

## INFLATION FORECASTING IN THE WESTERN BALKANS AND EU: A COMPARISON OF HOLT-WINTERS, ARIMA AND NNAR MODELS

Vesna Karadzic<sup>1\*</sup> and Bojan Pejovic<sup>2</sup>

<sup>1)2)</sup> University of Montenegro, Podgorica, Montenegro

### Please cite this article as:

Karadzic, V. and Pejovic, B., 2021. Inflation Forecasting in the Western Balkans and EU: A Comparison of Holt-Winters, ARIMA and NNAR Models. *Amfiteatru Economic*, 23(57), pp.517-532.

DOI: [10.24818/EA/2021/57/517](https://doi.org/10.24818/EA/2021/57/517)

### Article History

Received: 16 December 2020

Revised: 8 February 2021

Accepted: 15 March 2021

### Abstract

The purpose of this paper is to compare the accuracy of the three types of models: Autoregressive Integrated Moving Average (ARIMA) models, Holt-Winters models and Neural Network Auto-Regressive (NNAR) models in forecasting the Harmonized Index of Consumer Prices (HICP) for the countries of European Union and the Western Balkans (Montenegro, Serbia and Northern Macedonia). The models are compared based on the values of ME, RMSE, MAE, MPE, MAPE, MASE and Theil's U for the out-of-sample forecast. The key finding of this paper is that NNAR models give the most accurate forecast for the Western Balkans countries while ARIMA model gives the most accurate forecast of twelve-month inflation in EU countries. The Holt-Winters (additive and multiplicative) method proved to be the second best method in case of both group of countries. The obtained results correspond to the fact that the European Union has been implementing a policy of strict inflation targeting for a long time, so the ARIMA models give the most accurate forecast of inflation future values. In the countries of the Western Balkans the targeting policy is not implemented in the same way and the NNAR models are better for inflation forecasting.

**Keywords:** Inflation, Holt-Winters models, Autoregressive Integrated Moving Average models, Neural Network Auto-regression models, forecasting

**JEL Classification:** C51, C53, E31, E37

---

\* Corresponding author, **Vesna Karadzic** – e-mail: [vesnaka@ac.me](mailto:vesnaka@ac.me)

### Authors' ORCID:

Vesna Karadzic: [orcid.org/0000-0003-4979-7674](https://orcid.org/0000-0003-4979-7674)

Bojan Pejovic: [orcid.org/0000-0001-9220-9327](https://orcid.org/0000-0001-9220-9327)

## **Introduction**

Price stability is one of the goals of all countries, especially the countries of the European Union and its potential members, as documented through the political agendas of the European Union and the Maastricht convergence criteria (Golinelli and Orsi, 2002). Predicting the future value of inflation is of particular importance for all countries, whether or not they have clearly declared inflation targeting policies. Historically, the European Union has successfully pursued a policy of targeting inflation and has not had high inflationary developments in the past, while the countries of the Western Balkans aspiring to become members of the European Union had high inflation during the late twentieth century caused by the numerous factors. The countries of the Western Balkans have diverse foreign exchange systems, Montenegro is a dollarized (euroized) economy, while Serbia and Macedonia have their own currencies. The Harmonized Index of Consumer Prices (HICP) monitored for the European Union comprises a number of countries and their diverse foreign exchange systems.

The paper examines the possibilities for application of Auto-Regressive Integrated Moving Average (ARIMA), Holt-Winters and Neural Network Auto-Regressive (NNAR) models in the inflation forecasting for the three countries of the Western Balkans (Montenegro, Serbia, North Macedonia) and full member countries of the European Union during the observed 2010-2020 period. The dual possibility of comparing the models is considered. Firstly, models are compared within their own class, and then the best models from each class are mutually compared.

There are many models that can be used for inflation modeling and forecasting. In this paper the possibility of inflation modeling and accuracy of its forecasting with univariate models are tested. The forecast of future values is based only on historical data on inflation. This type of models often give a faster and more accurate forecast compared to more complex factor models. This is a consequence of the unpredictability and impossibility of accurate measurement and evaluation of numerous factors, as well as their erroneous specifications in the models of traditional econometric analysis. An additional limitation is the need to predict the values of all determinants of the observed series, which further complicates the work and leads to inaccuracies in forecasting.

The paper is organized as follows. A literature review is presented in the next section. The third section reviews the basic methodological bases of development and specifications of econometric models. The fourth section presents the empirical analyses and comparison of estimated models. Finally, conclusions are presented in the fifth section.

### **1. Literature review**

Today, in modern monetary theory and central banking practice price stability is usually associated with moderate price growth (Ascari and Sbordone, 2014). The level and degree of change in inflation have always been an interesting research topic for many researchers. Researchers' interest in the field reflected the current level of development of empirical apparatus for forecasting time series.

Meyler, Kenny and Quinn (1998) forecast inflation in Ireland by comparing the ARIMA models obtained using the Box-Jenkins methodology and objective penalty function methods. Pufnik and Kunovac (2006) give a forecast of short-term inflation based on the

ARIMA model by observing the consumer price index (CPI) in Croatia. Other papers also deal with inflation forecasting using the ARIMA model with the most frequent use of Box-Jenkins methodology in model evaluation (Alnaa and Ahiakpor, 2011; Okafor and Shaibu, 2013). A comparison of the power to predict inflation by using the Holt-Winters and ARIMA models is presented by Omane-Adjepong, Oduro and Oduro (2013).

In the literature special attention is given to comparisons of more advanced prognostic models, such as comparisons of ARMA, ARIMA and GARCH models (Nyoni, 2018), comparison of VAR and ARIMA models in HICP prognosis in Austria (Fritzer, Moser and Scharler, 2002), comparison of ARIMA, VAR and ECM models (Uko and Nkoro, 2012). Suhartono (2005) compares the prognostic performances of the Neural Networks, ARIMA and ARIMAX models in inflation forecasting in Indonesia, where it is concluded that the Neural Networks model gives a more accurate inflation forecast compared to traditional econometric time series models. Sari, Mahmudy and Wibawa (2016) give an inflation forecast using the Backpropagation Neural Network method. McNelis and McAdam (2004) apply linear and neural network-based “thick” models for inflation forecast in the USA, Japan and in the euro area. Hubrich (2005) studies inflation in the European Union measured by the change in HICP and investigates whether the forecasting accuracy of forecasting aggregate euro area inflation can be improved by aggregating forecasts of sub-indices of the HICP as opposed to forecasting the aggregate HICP directly.

