A forecasting the consumer price index using time series model

Volodymyr Shinkarenko[1,*], *Alexey* Hostryk[1,**], *Larysa* Shynkarenko[2,***], and *Leonid* Dolinskyi[3,****]

[1]Odessa National Economic University, 8 Preobrazhenskaya Str., 65082, Odessa, Ukraine
[2]International Humanitarian University, 33 Fountain Road Str., 65009, Odessa, Ukraine
[3]University of the State Fiscal Service of Ukraine, 31 Universytetska Str., Irpin city in Kiev Region, 08201, Ukraine

Abstract. This article examines the behavior of the consumer price index in Ukraine for the period from January 2010 to September 2020. The characteristics of the initial time series, the analysis of autocorrelation functions made it possible to reveal the tendency of their development and the presence of annual seasonality. To model the behavior of the consumer price index and forecast for the next months, two types of models were used: the additive ARIMA*ARIMAS model, better known as the model of Box-Jenkins and the exponential smoothing model with the seasonality estimate of Holt-Winters. As a result of using the STATISTICA package, the most adequate models were built, reflecting the monthly dynamics of the consumer price index in Ukraine. The inflation forecast was carried out on the basis of the Holt-Winters model, which has a minimum error.

1 Introduction

The pandemic caused by the spread of the COVID-19 virus around the planet has painfully affected the global and national economy [1–4]. According to the NBU [5], the gross domestic product of Ukraine decreased by 6% in 2020. In these conditions, one of the most pressing problems is the containment of price increases. In macroeconomics, the current situation is called stagflation [6], which is characterized by an increase in prices during a crisis in the economy. The consumer price index demonstrates the general level of inflation in the economy and is an indicator of the standard of living of the population and socio-economic development. To regulate price increases, it is necessary to model inflation based on a scientific analysis of the dynamics of the cost of goods and services.

M. Sarel [7] analyzed the possibility of a nonlinear impact of the CPI on economic growth, when this indicator is critical 108. Below this value, the CPI does not affect growth, or may even have a slightly positive effect.

M. Kuzheliev [8] examines the impact of inflation targeting and other key instruments of monetary policy on key economic indicators in Ukraine during periods of stability and crises. This study uses econometric methods (multidimensional regression and the model of the simultaneous equation), which are used for the overall and transmission impact of inflation on the assessment of economic growth. The results show that inflation does not affect (less than 0.46 linear correlation) the basic economic indicators during periods of real GDP growth and the quarterly CPI level is less than 2%. On the other hand, there are significant simultaneous regressions (over 0.8 coefficient of determination) between the unemployed, real final consumption expenditure, UAH exchange rate and monetary policy instruments (discount rate, international reserves, amount of government bonds) for periods when the quarterly CPI (consumer price index) is more than 2%.

In papers [9, 10] presented the influence of the consumer price index on the economic growth of Romania over a period of 28 years, based on the annual data, 1991-2018. The increase of the values of the gross domestic product ensures a sustainable economic growth in Romania. It is proposed to use the cubic analysis model. The results obtained from the analysis show the influence of the consumer price index (CPI) on the gross domestic product (GDP).

The mutual influence of the growth of consumer prices and inflationary expectations in Ukraine is studied in the work of D. Khokhych [11]. T. Gitis [12] raises the problem of the influence of the growth of consumer prices on the level of income of the population. The analysis of the use of the value unit index for curbing inflation in Latin American countries was carried out by Yu. Yereshko[13]. Forecasting the consumer price index in Ukraine for the period of the end of 2020, the beginning of 2021 was not previously carried out in the works.

2 Materials and methods

Inflationary processes are studied based on the monthly consumer price index in Ukraine for the period from January 2010 to September 2020 according to the website of the State Statistics Service of Ukraine [14]. The data array was 129 variants, the units of measurement are percentages up to the previous month.

The analysis of the dynamics of the series and the construction of models were carried out using the STATIS-

─────────────
[*]e-mail: shinkar@te.net.ua
[**]e-mail: alexeyGostrik@gmail.com
[***]e-mail: larysoncka@gmail.com
[****]e-mail: dolinsky.com@gmail.com

TICA package, which made it possible to choose the optimal model with a minimum error of the forecast.

The use of the STATISTICA package to construct time series is carefully described in [15, 16].

