Environ Monit Assess (2021) 193:84
https://doi.org/10.1007/s10661-020-08821-1

Prediction of tropospheric ozone concentration using artificial neural networks at traffic and background urban locations in Novi Sad, Serbia

Slavica Malinović-Milićević · Yaroslav Vyklyuk · Gorica Stanojević · Milan M. Radovanović · Dejan Doljak · Nina B. Ćurčić

Received: 21 September 2020 / Accepted: 27 December 2020

Abstract In this paper, we described generation and performances of feedforward neural network models that could be used for a day ahead predictions of the daily maximum 1-h ozone concentration ($1hO_3$) and 8-h average ozone concentration ($8hO_3$) at one traffic and one background station in the urban area of Novi Sad, Serbia. The six meteorological variables for the day preceding the forecast and forecast day, ozone concentrations in the day preceding the forecast, the number of the day of the year, and the number of the weekday for which ozone prediction was performed were utilized as inputs. The three-layer perceptron neural network models with the best performance were chosen by testing with different numbers of neurons in the hidden layer and different activation functions. The mean bias error, mean absolute error, root mean squared error, correlation coefficient, and index of agreement or Willmott's Index for the validation data for $1hO_3$ forecasting were 0.005 μg m^{-3}, 12.149 μg m^{-3}, 15.926 μg m^{-3}, 0.988, and 0.950, respectively, for the traffic station (Dnevnik), and − 0.565 μg m^{-3}, 10.101 μg m^{-3}, 12.962 μg m^{-3}, 0.911, and 0.953, respectively, for the background station (Liman). For $8hO_3$ forecasting, statistical indicators were − 1.126 μg m^{-3}, 10.614 μg m^{-3}, 12.962 μg m^{-3}, 0.910, and 0.948 respectively for the station Dnevnik and − 0.001 μg m^{-3}, 8.574 μg m^{-3}, 10.741 μg m^{-3}, 0.936, and 0.966, respectively, for the station Liman. According to the Kolmogorov–Smirnov test, there is no significant difference between measured and predicted data. Models showed a good performance in forecasting days with the high values over a certain threshold.

Keywords Air pollution · Tropospheric ozone · Neural networks · Novi Sad (Serbia)

S. Malinović-Milićević (✉)
ACIMSI - University Center for Meteorology and Environmental Modelling, University of Novi Sad, Novi Sad 21000, Serbia
e-mail: slawicam@gmail.com

Y. Vyklyuk
Bukovinian University, Chernivtsi, Ukraine

G. Stanojević · M. M. Radovanović · D. Doljak · N. B. Ćurčić
Geographical Institute "Jovan Cvijic", Serbian Academy of Sciences and Arts, Belgrade 11000, Serbia

M. M. Radovanović
Institute of Sports, Tourism and Service, South Ural State University, Chelyabinsk, Russia 454080

Introduction

Most cities around the world are dealing with serious air pollution problems, which have been receiving increasing attention in the recent decades. The main cause of the air pollution is the rapid growth of the urban population, accompanied by increased motor vehicle traffic. The ground-level or tropospheric ozone (O_3) is the basic component of photochemical smog and is considered a very dangerous pollutant. Unlike stratospheric O_3, which protects us from harmful UV radiation, it is located at altitudes inhabited by humans and represents a threat to human health. Ozone has a harmful influence

Published online: 26 January 2021

on the respiratory organs causing an increase in mortality, which is registered in areas with high O_3 concentration (Ito et al. 2005; Bates 2005; Zanobetti and Schwartz 2008). Studies in 23 European cities have shown that the daily mortality rate increases linearly with an increase of the average hourly concentration of O_3 by 50–60 μg m^{-3} (Gryparis et al. 2004). Similar studies have been carried out for 98 cities in the USA and have shown that concentrations lower than 40 μg m^{-3} do not significantly affect mortality and that mortality increases linearly with concentrations of 70–80 μg m^{-3} and above (Bell et al. 2006). Ozone also affects vegetation and materials (Lee et al. 2003; Cape 2008). High levels of O_3 can damage the leaves and reduce plant productivity. Exposure to O_3 over time leads to degradation of the quality of materials and shortens their life span.

Ozone formation in urban areas is a unique phenomenon, i.e., it is not emitted directly into the atmosphere, but it is the result of the complex interaction between nitrogen oxides and organic volatile compounds in favorable meteorological conditions. It has a relatively long lifetime in polluted urban regions and can be transported to areas hundreds of kilometers away from the place of formation (Stevenson et al. 2006). At a certain location, concentration depends on photochemical reactions that produce O_3, regional transport, descending of O_3 from stratosphere, and ozone-destroying reactions (Cape 2008). Meteorological conditions strongly influence O_3 concentrations. They play an essential role in the formation, dissemination, transportation, and also destruction of tropospheric O_3. The high amount of solar radiation and high atmospheric pressure increase tropospheric O_3 concentrations, while high relative humidity and rain reduce it (Lelieveld and Crutzen 1990; Zanis et al. 2014). Wind velocity can increase or decrease O_3 concentrations, depending on photochemical production in the boundary layer (Biancofiore et al. 2015).

In Europe, during hot and sunny summer weather conditions, O_3 concentrations significantly may exceed the allowable values (EPA 2009), ranging from 200 to 300 μg m^{-3}. The European standard for the daily maximum 8-h average O_3 concentration is 120 μg m^{-3} (not to be exceeded more than 25 days per calendar year; averaged over 3 years). In 2006, the WHO reduced the threshold limit of the daily maximum 8-h average concentration that poses health risks from 120 to 100 μg m^{-3} (WHO 2006). In addition to the daily maximum 8-h average ozone concentration, the European standards defined an "information threshold" of 180 μg m^{-3} and an "alarm threshold" of 240 μg m^{-3} for daily maximum 1-h O_3 concentration. There are many epidemiological studies dealing with short-term exposure threshold; however, their findings are not consistent. According to the WHO (2013), the threshold for a daily maximum 1-h O_3 concentration is likely to be below 90 μg m^{-3}. Therefore, in order to monitor and predict changes in tropospheric O_3 concentrations in the future and to provide the latest information on its current and future levels that can be used as a warning when the limit value is expected to be exceeded, it is necessary to understand its nature and conditions that contribute to its creation.

Predicting O_3 concentration is difficult due to very complex interactions between pollutants, and it can be done by using deterministic and statistical modeling. Deterministic models provide the output which is fully determined by the parameter values and the initial conditions, and they are very time-consuming. These models require a significant number of data on meteorological elements and O_3 precursors, which are rarely available. On the other hand, statistical models are much simpler. They are not aiming to establish cause–effect relationships, but to find a statistical relationship between predictors and output data. There are many studies using various statistical techniques to predict O_3 concentration such as multiple linear regression, principal component analysis, and clustering technique (Abdul-Wahab et al. 2005; Duenas et al. 2005; Ghazali et al. 2010; Sun et al. 2013). Although linear models are easy to use and acceptable for O_3 modeling, they do not take into account the nonlinear response of O_3 to precursor concentrations. These deficiencies can be compensated by the use of neural networks. Neural network models are based on the use of a set of the input variables that include meteorological parameters and concentrations of O_3 precursors. They have been used in air quality modeling since the 1990s. As input parameters, neural network models use measured air pollutant concentrations and meteorological variables. The comparison of linear and nonlinear methods showed better performance of neural network models compared to linear regression (Prybutok et al. 2000; Chaloulakou et al. 2003; Sousa et al. 2007; Nghiem and Kim Oanh 2009).