## 2. Research methodology

### 2.1. ARIMA models

ARIMA models are particularly suitable for short-term forecasts and the model evaluation methodology is the result of the work of Box and Jenkins (1976). ARIMA models have two components, the autoregressive component (AR) and the moving average component (MA). The AR component refers to the autocorrection coefficients of the previous data, while the MA component represents moving averages with a corresponding lag.  $I$  represents the level of integration of the series, i.e. is the series stationary with the original data, on the first or second difference. After model identification, evaluation of model parameters is tested to meet the statistical and econometric criteria of model validity. If the model meets all the validity criteria it can be used for forecasting. Seasonal ARIMA models can be represented by ARIMA (p, d, q) (P, D, Q)<sub>s</sub>, where (p, d, q) represents the non-seasonal part of the model, while (P, D, Q)<sub>s</sub> is the seasonal part where  $s$  represents the number of periods during one year.

$$(1 - \phi_1 B) (1 - \phi_1 B^{12})(1 - B)(1 - B^{12})y_t = (1 + \theta_1 B) (1 + \theta_1 B^{12})\varepsilon_t \quad (1)$$

In seasonal ARIMA models  $p$  represents the number of autoregressive elements,  $d$  is the level of series differentiation,  $q$  is the number of moving average elements, while  $P$  is the number of seasonal autoregressive elements,  $D$  is the number of seasonal differences and  $Q$  is the number of seasonal moving average elements.

### 2.2. Holt-Winters models

Holt (2004) and Winters (1960) have developed a method for forecasting time series that can successfully capture the level, trend and seasonality in the series. When forecasting

future values of a time series, values that are closer to the current one are more important than previous values that are further away. The method can be used for short, medium and long term forecasts. There are two types of methods in relation to the nature of time series, and these are additive and multiplicative models. The additive method is used when seasonal fluctuations are at approximately the same level during a time series, while the multiplicative method is used when seasonal variations change in proportion to the time series level.

Equations for the additive method:

$$\text{Forecast: } \hat{y}_{t+h|t} = \ell_t + hb_t + s_{t+h-m(k+1)} \tag{2}$$

$$\text{Level: } \ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \tag{3}$$

$$\text{Trend: } b_t = \beta * (\ell_t - \ell_{t-1}) + (1 - \beta *)b_{t-1} \tag{4}$$

$$\text{Seasonal: } s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m} \tag{5}$$

Equations for the multiplicative method:

$$\text{Forecast: } \hat{y}_{t+h|t} = (\ell_t + hb_t)s_{t+h-m(k+1)} \tag{6}$$

$$\text{Level: } \ell_t = \alpha \frac{y_t}{s_{t-m}} + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \tag{7}$$

$$\text{Trend: } b_t = \beta * (\ell_t - \ell_{t-1}) + (1 - \beta *)b_{t-1} \tag{8}$$

$$\text{Seasonal: } s_t = \gamma \frac{y_t}{(\ell_{t-1} + b_{t-1})} + (1 - \gamma)s_{t-m} \tag{9}$$

where  $y_t$  is the observed series,  $s$  is the length of the seasonal cycle,  $\ell_t$  gives the level of the series,  $b_t$  represents the trend,  $0 \leq \alpha \leq 1$ ,  $0 \leq \beta \leq 1$ ,  $0 \leq \gamma \leq 1$  and  $\hat{y}_{t+h|t}$  presents forecast for  $h$ -periods ahead.

**2.3. Neural Network Auto-Regression models**

NNAR models are a newer way that allows modeling of complex connections between inputs and outputs. In the case of the NNAR model, the previously lagged values of the observed time series serve as inputs for forecasting future values. Due to the seasonality of the observed HICP time series, the NNAR (p, P, k)m model will be used.

Figure no. 1 represents an NNAR model with one hidden layer,  $k$  hidden neurons, also known as a multilayer feed-forward network where each layer of nodes receives inputs from the previous one and sends it on to the next one (Hyndman and Athanasopoulos, 2018). The inputs of each node are obtained based on the linear combination function. The results are then modified by a nonlinear function and forwarded. The linear combination function for node  $j$  is formulated as:  $z_j = b_j + \sum_{i=1}^n w_{i,j}x_i$  it is then modified by a non-linear function, such as a sigmoid,  $s(z) = \frac{1}{1+e^{-z}}$  and it is sent to the next layer. This aims to reduce the effects of extreme values and to make networks more resistant to extreme values. The notation  $p$  represents the number of lagged autoregressive components of the model, while the notation  $P$  represents the number of lagged seasonal autoregressive components of order  $m$ . The  $k$  mark represents the number of nodes in the hidden layer.

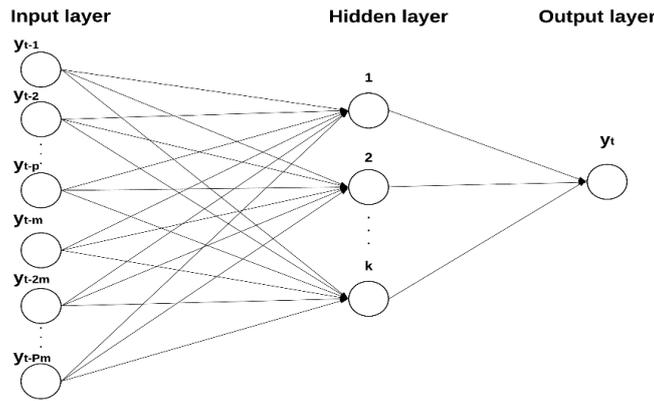


Figure no. 1. A diagrammatic representation of the NNAR (p, P, k)m model

### 2.4. Comparison of forecast accuracy

The comparison of the evaluated models for each of the observed time series will be performed on the basis of 7 criteria:

