

Mirza Pasic¹
Izet Bijelonja

Article info:
Received 18.12.2019
Accepted 19.03.2020

UDC – 613.15
DOI – 10.24874/IJQR14.02-15



NARX-SP NEURAL NETWORK MODELS FOR AIR QUALITY PREDICTION FOR THE 24TH AND 48TH HOUR AHEAD

Abstract: *Neural networks are important method of machine learning that can be used to predict air quality with high accuracy. Using NARX-SP neural network type, several neural network models are developed to predict concentration of air pollutants in Sarajevo for two prediction cases, for 24th and 48th hour ahead, with different combinations of inputs and outputs. The data used in this paper contain hourly values of meteorological parameters (air humidity, pressure and temperature, wind speed and direction) and concentrations of SO₂, PM₁₀, NO₂, O₃ and CO from 2016 to 2018. Optimal models are selected for both prediction cases. It is concluded that the optimal models have very good performances and can be used to predict concentration of pollutants in Sarajevo with great accuracy and contribute to improve quality of life. By adequate application of optimal models, concentration of air pollutants can be predicted for each hour over the next 48 hours.*

Keywords: *Neural networks; NARX-SP; Air quality; Concentration of air pollutants; Prediction*

1. Introduction

Neural networks are used in many fields like engineering, medicine, economics, artificial intelligence, as well as in the field of air quality prediction, where they are being established as an effective prediction tool. One of the measures to protect human health and to improve quality of life from the harmful effects of air pollutants is an early warning, which can be achieved by predicting the concentration of pollutants using neural networks.

Sarajevo, the capital of Bosnia and Herzegovina, faces a major problem of air pollution, so the prediction of the concentration of air pollutants in Sarajevo is very important to protect human health against the effects of harmful pollutants. In this paper neural network models are

developed that effectively predict the air quality in Sarajevo, for prevention and protection purposes.

The importance of predicting air quality is highlighted in many papers. Wen et al. (2019) emphasizes that the prediction of air pollution is of great importance for people in terms of day-to-day health monitoring, as well as for competent institutions in terms of making appropriate decisions. Li et al. (2017) also emphasized the importance of this problem, where it was pointed out that the prediction of air pollution is an effective method of protecting public health, as it provides early warnings for elevated concentration of air pollutants.

Air quality depends on meteorological conditions (Wang et al., 2013) and pollutant sources (Urbanski et al., 2011). Sources of pollutants produce pollutants, while

¹ Corresponding author: Mirza Pasic
Email: mirza.pasic@mef.unsa.ba

meteorological conditions affect the transfer and diffusion of these pollutants in the atmosphere (Bai et al., 2016).

The problem of air pollution needs to be considered very seriously, as pollutants have a very harmful effect on human health and the entire social community. According to the British Lung Foundation, pollutants that are most hazardous to human health are particulate matter (PM_{10} , $PM_{2.5}$), ozone at ground level (O_3), carbon monoxide (CO) and sulfur dioxide (SO_2). High concentrations of these pollutants can cause many diseases, including lung disease (asthma, bronchitis, chronic obstructive pulmonary disease, lung cancer, etc.), heart disease, heart attack and stroke (British Lung Foundation, 2017).

He et al. (2013) concluded that meteorological conditions play an essential role in the daily fluctuation of pollutants. As a result, more scientists predict air quality depending on meteorological parameters through various scientific approaches, such as statistical models (Ozel & Cakmakyapan, 2015), artificial neural networks (ANN) (Feng et al., 2015), the grey model (Pai et al., 2013) and others. The neural network, as a very efficient prediction method, has shown great dominance in the field of prediction and analysis of air quality (Bai et al., 2016).

