, firstly it was determined the form of stationary of the original time series by applying the "Unit root test" by correlogramtesting and the Augmented Dickey-Fuller test (ADF) on the original data.



	AC	PAC	Q-Stat	Prob
1	0.056	0.056	0.0820	0.775
2	0.238	0.236	1.6367	0.441
3	-0.034	-0.061	1.6695	0.644
4	0.282	0.246	4.0823	0.395
5	-0.117	-0.144	4.5170	0.478
6	-0.247	-0.391	6.5866	0.361
7	0.034	0.211	6.6274	0.469
8	-0.134	-0.142	7.3167	0.503
9	0.001	0.017	7.3167	0.604
10	-0.412	-0.233	14.816	0.139
11	0.006	-0.143	14.818	0.191
12	-0.133	0.074	15.745	0.203
13	-0.110	-0.195	16.437	0.226
14	-0.179	0.000	18.492	0.185
15	-0.096	-0.120	19.160	0.207
16	0.011	-0.218	19.169	0.260
17	-0.109	0.107	20.313	0.259
18	0.139	0.073	22.528	0.209
19	0.061	0.058	23.072	0.234
20	0.140	-0.095	26.804	0.141

Figure 1. Corellogram of production

Source: Authors, based on Statistical FAO and EViews 10 data

The input series for ARIMA needs to be stationary, i.e. it should have a constant mean, variance, and autocorrelation through time (Novkovic et al., 2019). Therefore, the series usually needs to be differenced until it is stationary (this also often requires log transforming the data to stabilize the variance).The time series of wheat production showed the non-stationarity and the 1stdifferencing should be done for the production data (Figure 2). Then the AR and MA values were determined by analyzing the correlogram.

Augmented Dickey-Fuller Unit Root Test on PRODUCTION__T_

Null Hypothesis: PRODUCTION__T_ has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=4)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.116007	0.0194
Test critical values:	1% level	-4.440739	
	5% level	-3.632896	
	10% level	-3.254671	

*MacKinnon (1996) one-sided p-values.

Figure 2. ADF test for production

Source: Authors, based on Statistical FAO and EViews 10 data

AR value is determined by a Partial Correlation (PAC) and MA value is determined by an Autocorrelation (AC). Next phase in the Box-Jenkins methodology was to estimate equation, in EViews10 too. In the command window was entered: D(production) c AR(p) MA(q). Four models were graded by comparing the R-squared values, Akaike info criterion (AIC) and Schwarz criterion (SIC). The best model is the one that has the maximum value of R-square and the minimum value of AIC and SIC. After that follows the residual diagnostics was tested. The last stage was the forecasting. It was done for the period of next five years, 2021, 2022, 2023, 2024 and for the 2025.

3. Results and Discussion

3.1. Unit root testing and determining AR value and MAvalue

Results of the first phase of the ARIMA modeling showed that the time series has one unit root. This means that the series is non-stationary and for the next phases of the applied model the time series should be stationary (*a time series is stationary if it remains unchanged in time and thus contains a pattern that is as a basis for predicting future values*).

Augmented Dickey-Fuller Unit Root Test on D(PRODUCTION__

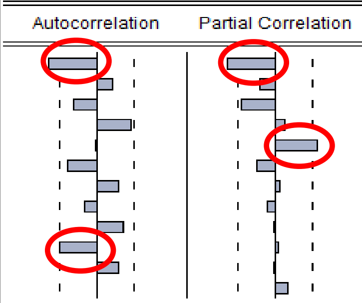
Null Hypothesis: D(PRODUCTION__T_) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Automatic - based on SIC, maxlag=4)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.823633	0.0008
Test critical values:	1% level	-4.532598	
	5% level	-3.673616	
	10% level	-3.277364	

*MacKinnon (1996) one-sided p-values.

