Time Series Prediction of Air Pollutants - A Case Study for Serbia, Bosnia and Herzegovina and Italy

Conference Paper · March 2019

DOI: 10.1109/INFOTEH.2019.8717778

3 authors, including:

Radmila Jankovic Babic
Mathematical Institute of the Serbian Academy of Sciences and Arts

Alessia Amelio
Università degli Studi G. d'Annunzio Chieti e Pescara

18th International Symposium INFOTEH-JAHORINA, 20-22 March 2019

Time Series Prediction of Air Pollutants

A Case Study for Serbia, Bosnia and Herzegovina and Italy

Radmila Janković
Mathematical Institute of
Serbian Academy of Sciences and Arts
Belgrade, Serbia
rjankovic@mi.sanu.ac.rs

Marijana Ćosović
University of East Sarajevo
Faculty of Electrical Engineering
Istočno Sarajevo, Bosnia and Herzegovina
marijana.cosovic@etf.ues.rs.ba

Alessia Amelio
DIMES, University of Calabria
Rende (CS), Italy
aamelio@dimes.unical.it

Abstract—**Pollution levels are highly dependent on the meteorological parameters, as the weather conditions dictate pollution dispersion and concentration. With the rise of global environmental protection initiatives, there is also a need for accurate prediction of pollution levels. This paper presents a time series prediction of NO_2 and CO given four meteorological parameters: (i) air pressure, (ii) relative humidity, (iii) average daily temperature, and (iv) wind speed, using a Nonlinear Autoregressive Exogenous (NARX) neural network. The research is a case study of three European countries: (i) Serbia, (ii) Bosnia and Herzegovina, and (iii) Italy, and involves data from 2014 to 2016 for a total of 1096 instances. The results show that the best prediction accuracy is obtained for CO for data regarding Italy and Bosnia and Herzegovina, and for NO_2 for data regarding Serbia. Moreover, the best predictor variables of NO_2 are air pressure and relative humidity, followed by the wind speed. The best predictor variables of CO are pressure and temperature for Bosnia and Italy, and wind speed for Serbia.**

Keywords-time series; air pollution; prediction; artificial neural network; data mining

I. INTRODUCTION

Outdoor or ambient air varies in its quality depending on the amount of the present air pollutants. Within every country there exists an environmental protection agency that performs air quality monitoring, implements programs for air quality control, informs the public about the air quality, assesses and categorizes the air quality based on results of available measurements. These tasks are done in accordance with limit and tolerance values of the USA and EU standards regarding ambient air pollutants.

It is well known fact that the air pollution has negative effects not only on humans and their health but also on ecosystem and the built environment. World Health Organization issues notifications regularly about deaths caused by the bad air quality. The normal functioning of the ecosystems, defined as non-obstructed functioning and growth, is affected by air pollution especially by sulphur and nitrogen emissions. Immovable cultural heritage such as historical/religious buildings as well as movable cultural heritage are also affected by air pollution. Besides ancient building structures contemporary buildings are subject to negative effects of the air pollution as well. Cultural heritage preservation is a very active research domain in which scientists and professionals address the air pollution and devise methods for minimizing its effects.

In general, most of the pollution comes from either man-made or natural sources (see Fig.1). Emissions from the industry, traffic or home coal-fired combustion systems are categorized as man-made sources while forest fires are categorized as natural sources. Serbia, and Bosnia and Herzegovina unlike Italy record an increase in concentrations of pollutants with the beginning of the heating season.

Air pollutants considered in this study are NO_2 and CO. In addition, we used atmospheric pressure, average temperature, relative humidity and wind speed. In this case study we consider data from three different countries, namely Serbia, Bosnia and Herzegovina and Italy. We perform time series prediction of NO_2 and CO future values given their past values and past values of one of the following meteorological variables: temperature, pressure, relative humidity, or wind speed.

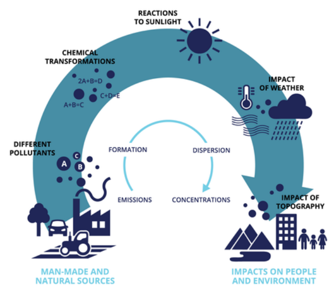

Figure 1. The emissions of air pollutants [1]

II. RELATED WORK

Pollutant prediction is one of the most important tasks of air pollution research. Carbon monoxide is a toxic air pollutant,

978-1-5386-7073-6/19/$31.00 ©2019 IEEE

byproduct of an incomplete combustion of carbon-based fuels. The most common source of CO pollution are the vehicle emissions. Oxides of nitrogen (NO – nitric oxide and NO_2 – nitrogen dioxide) are produced in the air as a reaction of nitrogen and oxygen gases during the combustion. As the temperature of combustion increases so does the nitrogen oxide. It is further involved in formation of additional pollutants: particulate matter and ground level ozone.