De Gennaro et al. (2013) developed an ANN to predict daily values of PM_{10} at two locations in the Western Mediterranean (Montseny and Barcelona) depending on local meteorological parameters. Wu et al. (2011) developed an Elman neural network for predicting air pollution index (API) values for an urban area of Wuhan, China, based on daily temperature, relative humidity, wind speed, pressure, precipitation, and duration of sunny periods. Prediction of SO_2 , NO_2 and O_3 depending on meteorological parameters (air temperature, wind speed, atmospheric pressure, air humidity) in the city of Konya, Turkey using the ANN and ANFIS method is presented by Dursun et al. (2015). Russo et al. (2013) performed air quality prediction using ANN for Lisbon in Portugal. In Kumar and

Goyal (2013), an ANN based on principal component analysis was developed to predict the air quality index based on different meteorological parameters for Delhi, India, for four annual periods (summer, monsoon, post-monsoon, and winter).

There are many types of neural networks used to predict air quality and most of them produce good prediction results. However, according to Dorffner (1996), Lin et al. (1996) and Boussaada et al. (2018), NARX neural network is recognized as one of the most powerful types of neural networks. It is the nonlinear autoregressive exogenous network (NARX). The NARX neural network is a type of neural network that is adequate in working with data where the values of the outputs that are predicted depend on their values from the previous period, as well as the values of exogenous inputs from that period. The prediction model and performance of NARX neural networks have been addressed by Dorffner (1996), Lin et al. (1996) and Boussaada et al. (2018). In Wang and Bai (2014) and Pisoni et al. (2009) the use of a NARX neural network for air quality prediction has shown very good results.

In this paper the original models of neural networks for air quality prediction are developed, which implies the prediction of concentration of air pollutants, with different structures and combinations of inputs and outputs of neural networks. The original structures and combinations of inputs and outputs are achieved by a different number of previous hours and a different hour ahead for which the pollutant concentrations were predicted, from which the prediction data are used. The original structures are defined by the number of neurons in the input, hidden and output layers of neural networks. With developed neural network models, it is possible to predict the concentration of pollutants SO_2 , PM_{10} , NO_2 , O_3 and CO for Sarajevo, which has its landscape, climate and economic specificities, with high accuracy. Models are developed for 2 prediction cases, for the 24th and 48th hour

ahead. For each prediction case, several neural network models have been developed that differ in the number of previous hours used for prediction. Based on the values of coefficient of determination R^2 and mean squared error (MSE) for the test data set, and based on the complexity of the model, the most appropriate prediction model is selected for each prediction case, which will be referred to as the optimal neural network model. Optimal models show high prediction accuracy. With optimal neural network model, it is possible to predict with high accuracy the concentration of air pollutants for each sequential hour ahead, including the hour for which the model is intended, relative to the current hour, by the appropriate choice of the referent hour. The referent hour is the hour relative to which the prediction is made. Data from the referent hour are treated as preliminary values for the prediction of pollutant concentration since they precede the hour for which the prediction is made.

2. Theoretical background

2.1. NARX neural network

In this paper, a non-linear auto-regressive neural network with exogenous inputs, NARX neural network, is used. In this type of neural network, previous values of outputs and exogenous inputs are used as inputs (Abrahamsen et al., 2018).

NARX neural network is represented by equation (1) (Horne et al., 2015):

$$y(t) = f[(u(t - n_u), \dots, u(t - 1), u(t), y(t - n_y), \dots, y(t - 1))] \quad (1)$$

where:

- $u(t)$ – input to neural network at time t ,
- $y(t)$ – output from neural network at time t ,
- n_u – input order,
- n_y – output order.

By using previous output values as inputs to the neural network, NARX network can be modeled as a network with serial-parallel

(NARX-SP) and parallel (NARX-P) neural network structure. In this paper, NARX-SP neural network is used to develop models for air quality prediction. For NARX-SP neural network structure, the previous output values, which are used as inputs to the neural network, are the actual output values. The backpropagation algorithm is used for training of this neural network structure (Al Hamaydeh, Choudhary and Assaleh, 2013).

