Communication Papers of the 19th Conference on Computer Science and Intelligence Systems (FedCSIS) pp. 177–184 DOI: 10.15439/2024F2570 ISSN 2300-5963 ACSIS, Vol. 41

Prediction of Temperature and Precipitation Changes for Serbia Using Time Series Models with Machine Learning

Miroslav Tomić, Slavica Kordić, Vladimir Dimitrieski
0000-0002-8987-6695
0000-0002-3992-0096
0000-0003-3234-6543
University of Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia
Email: {tmiroslav, slavica, dimitrieski}@uns.ac.rs

Marija Đukić, Ivan Luković
0000-0002-1136-4278
0000-0003-1319-488X
University of Belgrade, Faculty of Organizational Sciences, Jove Ilića 154, 11000 Belgrade, Serbia
Email: {marija.djukic, ivan.lukovic}@fon.bg.ac.rs

***Abstract*—In the past few decades, there has been an evident change in climatic conditions worldwide as well as on the territory of Serbia. Extremely high temperatures, heavy floods, and sudden changes in the weather are increasingly frequent occurrences that bring great social and material damage. Climate change affects many economic sectors, like tourism and agriculture, which are potentially at risk. In Serbia, one of the vital economic sectors is agriculture. In order to act preventive, the main goal of this research was to predict the mean monthly temperature and precipitation for Serbia for periods 2021-2050 and 2071-2100. We collected a dataset titled ERA5 monthly averaged data on single levels from 1940 to present from the Climate Data Store. The dataset was analyzed and prepared to be used with SARIMA(X) and ARIMA(X) methods, which are utilized for prediction. The results that we identified are presented in this paper.**

***Index Terms*—prediction, temperature, precipitation, time series, machine learning.**

I. Introduction

CLIMATE changes and weather conditions have a great impact on human life. They can bring great material and financial damage, and even human lives can be at risk. Also, many economic sectors, such as tourism, hospitality industry, and construction, are affected. Besides this, climate change negatively affects agriculture in a financial sense. With climate change, there is a severe risk of crop economic losses as well as food security [1]. In Serbia, agriculture proved to be the most vital economic sector [2]. It participated in the creation of GDP with around 8% in 2022 [3], while before, this number was even greater around 20% [2]. The results of a study created by Đuričin et al. [4] indicate that there is a high degree of impact of climate conditions and reduction in crop production in Serbia. Poor crop yields can be a consequence of insect pests that can be affected by changing climate conditions in several ways [1], [5], [6]. They can result in things like an increased number of generations, expansion of their geographic distribution, and increased survival during overwintering. Some attempts were made to detect plant pathogen infestation using AI [7] to increase corn income. Extending such an approach with the prediction of the climate conditions could bring even more benefits. Climate change in Serbia will directly cause drier and warmer regions where vineyards are located, which will have a negative impact on grape growth [8]. Due to climate change, crop yields in Serbia may be reduced, but this can be avoided with preventive actions [9]. One of the possible preventive actions is to predict how much weather conditions will be changed. Additionally, the impact of climate change is important in forest ecosystems. Orlović et al. [10] discussed this impact in their study, which is made for forest ecosystems in Serbia. Also, Stojanović et al. [11] discussed the impact of weather extremes on the forests in Serbia.

Climate change is a consequence of global warming, which is a serious problem that the world is facing today. Global warming has reached record-breaking levels, visible as an increase in atmospheric temperature and sea level [12]. The Earth could experience global warming of 1.4 to 5.8 °C over the next century, which is based on many global climate models and development scenarios [13].

The climate of Serbia can be described as moderate-continental with more or less local characteristics when taking the standard normal period of climatology 1961-1990 [9]. Countries in Southeastern Europe, including Serbia, are facing significant impacts from climate change according to results from all general circulation models. Lalić et al. [14] observed that there is a positive trend in the extreme winter and summer temperatures in Serbia. As shown in a study created by the Ministry of Environmental Protection [15], from 2000 to 2015, the total material damage in Serbia caused by extreme climatic and weather conditions was more than five billion euros. Among the direct consequences of climate change is an increase in temperature, followed by heavy rainfall, which further causes floods and landslides. The cat-

This research has been supported by the Ministry of Science, Technological Development and Innovation, Faculty of Technical Sciences, University of Novi Sad and Faculty of Organizational Sciences, University of Belgrade

astrophic floods that occurred in 2014 in Serbia were a direct consequence of heavy rainfall, and it was estimated that 1.35 billion euros would be needed to recover from those floods. This directly indicates that it is necessary to take climate change into account when planning sectoral development and infrastructure. According to this study, Serbia is more affected by rising temperatures than most places on Earth, which is a consequence of climate change. Besides this, national defense issues are also affected by climate change. It took a long time for the discussion regarding climate change to expand from the framework of environmental protection to the functioning of national defense [16].