Machine learning techniques are being used extensively for forecasting of various pollutants in the research domain. In [2] the authors use two deep learning forecasting models: Multilayer perceptron (MLP) feed forward artificial neural network (ANN) and Long Short-Term memory (LSTM) recurrent neural network in order to forecast the next-day value of particulate matter. The prediction task aimed to detect particulate matter based on hourly averaged values of meteorological data (temperature, relative humidity, pressure, and average wind speed) and air pollutants (PM_{10}, SO_2, NO_2, CO, and O_3) collected in Sarajevo, Bosnia and Herzegovina during 2017. Since the relationship between the variables is not linear, the authors used the Spearman's correlation coefficient amongst selected features and reported the largest correlation coefficients between PM_{10}, CO, NO_2, and SO_2, confirming the fact that these air pollutants are originating from the same sources. The authors obtained the best prediction results with one-day prior information, hence the performance measure values were the smallest in that case.

A research by [3] investigates a short-term prediction of NO_2 and NO_X air pollutants using MLP ANNs. The prediction task aims to detect NO_2 and NO_X based on hourly averaged values of meteorological data (temperature, relative humidity, pressure, and average wind speed) and air pollutants (NO_2 and NO_X) collected in Tabriz, Iran during two winter months of 2012. The authors determine the network architecture by trial and error and concluded that a 30 neuron ANN provides the best performance. In addition, the authors concluded that modeling the relationship between meteorological data and air pollutants using ANNs in contrast to using linear models is superior.

The authors in [4] use two models for predictions of nitrogen oxides: multiple linear regression and ANNs model and conclude that using ANNs can be applied for air pollution forecasting. The prediction task intends to detect NO_X based on hourly averaged values of meteorological data (temperature, relative humidity, pressure, average wind speed, wind direction, solar radiation, and rainfall) and air pollutants (NO and NO_2) collected in Braila, Romania during a five-year period.

In [5] the authors explore two models for prediction of daily CO levels based on Support Vector Machines (SVM): SVM and Partial Least Square (PLS) SVM model. The prediction task aimed to detect the CO concentration based on hourly averaged values of air pollutants (PM_{10}, SO_2, NO_X, CH_4, O_3, and total hydrocarbons) and meteorological data (temperature, relative humidity, pressure, wind speed, and wind direction) collected in Tehran, Iran during 2007-2011. SVM is used as predictor and PLS as a data selection tool based on the measured CO levels. The authors concluded that both models have good prediction but the latter has better accuracy.

The research community has proven that machine learning techniques can be reliably used as an air pollution time series modeling tool [2]. We propose to use time series prediction by a new neural network-based model for NO_2 and CO air pollutants given their past values and past values of one of the following meteorological variables: temperature, pressure, relative humidity, or wind speed. Data used in this case study has been collected during a period of three years from two urban locations (Serbia and Bosnia and Herzegovina) and one non-urban location (Italy). The obtained prediction results are compared with SVM regression and regression trees. This presents a first step towards the realization of a more detailed study of air pollution prediction in the mentioned countries.

III. DATA AND METHODOLOGY

A. Data

The experiment is performed on three different datasets, one for each selected country: (i) Serbia, (ii) Bosnia and Herzegovina, and (iii) Italy. Each dataset consists of continuous measurements (averaged hourly values) of the temperature (in Celsius degrees), relative humidity (in percentage), pressure (in mbar), and average wind speed (in m/s) in addition to NO_2 and CO air pollutants for three calendar years: 2014-2016. The total number of instances used in each dataset is 1096.

For each dataset, a normalization of its features was performed in order to adjust the values in the range [0-1] using Weka v.3-8-3 and its Normalize filter.

1) Data for Bosnia and Herzegovina

Air pollutants' data for Bosnia and Herzegovina are collected by the Federal Hydrometeorological Institute (FHMZ) BiH in the city of Zenica, located in the city center, about 15 meters above the street level, managed by the municipality of Zenica.