So far, no forecasting tools have been applied to predict tropospheric O_3 in the city of Novi Sad. This study aimed to examine the possibility of using artificial neural network models to make predictions of daily

maximum 1-h O_3 concentration ($1hO_3$) and daily maximum 8-h average O_3 concentration ($8hO_3$) for the next day at an urban location in north Serbia by using meteorological parameters and O_3 concentration from the previous day. The daily maximum 1-h O_3 concentration is the maximum value of the hourly average concentration on a particular day, while daily maximum 8-h average O_3 concentration is the maximum of moving averages of eight consequent hourly averaged concentrations calculated between 1 and 24 h on a particular day. Therefore, the aim of this study is to predict tropospheric O_3 concentration on day $t+1$ on the both monitoring stations, Dnevnik and Liman, using measurements of meteorological parameters on the day t, predictions of meteorological parameters on the day $t+1$, and measurements of O_3 concentration on the day t.

Materials and methods

Site description and data

The observations were carried out in the city of Novi Sad (45.15° N, 19.50° E; 80–130 m a.s.l.), the second largest urban area in the Republic of Serbia. It is situated in Central Europe in Pannonian pain (Fig. 1). The area is characterized by temperate continental climate with warm summers. The built-up area of the city covers 102 km^2 and has a population approximately 350,000 inhabitants (Statistical Office of the Republic of Serbia 2019). The main sources of air pollution are local transportation and residential heating (heating plants and individual furnaces). The main fuels used for residential heating are natural gas, wood, and coal (Official Gazette of the City of Novi Sad, No. 49 2018).

Air pollutants and meteorological data for five years (2010, 2011, 2012, 2013, and 2016) were used in the models developed in this study. During 2014 and 2015, O_3 measurements were not conducted in Novi Sad. Air pollutant data was collected from the Serbian National Air Monitoring network (SEPA 2019). The dataset consists of hourly average tropospheric O_3 concentrations collected from the two air quality monitoring stations, Dnevnik and Liman. The Dnevnik station is an urban traffic type, while Liman station is an urban background type. Both stations are located in a densely populated urban area, while the distance between them is about 2 km. Ozone concentrations were recorded every 60 min by an automatic detector (Teledyne API Model 400A O3 Analyzer). Meteorological data used in the study were obtained from the nearest meteorological station, Rimski Sancevi, operated by the Republic Hydrometeorological Service of Serbia (RHMZ 2019). The distance between meteorological and air pollution stations is 8.5 and 10.5 km, respectively. Used meteorological parameters are as follows: the daily mean temperature, °C (T_{av}); the daily maximum temperature, °C (T_{max}), global radiation, MJ m^{-2} (GR), wind speed, m s^{-1} (WS), relative humidity, % (RH), and atmospheric pressure, mbar (P). Daily values of global radiation were determined using data on daily duration of sunshine using Prescott's formula (Coulson 1975).

Structure of input parameters

To build the model that can be used to forecast O_3 concentration, historical meteorological and O_3 data, as well as forecasted values of meteorological parameters for a day ahead, are needed in the input layer. Although Kukkonen et al. (2003) showed that incorporating forecasted meteorological variables improves the performance of artificial neural network (ANN) models, we decided not to include forecasted values at this stage of the modeling. We considered that error in the forecast of the meteorological input parameters might influence the prediction model, so for the purpose of model construction instead of forecasted values, we used measured meteorological values for a day ahead. This usual approach assuming a perfectly accurate weather forecast is proposed for the use in the developing model stage by Klein et al. (1959) and Wilks (1995).

There were interruptions in the measurement of O_3 concentrations due to device malfunctions and maintenance; therefore, if the value is not measured more than 6 h during the day, the entire day was removed from the dataset. Hence, the total dataset used in this study was 704 at Dnevnik station and 902 at Liman station (1 per day).

The daily, weekly and annual pattern of ozone concentration at both measuring stations is clearly visible in (Fig. 2). The highest concentrations occur during the periods of the highest radiation, in the middle of the day, and during summer. Annual variations in ozone concentrations depend on different factors, such as the proximity of the original ozone precursor, latitude, meteorological factors, and transportation from the stratosphere. At both measuring stations, ozone concentrations show

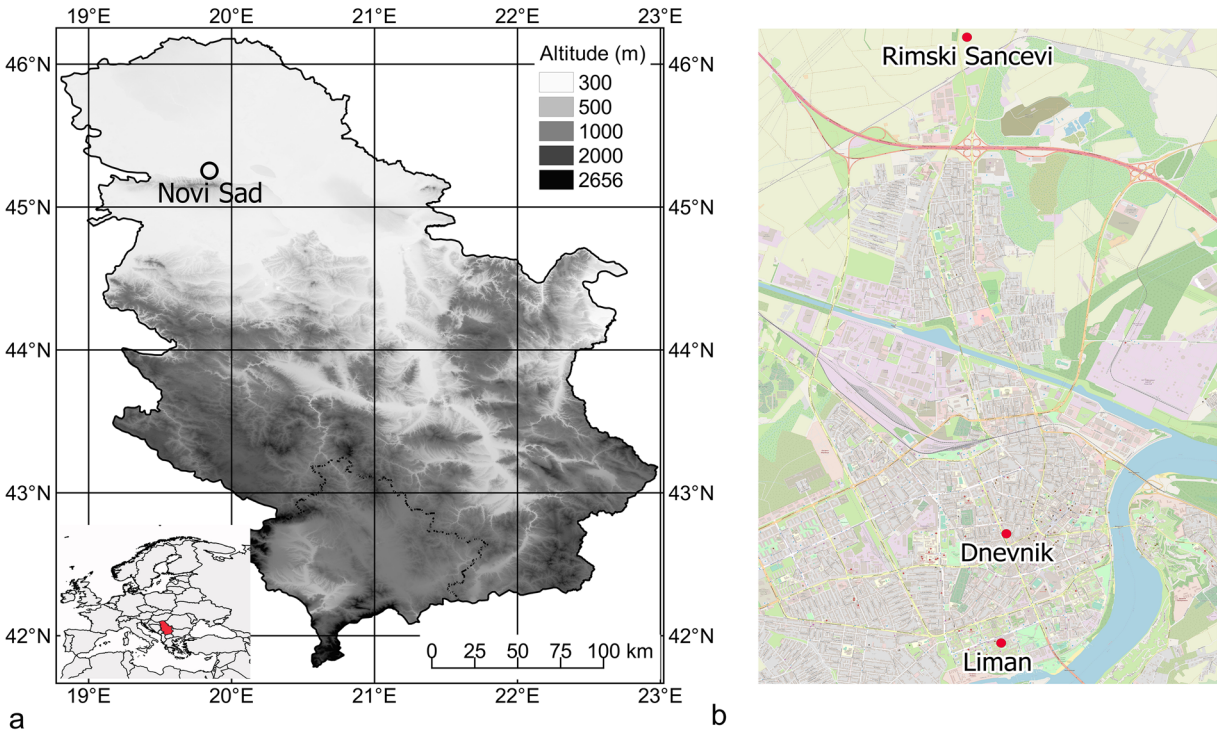

Fig. 1 **a** Location of Serbia in Europe, Novi Sad in Serbia and **b** the air quality monitoring stations (Dnevnik and Liman) and meteorological stations (Rimski Sancevi) in Novi Sad

summer maximum, and winter minimum, which is an annual cycle typical for polluted areas. The existence of a broad summer maximum is associated with the photochemical production of ozone which is formed in reactions involving the VOCs and NO_x under the influence of solar radiation (Logan 1985). At both measuring stations, an occurrence of higher maximum surface O_3 concentrations on weekends is also visible, so-called ozone weekend effect, caused by the lower NO_x emissions from motor vehicles indicating complex O_3 formation chemistry (Altshuller et al. 1995). Figure 3 reveals that relative humidity shows an inverse correlation with temperature, showing lower values in summer and higher values in winter, which confirms that production of O_3 is the most intense during sunny, warm, and dry meteorological conditions.