In order to reduce damage, it is possible to act preventive. One of the possible ways is to predict climate change. Quality predictions require the creation of quality predictive models. With this in place, the preparation for upcoming disasters would be much better, and the damage would be significantly reduced. The benefit of such a system is primarily social but also economic. Climate prediction at the state level (certainly at the global level) would help a state to determine its strategic goals. A concrete example of this is the prediction of how many dry years will occur in the coming period. If such predictions guided a country, it could set its strategic goals as early as possible and start with earlier investments in irrigation systems, which would certainly be more favorable from the financial side at this stage and would bring greater profit in later years. A concrete example of this is given by Ruml et al. [8], who stated that in the future, in a certain scenario, the vineyard yield would decrease if there were no irrigation.

Guided by such knowledge, we researched weather conditions and possible climate changes in Serbia. We aim to identify if there will be changes in temperature and precipitation in Serbia using mean monthly temperature and precipitation. The main goal of our research is to predict the mean monthly temperatures and precipitation of Serbia for the periods 2021-2050 and 2071-2100 using machine learning (ML) methods with weather variables that are expected to have the most impact on the prediction. These two periods of time have been choose because other authors also used these periods for prediction, and thus, we can compare our results with their. The obstacle to achieving this prediction with meteorological models is their complexity, which requires a deep understanding of mathematical equations. Also, the number of variables that could affect temperature and precipitation is huge, and it is almost impossible to include all of them. We expect to identify that there will be an increase in temperature and a possible change in the amount of precipitation. These results can help the country to identify earlier which preventive measures can be taken to reduce the damage that can be caused by climate change. In addition, the results could help civil services to better prepare for upcoming disasters, and farmers would be able to make long-term plans for growing crops.

This paper is organized into seven sections. Besides the introduction section, related work is given in the second section. A description of the materials that are used, as well as applied methods, can be found in the third section. The results are presented in the fourth section. The fifth section contains a discussion of the results. Threats to validity are given in the sixth section, and in the seventh section, a conclusion is given.

II. Related Work

In the literature review, we encountered many papers dealing with climate change and the prediction of weather conditions. We classified papers into two categories. The first category contains papers describing statistical and ML methods used to predict temperature and precipitation. The papers from the meteorological domain that focus on Serbia and the impact of climate change on agriculture are in the second category. We use them to compare their results with our findings.

Papacharalampous et al. [17] considered each continent on Earth as a whole, but they did not consider geographical differences within the continent, such as the proximity of the sea mountains. In this paper, the authors used different methods to predict monthly temperature and the monthly amount of precipitation. They used methods: naïve, random walk, AutoRegressive Fractionally Integrated Moving Average (ARFIMA), Box-Cox transformation, ARMA errors, Trend and Seasonal Components (BATS), Simple exponential smoothing, Theta, and Prophet. Data contains a sample of 985 40-year-long monthly temperatures and 1552 40-year-long monthly precipitation time series. For model evaluation, they used RMSE and NSE. In our research, we also have time series data, and we also consider the country as a whole, and we use different statistical and ML methods to achieve better results.

Camelo et al. [18] used Autoregressive Integrated Moving Average with eXogenous Variable (ARIMAX) and the Holt-Winters (HW), both combined with Artificial Neural Networks (ANN) to predict wind speed. As an input to their model, they used time series data, which contains weather-related variables. In our case, we also plan to use a similar combination of models and data. The authors used MAE, MAPE, and RMSE to evaluate models.

According to the literature [18], [19] that we have found, it is possible to apply different types of predictions with ARIMA models, such as long-term and short-term. In both cases, it is possible to predict multiple values or to predict only one value (one-step ahead). Short-term prediction is executed on data where data closer in time has a greater influence on the target variable. It is possible to use short-term and long-term predictions on climate data, which depends also on the amount of available data and the goal of prediction.

To the best of our knowledge, several studies have presented analyses of climate change in Serbia. From a meteorological perspective, the best way to simulate future climate change is by using global oceanic-atmosphere coupled models [20]. Since regional climate has its own specificity, the best way is dynamic downscaling to provide fine-scale information [21]. Because of this, many regional climate model systems have been developed, such as EBU-POM [22]. Kržić et al. [20] predicted changes in climate indices for Serbia using EBU-POM for SRES-A1B and SRES-A2 scenarios. Their results show an overall increase in the surface air temperature of about 2 and 4 °C and a decrease in seasonal precipitation. The number of days with absolute maximum temperature > 30 °C (tropical days) will increase, while the total number of days with absolute minimum temperature < 0 °C (frost days) will decrease in the future.