$$\text{Mean error: } ME = \frac{1}{g} \sum_{j=1}^g (x_{m+j} - \hat{x}_{m(j)}) \tag{10}$$

$$\text{Root mean square error: } RMSE = \sqrt{\frac{1}{g} \sum_{j=1}^g (x_{m+j} - \hat{x}_{m(j)})^2} \tag{11}$$

$$\text{Mean absolute error: } MAE = \frac{1}{g} \sum_{j=1}^g |x_{m+j} - \hat{x}_{m(j)}| \tag{12}$$

$$\text{Mean percentage error: } MPE = \frac{100}{g} \sum_{j=1}^g \frac{x_{m+j} - \hat{x}_{m(j)}}{x_{m+j}} = \frac{100}{g} \sum_{j=1}^g \left( 1 - \frac{\hat{x}_{m(j)}}{x_{m+j}} \right) \tag{13}$$

$$\text{Mean absolute percentage error: } MAPE = \frac{100}{g} \sum_{j=1}^g \left| \frac{x_{m+j} - \hat{x}_{m(j)}}{x_{m+j}} \right| \tag{14}$$

$$\text{Mean absolute scaled error: } q_j = \frac{e_j}{\frac{1}{T-m} \sum_{t=m+1}^T |x_t - x_{t-m}|}, \tag{15}$$

$$MASE = \text{mean}(|q_j|) \tag{16}$$

$$\text{Theil's U statistic: } U = \sqrt{\frac{\sum_{t=1}^{n-1} \left( \frac{\hat{x}_{t+1} - x_{t+1}}{x_t} \right)^2}{\sum_{t=1}^{n-1} \left( \frac{x_{t+1} - x_t}{x_t} \right)^2}} \tag{17}$$

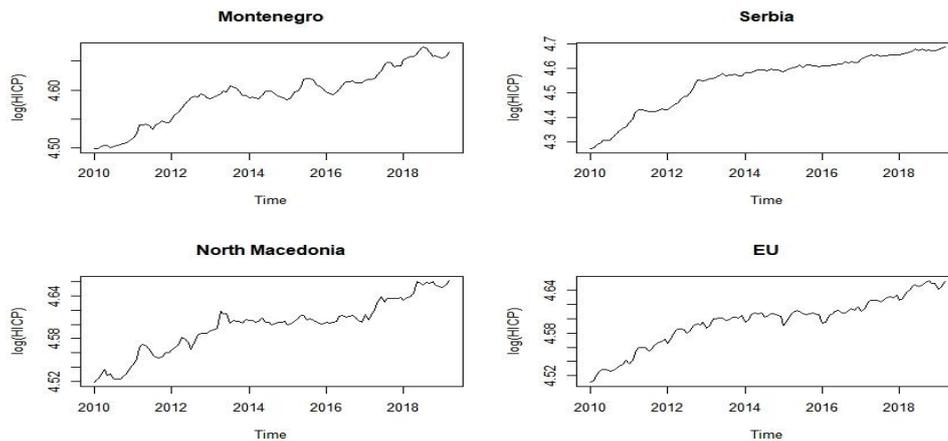
where  $x_{m+1} - \hat{x}_{m(j)}$  represents the forecast error, i.e. the difference between the actual and the predicted value of the time series. Smaller values of forecast accuracy statistics correspond to a better forecast model. In theory there are no exact limits for the values of forecast statistics that separate good from bad models. Therefore, we take the criterion that the best model has lower values of all statistics compared to competitors. The best model minimizes all ME, RMSE, MAE, MPE, MAPE, MASE and Theil's U values.

Data base for this research consist of time series of Harmonized Index of Consumer Prices (HICP). The data are presented on a monthly basis, for the period from January 2010 to March 2020.

R software package was used for analysis purposes.

**3. Results and discussion**

The analysis uses logarithmic data of the Harmonized Index of Consumer Prices (HICP) for the countries of the Western Balkans: Montenegro, Serbia, Northern Macedonia, as well as for the full member countries of the European Union in the observed period. The observed period is from January 2010 to March 2019, while the data from April 2019 to March 2020 are used to test the forecasting power of the model. The original values of the observed series are shown in figure no. 2.



**Figure no. 2. HICP log for the period from 2010 to 2019**

*Source: MONSTAT, 2020 and EUROSTAT, 2020*

**3.1. ARIMA forecasting**

From observing the ACF (Auto Correlation Function) and PACF (Partial Auto Correlation Function) of logarithmic HICP data, in figure no. 3, it is clear that all series are nonstationary and are characterized by the presence of a unit root. The graphical conclusion will be checked by ADF (Augmented Dickey-Fuller), PP (Phillips-Perron) and KPSS (Kwiatkowski–Phillips–Schmidt–Shin) empirical tests. (Table no. 1)