Correlogram of D(PRODUCTION__T_)

Date: 09/07/21 Time: 12:45
Sample: 1998 2025
Included observations: 22



	AC	PAC	Q-Stat	Pro
1	-0.559	-0.559	7.8698	0.0
2	0.184	-0.187	8.7678	0.0
3	-0.278	-0.393	10.920	0.0
4	0.398	0.099	15.566	0.0
5	-0.009	0.487	15.568	0.0
6	-0.344	-0.224	19.470	0.0
7	0.262	0.058	21.881	0.0
8	-0.140	-0.085	22.625	0.0
9	0.311	-0.008	26.554	0.0
10	-0.435	0.036	34.891	0.0
11	0.246	-0.021	37.799	0.0
12	0.003	0.138	37.800	0.0

Figure 3. ADF test and correlogram for 1st difference of production

Source: Authors, based on Statistical FAO and EViews 10 data

In order to apply certain techniques for identifying andARIMA model for the data, Alfaki and Masif, 2015determined the form ofstationary of the original time series and drawed theautocorrelation function

*Corresponding author: Nemanja Jalic; E-mail: nemanja.jalic@agro.unibl.org
(Accepted for publication 04.09.2021)
***ISSN**: 2218-2020, © Agricultural University of Tirana*

(ACF), and the partialautocorrelation function (PACF) of data and confidence interval of (ACF) and (PACF) to detect the stationary or non-stationary of time series. ADF test of 1st difference (Figure 3) shows the stationarity (the value of test critical value is higher than ADF test statistics). After that it made posible to identify the models.*Box-Jenkins* modelidentification involves the determination of theappropriate orders of AR and MA polynomials i.e.the values of p and q(Verma et al., 2012).Based on the analyzed correlogram of the first difference the models (1,1,1), (5,1,1), (5,1,10) and (1,1,10) were chosen for the comparing. The values PAC 1 and 5 are crossing the vertical lines and the AC 1 and 10 (Figure 3).

3.2. Estimation and model diagnostics stage

In the phase of estimation of the equation all the models that are identified are compared. Four models with different combinations of AR and MA variables were examined. In interpreting the criteria, we considered the more value of R^2 or 2R, the better is the fitness of the model. Smaller value of RMSE, AIC, *SIC*, BIC, MAE, and MAPPE are better fitness of the model(Karim et al., 2010).Based on the results of the indicators for comparing that are listedthe most acceptable model was AR (1) MA (10) i.e. (1,1,10). This model (Figure 4) has the biggest R-squared value(R^2=0,71) and the lowest Akaike info criterion (AIC=24,79) and Schwarz criterion (SIC=24,98).

Dependent Variable: D(PRODUCTION__T_)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 08/31/21 Time: 09:20
Sample: 1999 2020
Included observations: 22
Failure to improve objective (non-zero gradients) after 14 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2695.341	5298.241	-0.508724	0.6171
AR(1)	-0.550822	0.213222	-2.583328	0.0187
MA(10)	-1.000000	16935.82	-5.90E-05	1.0000
SIGMASQ	1.38E+09	1.17E+13	0.000118	0.9999

R-squared	0.713816	Mean dependent var	-869.5909
Adjusted R-squared	0.666119	S.D. dependent var	70996.17
S.E. of regression	41023.34	Akaike info criterion	24.78642
Sum squared resid	3.03E+10	Schwarz criterion	24.98479
Log likelihood	-268.6506	Hannan-Quinn criter.	24.83315
F-statistic	14.96554	Durbin-Watson stat	1.956622
Prob(F-statistic)	0.000039		

Figure 4. ARIMA (1,1,10) model statistics

Source: Authors, based on Statistical FAO and EViews 10 data

Proving that the chosen model is the best for the ARIMA forecast can be confirmed by the Durbin – Watson (DW) statistics which value should be approximately 2, significance of F – statistics (should be less than 0,05) and the coefficient of determination (R^2) which should be as high as possible. Based on the coefficient below it can be concluded that their value is completely acceptable.

$$Production = (-2695{,}34) - 0{,}55Production_{t-1} - 1\varepsilon_{t-10}$$

$R^2 = 0{,}71$

DW = 1,956

Prob. (F – statistic) = 0,000039

The validity of the chosen model is evaluated on the basis of predictive power and connectivity with the original time series. Validity is assessed by doing histogram and correlogram of residuals (Q-statistics and residual root) and it is sown that the model (1,1,10) has a normal distribution of residuals. The model is rated positively because autocorrelation and partial correlation indicates the random movement of residuals (Figure 5).