The dataset regarding the air pollutants (NO_2 and CO) was extracted from the FHMZ air quality reports located at the following link: http://www.fhmzbih.gov.ba/latinica/ZRAK/izvjestaji.php. The dataset regarding the meteorological data was released upon an official request to the FHMZ.

2) Data for Italy

Air pollutants' data for Italy are collected by the Italian Climate Observatory "O. Vittori" (ICO-OV) in Mt. Cimone, the highest peak of the Northern Apennines (2,165 mt), managed by the Institute of Atmospheric Sciences and Climate (ISAC) of the National Research Council of Italy (CNR). Collected data is available at the following link: http://www.isac.cnr.it/cimone/data-access.

Before starting the prediction task, all missing values for each dataset feature were substituted with the mean of the distribution of the remaining values. Managing of missing values was performed in Weka v.3-8-3 using the unsupervised filter for attributes removeMissingValues.

3) Data for Serbia

The data regarding the pollutants for Serbia are collected by the Serbian Environmental Protection Agency (SEPA) in Belgrade (Vracar) at an altitude of 141 m.

The meteorological data are collected by the Republic Hydrometeorological Service of Serbia at Belgrade Vračar measuring station.

Before the analysis, the missing values were analyzed in SPSS v.25 using the Little's MCAR test. The null hypothesis for this test is that the data are missing completely at random. The results of the analysis show that the data are not missing completely at random (p=0.000), so the multiple imputation method, in particular Fully Conditional Specification model, was used for imputing missing data. The Fully Conditional Specification model is a Markov Chain Monte Carlo method that imputes missing values in iterations starting from the first variable with missing data until the maximum number of iterations is attained, hence creating multiple imputations. The same process was performed on the dataset for Bosnia and Herzegovina.

B. Methodology

The aim of the analysis is the prediction of future values of air pollutants (NO_2 and CO) based on the past values of the same variables and another variable selected among: (i) relative humidity, (ii) pressure, (iii) temperature, and (iv) wind speed. Since the measurements are daily-level, this task corresponds to prediction of time series. This study allows to capture the relationship between external features and air pollutants and, consequently, to understand which factor mostly influences the air pollution.

Formally, let y(t) be the target time series (NO_2 or CO) that should be predicted, and x(t) be the other time series which is involved in the prediction (relative humidity, pressure, temperature or wind speed). This type of prediction is called nonlinear autoregressive with exogenous (external) input, or NARX, and can be expressed as follows:

$$y(t) = f(y(t-1),\ldots\ldots, y(t-d), x(t-1),\ldots\ldots, x(t-d)) \qquad (1)$$

The NARX model usually provides accurate prediction results for y(t) due to the past values which are considered in the model. The proposed model is realized using an ad-hoc ANN which is the NARX network [6]. It is a two-layer feedforward network, where the hidden layer is composed of a sigmoid transfer function, while the output layer is characterized by a linear transfer function. This network also includes tapped delay lines to save the past values of the x(t) and y(t) sequences. It is worth noting that, since y(t) depends on y(t–1), y(t–2), ..., y(t–d), the output of the NARX network, y(t), is sent back to the network input (through delays). For making the training more efficient, the feedback loop is opened in the network. In particular, due to availability of the true output during the network training, an open-loop architecture is adopted, where the estimated output is not sent back, but the true output is used. This provides a more accurate input for the network. Also, since the network has no loops, a more efficient method for training of feedforward networks can be adopted.

Fig. 2 shows a sample of the adopted NARX network architecture with 5 neurons in the hidden layer and a number of delays *d*=2.

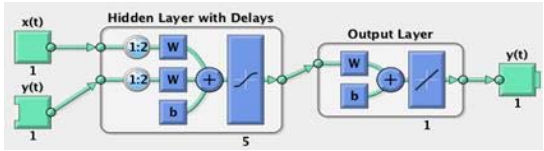

Figure 2. The adopted NARX network architecture

In the experiment, the number of neurons in the hidden layer is varied between 5 and 20 with intervals of 5, in order to select the network achieving the best performances. Also, the number of delays *d* is set to 2, since it provided the most accurate results.

The NARX model is learned from each of the three datasets for Serbia, Bosnia and Herzegovina and Italy. To build the model, one dataset is randomly divided into training, validation and test set. In the experiment, 70% of data is used for training, 15% for validation and 15% for test.