Artificial neural network models

Neural network-based models are complex and flexible, and they can model the expression of nonlinear dependencies arising from measurement data. Among the variants of the ANN model for urban air quality modeling, the multi-layer perceptron model (MLP) stands out (McKendry 2002; Rahman et al. 2020). This model has proven to be very accurate and reliable for the urban air quality forecast. The MLP model consists of a network of simple process elements and their links. Process elements, or neurons, take their place in the respective layers. There are several layers: the input layer, one or more hidden layers, and the output layer. For each neuron, the weighted sum of the inputs received from adjacent neurons is calculated, then is processed using the activation function, and the result obtained is submitted to the next layer. Finally, an output signal is obtained (Kukkonen et al. 2003).

We used a feedforward neural network, consisting of 15 input nodes: meteorological variables (T_{av}, T_{max}, GR, WS, RH, P) for the day preceding the forecast (t) and forecast day ($t+1$); O_3 concentrations for the day preceding the forecast, $1hO_3(t)$ or $8hO_3(t)$; number of day of year for which O_3 forecast was carried out, DOY($t+1$); and number of weekday for which O_3 forecast was carried out, DOW($t+1$). An output node, the target variables are daily maximum 1-h O_3 concentration, $1hO_3(t+1)$, or daily maximum 8-h average O_3 concentration, $8hO_3(t+1)$. The complete dataset was implemented in Statistica Automated Neural Networks

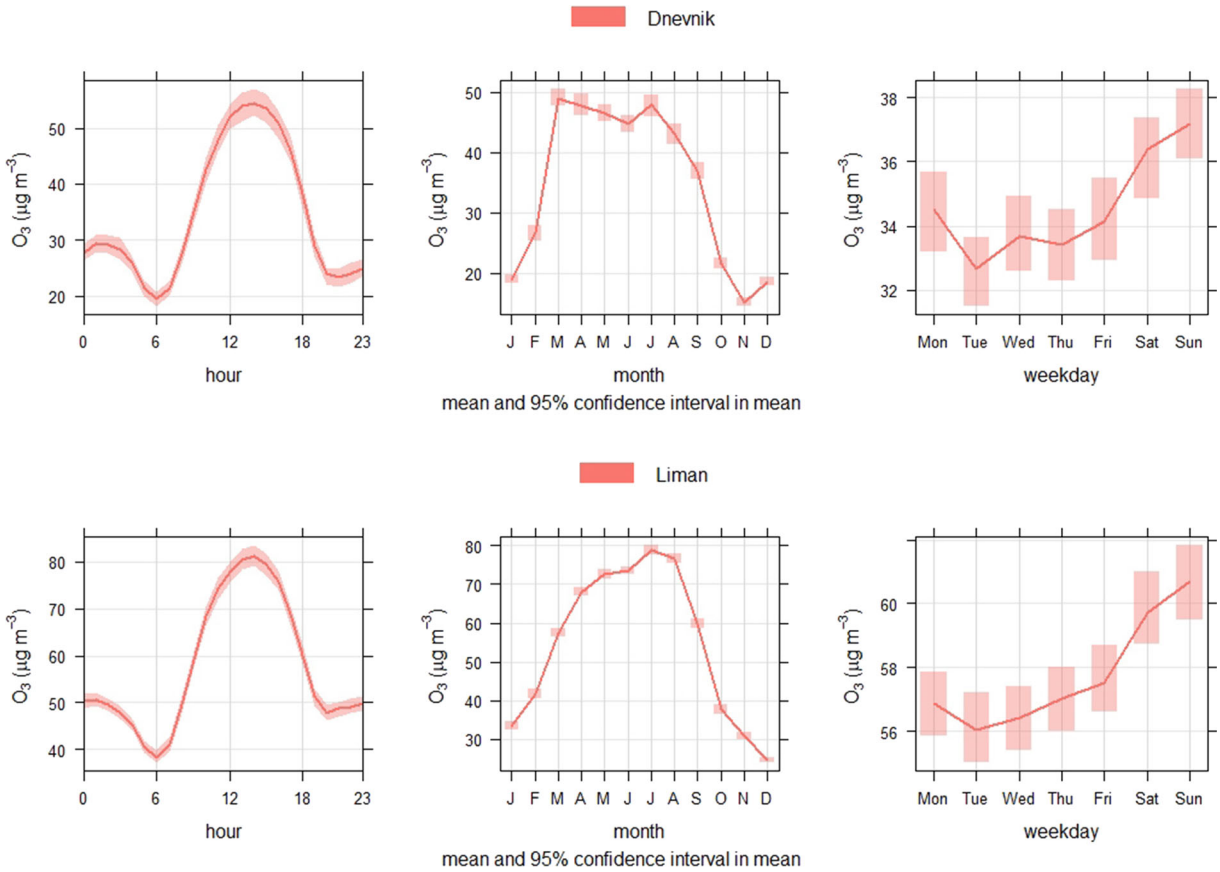

Fig. 2 Daily, monthly, and weekly pattern of hourly O_3 concentrations at measuring sites Dnevnik and Liman

(SANN) (TIBCO Software 2020) which tests a three-layer neural network model with different numbers of hidden layer neurons and various activation functions (Logistic, Tanh, Exponential, and Identity), and then chooses the networks with the best performance. All input variables were automatically normalized to provide values between 0 and 1. Guided by the claim of Hecht-Nielsen (1987) that any continuous function with n inputs in the range 0–1 can be implemented exactly by a three-layer neural network having $2n + 1$ neurons in the hidden layer, we have limited the maximum number of hidden layer neurons to 31. The entire dataset is automatically divided into subsets of training (70% of cases), testing (15%), and validation (15%) using the

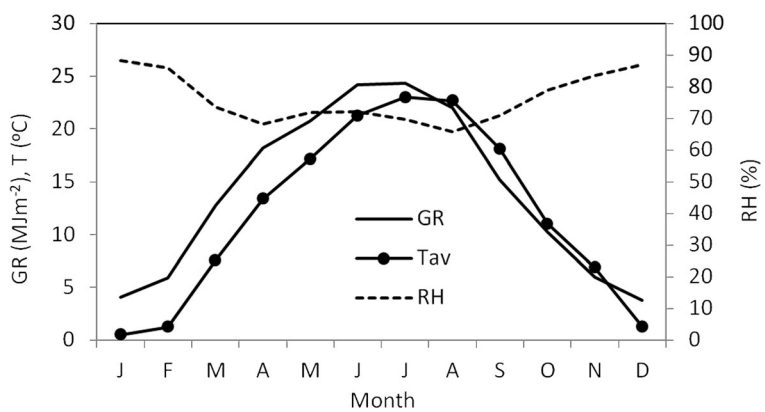

Fig. 3 Monthly averages of global sun radiation (GR), air temperature (T), and relative humidity (RH) in Novi Sad

"Random Sampling method." The training dataset are samples used to create the model, while the test and validation datasets are used to qualify performance. The validation set is used to evaluate the model during the training process, while the test set is used to evaluate the final trained model. For the network learning process, the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm was used. The error function used to find the best solution was the sum of the squares of errors (SOS). The selection of the optimal model is made on the basis of correlation coefficients of validation and test subsets.