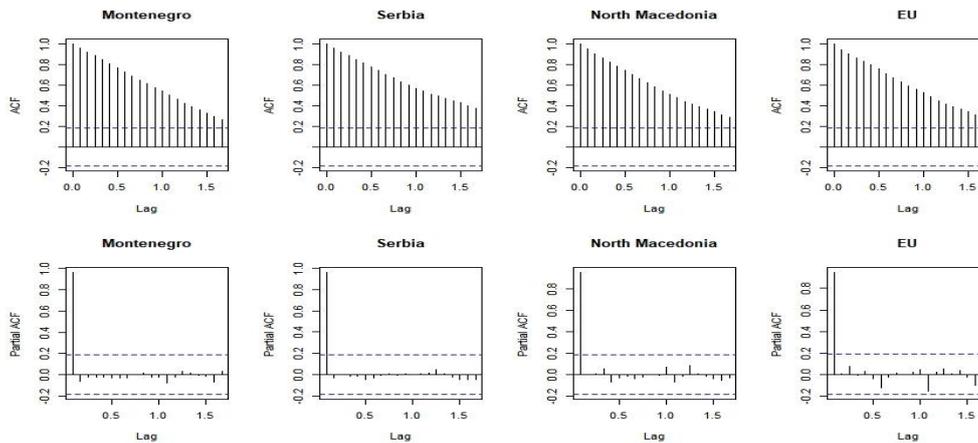


Figure no. 3. ACF and PACF of HICP

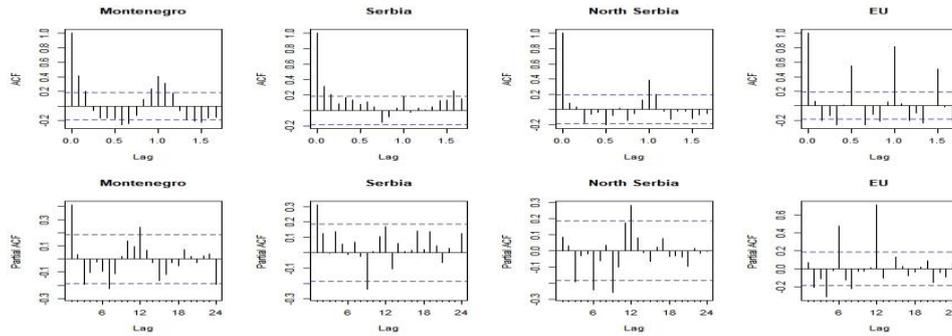
Table no. 1. Stationarity tests of HICP

Stationarity tests log(HICP)	Montenegro	Serbia	North Macedonia	EU
ADF	-2.322	-2.367	-2.426	-2.226
p-value	0.443	0.424	0.400	0.483
PP	-9.729	-4.212	-12.560	-8.878
p-value	0.553	0.873	0.388	0.602
KPSS	2.027	2.043	2.013	2.031
p-value	0.010	0.010	0.010	0.010

Based on the values of ADF and PP test statistics and calculated p probabilities for all four data series, it can be concluded that the null hypothesis is confirmed and that the series have at least one unit root, i.e. that they are not stationary. We reach the same conclusion on the basis of KPSS test statistics and p values, where we reject the null hypothesis that the series is stationary and provide an alternative that the series have at least one unit root. In order to obtain a stationary time series that can be used in further analysis, the first difference of all series was determined. The results of ADF and PP tests proved that the first differences of all series are stationary, and that as such they can be used in the further Box-Jenkins procedure. The results of the applied KPSS test in the case of Serbia are in contradiction with the obtained results of other unit root tests. In order to meet the requirements of all three unit root tests, the time series needs to be differentiated once again. Simultaneous use of a larger number of tests leads to better results due to the elimination of the shortcomings of the use of individual tests. In this research the strictest criterion is used for fulfilling the conditions of all three unit root tests.

The ACF and PACF given in figure no. 4 are used in determining the order of the AR and MA models for each of the observed series. The results of the unit root tests of the first difference data are given in table no. 2. From the correlogram of all series the presence of seasonal components is clearly noticed, which will influence the further evaluation of seasonal ARIMA models, i.e. SARIMA models. Depending on the observed series, SARIMA models will be of different order, with or without seasonal differentiation of the series. The appropriate SARIMA model is evaluated for each of the observed series, and information criteria is used as a criterion in selecting the optimal model. The best model for each of the

countries will minimize the values of all information criteria: Akaike information criterion (AIC), Akaike information criterion for small sample (AICc), and Bayesian information criterion (BIC).



**Figure no. 4. ACF and PACF of HICP (first differences)**

**Table no. 2. Stationarity tests of HICP (first differences)**

Stationarity tests diff(log(HICP))	Montenegro	Serbia	North Macedonia	EU
ADF	-4.973	-3.868	-5.027	-6.904
p-value	0.010	0.018	0.010	0.010
PP	-62.570	-87.560	-93.360	-78.040
p-value	0.010	0.010	0.010	0.010
KPSS	0.087	0.767	0.088	0.329
p-value	0.100	0.010	0.100	0.100

After the confirmed stationarity of the first difference of the time series HICP in Montenegro, the identification and evaluation of the best model was performed. The ARIMA (1,1,0) (1,0,0) [12] with drift model was chosen as the best model in the case of Montenegro. (Table no. 3)

**Table no. 3: Identification and evaluation of the model in the case of Montenegro**

Models	AICc	Best model: ARIMA(1,1,0)(1,0,0)[12] with drift
ARIMA(1,1,0)(1,0,0)[12] with drift	-895.574	Series: Montenegro ARIMA(1,1,0)(1,0,0)[12] with drift
ARIMA(1,1,0)(0,0,1)[12] with drift	-895.565	Coefficients:
ARIMA(2,1,0)(1,0,0)[12] with drift	-893.836	ar1 sar1 drift
ARIMA(1,1,1)(1,0,0)[12] with drift	-893.667	0.3236 0.3384 0.0015
ARIMA(1,1,0)(1,0,0)[12]	-894.617	s.e. 0.0945 0.0941 0.0008
		sigma^2 estimated as 1.62e-05: log likelihood = 451.98
		AIC=-895.95 AICc=-895.57 BIC=-885.15

In the case of data for Serbia, we previously concluded that the series on the first difference is not stationary, so the second difference of the original data series was determined and its stationarity examined. Based on the statistics values of all unit root tests, it is confirmed that the second difference of the series is stationary. (Table no. 4)

**Table no. 4. Stationarity test of HICP Serbia (second differences)**

Serbia	ADF	p-value	PP	p-value	KPSS	p-value
Stationarity test diff(diff(log(HICP)))	-9.151	0.01	-9.151	0.01	0.0299	0.1

The ARIMA (1,2,1) (1,0,0) [12] model is chosen as the best model in the case of Serbia. The observed series of the Harmonized Consumer Price Index for Northern Macedonia is stationary on the first difference, so this series was used in the evaluation and selection of models. (Table no. 5)

**Table no. 5. Identification and evaluation of the model in the case of Serbia**

Models	AICc	Best model: ARIMA(1,2,1)(1,0,0)[12]
ARIMA(1,2,1)	-788.295	Series: Serbia
ARIMA(1,2,1)(1,0,0)[12]	-789.811	ARIMA(1,2,1)(1,0,0)[12]
ARIMA(1,2,1)(0,0,1)[12]	-788.74	Coefficients:
ARIMA(2,2,1)(1,0,0)[12]	-788.906	ar1 ma1 sar1
ARIMA(1,2,2)(1,0,0)[12]	-788.929	0.268 -0.951 0.192
		s.e. 0.104 0.034 0.099
		sigma^2 estimated as 3.94e-05: log
		likelihood=399.1
		AIC=-790.2 AICc=-789.81 BIC=-779.43