The proposed NARX model is compared with other two well-known methods for time series prediction: (i) SVM for regression [7], and (ii) regression trees [8]. Both the methods are available in the Weka package for time series forecasting as SMOReg and REPTree, respectively. Also, the SVM uses a polynomial kernel of degree 2, allowing a nonlinear regression modeling.

IV. TIME SERIES ANALYSIS

The NARX model for time series prediction has been designed in Matlab R2017a. All experiment has been run on a laptop computer with Quad-Core CPU at 2.2 GHz, 16 GB RAM and Unix operating system.

A. Performance measures

For the evaluation of the prediction task, two performance measures are adopted: (i) Mean Squared Error (MSE), and (ii) Pearson's correlation coefficient (R).

The MSE measures the squared average difference between the target time series y(t) and the estimated time series ŷ(t). Specifically, it is defined as follows:

$$MSE = \frac{\sum_{t=1}^{n} (y(t) - \hat{y}(t))^2}{n}, \qquad (2)$$

where n represents the number of considered time points.

The R coefficient measures the linear relationship between the target time series y(t) and the estimated time series ŷ(t). If the R coefficient is higher than 0, there is a positive correlation between y(t) and ŷ(t). Otherwise, if R is less than 0, the correlation between y(t) and ŷ(t) is negative. Finally, if R is equal to 0, then y(t) and ŷ(t) are not correlated. When R is equal to 1 or -1, the correlation between y(t) and ŷ(t) is perfectly positive or negative.

B. Results

1) Results for Bosnia and Herzegovina

Tables I and II report the MSE and R values obtained from the prediction of NO_2 and CO, respectively (target time series y(t)) based on: (i) pressure (P), (ii) relative humidity (Rh), (iii) temperature (T), or (iv) wind speed (Ws) (other time series x(t)) in the test set. The performance measures are generated considering a different number of neurons in the hidden layer from 5 to 20 in the NARX model.

TABLE I. *MSE* AND *R* COEFFICIENT OBTAINED WHEN PREDICTING NO_2 OF THE TEST SET WITH A NUMBER OF NEURONS IN THE HIDDEN LAYER FROM 5 TO 20. THE BEST VALUES FOR EACH FEATURE ARE BOLD MARKED

	5 neurons		**10 neurons**		**15 neurons**		**20 neurons**	
NO_2	*MSE*	*R*	*MSE*	*R*	*MSE*	*R*	*MSE*	*R*
P	0.009	0.761	**0.007**	**0.837**	0.009	0.836	0.010	0.790
Rh	0.019	0.716	0.017	0.772	**0.017**	**0.774**	0.019	0.709
T	0.019	0.713	0.018	0.758	**0.018**	**0.773**	0.018	0.713
Ws	0.020	0.430	0.019	0.379	**0.007**	**0.767**	0.019	0.470

TABLE II. *MSE* AND *R* COEFFICIENT OBTAINED WHEN PREDICTING CO OF THE TEST SET WITH A NUMBER OF NEURONS IN THE HIDDEN LAYER FROM 5 TO 20. THE BEST VALUES FOR EACH FEATURE ARE BOLD MARKED

	5 neurons		**10 neurons**		**15 neurons**		**20 neurons**	
CO	*MSE*	*R*	*MSE*	*R*	*MSE*	*R*	*MSE*	*R*
P	**0.003**	**0.820**	0.005	0.684	0.005	0.841	0.005	0.718
Rh	0.004	0.764	0.004	0.837	0.005	0.749	**0.004**	**0.843**
T	0.003	0.818	0.006	0.737	**0.003**	**0.849**	0.005	0.794
Ws	0.005	0.779	0.006	0.783	0.004	0.727	**0.004**	**0.773**

It is worth noting that the best prediction accuracy of NO_2 given pressure and relative humidity is obtained with 10 and 15 neurons, respectively, in the hidden layer of the network. By contrast, the best performances are obtained with 15 neurons for temperature and the wind speed. In these network configurations, we can observe that all variables are helpful for predicting the future trend of NO_2, since the MSE is very low and the R coefficient is higher than 0.75, which is a promising result. In particular, the best predictor variables in terms of MSE are the pressure and wind speed, with a value of 0.007, while in terms of R it is the pressure, with a value of 0.837.