Performance indicators

Evaluation of the model performances was done using the following statistical indicators: mean bias error (MBE), mean absolute error (MAE), root mean square error (RMSE), correlation coefficient (R), and index of agreement (d).

The MBE, MAE, and RMSE are the most commonly used statistical indicators. The MBE shows overestimation or underestimation of the model and MAE measures the size of the error without considering their direction, while RMSE quantifies the average dispersion.

$$MBE = \frac{1}{N}\sum(P-O)$$
$$MAE = \frac{1}{N}\sum|P-O|$$
$$RMSE = \frac{1}{N}\sqrt{\sum(P-O)^2}$$

The correlation coefficient (R) shows the strength of the connection between measured and modeled data and it is calculated as:

$$R = \frac{\sum\left(O-\overline{O}\right)\left(P-\overline{P}\right)}{\sqrt{\sum\left(O-\overline{O}\right)^2}\sqrt{\left(P-\overline{P}\right)^2}}$$

The Willmott index of agreement (d) is a dimensionless measurement of model accuracy. It represents the ratio of the mean square error and the potential error which varies between 0 and 1.

$$d = 1-\frac{\sum(O-P)^2}{\sum\left(\left|P-\overline{O}\right|+\left|O-\overline{O}\right|\right)^2}$$

The O and P are observed and predicted values and an overbar indicates an arithmetic average, while N is the number of data points.

Results and discussion

Input variables arranged according to their relative importance in the respective model are shown in Table 1. The ratio sensitivity quotients (RSQ) in Table 1 were calculated using SANN as the sums of squares residuals or misclassification rates for the model when the corresponding predictor was eliminated from the neural network (TIBCO Software 2020). According to RSQ, the predictors are sorted by their relevance in a particular model. A higher RSQ value indicates a greater ability of a particular predictor to influence better model performance. The RQS values are 1 and above, which indicate that all parameters are important for the learning process. Generally, the most important predictors are maximum air temperature on the forecast day, $T_{max}(t+1)$; O_3 concentration on the day preceding the forecast, $1hO_3(t)$ or $8hO_3(t)$; and the global radiation on forecast day, $GR(t+1)$. The positive relationship between ozone and global radiation is the result of the generation OH radicals from the photolysis of ozone at wavelengths < 319 nm that leads to cycles of reactions that result in the photochemical degradation of organic compounds and the enhanced formation of ozone (WMO 1986). Another reason for the general increase in O_3 concentrations in the troposphere due to increased temperature and increased short-wave radiation is increasing emissions of biogenic isoprene which reacts with OH and producing peroxy radicals. In the presence of NO, these radicals can react to form NO_2, which is then photolyzed to produce O_3 precursor, O(3P)(Han et al. 2005; Watson et al. 2006).

To achieve the best forecasting performance, hundreds of networks have been tested. The results of the performance indicators for the training, testing, and validation datasets data are given in Table 2. Examination of combinations of activation functions showed that the exponential functions in the hidden layer give the best results. The MBE is negative and small (up to −2.98 μg m^{-3}) and indicates that all models have a tendency to slightly underestimate O_3 values. The RMSE varies from 10.278 to 15.553 μg m^{-3} and is higher than MAE for all models, while R and d are relatively high at both sites. The obtained indicators

Table 1 Input parameters for the final models ranked according to ratio sensitivity quotients (RSQ), where $t+1$ means values on the forecast day, while t presents values on the day preceding the forecast

Rank	$1hO_3$				$8hO_3$			
	Dnevnik		Liman		Dnevnik		Liman	
	Input variable	RSQ	Input variable	RSQ	Input variable	RSQ	Input variable	RSQ
1	$T_{max}(t+1)$	2.67	$T_{max}(t+1)$	3.25	$8hO_3(t)$	3.48	$T_{max}(t+1)$	3.85
2	$1hO_3(t)$	2.64	GR($t+1$)	2.29	GR($t+1$)	1.79	$8hO_3(t)$	2.14
3	$T_{av}(t+1)$	1.44	$1hO_3(t)$	1.69	GR(t)	1.46	GR($t+1$)	1.94
4	GR($t+1$)	1.43	$T_{max}(t)$	1.39	RH($t+1$)	1.27	$T_{max}(t)$	1.57
5	GR(t)	1.21	RH($t+1$)	1.19	$T_{max}(t+1)$	1.21	RH($t+1$)	1.35
6	RH($t+1$)	1.20	$T_{av}(t+1)$	1.14	$T_{av}(t+1)$	1.08	$T_{av}(t+1)$	1.34
7	$P(t)$	1.05	$T_{av}(t)$	1.07	$T_{av}(t)$	1.05	DOY($t+1$)	1.07
8	$T_{max}(t)$	1.04	$P(t+1)$	1.07	$P(t+1)$	1.04	WS($t+1$)	1.07
9	DOY($t+1$)	1.04	WS($t+1$)	1.06	WS($t+1$)	1.03	$P(t+1)$	1.06
10	WS($t+1$)	1.03	$P(t)$	1.03	$P(t)$	1.03	GR(t)	1.05
11	$P(t+1)$	1.03	DOY($t+1$)	1.03	RH(t)	1.02	RH(t)	1.03
12	RH(t)	1.03	DOW($t+1$)	1.02	DOY($t+1$)	1.02	$T_{av}(t)$	1.03
13	$T_{av}(t)$	1.01	WS (t)	1.01	DOW($t+1$)	1.02	$P(t)$	1.02
14	DOW($t+1$)	1.01	RH(t)	1.01	$T_{max}(t)$	1.00	DOW($t+1$)	1.01
15	WS (t)	1.00	GR(t)	1.01	WS (t)	1.00	WS (t)	1.01

show that prediction models for Liman station have lower error values regarding MBE, MAE, and RMSE, and higher dimensionless measures (R and d) than prediction models for Dnevnik station. Also, the prediction of $8hO_3$ is slightly better than the prediction of $1hO_3$ at both stations. The small differences between the performance indicators for the subsets of training, testing, and validation indicate good generalization performance of the models.

To test the performance of used neural network models, we compared the statistical parameters RMSE and MAE with other neural network models that use meteorological parameters and O_3 values from previous days as input parameters. Chaloulakou et al. (2003) made neural network models to forecast the next day's $1hO_3$ using eight meteorological predictors and $1hO_3$ values for three previous days in the Athens basin at four monitoring stations and reported RMSE between 18.4 and 39.4 μg m^{-3} and MAE between 14.4 and 32.6 μg m^{-3}. Pawlak and Jaroslawski (2019) used neural network models to predict $1hO_3$ for the following day in rural and urban locations in central Poland, using six meteorological parameters and $1hO_3$ recorded the previous day. The values of RMSE found by Pawlak and Jaroslawski (2019) ranged from 9.9 to 16.3 μg m^{-3}, while MAE ranged from 8.6 to 13.0 μg m^{-3}. When we compared MBE and RMSE by our models, we found better results than Chaloulakou et al. (2003), and similar results with Pawlak and Jaroslawski (2019). We also compared our models with models where measured concentrations of other air pollutants were used, as an input, besides meteorological parameters and O_3 concentrations for the previous day. Inal (2010) used, as additional input parameters, measured concentrations of seven air pollutants (SO_2, PM_{10}, CO, NO, NO_2, CH_4, NMHCs) as additional input parameters to forecast $1hO_3$ in Istanbul, Turkey, while Kocijan et al. (2016) additionally used four air pollutants (NO_x, NO_2, SO_2, and CO) to forecast $1hO_3$ and $8hO_3$ at mobile stations in five Slovenian cities. Inal (2010) reported RMSE between 10.9 and 11.6 μg m^{-3}, MAE between 8.1 and 9.3 μg m^{-3}, R between 0.89 and 0.90, and d between 0.94 and 0.95. The values of RMSE found by Kocijan et al. (2016) ranged from 14.6 to 17.4 μg m^{-3} for $1hO_3$, and from 13.4 to 15.9 μg m^{-3} for $8hO_3$. When we compared the RMSE and MBE by our models to the RMSE and MBE given by Inal (2010), we found slightly lower but comparable results, and similar R and d.