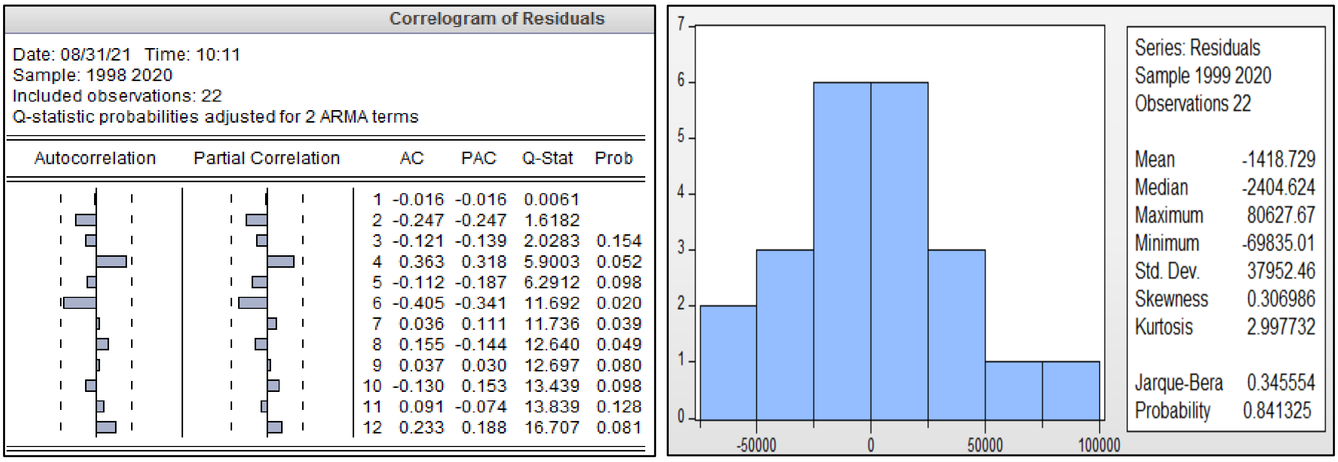


Figure 5. Residual normality test and correlogram of residuals

Source: Authors, based on Statistical FAO and EViews 10 data

Also for the evaluation of model heteroscedasticity (White) test was included. Test value (Obs*R-squared = 19,17) of chosen modelis less than the critical value of $\chi^2_{(0,05;18)}$ which is 28,869. The assumption about the model variance homoscedasticity is accepted.

Heteroskedasticity Test: White			
F-statistic	40.59755	Prob. F(3,18)	0.0000
Obs*R-squared	19.16723	Prob. Chi-Square(3)	0.0003
Scaled explained SS	12.51589	Prob. Chi-Square(3)	0.0058

Figure 6. White test – heteroscedasticity test

Source: Authors, based on Statistical FAO and EViews 10 data

3.3. Forecasting future trend of wheat production

In order to forecast the future trend of wheat production in tons for Bosnia and Herzegovina the original time series is modified. The length of time series is extrapolated for five years. The Figure below shows the best and worst scenario between ±2 sandard errors of real production time series. Also on the right side is the trend of the forecasted production for next five years. Red line shows the data values available on FAOSTAT data and the blue line shows the future values trend.