2.1 The Holt-Winters model or exponential smoothing

The Holt-Winters model or exponential smoothing with a linear trend and seasonality assessment is the progenitor of adaptive forecasting methods [17, 18]. The complete procedure consists of four steps:

1. Calculation of exponentially smoothed series:

$$L_t = \alpha \cdot \frac{Y_t}{S_{t-s}} + (1 - \alpha) * (L_{t-1} + T_{t-1}) \quad (1)$$

2. Determining the value of the trend:

$$T_t = \beta \cdot (L_t - L_{t-1}) + (1 - \beta) * T_{t-1} \quad (2)$$

3. Seasonality assessment:

$$S_t = \gamma \cdot \frac{Y_t}{L_t} + (1 - \gamma) * S_{t-s} \quad (3)$$

4. The forecast is carried out:

$$Y_{t+p} = (L_{t+p} \cdot T_t) + S_{t-s+p} \quad (4)$$

where are L_t – smoothed value at the time of calculation; α is the smoothing coefficient of the series; S_{t-s} – coefficient seasonality of the previous series; Y_t is the value of the period; L_{t-1} – smoothing value for the previous period; T_{t-1} – trend value for the previous period; T_t – the value of the trend at the time of calculation; β – trend smoothing coefficient; S_0 – seasonality factor at the time of calculation; γ is the coefficient seasonality smoothing; Y_{t-p} – Holt-Winters model forecast for p periods ahead; p is ordinal the number to which we make a forecast; S_{t-s+p} – seasonality factor for the same period in the last season.

Thus, on the one hand, it is necessary to increase the statistical range and more correctly choose the smoothing coefficients. Finding compromise values of α, β, γ is the task of model optimization. One of the important components is the choice of coefficients that affect the forecasting error and subsequently the accuracy of the forecast .

2.2 Box-Jenkins model

ARIMA * ARIMA or Box-Jenkins model is an integrated model of auto regression – variable mean and time series analysis. Is an extension of ARMA models for non- stationary time series that can be made stationary by calculating the difference of some order from the original time series. The ARIMA model (p, d, q) means that the time series differences of order d obey the ARIMA model (p, q).

The methodology of the model is described in detail in [19–21].

We use the following Box-Jenkins additive models of time series:

$$Y_t = f(t) + S_t + e_t \quad (5)$$

where Y_t is a level of time series at time $t = 1, 2, ...$; $f(t)$ – trend is traced the long-term and evolution is deterministic in time; S_t – seasonal fluctuations; e_t – random fluctuations (white noise).

3 Results and discussion

3.1 Determining the dynamics of a series

In contrast to [22], the time series should not be reduced to a special form, but its main characteristics should be determined immediately.

To build a optimal model and to forecast inflationary processes, first, the main characteristics of the dynamic series of the consumer price index are found. They are shown in table 1.

Table 1. Dynamic series characteristics

Descriptive Statistics (CI.sta)				
Variable	Mean	Std.Dv.	Min	Max	N
CI	100.89	1.82	98.7	114.00	129

The characteristics of the series show that the indicator ranges from 98.7% to 114%, the standard deviation of 1.82% demonstrates that the spread of the consumer price index is quite small.

The presence of anomalous values of 112% and 114% in general does not affect the overall dynamics of the consumer price index, so the study was conducted without removing them.

To determine the general trend in the behavior of a series of data, a graph was built showing the consumer price index and the polynomial trend $n = 3$ (figure 1).
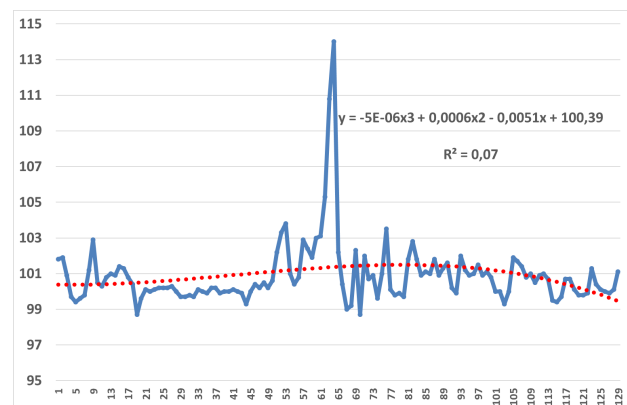
Figure 1. Graph the consumer price index and the polynomial