Mihailović et al. [9] analyzed climate change effects on crop yields in Serbia. They used different meteorological models to predict temperature and precipitation. In meteorology, a 30-year period is used as a reference period, and then an additional period of 30 years should be used for the evaluation of models that are created based on the reference period. This was applied by Mihailović et al. [9]. They used the period 1961-1990 as a reference period, and they gave predictions for the periods 2001-2030 and 2071-2100. In their work, they also included CO2 emissions besides weather variables. Vukovic et al. [23] analyzed the effect of global warming on climate change in Serbia. They analyzed the period 1961-2100 and presented concrete changes in temperature and precipitation. Both papers conclude that Serbia will be affected by climate change due to global warming and that there will be an increase in temperature and an increased amount of precipitation. We use these two papers and paper by Kržić et al. [20] to compare concrete values of predicted monthly temperatures and precipitation with the results that we get.

Gocic and Trajkovic [24] analyzed data for twelve stations in Serbia during 1980-2010 with non-parametric Mann-Kendall and Sen's methods. Their results indicate an increasing trend in temperature. One of the most recent studies on climate conditions in Serbia by Burić et al. [25] shows a tendency towards an arid climate with a significant increase in temperature and changes in precipitation patterns in Serbia.

III. Material and Methods

In this section we first describe the dataset, data preprocessing, data analysis and software tools that we used. Then, we present the method being used in our research.

A. Dataset

The dataset was downloaded from the Climate Data Store (CDS), which provides information on Earth's climate. In our research, we used a dataset titled ERA5 monthly averaged data on single levels from 1940 to present [5]. This dataset is created and maintained by the European Center for Medium-Range Weather Forecasts (ECMWF). The abbreviation ERA5 stands for ECMWF Reanalysis 5th Generation. ERA5 is the climate reanalysis that offers data on an hourly or monthly temporal resolution about atmospheric, land, and ocean parameters and uncertainty estimates. Climate reanalysis combines past observations with meteorological models to generate a consistent time series of multiple climate variables. The data is gridded to a regular latitude-longitude grid of 0.25 degrees and contains 261 variables. After conversation with domain experts, we decided to use five variables in the first phase of our research. We aim to determine if it is possible to create smaller models that don't need many variables but can still make accurate predictions. Variables are described in Table I.

Table I.
Description of variables from the source dataset

Name	Units	Description
10m wind speed	ms^{-1}	The horizontal speed of the wind at a height of 10 meters above the surface of the Earth.
2m temperature	K	The temperature of the air at 2 meters above the surface of the land, sea or inland waters.
2m dewpoint temperature	K	The temperature to which the air, at 2 meters above the surface of the Earth, would have to be cooled for saturation to occur. It can be used to calculate relative humidity.
total precipitation	m	The accumulated liquid and frozen water, comprising rain and snow, falls to the Earth's surface.
evaporation	m of water equivalent	The accumulated amount of water that has evaporated from the Earth's surface.

To analyze data and use it for statistical and ML models, we had to preprocess data, which we did in three steps. In the first step, we downloaded data via CDS Application Programming Interface (API) in Gridded Binary of General Reqularly-distributed Information in Binary form (GRIB). Before downloading data, we had to define the area for the data to be downloaded. We set latitude boundaries between 41 and 47 degrees and longitude boundaries between 18 and 23 degrees which approximately corresponds to Serbia's borders. Also, we defined desired variables, years, and months. In the second step, we went sequentially through the downloaded data and extracted the desired values of variables from GRIB format. Then, we transformed data as a part of the third step and created the dataset that is described in Table II.

Variables 2m temperature and 2m dewpoint temperature were transformed from Kelvin degrees to Celsius degrees. Total precipitation and evaporation were transformed from meters to millimeters. Also, the dataset was appended with additional variable humidity, as the domain expert advised. According to found literature humidity can be calculated as humidity = 100*es(2m dewpoint temperature)/es(2m temperature) where es(t)=610.94*e^(17.625*t/(243.04+t)). Since we used the variable 2m dewpoint temperature only to calculate humidity, we dropped this variable from the dataset. After the data preprocessing we got the dataset which have 6 variables where each variable has 720 values.

TABLE II.
DESCRIPTION OF VARIABLES FROM THE PREPROCESSED DATASET

Name	Units	Description
date	YYYY-MM (format)	Year and month
wind_speed	ms-1	Average monthly wind speed
temp_mean	°C	Average monthly air temperature
prec	mm	Average monthly precipitation
evap	mm of water equivalent	Average monthly evaporation

Since the dataset does not contain any missing values, we analyzed the correlation between data to see if there are dependent variables. Fig 1. shows that the highest correlation is between temperature and humidity. Since a domain expert advised including this variable, we decided to preserve it. Also, there is a correlation between temperature and evaporation. A correlation between temperature and wind speed exists, but according to Wooten , it is only near the Earth's surface, and because of this, we can't say that temperature and wind speed are correlated in general. Precipitation is not correlated with any other variable in the dataset.