The ARIMA (0,1,0) (1,0,1) [12] with drift model is chosen as the best model in the case of North Macedonia. Analyzing the results of the unit root test of the first time series difference for the European Union, it was noticed that the series is characterized by the presence of a unit seasonal root. In order to minimize the assessment of the standard deviation in the time series, in addition to the first difference of the series, the seasonal difference of order 12 was determined. (Table no. 6)

**Table no. 6. Identification and evaluation of the model in the case of North Macedonia**

Models	AICc	Best model: ARIMA(0,1,0)(1,0,1)[12] with drift
ARIMA(1,1,0)(1,0,1)[12] with drift	-849.391	Series: North Macedonia
ARIMA(0,1,0)(1,0,1)[12] with drift	-851.086	ARIMA(0,1,0)(1,0,1)[12] with drift
ARIMA(0,1,0)(1,0,0)[12] with drift	-849.487	Coefficients:
ARIMA(0,1,0)(2,0,0)[12] with drift	-850.123	sar1 sma1 drift
ARIMA(0,1,0)(1,0,1)[12]	-850.867	0.785 -0.443 0.001
		s.e. 0.125 0.178 0.001
		sigma^2 estimated as 2.37e-05: log
		likelihood=429.73
		AIC=-851.47 AICc=-851.09 BIC=-840.67

The ARIMA (1,1,0) (0,1,1) [12] model is chosen as the best model in the case of European Union. (Table no. 7)

**Table no. 7. Identification and evaluation of the model in the case of European Union**

Model	AICc	Best model: ARIMA(1,1,0)(0,1,1)[12]
ARIMA(0,1,1)(0,1,1)[12]	-950.854	Series: EU ARIMA(1,1,0)(0,1,1)[12] Coefficients: ar1  sma1 0.37 -0.693 s.e. 0.10 0.152 sigma^2 estimated as 5.8e-06: log likelihood=-480.18 AIC=-954.36 AICc=-954.11 BIC=-946.61
ARIMA(0,1,1)(0,1,2)[12]	-950.837	
ARIMA(1,1,0)(0,1,1)[12]	-954.105	
ARIMA(1,1,0)(0,1,2)[12]	-953.569	
ARIMA(1,1,0)(1,1,2)[12]	-953.257	

After selecting the best models for each country separately, minimizing the values of the information criteria, it is necessary to examine the existence of an autocorrelation between the residuals and the normality of their distribution. Based on the value of Ljung-Box statistics and the corresponding p value with a risk of error of 5%, the null hypothesis can be confirmed that all autocorrelation coefficients are statistically equal to zero, even up to 22th lag. The residual distributions from all the described models can be approximated by the normal distribution. Hence, the models meet both required characteristics of model validity (absence of autocorrelation and normality of residual distribution), and can be used in further analysis as a benchmark in model comparison. (Table no. 8)

**Table no. 8. Ljung-Box residual test for all ARIMA model**

Montenegro Ljung-Box test	Serbia Ljung-Box test	North Macedonia Ljung-Box test	EU Ljung-Box test
data: Residuals from ARIMA(1,1,0)(1,0,0)[12] with drift Q* = 20.86, df = 19, p-value = 0.345 Model df: 3. Total lags used: 22	data: Residuals from ARIMA(1,2,1)(1,0,0)[12] Q* = 22.19, df = 19, p-value = 0.275 Model df: 3. Total lags used: 22	data: Residuals from ARIMA(0,1,0)(1,0,1)[12] with drift Q* = 12.64, df = 19, p-value = 0.857 Model df: 3. Total lags used: 22	data: Residuals from ARIMA(1,1,0)(0,1,1)[12] Q* = 11.97, df = 20, p-value = 0.917 Model df: 2. Total lags used: 22

Models that meet all the required properties can be used to predict future time series values. Figure no. 5 gives 12-month forecasts of all four selected ARIMA models with 80% and 95% confidence intervals. Forecasts of the future value of the logarithm of the Harmonized Index of Consumer Prices represent the basis for calculating the statistics of the evaluation of the forecast in comparison with the actually realized values. The predicted values will be used not only for comparison with the actual values but also for comparison with the predicted values obtained by the Holt-Winters method and the use of neural networks. A comparison of prognostic methods will be given separately for each of the observed countries.