The coefficient of determination $R^2 = 0.0732$ shows that there is no obvious dependence of the consumer price index on the time of its measurement. The absence of a
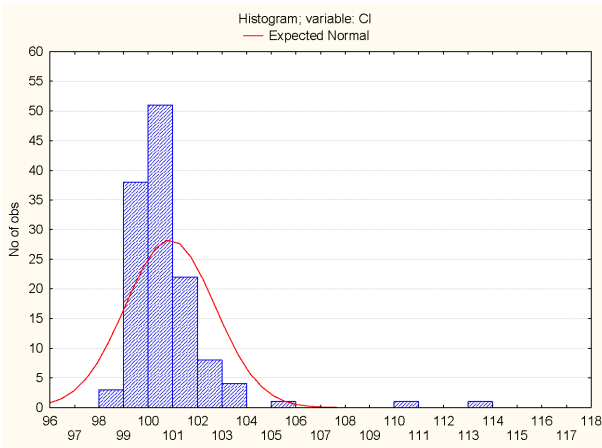
Figure 2. Histogram of the variable

Figure 3. Summary of error Holt-Winters model

Figure 4. Prediction results using the Holt-Winters model

trend relationship is also shown by the constructed histogram of the variable (figure 2).

Therefore, to model inflation and determine the indicator for subsequent periods, it is necessary to use the methods of Holt-Winters and Box-Jenkins.

3.2 The Holt-Winters model

To build the model, we use the "Time series and forecasting" module of the STATISTICA package. To find the optimal values of the model parameters, we use the "Search on the grid" command.

The method of random search is effective with the use of computers to solve optimization problems. The method is based on the generation of random numbers that have a uniform distribution in the interval [0, 1]. Random numbers obtained on a computer are called pseudo-random numbers. At the first stage, a step width of 0,1 is selected and all values of the coefficients α, β, γ are selected as $0, 1n, n = 0, 1, \ldots, 10$. Determine the model values: among the set of values of the coefficients are solutions α, β, γ for which the objective function takes a high accuracy of value. If the forecast accuracy is insufficient, you should reduce the step width and find new optimal values of the coefficients.

As a result, the following optimal parameters were obtained: exponential smoothing $\alpha = 0, 8$; trend component $\beta = 0, 1$ and seasonality estimate $\gamma = 0, 1$. A summary of the error is shown in the figure 3.

Mean absolute error

$$MAPE = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{|y_i - \bar{y}_i|}{y_i} = 0.81 \qquad (6)$$

Mean squared error

$$MSE = \frac{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2}{n} = 2.1 \qquad (7)$$

is small, which indicates the adequacy of the model and high prediction accuracy. The forecast for 10 subsequent periods is shown in figure 4.

To assess the forecasting accuracy, let us compare the forecasted consumer price indices with the real ones for October and November 2020 and calculate the relative forecast error:

$$\delta_{10,2020} = \left| \frac{y_{10,2020} - \bar{y}_{10,2020}}{y_{10,2020}} \right| = \qquad (8)$$

$$= \left| \frac{100, 8 - 100, 7576}{100, 8} \right| = 0,0004$$

$$\delta_{11,2020} = \left| \frac{y_{11,2020} - \bar{y}_{11,2020}}{y_{11,2020}} \right| = \qquad (9)$$

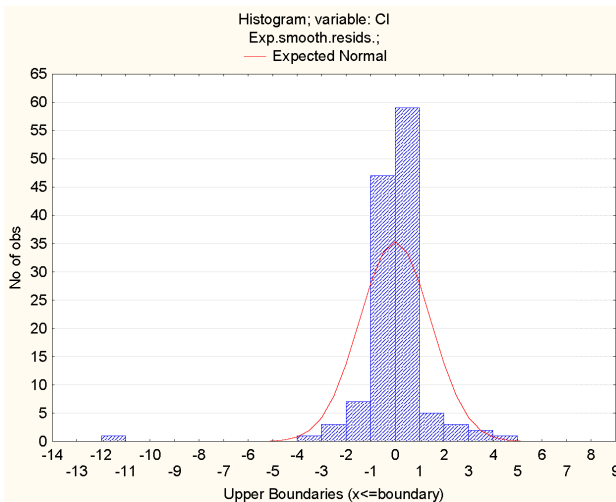$$= \left| \frac{100, 7 - 100, 7824}{100, 7} \right| = 0,0008$$

It should be noted that the forecasting accuracy of the constructed model decreases every month, since it is intended for short-term forecasts.