By contrast, the best prediction accuracy of CO is obtained with 5 hidden neurons for pressure, 20 neurons for relative humidity and wind speed, and 15 neurons for temperature. In these cases, we can observe that the prediction is more accurate than for NO_2, with lower MSE below 0.004. Also, a promising result is obtained for all predictor variables in terms of MSE and R. In particular, the best predictor variables, considering the combination of MSE and R values, are the pressure (MSE = 0.003 and R=0.820) and the temperature (MSE = 0.003 and R=0.849).

For the best number of hidden neurons (bold marked in Tables I-II), Fig. 3 shows the comparison results in terms of MSE for predicting NO_2 and CO given: (i) pressure, (ii) relative humidity, (iii) temperature, and (iv) wind speed between NARX, SMOReg and REPTree. It is worth noting that the NARX model achieves the lowest MSE for all predictor variables when CO is predicted (all cases). Also, it overcomes the competing methods when relative humidity and wind speed are used for predicting NO_2 (2 out of 4 cases).

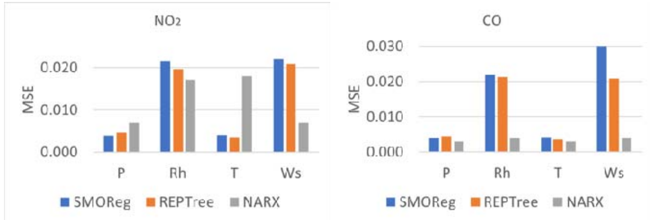

Figure 3. Comparison results in terms of MSE for predicting NO_2 and CO given: (i) pressure, (ii) relative humidity, (iii) temperature, and (iv) wind speed between NARX, SMOReg and REPTree

2) *Results for Italy*

Tables III and IV report the MSE and R values obtained from prediction of NO_2 and CO, as in Tables I and II.

TABLE III. *MSE* AND *R* COEFFICIENT OBTAINED WHEN PREDICTING NO_2 OF THE TEST SET WITH A NUMBER OF NEURONS IN THE HIDDEN LAYER FROM 5 TO 20. THE BEST VALUES FOR EACH FEATURE ARE BOLD MARKED

	5 neurons		**10 neurons**		**15 neurons**		**20 neurons**	
NO_2	*MSE*	*R*	*MSE*	*R*	*MSE*	*R*	*MSE*	*R*
P	**0.016**	**0.628**	0.034	0.497	0.037	0.564	0.027	0.625
Rh	**0.019**	**0.754**	0.031	0.403	0.031	0.416	0.017	0.600
T	0.028	0.693	0.089	0.898	**0.024**	**0.737**	0.051	0.440
Ws	0.050	0.550	0.061	0.273	0.020	0.302	**0.018**	**0.663**

TABLE IV. *MSE* AND *R* COEFFICIENT OBTAINED WHEN PREDICTING CO OF THE TEST SET WITH A NUMBER OF NEURONS IN THE HIDDEN LAYER FROM 5 TO 20. THE BEST VALUES FOR EACH FEATURE ARE BOLD MARKED

	5 neurons		**10 neurons**		**15 neurons**		**20 neurons**	
CO	*MSE*	*R*	*MSE*	*R*	*MSE*	*R*	*MSE*	*R*
P	0.007	0.769	**0.005**	**0.828**	0.008	0.784	0.004	0.818
Rh	0.010	0.712	**0.005**	**0.821**	0.008	0.763	0.006	0.823
T	**0.006**	**0.841**	0.007	0.804	0.007	0.814	0.010	0.720
Ws	0.007	0.795	**0.007**	**0.808**	0.007	0.771	0.008	0.691

It is worth noting that the best prediction accuracy of NO_2 given pressure and relative humidity is obtained with 5 neurons in the hidden layer of the network. By contrast, the lowest MSE is obtained with 15 neurons for the temperature, and 20 neurons for the wind speed.

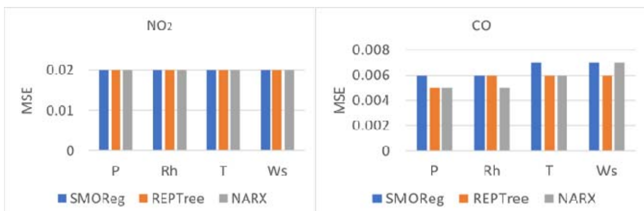

Figure 4. Comparison results in terms of MSE for predicting NO_2 and CO given: (i) pressure, (ii) relative humidity, (iii) temperature, and (iv) wind speed between NARX, SMOReg and REPTree