2.2. MLP

The structure of NARX-SP neural network is based on the principle of MLP functioning. A single hidden layer MLP is shown in Figure 1.

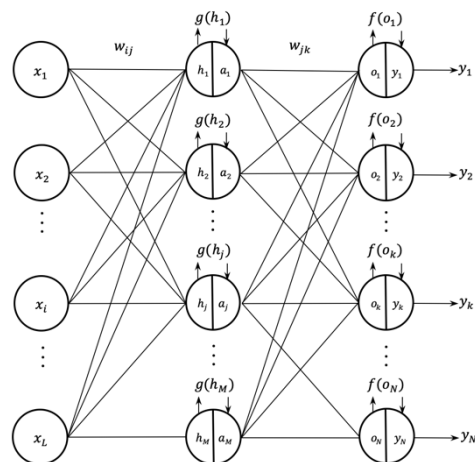


Figure 1. MLP with one hidden layer

Adapted from: Marsland (2015)

The principle of functioning of MLP having one hidden layer can be represented by equations (2) to (5).

Equation (2) shows the value entering each hidden layer neurons and represents the sum of the products of the weights and the output values from the input layer neurons (Marsland, 2015).

$$h_j = \sum_{i=1}^L x_i \cdot w_{ij} \quad (2)$$

Hidden layer neurons have a defined activation function $g(h_j)$ in which h_j is the input value. The output value of the hidden layer neurons is the result of the activation function $g(h_j)$:

$$a_j = g(h_j) \quad (3)$$

As for the output layer neurons, their input value o_k is the sum of the products of the weights and the output value of the hidden layer neurons a_j :

$$o_k = \sum_{j=1}^M a_j \cdot w_{jk} \quad (4)$$

The output value from the output layer neurons is the result of the activation function $f(o_k)$.

$$y_k = f(o_k) \quad (5)$$

where:

x_i – output value from input layer neurons,

h_j – input value to hidden layer neurons,

w_{ij} – weights between input and hidden layer neurons,

$g(h_j)$ – activation function in hidden layer neurons,

a_j – output value from hidden layer neurons,

o_k – input value to output layer neurons,

w_{jk} – weights between hidden and output layer neurons,

$f(o_k)$ – activation function in output layer neurons,

y_k – output value from output layer neurons,

i – the number of neurons in the input layer ($i = 1, \dots, L$ if there is no bias, and if there is $i = 0, \dots, L$),

j – the number of neurons in the hidden layer ($j = 1, \dots, M$ if is no bias, and if there is $j = 0, \dots, M$),

k – the number of neurons in the output layer ($k = 1, \dots, N$).

In order to obtain the most accurate output value from the output layer neurons, a correction of the results must be performed using the backpropagation algorithm, which corrects the weights (Marsland, 2015).

2.3. Air pollutants

There are many different pollutants in the air that have a negative impact on human health, environment, climate and many other areas of life. According to WHO report, the most evidence for adverse effects on human health exists for PM , SO_2 , NO_2 and O_3 pollutants. In this paper, basic informations about pollutants SO_2 , PM , NO_2 , O_3 , and CO that are the subject of prediction in this paper are given according to European Environment Agency (EEA).

There are particles of different sizes in the air, but the most dangerous for human health are those whose diameter is equal to or less than $10 \mu m$, which are PM_{10} and $PM_{2.5}$. It is extremely dangerous for lung, bloodstream and entire human body.

Sulfur dioxide (SO_2) is a colorless gas with an extremely pungent odor. It is threat to human health because it causes asthma, cough and bronchitis.

Nitrogen dioxide (NO_2) is most prevalent in urban areas and causes and exacerbates asthma, pneumonia and other lung diseases.

Carbon monoxide (CO) has no color or odor and is very dangerous because, in increased concentrations, it causes death, while in smaller concentrations it negatively affects the brain, causes headache and visual disturbance and reduces cognitive ability.