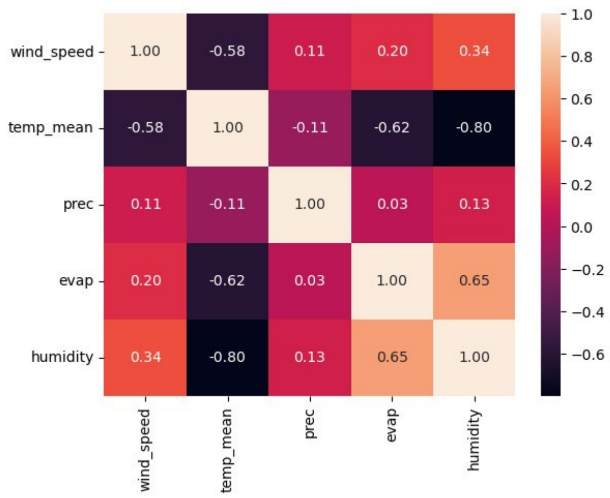

Fig 1. Correlation matrix

Next, we analyzed outlier values with interquartile range. For wind speed, temperature, and evaporation there were no outlier values. For humidity, there are 3 outlier values that are below the lower whisker. In the case of precipitation, 3 outlier values are above the upper whisker. From this analysis, it is obvious that there is a very small number of outliers and we decided to preserve all values.

Since time series data can have seasonality, it was necessary to check if this data has such a component. We applied AutoCorrelation Function (ACF) and Partial AutoCorrelation Function (PACF) to check seasonality. The outcome of this check is important for choosing parameters for statistical models. ACF and PACF were applied to the subset of data to be able to visualize it in Fig 2. and Fig 3. These diagrams can be seen for temperature and precipitation, respectively. In the case of temperature, we can notice a seasonality since data repeat every 12 months. This is expected because Serbia has a moderate continental climate condition. On the other hand, for precipitation, seasonality can't be observed, which is typical for precipitations in Serbia, according to the Republic Hydrometeorological Service of Serbia .

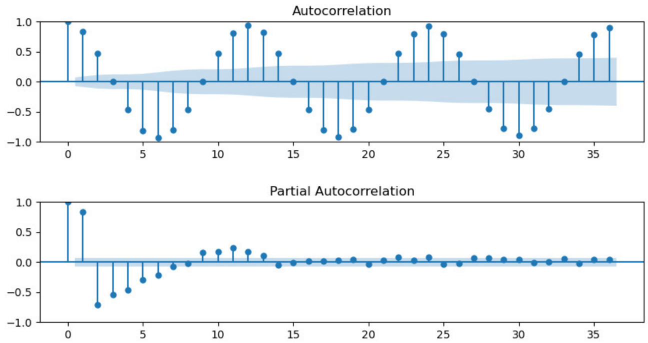

Fig 2. ACF and PACF for temperature

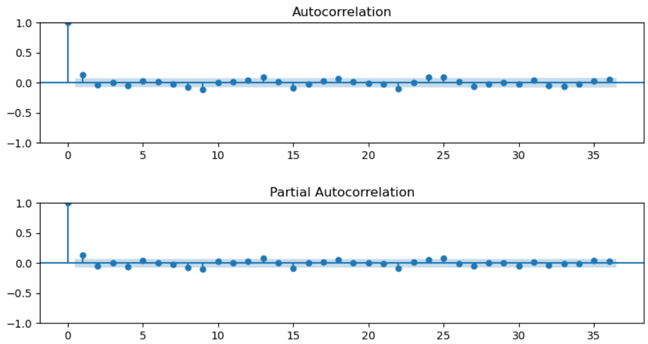

Fig 3. ACF and PACF for precipitation

The statistical models that we used are sensitive to data that is not stationary, which was checked with the Augmented Dickey-Fuller test (ADF). With this test, we get the p-value, and if the p-value value is greater than 0.05, the data is not stationary, and if the p-value is lower or equal to 0.05, the data is stationary. After running this test, we found that all the data from the dataset is stationary. Also, we found that there is no significant trend in the data. After data analysis, we defined methods that will be used for prediction, which is described in the following subsection.

To implement the previously described preprocessing of the dataset, we used the Python programming language with the libraries OS, pandas, matplotlib, seaborn, plotly, and statsmodel. We also used the Python programming language with the libraries OS, itertools, numpy, pandas, math, time, sklearn, and statsmodels to implement methods that will be described in the following subsection.

B. Methods

To achieve the defined goal of this research we decided to use the following models: ARIMAX and Seasonal ARIMA(X) (SARIMAX) with and without exogenous variables. The ARIMA method is widely used in statistical technique for time series analysis and forecasting. It combines three key components: autoregression (AR), differencing (I), and moving average (MA), to model a variety of temporal data. The autoregressive part leverages the dependency between an observation and several lagged observations, while differencing involves subtracting an observation from a previous observation to make the time series stationary. The moving average part models the relationship between an observation and a residual error from a moving average model applied to lagged observations. These components make ARIMA a robust tool for handling non-stationarity data by converting it into a stationary form through differencing. The model parameters (p, d, q) are optimized to minimize forecast errors, where p denotes the number of lag observations, d is the number of times the data needs to be differenced to achieve stationarity, and q is the size of the moving average window. The SARIMA method extends the ARIMA model to handle seasonal variations in time series data. By incorporating seasonal components along with non-seasonal ones, SARIMA can effectively model and forecast data exhibiting periodic patterns. The model is extended by a seasonal part of parameters (P, D, Q, s) where s represent the length of the seasonal cycle. SARIMA offers a framework well-suited for applications such as weather forecasting and other fields where seasonality is a significant factor. Also, we have data that cover a long period of time and since we aim to predict temperature and precipitation for a long period of time in the future we decided to use long-term predictions from SARIMAX model.