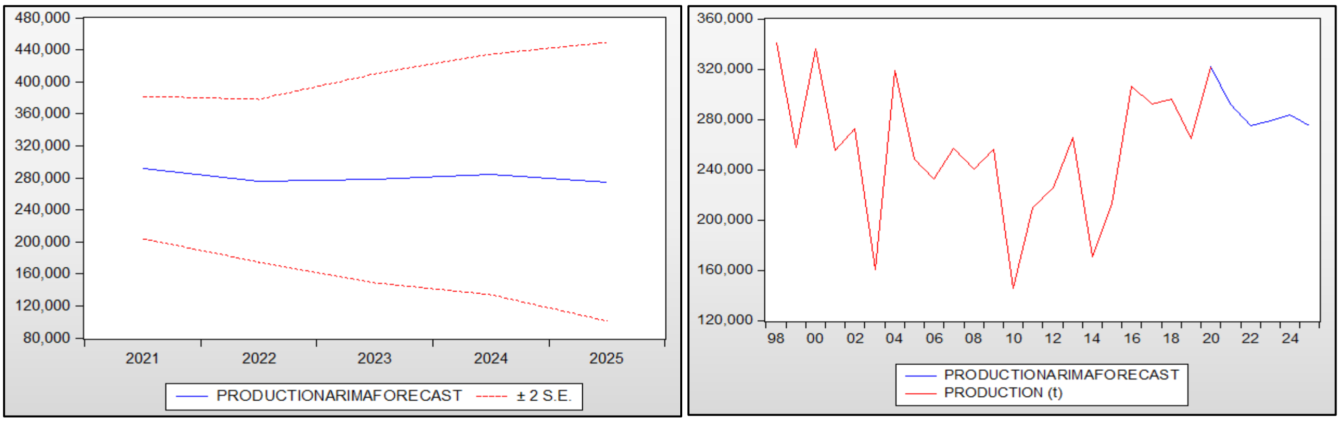


Figure 7. Forecast of future wheat production 2021-2025 for Bosnia and Herzegovina

Source: Authors, based on Statistical FAO and EViews 10 data

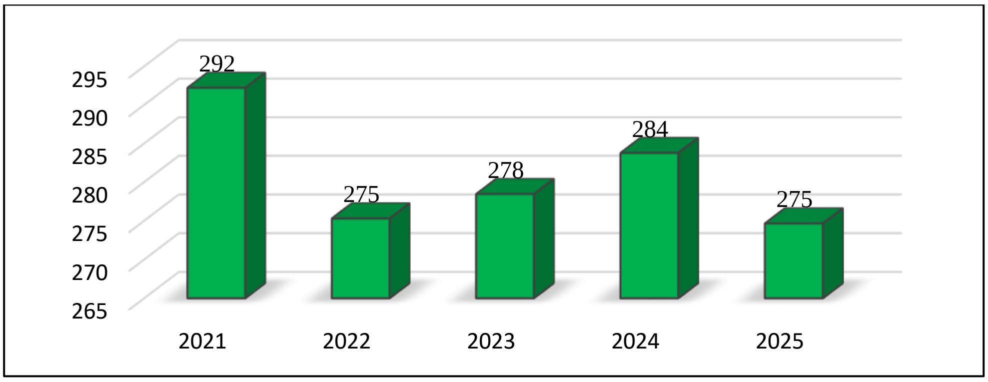


Figure 8. Wheat production trend 1998.– 2025. for Bosnia and Herzegovina(000 t)

Source: Authors, based on Statistical FAO and EViews 10 data

The forecasted future scenario shows that the Bosnia and Herzegovina wheat production will decrease due to many effects, primarily climatic and market factors. The production should be in 2021. about 292.208 tons, then the fall is expected to 275.292 in 2022. In next two years the growth of production amount is predicted. In 2025. based on the created model the production value is the smallest off forecasted period and will be 274.650 tons.

4. Conclusions

Bosnia and Herzegovina is a country with a low level of economic development and of course with a large percentage of agricultural population (about 30%). Agriculture share in GDP is about 7% (MOFTER, 2018). Because of that the agriculture and agriculture production is very important. The second largest production by area is wheat production. Despite that, Bosnia and Herzegovina does not have the self sufficiency for this product, because it is used as human and also animal food. The authors wanted to forecast this production and to perceive and analyze it. Based on the results of this paper and applied method, chosen ARIMA (1,1,10) model it can be concluded that it is expected a downward trend of wheat production in Bosnia and Herzegovina. The falling trend can be accepted because the climatic and market factors that are expected in future and happens now are not good for agricultural production. The values of production in future will be on the average level about 280.000 tons (25.000 tons higher than average 1998-2020).