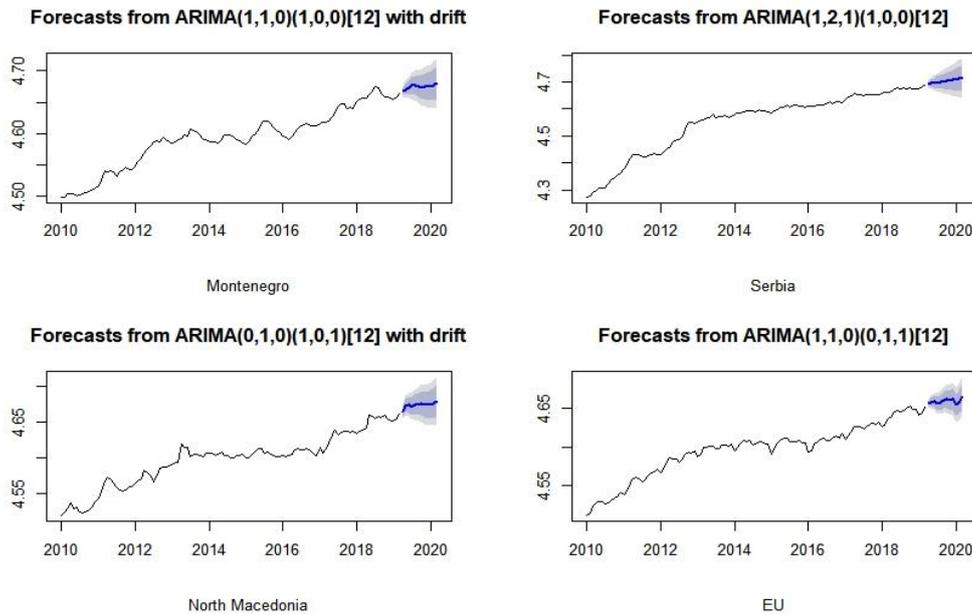


Figure no. 5. Forecast based on ARIMA models

### 3.2. Holt-Winters forecasting

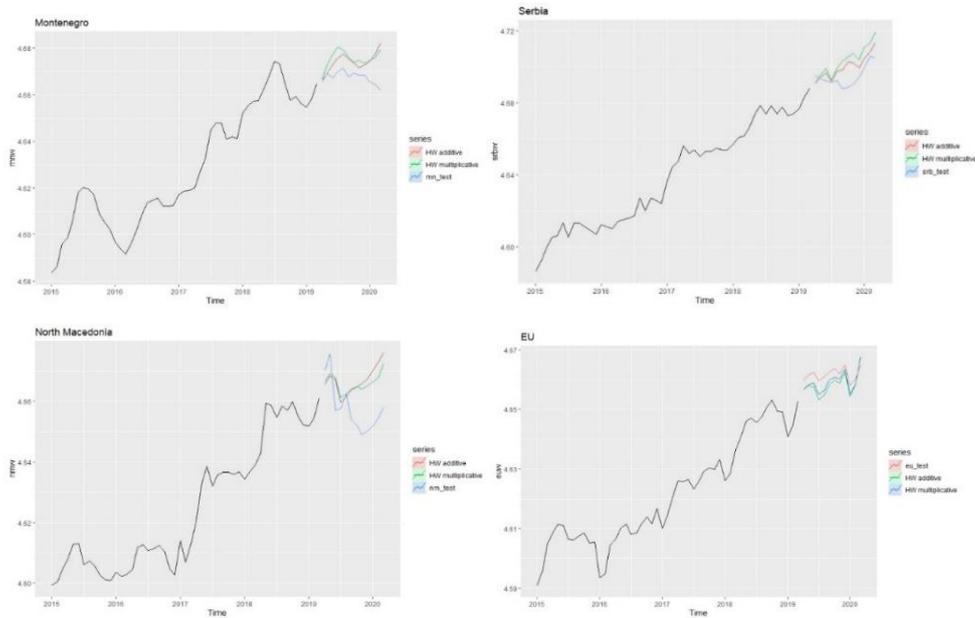
The type of time series is of special importance for the forecast of future values of the observed time series when using the Holt-Winters method. Forecasts for both types of methods (additive and multiplicative) are given in the analysis and the choice of the better method was made on the basis of forecast statistics. Data for all countries were observed separately, and the selection of the best forecast model was made based on minimizing the forecast error. (Table no. 9)

Table no. 9. Forecast statistics HW Additive and HW Multiplicative methods

		ME	RMSE	MAE	MPE	MAPE	MASE	Theil's U
Montenegro	<b>HW A</b>	<b>-0.007</b>	<b>0.008</b>	<b>0.007</b>	<b>-0.140</b>	<b>0.142</b>	<b>0.315</b>	<b>3.996</b>
	HW M	-0.008	0.009	0.008	-0.170	0.170	0.377	4.230
Serbia	<b>HW A</b>	<b>-0.005</b>	<b>0.007</b>	<b>0.006</b>	<b>-0.115</b>	<b>0.130</b>	<b>0.139</b>	<b>2.332</b>
	HW M	-0.009	0.011	0.009	-0.187	0.202	0.217	3.479
North Macedonia	HW A	-0.009	0.013	0.011	-0.199	0.242	0.670	1.987
	<b>HW M</b>	<b>-0.008</b>	<b>0.011</b>	<b>0.010</b>	<b>-0.179</b>	<b>0.214</b>	<b>0.592</b>	<b>1.714</b>
European Union	HW A	0.003	0.004	0.004	0.072	0.082	0.257	1.383
	<b>HW M</b>	<b>0.002</b>	<b>0.003</b>	<b>0.003</b>	<b>0.053</b>	<b>0.063</b>	<b>0.196</b>	<b>1.033</b>

Based on the calculated statistics, we notice that for Montenegro and Serbia the Holt-Winters Additive method gives better results, while for Northern Macedonia and the countries of the European Union the HW Multiplicative method is better. In the further analysis, the predicted values based on better methods will be used as a benchmark. Figure no. 6 shows the values of time series for the observed period, forecasts based on two Holt-Winters methods and actual series values for the out-of-the-sample period and which serves as a test of the model's

prognostic performance. Previously calculated forecast statistics and the choice of a better forecast method can also be clearly seen in the figure no. 6 because these values are closer to the actual observed data that serve as test data.



**Figure no. 6. Forecasts based on the Holt-Winters methods**

**3.3. Neural Network Auto-Regression forecasting**

Using artificial neural networks, i.e. a special type of Neural Network Auto-Regression model (NNAR), the model was evaluated, and then a forecast was made for out of the sample periods. NNAR (1,1,2) [12] models were evaluated for all observed time series, with one autoregressive element, one seasonal autoregressive element of order 12 and two neurons in the hidden layer. The estimated models are presented in table no. 10.

**Table no. 10. Selected NNAR models**

Montenegro	Model: NNAR(1,1,2)[12] sigma^2 estimated as 1.84e-05
Serbia	Model: NNAR(1,1,2)[12] sigma^2 estimated as 2.74e-05
North Macedonia	Model: NNAR(1,1,2)[12] sigma^2 estimated as 2.16e-05
European Union	Model: NNAR(1,1,2)[12] sigma^2 estimated as 8.99e-06

Data from January 2010 to March 2019 are used to evaluate the model, while data for the next 12 months until March 2020 are used to test predictive power of the model. For the forecast of one period out-of-the sample, all available data are used, for the forecast of two periods in advance, in addition to the data from the sample, the first forecast values are used. A twelve-month forecast was formed in the same way. The distribution of the residuals of the evaluated models can be approximated by the normal distribution, so the models can be

used as adequate in the further analysis. Forecasts based on the evaluated models, for all observed time series are given in figure no. 7.