Ozone (O_3) is a secondary pollutant formed by the chemical reaction of primary pollutants NO_2 , $VCOs$ and CH_4 in the presence of sunlight. Ozone is a very aggressive gas that causes chest pain, cough, exacerbation of respiratory diseases and chronic lung disease in humans (European Environment Agency, 2019).

3. Methodology

In this paper, neural network models are developed to predict the concentration of pollutants for the 24th and 48th hour ahead, optimal models are selected and performances of models are analyzed.

For this study data obtained from the Federal Meteorological Institute of BiH are used for this research. Those are:

- values of meteorological parameters – air temperature (T), pressure (p) and humidity (H), wind speed (v) and wind direction (wd) and

- concentration of air pollutants (SO_2 , PM_{10} , NO_2 , O_3 and CO).

These values are measured for every hour between the beginning of 2016 and the end of 2018. For each meteorological parameter and for each pollutant in the air 26304 measured hourly values are obtained. These data are grouped into samples and divided into training, validation and test data sets.

In order to gain a better insight into the data used in this work, the average monthly values of air pollutant concentrations for the period from the beginning of 2016 to the end of 2018 are shown in Figures 2a-2e.

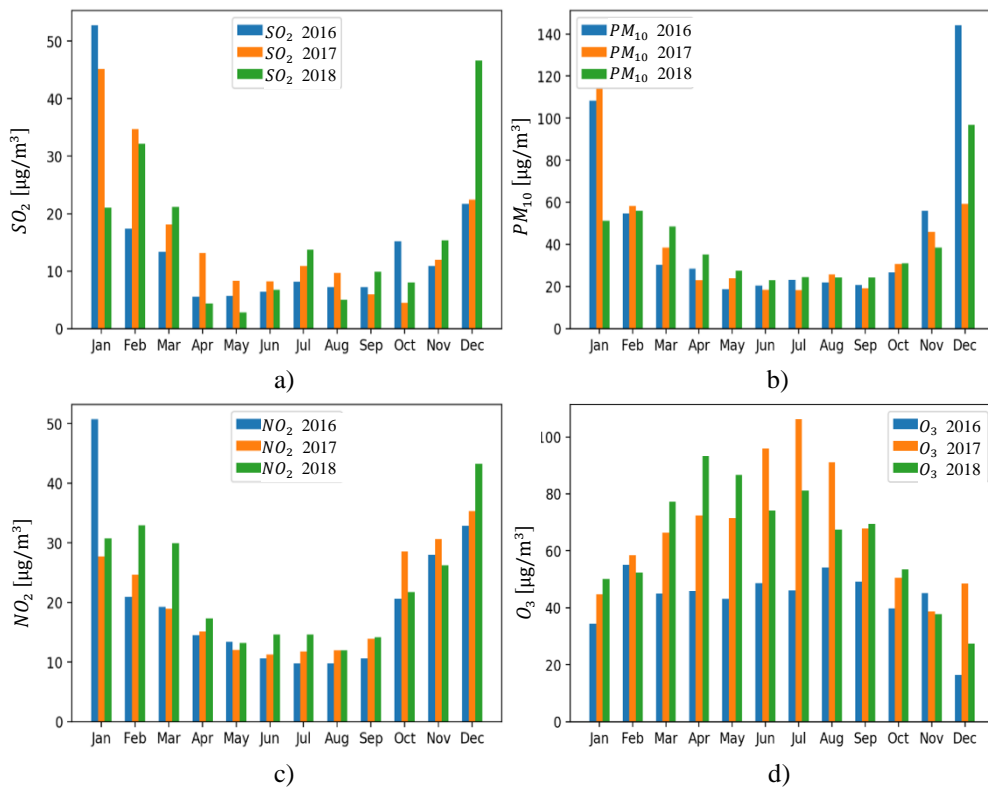
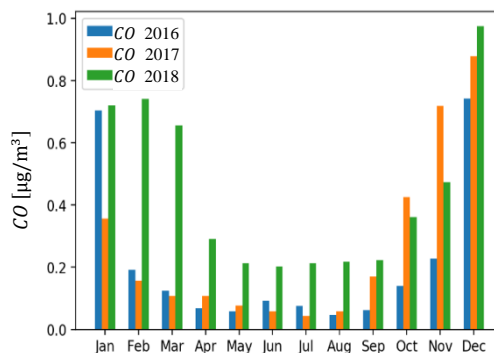