In these network configurations, we can observe that all variables are helpful for predicting the future trend of NO_2, since the MSE is very low and the R coefficient is higher than 0.60, which is a promising result. In particular, the best predictor variable in terms of MSE is the pressure, with a value of 0.016, while in terms of R it is the relative humidity, with a value of 0.754. By contrast, the best prediction accuracy of CO is obtained with 10 hidden neurons for pressure, wind speed and relative humidity, and 5 hidden neurons for temperature. In these cases, we can observe that the prediction is more accurate than NO_2, with higher R above 0.8 and lower MSE below 0.01. Also, a promising result is obtained for all predictor variables in terms of MSE and R. In particular, the best predictor variables, considering the combination of MSE and R values, are the pressure (MSE = 0.005 and R=0.828) and the temperature (MSE = 0.006 and R=0.841).

For the best number of hidden neurons (bold marked in Tables III-IV), Fig. 4 shows the comparison results in terms of MSE for predicting NO_2 and CO given: (i) pressure, (ii) relative humidity, (iii) temperature, and (iv) wind speed between NARX, SMOReg and REPTree. It is worth noting that the NARX model achieves lower MSE than its competing methods when pressure, relative humidity and temperature are used for predicting CO (3 out of 4 cases), whereas in the other cases NARX performs consistently with the competing methods.

3) *Results for Serbia*

Tables V and VI show the MSE and R values obtained from prediction of NO_2 and CO, as in the previous tables.

TABLE V. *MSE* AND *R* COEFFICIENT OBTAINED WHEN PREDICTING NO_2 OF THE TEST SET WITH A NUMBER OF NEURONS IN THE HIDDEN LAYER FROM 5 TO 20. THE BEST VALUES FOR EACH FEATURE ARE BOLD MARKED

NO_2	**5 neurons**		**10 neurons**		**15 neurons**		**20 neurons**	
	MSE	*R*	*MSE*	*R*	*MSE*	*R*	*MSE*	*R*
P	0.005	0.866	0.003	0.871	0.005	0.870	**0.003**	**0.908**
Rh	0.005	0.880	**0.003**	**0.859**	0.005	0.868	0.006	0.850
T	0.007	0.824	0.005	0.817	**0.004**	**0.874**	0.007	0.757
Ws	0.006	0.871	**0.005**	**0.867**	0.006	0.851	0.006	0.857

TABLE VI. *MSE* AND *R* COEFFICIENT OBTAINED WHEN PREDICTING CO OF THE TEST SET WITH A NUMBER OF NEURONS IN THE HIDDEN LAYER FROM 5 TO 20. THE BEST VALUES FOR EACH FEATURE ARE BOLD MARKED

CO	**5 neurons**		**10 neurons**		**15 neurons**		**20 neurons**	
	MSE	*R*	*MSE*	*R*	*MSE*	*R*	*MSE*	*R*
P	0.008	0.644	**0.007**	**0.776**	0.010	0.643	0.010	0.706
Rh	0.012	0.611	0.010	0.675	0.011	0.780	**0.007**	**0.747**
T	0.009	0.772	**0.009**	**0.749**	0.009	0.727	0.016	0.568
Ws	0.007	0.768	0.010	0.667	**0.005**	**0.852**	0.008	0.821

From the tables, it can be observed that the highest prediction accuracy of NO_2 based on the air pressure parameter is achieved with 20 neurons in the hidden layer. Moreover, for both relative humidity and wind speed, the highest prediction accuracy of NO_2 is achieved for 10 neurons in the hidden layer, while considering the temperature, best prediction accuracy is obtained for 15 neurons. It can also be noted that the MSE for all variables is very low (below 0.008), while the R coefficient is higher than 0.7, hence all variables are important and contribute to the prediction of NO_2 emissions. In terms of MSE, the best predictor variables are the pressure and relative humidity (MSE=0.003), and in terms of R coefficient, it is also the pressure (R=0.908).

Considering CO, the best prediction accuracy is achieved with 10 neurons for pressure and temperature, 15 neurons for wind speed, and 20 neurons for relative humidity. In terms of prediction accuracy, the MSE values are all below 0.02, while the R values are higher than 0.7. The best predictor variable in terms of both MSE and R for predicting the value of CO is wind speed (MSE=0.005, R=0.852).

Figure 5 shows the comparison results in terms of MSE for predicting NO_2 and CO given: (i) pressure, (ii) relative humidity, (iii) temperature, and (iv) wind speed between NARX (for the best number of hidden neurons - bold marked in Tables V-VI), SMOReg and REPTree.