The relative forecast error is very small (less than 1%), therefore, forecasting the indicator using the Holt-Winters model shows good results. To illustrate the adequacy of the model, graphs of real and simulated consumer price indices are built. They are shown in figure 5.
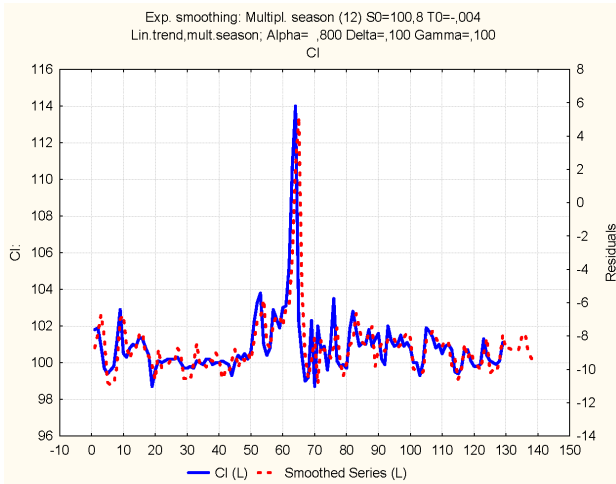
Figure 5. Graph of real indices and obtained by the Holt-Winters model

We are convinced of the adequacy of the obtained the Holt-Winters model of exponential smoothing by analyzing the histogram of residuals and Normal curve (figure 6).


Figure 6. Histogram of residuals the Holt-Winters model

It should be noted that the forecasting accuracy of the constructed model decreases every month, since it is intended for short-term forecasts.

3.3 Box-Jenkins model

We will first define the time series. Analyzing the graph of autocorrelation (figure 7) and the partial autocorrelation function presented below, we conclude that there is a first-order autocorrelation with a lag of 1 and an annual seasonality with a delay of 12.
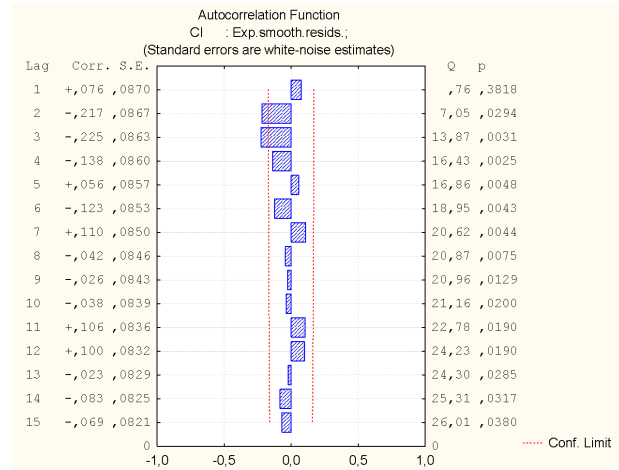
Figure 7. Graph of autocorrelation function

Partial autocorrelation function is presented below (figure 8).
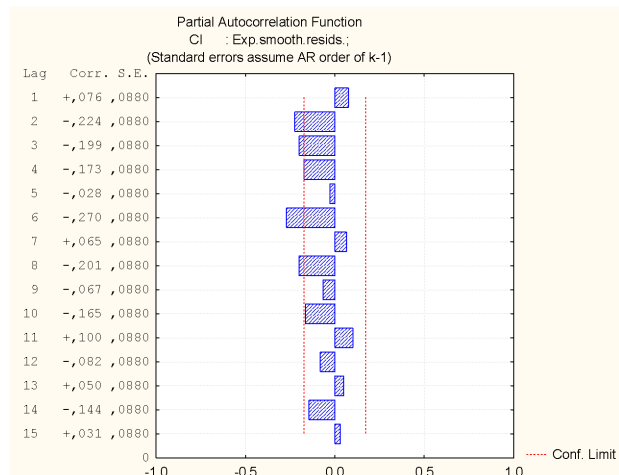
Figure 8. Graph of partial autocorrelation function

We will build an ARIMA*ARIMAS model using the "Automatic search for ARIMA model parameters". As a result, the model

$$p = 1, q = 1, P_s = 1 \qquad (10)$$

was proposed as the best program, seasonal lag = 12. The calculated parameters are shown in table 2.

Table 2. Estimates of parameters of the ARIMA model

Input: CI (CI2.sta) Transformations: D(1),D(12) Model:(1,1,1)(1,1,0) Seasonal lag: 12 MS Residual = 2.8764		
Parameter	Estimate	Standard deviation
p(1)	0.672278	0.075567
q(1)	0.986673	0.017673
Ps(1)	-0.342172	0.089743

For the built model $MSE = 2.8764$, which is slightly higher than the MSE for the previous model. Table 3

shows the forecasting results for the next 12 periods (for a year).