Figure 2. Average monthly concentrations of air pollutants for the period from 2016 to 2018: a) SO_2 , b) PM_{10} , c) NO_2 , d) O_3



e)

Figure 2. (continued) Average monthly concentrations of air pollutants for the period from 2016 to 2018: e) CO

The aim of this paper is to develop neural network models to predict the concentration of air pollutants SO_2 , PM_{10} , NO_2 , O_3 and CO with the original structures of neural networks and different combinations of inputs and outputs for Sarajevo. Models differ from each other according to the selected number of previous hours and the hour for which the prediction is made. The number of inputs selected and therefore the number of neurons in the hidden layer, depends on the number of previous hours selected.

As input data to the neural networks values of meteorological parameters, concentrations of pollutants, month of the year, day of week and time of day from the referent hour or the referent hour and a certain number of previous hours as well as the values of meteorological parameters, month, day and time for the hour for which prediction is made are used.

For each developed neural network model, the values of R^2 and MSE for the model and for each pollutant individually for all data sets are calculated as indicators of the performance of the developed models.

Based on the values of R^2 and MSE of the test data set, and the complexity of the model, for both cases the most appropriate model is selected, which is referred to as an optimal

model. If there are more models whose value of R^2 is up to 1% less than the maximum value of R^2 and whose values of MSE are of the same order of magnitude for the same prediction case, the one with the less complex structure is chosen as the optimal model.

Based on the values of R^2 and MSE for the test data set, analysis and conclusions are made about the accuracy of prediction of optimal models, thereby verifying the model. For each optimal neural network model, the relationship between predicted and observed values for the training, validation, test and total data sets are graphically displayed and negligible deviations between predicted and observed values can be noticed. Based on that, conclusions about the accuracy of prediction models are confirmed.

Values of R^2 and MSE are calculated according to the formulas (6) and (7) respectively (Levine et al., 2017):

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

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

where:

y_i – actual output value,

\bar{y} – the mean of output value,

\hat{y}_i – the output value predicted by the model,

\hat{y}_i – sample size.

Values of the meteorological parameters other than wind direction and concentration of air pollutants are normalized to interval [0,1] as shown in equation (8) (Marsland, 2015):

$$x'_i = \left(\frac{X_i - X_{min}}{X_{max} - X_{min}} \right) \quad (8)$$

Values referring to the month of the year, day of the week, time of the day and the wind direction are normalized using sine and cosine functions to the interval [-1,1] as shown in equations (9) and (10) (Feng et al., 2015; Russo et al., 2013):

$$x'_{i,sin} = \sin\left(\frac{2 \cdot \pi \cdot X_i}{n}\right) \quad (9)$$

$$x'_{i,cos} = \cos\left(\frac{2 \cdot \pi \cdot X_i}{n}\right) \quad (10)$$

In order to reduce noise in the data, filtering of the data is done using the moving average method as shown in equation (11).

$$(y_k)_s = \frac{\sum_{i=-n}^{i=n} y_{k+i}}{2n + 1} \quad (11)$$

where:

$(y_k)_s$ – filtered data at point k (s – smoothing),

y_k – data at point k before settlement,

n – number of data to the left and right of the midpoint.

In this paper for the filtering of the data, the value of $n = 2$ is chosen.