Table 2 Performance measures at two air quality monitoring stations for predictions of the 1-h daily maximum and the 8-h daily maximum concentrations

		Dnevnik				Liman			
		Training	Testing	Validation	All	Training	Testing	Validation	All
1-h daily maximum concentration ($1hO_3$)									
Performance indicator	MBE	0.000	− 0.004	0.005	− 0.289	0.116	− 0.383	− 0.565	− 0.060
	MAE	12.114	11.946	12.149	12.094	9.122	9.299	10.101	9.295
	RMSE	15.535	15.254	15.926	15.553	12.163	11.716	12.962	12.221
	R	0.889	0.898	0.899	0.891	0.930	0.936	0.911	0.981
	d	0.938	0.937	0.940	0.939	0.962	0.966	0.953	0.962
Architecture		15-29-1				15-4-1			
Activation function	Hidden	Exponential				Exponential			
	Output	Tanh				Logistic			
8-h daily maximum concentration ($8hO_3$)									
Performance indicator	MBE	0.021	− 0.442	− 1.126	− 0.218	0.000	− 0.005	− 0.001	− 0.152
	MAE	10.495	10.915	10.614	10.575	8.026	8.296	8.574	8.148
	RMSE	12.163	11.716	12.962	13.445	10.201	10.158	10.741	10.278
	R	0.900	0.900	0.910	0.904	0.946	0.951	0.936	0.993
	d	0.947	0.941	0.948	0.948	0.971	0.973	0.966	0.971
Architecture		15-26-1				15-7-1			
Activation function	Hidden	Exponential				Exponential			
	Output	Identity				Logistic			

When we compared the RMSE by our models to the RMSE reported by Kocijan et al. (2016), we found better results. Overall, the comparative results discussed above show that the forecast quality of our models at both sites is in accordance with those reported in the similar studies.

We also compared our models with other non-neural network models. Moustris et al. (2012) applied a multiple linear regression model (MLR) against an ANN to forecast $1hO_3$ of the forthcoming 24 h in Athens, Greece, using meteorological parameters and observed ozone levels of the previous day. The authors concluded that ANN has limited precedence against MLR. Performance statistics, *R*, MBE, RMSE, and *d*, for developed MLR were 0.8, − 1.6 μg m^{-3}, 25.5 μg m^{-3}, and 0.89, respectively. Salazar-Ruiz et al. (2008) applied three non-neural network models and three non-parametric ANN models for forecasting $1hO_3$ in Mexico and the USA, using meteorological parameters, and values of several air pollutants for the previous day (O_3, NO_x, NO_2, NO, and CO_2). The authors also concluded that the results of non-parametric models are better than those obtained with parametric and semiparametric techniques. Salazar-Ruiz et al. (2008) reported RMSE between 16.0 and 19.6 μg m^{-3}, MAE between 11.3 and 13.1 μg m^{-3}, *R* between 0.52 and 0.59, and *d* between 0.71 and 0.72, for non-neural-network models. When we compared statistical parameters of our models with parameters of non-neural network models, we found significantly better results, which confirm the advantages of neural network. One of the main advantages is that ANN enables the modeling of non-linear complex relationships without the need for mathematical representations. However, neural ANN requires a large amount of input parameter data to cover a wide range of results' prediction possibilities.

Scatter plots of measured maximum daily O_3 concentrations against predicted concentrations are given in Fig. 4 for the training and testing datasets, while the courses of measured and predicted values during the summer season are given in Fig. 5. The time series of measured versus predicted concentrations show that the predicted values are in a good agreement with the observed ones. According to Fig. 4, the measured daily maximums, $1hO_3$ and $8hO_3$, are higher at the

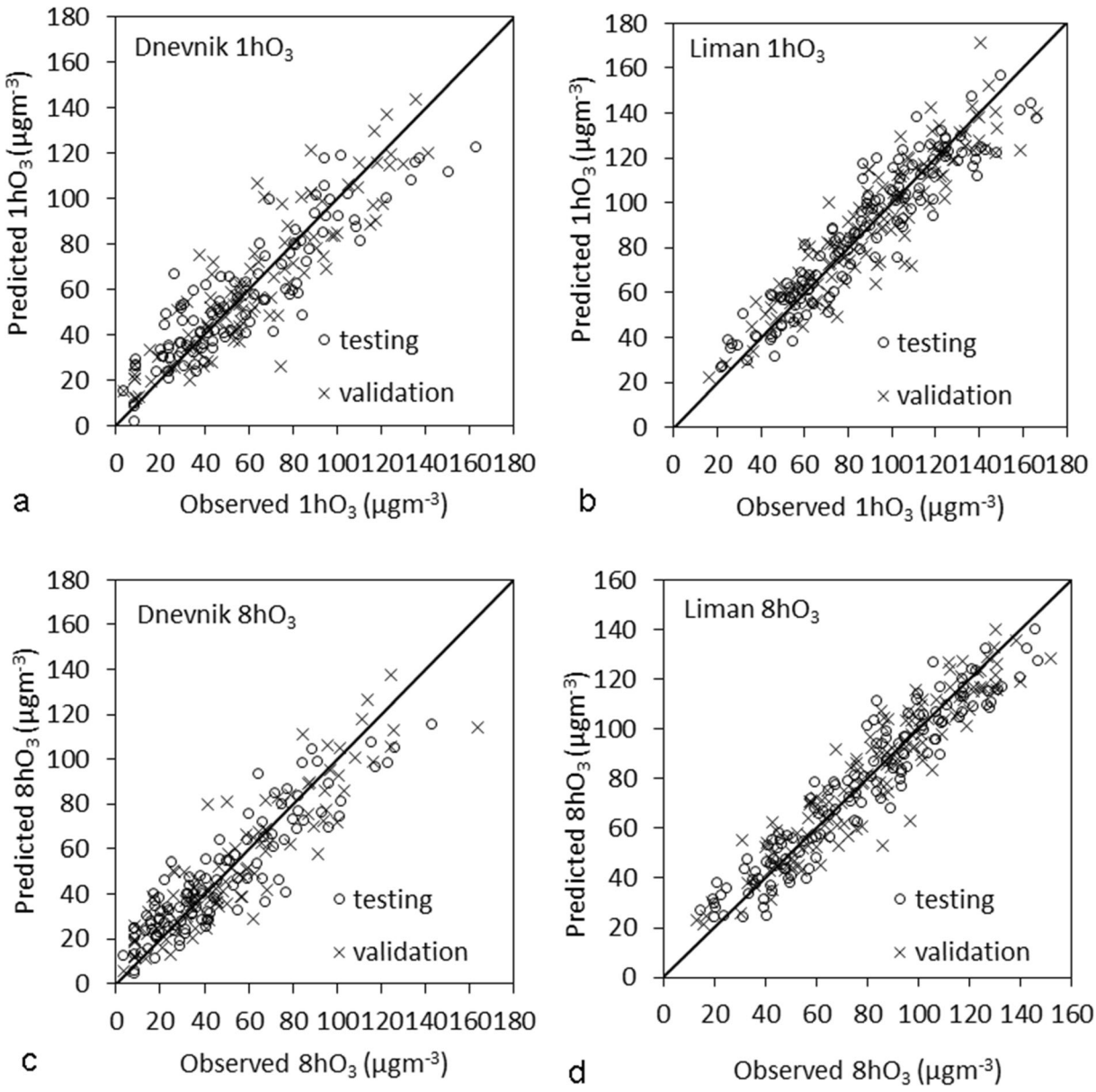

Fig. 4 Scatter plots of observed versus predicted O_3 levels. **a** Dnevnik $1hO_3$. **b** Liman $1hO_3$. **c** Dnevnik $8hO_3$. **d** Liman $8hO_3$