IV. Results

For this research, it is important to find appropriate values for all (S)ARIMA(X) parameters to increase the quality of the model. In order to find these values, we used a generic approach called grid search, where all parameter combinations from a limited set of parameter values are exhaustively considered. For parameters related to SARIMAX components we created a set of possible values and for exogenous variables, we defined partitive set of available variables. Each iteration of the grid search created model is evaluated on the test data, and its evaluation is saved to determine the best parameter combination. We use Root Mean Square Error (RMSE) as a metric for evaluation. At the same time, Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are calculated to be used potentially by other researchers to compare their results.

Since we aimed to predict the mean monthly temperature and precipitation, we divided the models into these two categories. Thus, for the temperature model, we will use long-term SARIMA with and without exogenous variables, and we will name these models *temp_model_1* and *temp_model_2*, respectively. For the precipitation model, we will use long-term ARIMA because there is no seasonality in precipitation data. This model will also be used with and without exogenous variables, and the names of these models will be *prec_model_1* and *prec_model_2*. For all of these models, for the training set, we used data from the reference period 1961-1990, and for the test set, we used data from the period 1991-2020. For the future predictions we chose the periods 2021-2050 and 2071-2100.

After applying the models *temp_model_1*, *temp_model_2*, *prec_model_1*, and *prec_model_2*, we got the results that indicate that there will be an increase in the mean month temperature. Also, results for precipitation indicate that there will be no change in mean month precipitation. Concrete values are depicted in this section, while details about the results are discussed in the next section.

In the case of *temp_model_1* and *prec_model_1*, we set exogenous variables, and for *temp_model_2* and *prec_model_2*, we didn't use exogenous variables. Since precipitation data does not show seasonality, we didn't use seasonal parameter in hyper-parameter optimization for the precipitation model. After hyper-parameter optimization, we got results which are shown in Table III.

Table III.
Hyper-parameters for temperature and precipitation models

Model	(p, d, q)	(P, D, Q, s)	trend	exog. variables
temp_model_1	(1, 1, 0)	(1, 1, 1, 12)	/	evap and humidity
temp_model_2	(1, 1, 0)	(0, 1, 1, 12)	/	/
prec_model_1	(1, 0, 0)	/	constant	evap and humidity
prec_model_2	(0, 0, 0)	/	constant	/

The values obtained in hyper-parameter optimization align with data analysis outcomes. We have evaluated these models over the period 1991-2020. and obtained values that are shown in Table IV. From these values, we can conclude that a temperature model with exogenous parameters gives better results than a model without exogenous variables.

Table IV.
Evaluation of temperature and precipitation models for the period 1991-2020

Model	RMSE	MAE	MAPE
temp_model_1	1.5195	1.1244	0.3857
temp_model_2	1.7007	1.3064	0.3369
prec_model_1	1.0710	0.8893	0.5578
prec_model_2	1.0803	0.8967	0.5773

The other part of our goal was to investigate whether temperature and precipitation increased or decreased when we compare the mean temperature of the reference period and the period 1991-2020. Real values of mean temperature for these periods will be shown in Table V. as well as predictions from temperature models. For precipitation, real values with predictions from precipitation models are shown in Table VI. Also, we created predictions for the mean monthly temperature and precipitation for periods 2021-2050 and 2071-2100, which are shown respectively in Table VII and Table VIII. For this prediction, since exogenous variables are not known for the future period, we used *temp_model_2* and *prec_model_2*, which don't require exogenous variables.

TABLE V.
REAL VS PREDICTED MEAN TEMPERATURE FOR THE PERIOD 1991-2020

Model	Variable	Value (°C)
/	Temp_true_1961_1990	10.6090
	Temp_true_1991_2020	11.4543
	ΔTemp_true_1991_1961	0.8453
temp_model_1	Temp_pred_1991_2020	11.6192
	ΔTemp_pred_1991_1961	0.1649
temp_model_2	Temp_pred_1991_2020	11.1005
	ΔTemp_pred_1991_1961	-0.3537