From the available database, for each neural network model, appropriate data samples are created, depending on the number of previous hours and the hour for which the prediction is made in that model, from which the prediction data are used. Each sample consists of meteorological parameter values, concentration of pollutants, month, day and time of day from the referent hour or referent hour and the certain number of its previous hours defined for that model, such as meteorological parameter values, month, day and time of day and the concentration of pollutants for the predicted hour, which is defined by the neural network model.

To allow the proper learning of neural networks, samples are permuted before the development of the neural network model.

After permutation, samples are divided into training, validation and test data sets in the proportion of 60:20:20 respectively.

For both cases, 5 neural network models are developed. In the development of the models for the prediction for 24th hour ahead, the data from the referent hour or referent hour and its previous 11, 23, 35 and 47 hours depending on the model are used as inputs as well as the data from the hour for which the prediction is made. The prediction for the 48th hour ahead used data from the referent hour or referent hour and its previous 23, 35, 47 and 71 hours depending on the model and data from the hour for which the prediction is made. In the prediction case for the 48th hour ahead, model that uses, among other data, the data from the referent hour and its previous 11 hours is not developed, because for this model values of R^2 and MSE are not satisfactory like in the prediction case for the 24th hour ahead.

Tables 1 and 2 show the structures of the developed neural network models for prediction of pollutant concentrations for 24th and 48th hour ahead.

Table 1. Structures of neural network models for the 24th hour ahead

Model	Number of previous hours	Structure
1	Referent hour	29-59-5
2	Referent hour + 11 previous hours	216-433-5
3	Referent hour + 23 previous hours	432-865-5
4	Referent hour + 35 previous hours	636-1273-5
5	Referent hour + 47 previous hours	864-1729-5

Table 2. Structures of neural network models for the 48th hour ahead

Model	Number of previous hours	Structure
1	Referent hour	29-59-5
2	Referent hour + 23 previous hours	432-865-5
3	Referent hour + 35 previous hours	636-1273-5
4	Referent hour + 47 previous hours	864-1729-5
5	Referent hour + 71 previous hours	1296-2593-5

During the training of neural networks 500 epochs with the early stopping are used as well as the batch size of 10, sigmoid activation function and Adam optimizer. The structure of neural networks implies the number of neurons in the input, hidden, and output layers. One hidden layer is used based on the Universal Approximation Theorem (Marsland, 2015), with the number of neurons

determined by the Kolomogorov rule (Mu et al., 2017). For the development of neural network models and R^2 and MSE value calculation, code is created in Python using the Keras library. R^2 and MSE values for developed models for training, validation and test data sets are presented in tables 3, 4, 5 and 6.

Table 3. R^2 values [%] for training, validation and test data sets for the 24th hour ahead

Model	R^2_{train}	$R^2_{validate}$	R^2_{test}
1	69,094	68,338	66,939
2	94,057	88,975	89,246
3	97,557	93,781	93,276
4	97,063	93,052	94,131
5	97,844	94,386	94,013

Table 4. MSE values for training, validation and test data sets for the 24th hour ahead

Model	MSE_{train}	$MSE_{validate}$	MSE_{test}
1	0,001858922	0,001939995	0,001912858
2	0,000349619	0,000651784	0,000670696
3	0,000139138	0,000384891	0,000381512
4	0,000169661	0,000396399	0,000380008
5	0,000123130	0,000352282	0,000356418

Table 5. R^2 values [%] for training, validation and test data sets for the 48th hour ahead

Model	R^2_{train}	$R^2_{validate}$	R^2_{test}
1	66,599	64,629	63,304
2	95,636	93,501	92,683
3	95,972	93,736	94,021
4	95,668	93,978	93,901
5	95,273	93,054	93,098

Table 6. *MSE* values for training, validation and test data sets for the 48th hour ahead

Model	<i>MSE_train</i>	<i>MSE_validate</i>	<i>MSE_test</i>
1	0,002124366	0,002014800	0,002079332
2	0,000401532	0,000262467	0,000437570
3	0,000372393	0,000238323	0,000376697
4	0,000355566	0,000250491	0,000354447
5	0,000409285	0,000283736	0,000411375

4. Result analysis and discussion

Optimal models for both prediction cases are selected based on the values of R^2 and *MSE* for the test data set, and based on the complexity of the model.