background station Liman (averages are 89.1 μg m^{-3} and 79.6 μg m^{-3}, respectively) than at the traffic station Dnevnik (averages are 65.2 μg m^{-3} and 54.4 μg m^{-3}, respectively). This situation is a result of the effect of intense local conversion (NO into NO_2 by O_3) in areas with high concentrations of freshly emitted NO, such as areas with heavy traffic (Pison and Menut 2004; Liu et al. 2007) The average values of forecasted O_3 concentrations at both stations are similar to measured averages, with differences lower than 0.3 μg m^{-3}.

To determine if the differences between observed and predicted data are statistically significant, we performed the D'Agostino normality test which describes normality in a way that combines the test for skewness and kurtosis. The analysis of the data given in Table 3 shows that the distribution of data at both stations is not normal because the p value of the test is lower than the significance of $\alpha = 0.05$. Therefore, to determine if there is a statistically significant difference between observed and predicted O_3 values, we used non-parametric two-

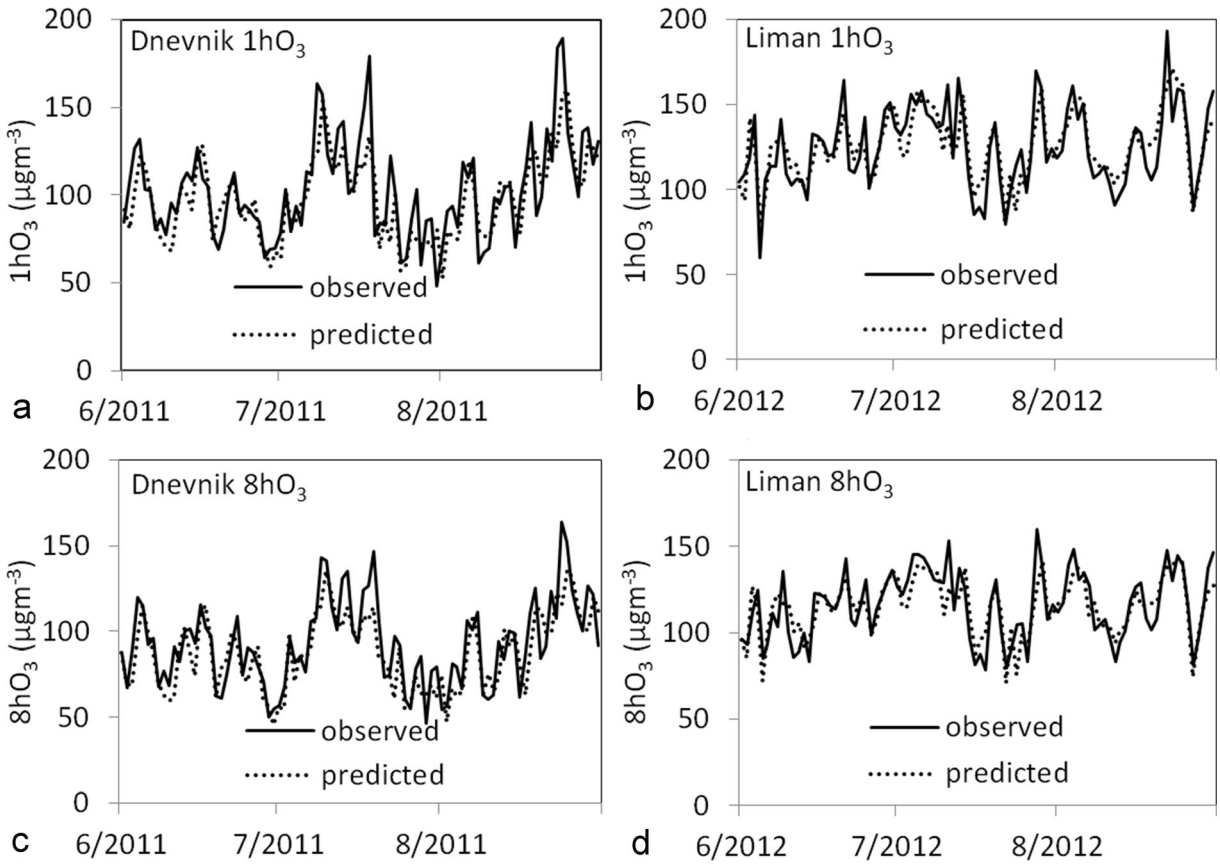

Fig. 5 The courses of measured and predicted ozone concentration values for the summer season. **a** Dnevnik $1hO_3$. **b** Liman $1hO_3$. **c** Dnevnik $8hO_3$. **d** Liman $8hO_3$

sample Kolmogorov–Smirnov test with significance level 0.05. The statistical comparison between observed and predicted data presented in Table 3 reveals that the difference between measured and predicted data is not significant.

Air pollution models commonly have fairly good ability to predict middle-range values of O_3 concentrations. However, many of them have a tendency to overestimate low values and underestimate high values (Comrie 1997; Niska et al. 2004). Therefore, we decided to examine the performance of our models to predict days with values over a certain threshold, modeled on the manuscript by Tsai et al. (2009). We have defined thresholds based on WHO guidelines and recommendations, which aim to provide guidelines for reducing adverse health impacts. The value of 100 µg m^{-3} was used for $8hO_3$ threshold, while the value of 90 µg m^{-3} was used for $1hO_3$ threshold. We classified all days into those in which the threshold was exceeded and those in which there was no exceedance, and then, we calculated the ability of the model to predict whether the threshold will be exceeded. Results are presented in Table 4. It can be seen that all four models have high accuracy to classify days on those with threshold exceedances and those with non-exceedances. When it comes to the ability of a model to predict days with threshold exceedance, three out of the four models can

Table 3 Results of the D'Agostino normality test and the two-sample Kolmogorov–Smirnov test

		D'Agostino normality test		Kolmogorov–Smirnov test
		Observed *p* value	Predicted *p* value	Observed vs predicted *p* value
$1hO_3$	Dnevnik	0.000	0.000	0.422
	Liman	0.002	0.000	0.814
$8hO_3$	Dnevnik	0.000	0.000	0.164
	Liman	0.000	0.000	0.582

Table 4 Statistics of the models' performance in prediction of threshold exceedances

		Observed exceedances	Predicted exceedances	Correctly predicted exceedances	Correctly predicted non-exceedances	False alarms	Accuracy
$1hO_3$	Dnevnik	176 (25.0%)	166 (23.6%)	133 (75.5%)	495 (93.7%)	33 (19.9%)	89.2%
	Liman	460 (50.1%)	453 (50.2%)	408 (88.7%)	397 (89.8%)	45 (9.9%)	88.9%
$8hO_3$	Dnevnik	70 (9.9%)	49 (6.9%)	36 (51.4%)	621 (97.9%)	13 (26.5%)	93.3%
	Liman	258 (28.6%)	270 (29.9%)	223 (86.4%)	598 (92.8%)	12 (4.4%)	91.0%

accurately forecast more than 75% cases. The lowest ability to predict exceedance has a model for $8hO_3$ at station Dnevnik (51.4%). The relatively small number of days with exceedances of the O_3 threshold at this station (only 9.9%) is the reason for the lower accuracy because rarer events are more difficult to predict.