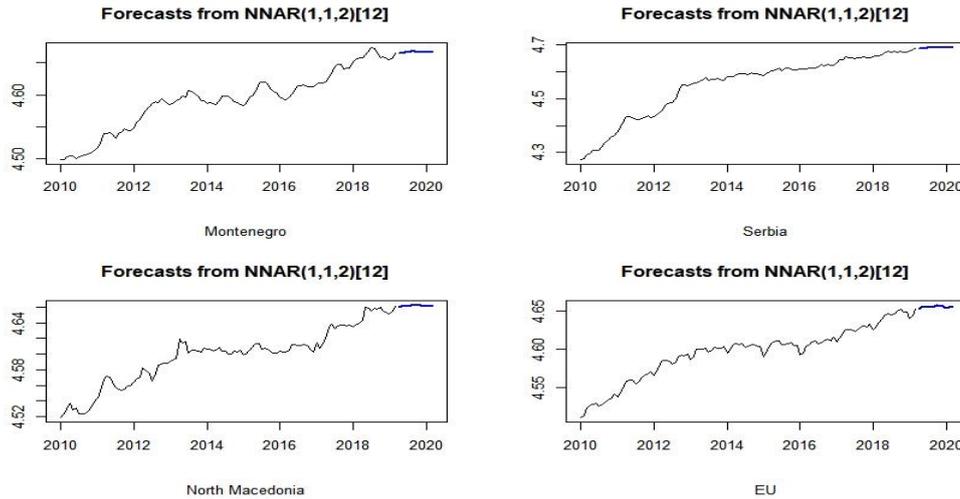


Figure no. 7. Forecast based on NNAR models

3.4. Comparison of forecast accuracy ARIMA, Holt-Winters and NNAR models

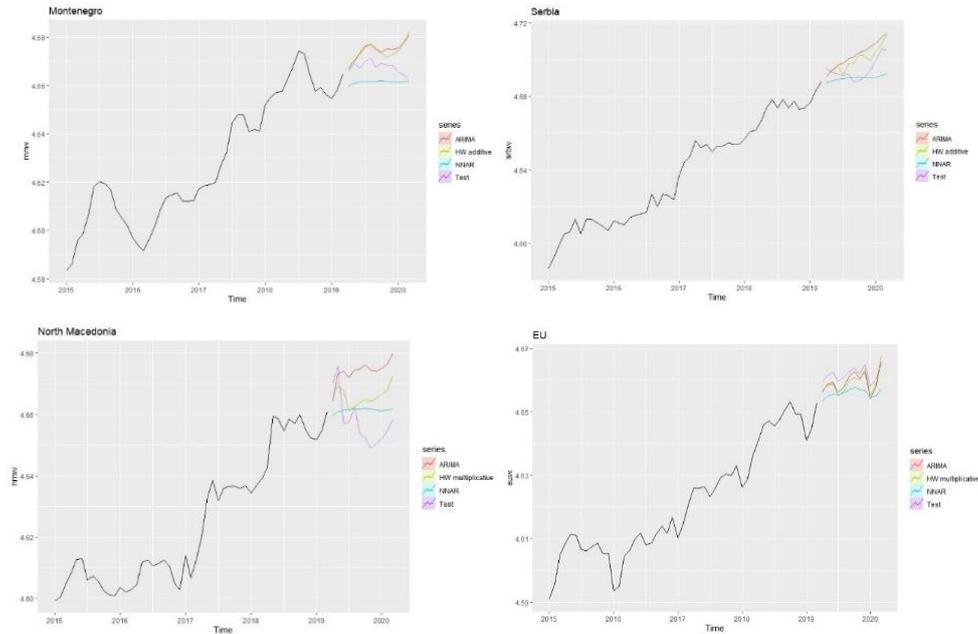
In order to compare the estimated models, it is necessary to calculate forecast errors for the out-of-the sample period. Forecast errors were calculated by comparing the actual (observed) values of the time series of the Harmonized Consumer Price Index (HICP) logarithm and the predicted values based on three types of models. The aim of this paper is to determine the class of models that gives the most accurate forecast individually for each observed series. The values of the calculated forecast statistics for all countries are shown in table no. 11.

Table no. 11. Comparison of forecasting accuracy measurement statistics

		ME	RMSE	MAE	MPE	MAPE	MASE	Theil's U
Montenegro	ARIMA	-0.007	0.009	0.007	-0.156	0.156	0.345	4.096
	HW A	-0.007	0.008	0.007	-0.140	0.142	0.315	3.996
	<b>NNAR</b>	<b>0.001</b>	<b>0.002</b>	<b>0.002</b>	<b>0.016</b>	<b>0.044</b>	<b>0.098</b>	<b>1.151</b>
Serbia	ARIMA	-0.008	0.010	0.009	-0.175	0.189	0.203	3.153
	HW A	-0.005	0.007	0.006	-0.115	0.130	0.139	2.332
	<b>NNAR</b>	<b>0.005</b>	<b>0.007</b>	<b>0.005</b>	<b>0.099</b>	<b>0.114</b>	<b>0.123</b>	<b>2.131</b>
North Macedonia	ARIMA	-0.016	0.019	0.018	-0.351	0.381	1.052	2.941
	HW M	-0.008	0.011	0.010	-0.179	0.214	0.592	1.714
	<b>NNAR</b>	<b>-0.004</b>	<b>0.009</b>	<b>0.008</b>	<b>-0.080</b>	<b>0.171</b>	<b>0.473</b>	<b>1.297</b>
European Union	<b>ARIMA</b>	<b>0.002</b>	<b>0.003</b>	<b>0.002</b>	<b>0.049</b>	<b>0.053</b>	<b>0.166</b>	<b>0.860</b>
	HW M	0.002	0.003	0.003	0.053	0.063	0.196	1.033
	NNAR	0.006	0.006	0.006	0.125	0.125	0.392	1.987

In case of inflation forecasts for Montenegro, Serbia and North Macedonia Neural Network Auto-Regressive (NNAR) models give the most accurate forecasts. The application of the

Holt-Winters model for forecasting (multiplicative or additive) gives slightly worse results, while the use of autoregressive moving averages (ARIMA) models give the most inaccurate forecast of inflation. In the case of the inflation forecast for the European Union, the autoregressive moving averages (ARIMA) model gives the most accurate forecast, followed by the Holt-Winters multiplicative model, while the Neural Network Auto-Regressive (NNAR) model gives the most inaccurate forecast of twelve-month inflation. (Figure no. 8)



**Figure no. 8. Comparison of the models for the 12-month inflation forecast**

**Conclusions**

Inflation targeting is one of the key goals of all European Union countries, as well as of the countries aspiring to become future members of EU. It is of particular importance to all Western Balkan countries that experienced severe hyper-inflation at the end of the twentieth century. Having in mind the convergence criteria that EU has put before its future member states, the dynamics of inflation measured by the change of the harmonized index of consumer prices (HICP) and its forecast is very important topic.