**Table 3.** Estimates of parameters of the ARIMA model

Forecasts; Model:(1,1,1)(1,1,0)
Seasonal lag: 12
Input: CI Start of origin: 1 End of origin: 129

Case No.	Forecast	Lower	Upper	Std.Err.
130	100.92	98.11	103.74	1.70
131	100.35	96.94	103.76	2.06
132	99.90	96.24	103.56	2.21
133	99.94	96.16	103.71	2.28
134	99.81	95.97	103.65	2.31
135	100.85	96.98	104.72	2.33
136	100.28	96.40	104.17	2.34
137	99.98	96.08	103.87	2.35
138	99.49	95.59	103.40	2.55
139	99.39	95.49	103.30	2.36
140	99.62	95.71	103.53	2.36
141	100.62	96.71	104.54	2.36

Calculate the relative forecast error:

$$\delta_{10,2020} = \left| \frac{y_{10,2020} - \overline{y}_{10,2020}}{y_{10,2020}} \right| = \qquad (11)$$

$$= \left| \frac{100,8 - 100,92}{100,8} \right| = 0,001$$

$$\delta_{11,2020} = \left| \frac{y_{11,2020} - \overline{y}_{11,2020}}{y_{11,2020}} \right| = \qquad (12)$$

$$= \left| \frac{100,7 - 100,35}{100,7} \right| = 0,0035$$

Comparing the predicted values with the real values, we obtain the relative forecast errors of 0.1% and 0.35%, which is worse than the Holt-Winters model, but it is quite acceptable for analyzing the dynamics of inflation. Figure 9 shows a plot of the ARIMA model with a predicted interval.
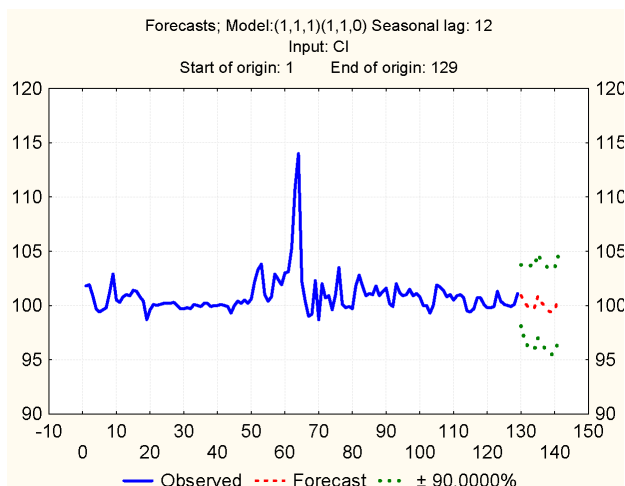

Figure 9. Plot of ARIMA model and 95% forecast interval

To clearly prove the adequacy of the selected model in figure 10, a graph of ARIMA model residues is constructed. Only in anomalous values the residues acquire rather large values, in others they do not exceed in absolute value 2%.
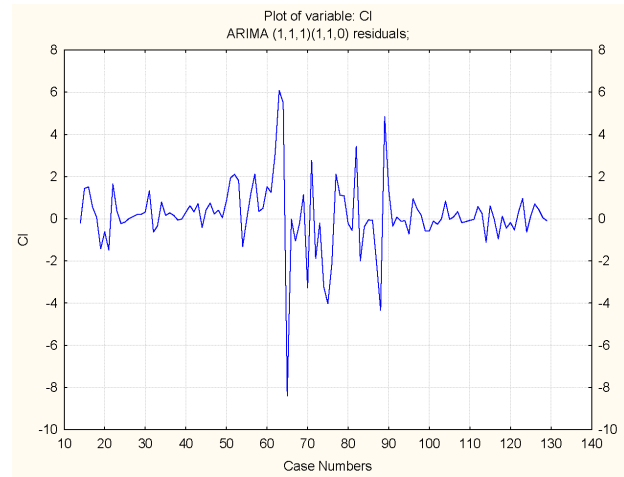
Figure 10. Graph of ARIMA model residues