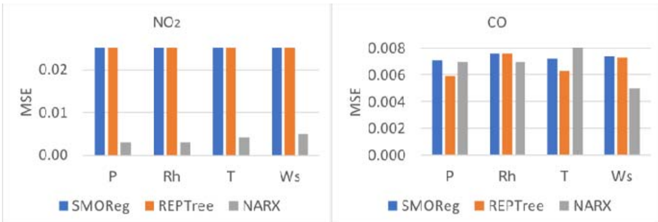

Figure 5. Comparison results in terms of MSE for predicting NO_2 and CO given: (i) pressure, (ii) relative humidity, (iii) temperature, and (iv) wind speed between NARX, SMOReg and REPTree

It is worth noting that NARX obtains a lower MSE than its competing methods for every variable used to predict NO_2 (all cases). By contrast, NARX overcomes the other methods when pressure, relative humidity and wind speed are used to predict the CO emission (3 out of 4 cases).

CONCLUSION

This paper introduced a NARX neural network model in order to predict the values of NO_2 and CO, in terms of meteorological parameters. About Serbia, a more accurate prediction is obtained for NO_2, while in terms of Bosnia and Herzegovina and Italy, more accurate predictions are obtained for CO. In terms of prediction strength of each variable, the best predictor variables of NO_2 are pressure and humidity (for Serbia and Italy), and pressure and wind speed (for Bosnia). The best predictor variables of CO are pressure and temperature (for Bosnia and Italy), and wind speed (for Serbia). A comparison of the NARX model with SVM regression and regression trees proved that the proposed model is able to overcome other well-known regression methods for time series prediction of CO emission in all countries, and time series prediction of NO_2 in Bosnia and Herzegovina.

In spite of the promising results obtained by the model, investigating about the effect of interaction among the used meteorological variables (pressure, relative humidity, temperature, wind speed) on predicting the air pollutants (NO_2 and CO) is beyond the scope of this work. Preliminary tests proved that additional effort in the NARX model is needed to obtain competitive results when considering such interaction, due to creation of instable solutions when the meteorological variables are used in conjunction. Accordingly, an extension of the proposed model will be provided as a future work direction for exploring and evaluating such interaction in the prediction process.

ACKNOWLEDGMENT

This work was partially supported by the Mathematical Institute of the Serbian Academy of Sciences and Arts (Project III44006). The authors gratefully acknowledge FHMZ BiH, MOVIDA-Multistations implemented in the framework of the Project of National Interest NextDATA funded by MIUR, operated and maintained by the Institute of Atmospheric Science and Climate (ISAC) by the National Research Council of Italy (CNR), and SEPA Serbia for providing the data.

REFERENCES

[1] European Environment Agency, [Online]: https://www.eea.europa.eu/themes/air/intro

[2] M. Ćosović and E. Junuz, "Air Pollution Forecasting using Machine Learning Techniques", in *Proc. Int. Conf. On Time Series and Forecasting*, Granada, Spain, Sept. 2018, pp. 264-273.

[3] A. Rahimi, "Short-term prediction of NO_2 and NOx concentrations using multilayer perceptron neural network: A case study of Tabriz, Iran," *Ecological Processes*, vol. 6, no. 1, Jan. 2017.

[4] L. Dragomir, C.M. Voiculescu, M. Constantin, and D. Georgescu. (Dec. 8, 2015). Prediction of the NO_2 concentration data in an urban area using multiple regression and neuronal networks. [Online]. Available: https://aip.scitation.org/doi/abs/10.1063/1.4937255

[5] B. Yeganeh, M. S. P. Motlagh, Y. Rashidi, and H. Kamalan, "Prediction of CO concentrations based on a hybrid partial least square and support vector machine model," *Atmos. Environ.*, vol. 55, pp. 357-365, Aug. 2012.

[6] H. Xie, H. Tang, and Y.-H. Liao, "Time series prediction based on NARX neural networks: An advanced approach," in *Proc. Int. Conf. on Mach. Learning and Cybern.*, Hebei, China, July 2009, pp. 1275-1279.

[7] C. Cortes, V.N. Vapnik, "Support-vector networks", *Machine Learning*, vol. 20, no. 3, pp. 273-297.

[8] Regression Trees, [Online]: http://www.stat.cmu.edu/~cshalizi/350-2006/lecture-10.pdf, 2006.