The purpose of this paper was to estimate adequate models for inflation forecasting and to compare their forecasting performances for the case of the three Western Balkans countries (Montenegro, Serbia, North Macedonia) and EU countries. This analysis has been carried out considering three methodologies, ARIMA, Holt-Winters and NNAR models. A comparison of the models for forecasting monthly inflation at several levels was performed. When comparing the models from the ARIMA class, the AICs information criterion was used. The Holt-Winters method comparison was performed based on the forecast error. An automatic evaluation procedure was used to evaluate the NNAR model. After selecting the best model from each of the model classes, a final comparison of the prognostic performances of the models is done on the bases of the forecast errors. For all analyzed countries of the Western

Balkans, the NNAR models give the best results in forecasting inflation. In case of the European Union the evaluated ARIMA model gave the best results. The Holt-Winters (multiplicative or additive) method is the second best in forecasting for the case of all analyzed countries.

The results of the research represent a framework for further analysis and do not provide final solutions to this problem in the observed countries. It is interesting to compare the possibilities of forecasting inflation for the countries that have different legacies that act through psychological factors, regardless of the fact that they are not included in the analysis empirically. The models do not take into account other factors that determine inflation, and therefore represent the forecast of future values only on the basis of previous values of the observed phenomenon. Although the models give a very accurate forecast, for long-term forecasts they may show certain shortcomings due to the univariate nature of the model. In the world that is characterized by numerous, rapid and sudden changes and a large number of factors and influences, even forecasts of the best model that can be evaluated can only give a rough picture of the always uncertain and challenging future.

### References

- Alnaa, S.E. and Ahiakpor, F., 2011. ARIMA (autoregressive integrated moving average) approach to predicting inflation in Ghana. *Journal of economics and international finance*, 3(5), pp.328-336.
- Ascari, G. and Sbordone, A.M., 2014. The macroeconomics of trend inflation. *Journal of Economic Literature*, 52(3), pp.679-739.
- Box, G.E., Jenkins, G.M. and Reinsel, G.C., 2011. *Time series analysis: forecasting and control*. S.I: John Wiley & Sons.
- EUROSTAT, 2020. *European Statistics*. [online] Available at: <<https://ec.europa.eu/eurostat>> [Accessed 3 October 2020].
- Fritzer, F., Moser, G. and Scharler, J., 2002. *Forecasting Austrian HICP and its components using VAR and ARIMA models*. Working Paper No. 73. Wien: Oesterreichische Nationalbank.
- Golinelli, R. and Orsi, R., 2002. *Modelling inflation in EU accession countries: the case of the Czech Republic, Hungary and Poland*. Ezoneplus working paper No. 9.
- Holt, C.C., 2004. Forecasting seasonals and trends by exponentially weighted moving averages. *International journal of forecasting*, 20(1), pp.5-10.
- Hubrich, K., 2005. Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy?. *International Journal of Forecasting*, 21(1), pp.119-136.
- Hyndman, R.J. and Athanasopoulos, G., 2018. *Forecasting: principles and practice*. S.I: OTexts.
- McNelis, P.D. and McAdam, P., 2004. *Forecasting inflation with thick models and neural networks*. European Central Bank Working Paper Series No. 352.
- Meyler, A., Kenny, G. and Quinn, T., 1998. *Forecasting Irish inflation using ARIMA models*. Central Bank and Financial Services Authority of Ireland Technical Paper Series , Vol. 1998, No. 3/RT/98.

- MONSTAT, 2020. *Statistical Office of Montenegro*. [online] Available at: <<https://www.monstat.org/eng/index.php>> [Accessed 3 October 2020]
- Nyoni, T., 2018. Modeling and forecasting inflation in Kenya: Recent insights from ARIMA and GARCH analysis. *Dimorian Review*, 5(6), pp.16-40.
- Okafor, C. and Shaibu, I., 2013. Application of ARIMA models to Nigerian inflation dynamics. *Research Journal of Finance and Accounting*, 4(3), pp.138-150.
- Omane-Adjepong, M., Oduro, F.T. and Oduro, S.D., 2013. Determining the better approach for short-term forecasting of Ghana's inflation: Seasonal ARIMA Vs holt-winters. *International Journal of Business, Humanities and Technology*, 3(1), pp.69-79.
- Pufnik, A. and Kunovac, D., 2006. *Short-term Forecasting of Inflation in Croatia with Seasonal ARIMA processes*. Croatian National Bank Working Papers No. 16.
- Sari, N.R., Mahmudy, W.F. and Wibawa, A.P., 2016. Backpropagation on neural network method for inflation rate forecasting in Indonesia. *Int. J. Soft Comput. Its Appl.*, 8(3), pp.70-87.
- Suhartono, S., 2005. Neural Networks, ARIMA and ARIMAX Models for Forecasting Indonesian Inflation. *Widya Journal of Management and Accounting*, 5(3), p.219732.
- Uko, A.K. and Nkoro, E., 2012. Inflation forecasts with ARIMA, vector autoregressive and error correction models in Nigeria. *European Journal of Economics, Finance & Administrative Science*, 50, pp.71-87.
- Winters, P.R., 1960. Forecasting sales by exponentially weighted moving averages. *Management science*, 6(3), pp.324-342.