Conclusions

This work demonstrates the generation and use of neural network models for predicting one day ahead of daily maximum 1-h O_3 concentration and daily maximum 8-h average ozone concentrations at the two monitoring sites in the city of Novi Sad, Serbia. Neural network models were generated using historical meteorological parameters and O_3 concentrations for the day before forecasting day and predicted values of meteorological parameters for the forecast day. This study is based on observational data from two urban stations (traffic and background) in a period of five years.

The analysis revealed that the most important predictors for O_3 forecasting are air temperature and global radiation on the forecast day, as well as O_3 concentration on the day preceding the forecast. The modeled data and actual observations were found to be consistent in both traffic and background stations. The statistical test showed that there is no statistically significant difference between measured and predicted data. The results showed that the quality of our models' comparable with the quality of models of similar studies around the world. The forecasts of days with values over a certain threshold revealed very good statistical consistency with the measured data.

This investigation shows that the feedforward neural network approach is well-suited for modeling daily maximum 1-h O_3 concentration neural and daily maximum 8-h average O_3 concentration at traffic and background urban stations in Novi Sad, Serbia. Future work in this area should focus on examining the merit of including O_3 precursors in modeling and increasing the time frame for the forecast to 48 h ahead.

Data Availability Data and material are available at Serbian National Air Monitoring network and Republic Hydrometeorological service of Serbia

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interest.

Code availability The software Statistica Automated Neural Networks (SANN) (TIBCO Software 2020) is applied.

References

Abdul-Wahab, S. A., Bakheit, C. S., & Al-Alawi, S. M. (2005). Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. *Enviromental Modelling and Software, 20*, 1263–1271. https://doi.org/10.1016/j.envsoft.2004.09.001.

Altshuller, S. I., Arcado, T. D., & Lawson, D. R. (1995). Weekday versus weekend ambient ozone concentrations: discussion and hypothesis with focus on northern California. *Journal of the Air and Waste Management Association, 45*, 161–180. https://doi.org/10.1080/10473289.1995.10467428.

Bates, V. D. . (2005). Ambient ozone and mortality. *Epidemiology, 16*, 427–429. https://doi.org/10.1097/01.ede.0000165793.71278.ec.

Bell, M., Peng, R., & Dominici, F. (2006). The exposure-response curve for ozone and risk of mortality and the adequacy of current ozone regulations. *Environtal Health Perspective, 114*, 532–536. https://doi.org/10.1289/ehp.8816.

Biancofiore, F., Verdecchia, M., Di Carlo, P., Tomassetti, B., Aruffo, E., Busilacchio, M., Bianco, S., Di Tommaso, S., & Colangeli, C. (2015). Analysis of surface ozone using a recurrent neural network. *Science of Total Environment, 514*, 379–387. https://doi.org/10.1016/j.scitotenv.2015.01.106.

Cape, J. N. (2008). Surface ozone concentrations and ecosystem health: past trends and a guide to future projections. *Science*

of Total Environment, 400, 257–269. https://doi.org/10.1016/j.scitotenv.2008.06.025.

Chaloulakou, A., Saisana, M., & Spyrellis, N. (2003). Comparative assessment of neural networks and regression models for forecasting summertime ozone in Athens. *Science of Total Environment, 313*, 1–13. https://doi.org/10.1016/S0048-9697(03)00335-8.

Comrie, A. C. (1997). Comparing neural networks and regression models for ozone forecasting. *Journal of Air and Waste Management Association, 47*, 653–663. https://doi.org/10.1080/10473289.1997.10463925.

Coulson, K. L. (1975). *Solar and terrestrial radiation*. New York, NY: Academic Press.

Duenas, C., Fernandez, M., Canete, S., Carretero, J., & Liger, E. (2005). Stochastic model to forecast ground-level ozone concentration at urban and rural areas. *Chemosphere, 61*, 1379–1389. https://doi.org/10.1016/j.chemosphere.2005.04.079.

EPA (2009). *Provisional assessment of recent studies on health and ecological effects on ozone exposure*. U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-09/101.

Ghazali, N. A., Ramli, N. A., Yahaya, A. S., Yusof, N. F., Sansuddin, N., & Madhoun, W. A. (2010). Transformation of nitrogen dioxide into ozone and prediction of ozone concentrations using multiple linear regression techniques. *Environmental Monitoring and Assessment, 165*, 475–489. https://doi.org/10.1007/s10661-009-0960-3.

Gryparis, A., Forsberg, B., Katsouyanni, K., Analitis, A., Touloumi, G., Schwartz, J., Samoli, E., Medina, S., et al. (2004). Acute effects of ozone on mortality from the "Air pollution and health: a European approach" project. *American Journal of Respiratory and Critical Care Medicine, 170*, 1080–1087. https://doi.org/10.1164/rccm.200403-333OC.

Han, Z. W., Ueda, H., & Matsuda, K. (2005). Model study of the impact of biogenic emission on regional ozone and the effectiveness of emission reduction scenarios over eastern China. *Tellus: Series B, Chemical and Physical Meteorology, 57*(*1*), 12–27. https://doi.org/10.1111/j.1600-0889.2005.00132.x.

Hecht-Nielsen, R. (1987). Kolmogorov's mapping neural network existence theorem. In *Proceedings of the 1st IEEE International Joint Conference on Neural Networks* (pp. 11–14). New York: IEEE Press.

Inal, F. (2010). Artificial neural network prediction of tropospheric ozone concentrations in Istanbul, Turkey. *Clean–Soil Air Water, 38*(*10*), 897–908. https://doi.org/10.1002/clen.201000138.

Ito, K., De Leon, S. F., & Lippmann, M. (2005). Associations between ozone and daily mortality: analysis and meta-analysis. *Epidemiology, 16*(*4*), 446–457. https://doi.org/10.1097/01.ede.0000165821.90114.7f.

Klein, W. H., Lewis, B. M., & Enger, I. (1959). Objective prediction of five-day mean temperature during winter. *Journal of Applied Meteorology, 16*, 672–682.

Kocijan, J., Gradišar, D., Božnar, M. Z., Grašič, B., & Mlakar, P. (2016). On-line algorithm for ground-level ozone prediction with a mobile station. *Atmospheric Environment, 131*, 326–333. https://doi.org/10.1016/j.atmosenv.2016.02.012.

Kukkonen, J., Partanen, L., Karppinen, A., Ruuskanen, J., Junninen, H., Kolehmainen, M., Niska, H., Dorling, S., Chatterton, T., Foxall, R., & Cawley, G. (2003). Extensive evaluation of neural network models for the prediction of NO2 and PM10 concentrations, compared with a deterministic modelling system and measurements in central Helsinki. *Atmospheric Environment, 37*, 4539–4550. https://doi.org/10.1016/S1352-2310(03)00583-1.