5. Acknowledgements

First limit of this study is that there is no the best and completely accurate ARIMA model. Based on the indicators (R^2, AIC, SIC, DW, Prob – F statistics, White test, histogram and correlogram of residuals)that are described in material and method and results and discussion chapterone model was chosen but always there is stochastic process in every time series. Especially in non-biological processes there is a big variations and the prediction is uncertain. Second limit of the paper is the data as an input for the analysis. In Bosnia and Herzegovina there is a problem with a data collection and the validity of them.

6. References

1. Alfaki M M A, Masih S B: **Modeling and Forecasting by using Time Series ARIMA Models**. International Journal of Engineering Research & Technology (IJERT) 2015, **4**(3): 914-918.
2. Box G, Jenkins G: **Time Series Analysis: Forecasting and Control, San Francisco (1970): Holden-Day**
3. **FAOSTAT**, Available on: http://www.fao.org/faostat/en/ , accessed 05.08.2021.
4. Hamulczuk M, Klimkowski C: **Response of the Polish wheat prices to the world's crude oil prices**. Acta oecomonika et informatica 2012, **2**: 50-56.
5. Ilic I, Jovanovic S, Jankovic-Milic V: **Forecasting corn production in Serbia using ARIMA model,** Economics of Agriculture 4/2016: 141-1156.
6. Karim Ahmadzai M, Eliw M: **Using ARIMA Models to Forecasting of Economic Variables of Wheat Crop in Afghanistan.** Asian Journal of Economics, Business and AccountinG 2019, **13**(4): 1-21.
7. Karim R, Awal A, Akhter M: **Forecasting of wheat production in Bangladesh.**Journal of Agriculture & Social Sciences 2005, **1**: 120-122.

8. Karim R, Awal A, Akter M: **Forecasting of wheat production in Bangladesh.**Bangladesh J. Agril. Res. 2010,**35**(1): 17-28.

9. Kovačević S: **Box Jenkinsov model.** Master thesis, University of Novi Sad, Faculty of Sciences, Serbia. 2016.

10. Ministry of Foreign Trade and Economic Relations of Bosnia and Herzegovina, Godišnji izvještaj iz oblasti Poljoprivrede, ishrane, i Ruralnog Razvoja za 2017. godinu 2018: 8

11. Nedeljkovic M: **Predvidjanje proizvodno ekonomskih pokazatelja ratarske proizvodnje u Republici Srpskoj**. PhD Dissertation. University of Novi Sad, Faculty of Agriculture, Serbia. 2019.

12. Novković N, Mutavdžić B, Ivanišević D, Drinić Lj, Vukelić N: **Models for forecasting the price of wheat and maize in Serbia.** Journal on Processing and Energy in Agriculture 2019, **23**(3): 138-141.

13. Ostojic A, Vasko Z, Brkovic D: **Assessment of wheat self-sufficiency in Bosnia and Herzegovina.** IX International Symposium on Agricultural Sciences AgroReS 2020 Book of proceedings 2020: 171-184.

14. Prabakaran K, Sivapragasam K, Jeevapriya C, Narmatha A: **Forecasting Cultivated Areas And Production Of Wheat In India Using ARIMA Model**. Golden Research Thoughts 2013, **3**(3): 1-7.

15. Sunil J, Sindhu D: **Growth Status and ARIMA Model Forecast of Area, Production of Wheat Crop in India.**Journal of Scientific Research & Reports2020,**26**(5):117-125.

16. Verma U, Koehler W, Goyal M: **A Study on Yield Trends of Different Cropsusing ARIMA Analysis.** Environment & Ecology 2012, **30**(4A) : 1459-1463.