TABLE VI.
REAL VS PREDICTED MEAN PRECIPITATION FOR THE PERIOD 1991-2020

Model	Variable	Value (mm)
/	Prec_true_1961_1990	2.5624
	Prec_true_1991_2020	2.5178
	ΔPrec_true_1991_1961	-0.0446
prec_model_1	Prec_pred_1991_2020	2.5198
	ΔPrec_pred_1991_1961	0.0020
prec_model_2	Prec_pred_1991_2020	2.5657
	ΔPrec_pred_1991_1961	0.0480

TABLE VII.
THE PREDICTED MEAN TEMPERATURE FOR THE PERIODS 2021-2050 AND 2071-2100

Model	Variable	Value (°C)
/	Temp_true_1961_1990	10.6090
	Temp_true_1991_2020	11.4543
temp_model_2	Temp_pred_2021_2050	11.5712
	Temp_pred_2071_2100	12.3568
	ΔTemp_pred_2021_1961	0.9623
	ΔTemp_pred_2021_1991	0.1170
	ΔTemp_pred_2071_1961	1.7478
	ΔTemp_pred_2071_1991	0.9025

TABLE VIII.
THE PREDICTED MEAN PRECIPITATION FOR THE PERIOD 2021-2050 AND 2071-2100

Model	Variable	Value (mm)
/	Prec_true_1961_1990	2.5624
	Prec_true_1991_2020	2.5178
prec_model_2	Prec_pred_2021_2050	2.5657
	Prec_pred_2071_2100	2.5657
	ΔPrec_pred_2021_1961	0.0034
	ΔPrec_pred_2021_1991	0.0480
	ΔPrec_pred_2071_1961	0.0034
	ΔPrec_pred_2071_1991	0.0480

V. DISCUSSION

In this section, we discuss our achieved results and compare them with the results found in the literature review.

All models that we created achieved good performances when we consider the RMSE metric, as it is shown in Table IV. As we expected, models with exogenous variables gave better results in the case of temperature prediction, while in the case of precipitation, we got identical results. There is one downside of models with exogenous variables they can't be easily used for future predictions. To use these models, exogenous variables values must be known, which means that the values of these variables must be predicted with another model.

A comparison between the period 1991-2020 and the reference period 1961-1990 for both temperature and precipitation is presented in Table V. and Table VI, respectively. In the case of temperature, we can see that the *temp_model_1* predicts an increase in temperature while *temp_model_2* predicts a small decrease. The true value indicates that there was a small increase in temperature. Concrete values from *temp_model_1* are close to values from papers [9], [23]. For precipitation, both models predicted that there would be a small increase in precipitation, while the true value indicates that there was a small decrease.

Since Vuković et al. [23] and Mihailović et al. [9] also predicted mean temperature for the period 2071-2100, we decided to use *temp_model_2* to predict mean temperature for that period and to compare results. The *temp_model_2* is used since we don't have exogenous variables for this period. Besides this period, we predicted a mean temperature for the period 2021-2050 in order to see if the trend of temperature increase will continue. The values that we got from our prediction indicate that there will be an increase in temperature. These results are shown in Table VII, where we can see that models predicted an increase in the temperature. When we compare this prediction with the prediction by Vuković et al. [23] and Mihailović et al. [9], we conclude the same that there will be an increase in temperature over the years. Kržić et al. [20] estimated that the overall temperature will increase by about 2 and 4 °C for the period 2071-2100 compared to the reference period 1961-1990. We also identified an increase in temperature, but our results show that the overall increase in temperature when we compare the period 2071-2100 with the reference period 1961-1990 will be around 1.75 °C. Our results in Table VIII. show that precipitation will be almost constant, which is different from the results in the papers [9], [20], [23], where it is said that precipitation will change. Precisely, Kržić et al. [20] identified a decrease in precipitation of about 13 and 6 mm while our results show that there will be no change in precipitation.

VI. Threats to Validity

In this section, we express threats to the validity of the proposed work:

- Due to limited hardware resources, we couldn't create a bigger set of parameter combinations that could be used in order to find the best combination with which models could be trained and evaluated. An example is the order of integration where we used range 0-2 which might not be enough. Because of this limitation, we can't be certain if these models could show better performances for temperature and precipitation prediction. To overcome this threat, we could find appropriate hardware to run these models.
- Since the used models are methods that utilize maximum likelihood, we noticed that after multiple runs with the same data and parameters, results can differ. Such a difference is small and appears at the fifth or sixth decimal place in the result. This was determined empirically after multiple tries. In every try, we didn't get results that changed a trend in the temperature and precipitation data that we presented in this paper.

VII. Conclusion

In this paper, we presented an approach to show that it is possible to utilize SARIMA(X) models for climate prediction. During the literature review that we conducted, we didn't find any study that utilizes ML methods to predict climate changes in Serbia. Even though there are many studies that describe the application of ML methods for climate prediction for different areas, there still could be differences in application due to area specificity. The concrete code that is used for this research should be adapted and packaged in an application for use by any user. Since this approach serves more for future predictions in this form, it can only be used for strategic planning. For this approach to be used for operative purposes, a data collection module and a similar model that could be used for short-term predictions are required. Finally, domain experts should validate the results that are obtained from our approach in order to use this approach.