For the 24th hour prediction case, model 4 is chosen as the optimal model. The value of $R^2 = 66,939\%$ for model 1 is significantly lower than for other models, and the value of *MSE* = 0,001912858 is significantly higher. Model 2 has values of $R^2 = 89,246\%$ and *MSE* = 0,000670696 and its performances are much better than performances of model 1. Models 3, 4 and 5 have high values of R^2 and low values of *MSE*. Model 4 has $R^2 = 94,131\%$ and *MSE* = 0,000380008 slightly

better than $R^2 = 94,013\%$ and *MSE* = 0,000356418 of model 5. Model 4 has a slightly higher R^2 value and a slightly higher value of *MSE* than model 5. Since these values are approximately the same and very good for models 4 and 5, model 4 is chosen as an optimal model because it has a simpler structure than model 5 and is more practical to use. When choosing the optimal neural network model, if values of R^2 and *MSE* models do not differ significantly, it is necessary to select a model that has a simpler structure.

R^2 and *MSE* values for all 5 models for the test data set are depicted graphically in figures 3 and 4.

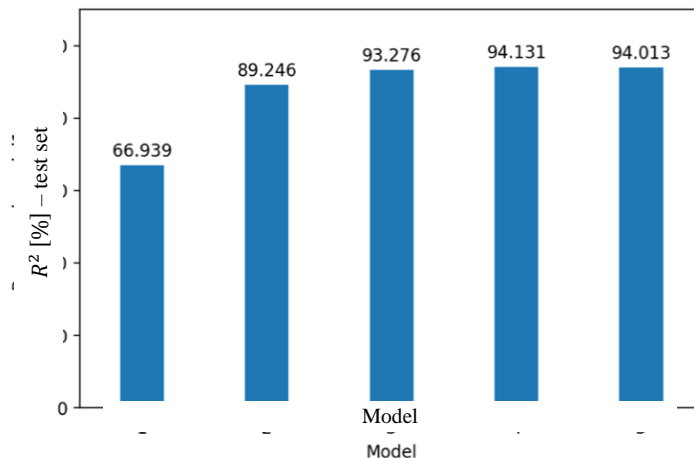


Figure 3. R^2 values [%] for test set by models for 24th hour ahead

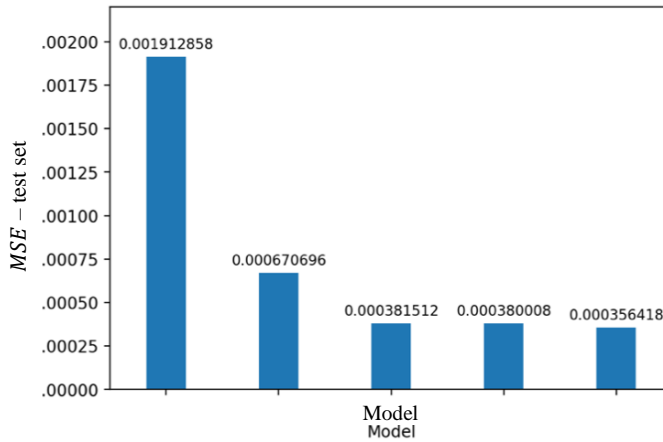


Figure 4. *MSE* values for test set by models for 24th hour ahead

Values of R^2 and *MSE* for all 5 models for test set for 48th hour prediction case are plotted in Figures 5 and 6. In 48th hour prediction case a sharp increase in R^2 can be observed from model 1 to model 2, with $R^2 = 63,304\%$ for model 1, to $R^2 = 92,683\%$ for model 2, and a sudden decrease in *MSE* with *MSE* = 0,002079332 for model 1, to *MSE* = 0,000437570 for model 2. Models 4 and 5 which have a more complex neural network structure, have a slightly lower value of R^2 than model 3, that can be seen from Figure 5. However, the