We are convinced of the adequacy of the obtained the ARIMA*ARIMAS Box-Jenkins model by analyzing the histogram of residuals and Normal curve (figure 11).
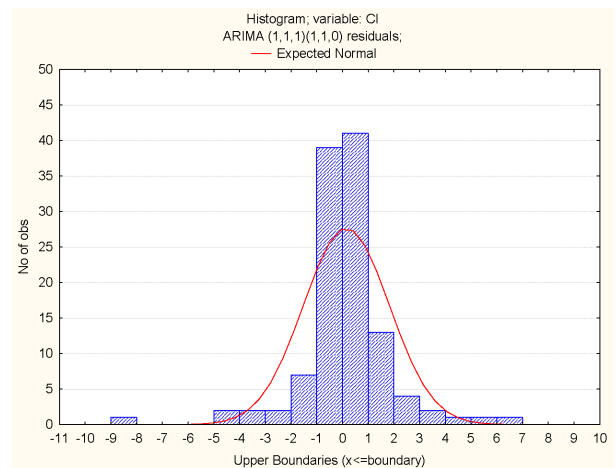
Figure 11. Histogram of residuals the ARIMA*ARIMAS model

4 Conclusion

Analyzing the dynamics of the consumer price index in the period from January 2010 to September 2020, it can be concluded that inflation at the level of 5-10% per year is typical for the Ukrainian economy. 2014 and 2015 were especially critical, after which the government managed not only to reduce the inflation rate, but also to maintain the growth of the consumer price index at 5% per annum.

The processing of the consumer price index array with the help of applied programs proved that the indicator is

characterized by high levels of seasonality and trend component and cannot be approximated by elementary functions. At the same time, the use of Holt-Winters and Box-Jenicks methods made it possible to build models of inflationary processes with high accuracy.

Forecasting the indicator for the beginning of 2021 showed that, despite the crisis in the economy associated with the COVID-19 pandemic, there are no reasons for increase the inflation rate.

For a deeper study of the reasons for the increase in the consumer price index, one should consider the change in the value of its individual components, the relationship of the indicator with the exchange rate of the national currency and the gross domestic product.

References

[1] S. Semerikov, S. Chukharev, S. Sakhno, A. Striuk, V. Osadchyi, V. Solovieva, T. Vakaliuk, P. Nechypurenko, O. Bondarenko, H. Danylchuk, E3S Web of Conferences **166** (2020)

[2] A. Bielinskyi, I. Khvostina, A. Mamanazarov, A. Matviychuk, S. Semerikov, O. Serdyuk, V. Solovieva, V. Soloviev, IOP Conference Series: Earth and Environmental Science **628** (2021)

[3] V.N. Soloviev, A.O. Bielinskyi, N.A. Kharadzjan, CEUR Workshop Proceedings **2832**, 24 (2020)

[4] S. Semerikov, H. Kucherova, V. Los, D. Ocheretin, CEUR Workshop Proceedings **2845**, 22 (2021)

[5] NBU, *National Bank of Ukraine* (2020), https://bank.gov.ua/ua/news/all/rishennya-oblikova-stavka

[6] OED, *Online etymology dictionary* (2020), https://www.dictionary.com/browse/stagflation#

[7] M. Sarel, *Nonlinear Effects of Inflation on Economic Growth* (IMF Working Paper, IMF, 1995)

[8] M. Kuzheliev, D. Zherlitsyn, I. Rekunenko, A. Nechyporenko, G. Nemsadze, Banks and Bank Systems **15**, 94 (2020)

[9] A.G. Macovei, L. Scutaru, Academic Research International **7**, 95 (2016)

[10] A.G. Macovei, Ecoforum **9** (2020)

[11] D. Khokhych, Finance of Ukraine **4**, 64 (2020)

[12] T. Gitis, Y. Chemerys, V. Antonova, A. Nosanyova, Economic Bulletin of Donbass **1**, 64 (2020)

[13] J. Yereshko, E. Hafarov, Efektyvna Ekonomika **5** (2020)

[14] SSSU, *The state statistics service of ukraine* (2020), https://www.ukrstat.gov.ua

[15] C. Dzitsaki, American Journal of Applied Mathematics and Statistics **4**, 136 (2016)

[16] N. Dzitsakis, P. Klareglon, Internathional Journal of Economics and Financial Issues **8**, 9 (2018)

[17] S. Everette, J. Gardner, Journal of forecasting **4**, 1 (1985)

[18] G. Box, G. Jenkins, *Time Series Analysis, Forecasting and Control* (1976)

[19] Y. Kozak, V. Matskul, T. Shengelia, *Mathematical methods and models for master of economics. Practical applications* (2017)

[20] L. Dvořáková, SHS Web of Conferences **39**, 01005 (2017)

[21] V. Matskul, D. Okara, N. Podvalna, SHS Web of Conferences **73**, 01020 (2020)

[22] V. Shinkarenko, M. Matskul, D. Linok, CEUR Workshop Proceedings **2422**, 147 (2019)