The models we created are simple to implement and utilize on any hardware, which is one benefit. Even if these models give predictions with a small error, they can still help create future plans. Also, because of low hardware requirements, they can be used on farms, where a farmer can use their predictions to create more accurate plans for growing crops. Since these models give accurate predictions for a temperature trend in the future, they can be used for government strategic planning. As an example, agriculture authorities could better allocate funding for irrigation systems. During creation of migrational and socio-economic politics authorities should consider climate change. Additionally, this prediction could influence plans for infrastructure projects as well as companies that want to start a business in Serbia.

It is observed by analyzing models that the best results for predicting mean monthly temperature give a long-term SARIMA model with exogenous variables. The downside of this model is that it can only be used with exogenous variables for future prediction. Even with such a shortage, a temperature model predicted an increase in temperature over the years, which aligns with the outcomes we found during the literature review. In the case of precipitation, our models didn't predict a significant change, which is not in line with outcomes from the literature that we found.

Our plan for future work is to adapt these models for short-term prediction, which could be useful for operative purposes. Since these models could be used in different geographical locations in Serbia, models created for geographical locations with similar climate conditions could show better performances than general models for the whole country described in this paper. It could be helpful to create a few different models for such purposes. Besides this, we would like to increase the number of weather variables used in prediction models, which could contribute to a more accurate prediction. Additionally, using a walk-forward validation technique could give better results. Also, more advanced methods than SARIMA(X), like Recurrent Neural Networks (RNNs), could be used with time-series data. We plan to implement this method and compare the results with the one we presented here. Also, we plan to provide a study on using exogenous variables. Last but not least, we would like to create a different kind of prediction where we will predict the distribution of temperature and precipitation. This is of great importance to avoid disasters caused by sudden changes in climatic conditions.

Acknowledgment

This research has been supported by the Ministry of Science, Technological Development and Innovation (Contract No. 451-03-65/2024-03/200156) and the Faculty of Technical Sciences, University of Novi Sad through project “Scientific and Artistic Research Work of Researchers in Teaching and Associate Positions at the Faculty of Technical Sciences, University of Novi Sad” (No. 01-3394/1); and by Faculty of Organizational Sciences, University of Belgrade.

References

[1] S. Skendžić, M. Zovko, I. P. Živković, V. Lešić, and D. Lemić, "The Impact of Climate Change on Agricultural Insect Pests," *Insects*, vol. 12, no. 5, Art. no. 440, 2021, doi: http://dx.doi.org/10.3390/insects12050440.

[2] P. Radovan, "Razvojni problemi i prioriteti poljoprivrede Republike Srbije," in *Performanse i razvojne mogućnosti agrarne industrije u AP Vojvodini : prilog regionalnoj strategiji unapređenja agrarnog biznisa*, vol. 15, Subotica, Serbia: Univerzitet u Novom Sadu, Subotica: Ekonomski fakultet, 2014, pp. 29–66. doi: http://dx.doi.org/10.2298/PRMAIV2013029P.

[3] Statistical Office of the Republic of Serbia, "Statistical release | Gross Domestic Product, 2022." Accessed: May 25, 2024. [Online]. Available: https://www.stat.gov.rs/en-US/vesti/statisticalrelease/?p=13820

[4] S. Đuričin, I. Beraha, and M. Đukić, "The impact of climate extremes on agricultural output in the Republic of Serbia," *Ecol. Nauka Privreda Iskustva*, vol. 20, no. 72, Art. no. 72, 2013.

[5] V. Trkulja, A. Tomić, T. Popović, and R. Iličić, "The Impact of Climate Change on the Occurrence of Diseases and Pests of Agricultural Plants and Forest Trees," in *Natural Resources Management in a Changing Climate*, vol. 8, 2023. doi: http://dx.doi.org/10.7251/EORU2308477T.

[6] R. Jevtić, S. Maširević, and D. Vajgand, "The impact of climate change on diseases and pests of small grains and sunflower in the Vojvodina Region (Serbia). In Essays on Fundamental and Applied Environmental Topics," in *Essays on Fundamental and Applied Environmental Topics*, New York, NY: Nova Science Pub Inc, 2012, pp. 277–306.

[7] J. Stypułkowska, "The Use of AI to Determine the Condition of Corn in a Field Robot that Meets the Requirements of Precision Farming," in *Annals of Computer Science and Information Systems*, 2023, pp. 319–327. doi: http://dx.doi.org/10.15439/2023F3649.

[8] M. Ruml *et al.*, "On the use of regional climate models: Implications of climate change for viticulture in Serbia," *Agric. For. Meteorol.*, vol. 158–159, pp. 53–62, 2012, doi: http://dx.doi.org/10.1016/j.agrformet.2012.02.004.