Lee, D. S., Lewis, P. M., Cape, N., Leith, D., & Espenhahn, S. E. (2003). The effects of ozone on materials – experimental evaluation of the susceptibility of polymeric materials to ozone. In P. Brimblecombe (Ed.), *The effects of air pollution on the built environment* (pp. 267–287). London: Imperial College Press. https://doi.org/10.1142/9781848161283_0009.

Lelieveld, J., & Crutzen, P. J. (1990). Influence of cloud and photochemical processes on tropospheric ozone. *Nature, 343*, 27–233. https://doi.org/10.1038/343227a0.

Liu, L., Andreani-Aksoyoglu, S., Keller, J., Braathen, G. O., Schultz, M., Prevot, A., & Isaksen, I. (2007). *A photochemical modelling study of ozone and formaldehyde generation and budget in the Po Basin*. Limassol, Cyprus: Proceedings of the 6th International Conference on Urban Air Quality.

Logan, J. A. (1985). Tropospheric ozone: seasonal behavior, trends, and anthropogenic influence. *Journal of Geophysical Research, 90*(D6), 10463–10482.

McKendry, I. G. (2002). Evaluation of artificial neural networks for fine particulate pollution (PM_{10} and $PM_{2.5}$) forecasting. *Journal of the Air and Waste Management Association, 52*(*9*), 1096–1101. https://doi.org/10.1080/10473289.2002.10470836.

Moustris, K. P., Nastos, P. T., Larissi, I. K., & Paliatsos, A. G. (2012). Application of multiple linear regression models and artificial neural networks on the surface ozone forecast in the greater Athens Area, Greece. *Advances in Meteorology, 894714*. https://doi.org/10.1155/2012/894714.

Nghiem, L. H., & Kim Oanh, N. T. (2009). Comparative analysis of maximum daily ozone levels in urban areas predicted by different statistical models. *Science Asia, 35*, 276–283. https://doi.org/10.2306/scienceasia1513-1874.2009.35.276.

Niska, H., Hiltunen, T., Karppinen, A., Ruuskanen, J., & Kolehmainen, M. (2004). Evolving the neural network model for forecasting air pollution time series. *Engineering Applications of Artificial Intelligence, 17*, 159–167. https://doi.org/10.1016/j.engappai.2004.02.002.

Official Gazette of the City of Novi Sad, No. 49 (2018). Air quality plan for Novi Sad agglomeration for the period 2017-2021. City of Novi Sad, City Administration for Regulations of the City of Novi Sad. [In Serbian].

Pawlak, I., & Jaroslawski, J. (2019). Forecasting of surface ozone concentration by using artificial neural networks in rural and urban areas in Central Poland. *Atmosphere, 10*, 52. https://doi.org/10.3390/atmos10020052.

Pison, I., & Menut, L. (2004). Quantification of the impact of aircraft traffic emissions on tropospheric ozone over Paris area. *Atmosperic Environment, 38*, 971–983. https://doi.org/10.1016/j.atmosenv.2003.10.056.

Prybutok, V. R., Yi, J., & Mitchell, D. (2000). Comparison of neural network models with ARIMA and regression models for prediction of Houston's daily maximum ozone concentrations. *European Journal of Operational Research, 122*(*1*), 31–40. https://doi.org/10.1016/S0377-2217(99)00069-7.

Rahman, M. M., Shafiullah, M., Rahman, S. M., Khondaker, A. N., Amao, A., & Zahir, M. H. (2020). Soft computing applications in air quality modeling: past, present, and future. *Sustainability, 12*, 4045.

RHMZ (Republic Hydrometeorological service of Serbia) (2019) http://www.hidmet.gov.rs/ciril/meteorologija/klimatologija_godisnjaci. Php. Accessed 20 December 2019.

Salazar-Ruiz, E., Ordieres, J. B., Vergara, E. P., & Capuz-Rizo, S. F. (2008). Development and comparative analysis of tropospheric ozone prediction models using linear and artificial intelligence-based models in Mexicali, Baja California (Mexico) and Calexico, California (US). *Environmental Modelling and Software, 23*(*8*), 1056–1069. https://doi.org/10.1016/j.envsoft.2007.11.009.

SEPA (Serbian Environmental Protection Agency) (2019). www.sepa.gov.rs;www.amskv.sepa.gov.rs. Accessed 20 December 2019 (in Serbian).

Sousa, S. I. V., Martins, F. G., Alvim-Ferraz, M. C. M., & Pereira, M. C. (2007). Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Enviromental Modelling and Software, 22*, 97–103. https://doi.org/10.1016/j.envsoft.2005.12.002.

Statistical Office of the Republic of Serbia. (2019). *Municipalites and regions of the Republic of Serbia 2019*. Belgrade, Serbia: Statistical Office of the Republic of Serbia Retrieved from https://publikacije.stat.gov.rs/G2019/PdfE/G201913046.pdf. Accessed 16 September 2020.

Stevenson, D. S., Dentener, F. J., Schultz, M. G., Ellingsen, K., van Noije, T. P. C., Wild, O., Zeng, G., Amann, M., Atherton, C. S., Bell, N., et al. (2006). Multimodel ensemble simulations of present-day and near-future tropospheric ozone. *Journal of Geophysical Research, 111*, 1–23. https://doi.org/10.1029/2005JD006338.

Sun, W., Zhang, H., & Palazoglu, A. (2013). Prediction of 8 h-average ozone concentration using a supervised hidden Markov model combined with generalized linear models. *Atmospheric Environment, 81*, 199–208. https://doi.org/10.1016/j.atmosenv.2013.09.014.

TIBCO Software (2020) STATISTICA Automated Neural Networks (SANN) https://community.tibco.com/wiki/tibco-statistica-automated-neural-networks. Accessed 09 Februar 2020.

Tsai, C. H., Chang, L. C., & Chiang, H. C. (2009). Forecasting of ozone episode days by cost-sensitive neural network methods. *Science of Total Environment, 407*(6), 2124–2135. https://doi.org/10.1016/j.scitotenv.2008.12.007.

Watson, L. A., Wang, K.-Y., Paul, H., & Shallcross, D. E. (2006). The potential impact of biogenic emissions of isoprene on urban chemistry in the United Kingdom. *Atmospheric Science Letters, 7*, 96–100. https://doi.org/10.1002/asl.140.

WHO. (2006). *Air quality guidelines global update 2005: particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Summary of risk assessment. Copenhagen*. WHO Regional Office for Europe. https://apps.who.int/iris/handle/10665/107823.

WHO. (2013). *Review of evidence on health aspects of air pollution – REVIHAAP Project: Technical Report*. Copenhagen: WHO Regional Office for Europe.

Wilks, D. S. (1995). *Statistical methods in the atmospheric sciences*, *International Geophysics Series* (Vol. 59). Academic Press.

WMO. (1986). *Atmospheric Ozone 1985. Atmospheric Ozone Assessment of Our Understanding of the Processes Controlling its Present Distribution and Change. Report No. 16, Vol. 1. Global Ozone Research and Monitoring Project*. Geneva: World Meteorological Organization.

Zanis, P., Hadjinicolaou, P., Pozzer, A., Tyrlis, E., Dafka, S., Mihalopoulos, N., & Lelieveld, J. (2014). Summertime free-tropospheric ozone pool over the eastern Mediterranean/ Middle East. *Atmospheric Chemistry and Physics, 14*, 115–132. https://doi.org/10.5194/acp-14-115-2014.

Zanobetti, A., & Schwartz, J. (2008). Mortality displacement in the association of ozone with mortality. An analysis of 48 cities in the United States. *American Journal of Respiratory and Critical Care Medicine, 177*, 184–189. https://doi.org/10.1164/rccm.200706-823OC.