[9] D. T. Mihailović, B. Lalić, N. Drešković, G. Mimić, V. Djurdjević, and M. Jančić, "Climate change effects on crop yields in Serbia and related shifts of Köppen climate zones under the SRES-A1B and SRES-A2," *Int. J. Climatol.*, vol. 35, no. 11, pp. 3320–3334, 2015, doi: http://dx.doi.org/10.1002/joc.4209.

[10] S. Orlović, Z. Galić, S. Stojnić, and B. Klašnja, "Monitoring of forest ecosystems in Serbia," in *Essays on Fundamental and Applied Environmental Topics*, New York, NY: Nova Science Pub Inc, 2012, pp. 253–276.

[11] D. B. Stojanović *et al.*, "Climate change within Serbian forests: Current state and future perspectives," *Topola*, no. 208, pp. 39–56, 2021, doi: http://dx.doi.org/10.5937/topola2108039S.

[12] "Climate change 2014: impacts, adaptation, and vulnerability: Working Group II contribution to the fifth assessment report of the Intergovernmental Panel on Climate Change," Cambridge University Press, New York, NY, 2014.

[13] "Climate change 2007: the physical science basis: contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change," Cambridge University Press, Cambridge ; New York, 2007.

[14] B. Lalic, J. Eitzinger, D. T. Mihailovic, S. Thaler, and M. Jancic, "Climate change impacts on winter wheat yield change – which climatic parameters are crucial in Pannonian lowland?," *J. Agric. Sci.*, vol. 151, no. 6, pp. 757–774, 2013, doi: http://dx.doi.org/10.1017/S0021859612000640.

[15] D. Bozanic and D. Mitrovic, "Study on the Socio-economic Aspects of Climate Change in the Republic of Serbia," United Nations Development Programme, 2019.

[16] V. M. Cvetković, S. Z. Vučić, and J. Gačić, "Climate change and national defense," *Vojno Delo*, vol. 67, no. 5, pp. 181–203, 2015, doi: http://dx.doi.org/10.5937/vojdelo1505181C.

[17] G. Papacharalampous, H. Tyralis, and D. Koutsoyiannis, "Predictability of monthly temperature and precipitation using automatic time series forecasting methods," *Acta Geophys.*, vol. 66, no. 4, pp. 807–831, 2018, doi: http://dx.doi.org/10.1007/s11600-018-0120-7.

[18] H. do N. Camelo, P. S. Lucio, J. B. V. Leal Junior, P. C. M. de Carvalho, and D. von G. dos Santos, "Innovative hybrid models for forecasting time series applied in wind generation based on the combination of time series models with artificial neural networks," *Energy*, vol. 151, pp. 347–357, 2018, doi: http://dx.doi.org/10.1016/j.energy.2018.03.077.

[19] G. K. Venayagamoorthy, K. Rohrig, and I. Erlich, "One Step Ahead: Short-Term Wind Power Forecasting and Intelligent Predictive Control Based on Data Analytics," *IEEE Power Energy Mag.*, vol. 10, no. 5, pp. 70–78, 2012, doi: http://dx.doi.org/10.1109/MPE.2012.2205322.

[20] A. Kržič, I. Tošić, V. Djurdjević, K. Veljović, and B. Rajković, "Changes in climate indices for Serbia according to the SRES-A1B and SRES-A2 scenarios," *Clim. Res.*, vol. 49, pp. 73–86, 2011, doi: http://dx.doi.org/10.3354/cr01008.

[21] F. Giorgi and L. O. Mearns, "Introduction to special section: Regional Climate Modeling Revisited," *J. Geophys. Res. Atmospheres*, vol. 104, no. D6, pp. 6335–6352, 1999, doi: http://dx.doi.org/10.1029/98JD02072.

[22] V. Djurdjević and B. Rajković, "Development of the EBU-POM coupled regional climate model and results from climate change experiments," in *Advances in Environmental Modeling and Measurements*, UK ed. edition., New York, NY: Nova Science Pub Inc, 2010, pp. 23–32.

[23] A. Vukovic *et al.*, "Global warming impact on climate change in Serbia for the period 1961-2100," *Therm. Sci.*, vol. 22, no. 6 Part A, pp. 2267–2280, 2018, doi: http://dx.doi.org/10.2298/TSCI180411168V.

[24] M. Gocic and S. Trajkovic, "Analysis of changes in meteorological variables using Mann-Kendall and Sen's slope estimator statistical tests in Serbia," *Glob. Planet. Change*, vol. 100, pp. 172–182, 2013, doi: http://dx.doi.org/10.1016/j.gloplacha.2012.10.014.

[25] D. Burić, J. Mihajlović, V. Ducić, M. Milenković, and G. Anđelković, "Contribution to the study of climate change in Serbia using continentality, oceanity, and aridity indices," *Időjárás*, vol. 127, no. 3, pp. 379–399, 2023, doi: http://dx.doi.org/10.28974/idojaras.2023.3.6.