MSE value is the lowest for model 4. Model 3 has a slightly higher value of $R^2 = 94,021\%$ than the value of $R^2 = 93,901\%$ of model 4 and has a slightly higher value of *MSE* = 0,000376697 than the value of *MSE* = 0,000354447 of model 4. Values of R^2 and *MSE* for models 3 and 4 are about the same and very good, and because model 3 has a simpler structure, it is chosen as the optimal model. It can be concluded that the values of R^2 and *MSE* of model 3 for the test data set are very good.

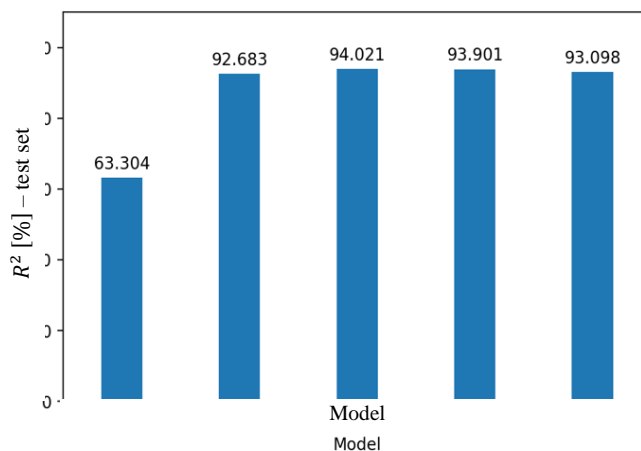


Figure 5. R^2 values [%] for test set by models for 48th hour ahead

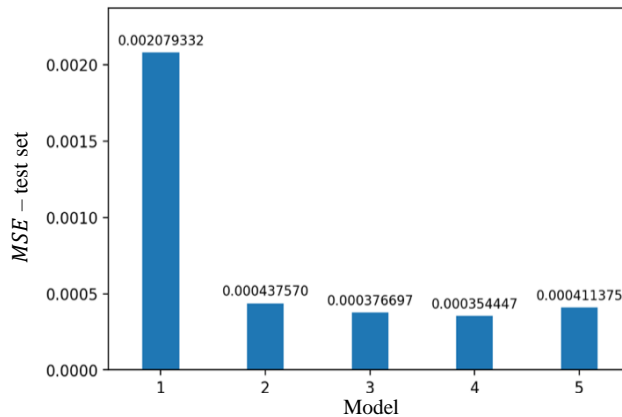


Figure 6. MSE values for test set by models for 48th hour ahead

Optimal models for both prediction cases have high R^2 values and low MSE values, so it can be concluded that they have high prediction accuracy. This can be confirmed in the figures where a graphical presentations of predicted and observed concentration values of all 5 pollutants is given. By analyzing Figure 7, it can be seen that the predicted values follow observed values with negligible deviations, thus confirming high prediction accuracy of optimal neural network models for the 24th and 48th hours prediction cases. For both prediction cases for the selected optimal neural network models, a graphical presentation of the predicted and observed values for a randomly selected period of 72 hours from a total period of three years is given, in order to give a clearer picture of how the values predicted by the model follow the observed values of concentration of pollutants, which is shown in Figure 7.

Comparing the values of R^2 and MSE for the test data set for models that predict the concentration of pollutants for the 24th and 48th hour ahead, which, in addition to other inputs, use data from referent hour or referent hour and its previous 23, 35 and 47 hours, slightly worse values can be observed for the 48th hour prediction case.

Optimal model 4 for the 24th hour prediction case has $R^2 = 94,131\%$ and $MSE = 0,000380008$ and uses the data from the referent hour and its previous 35 hours from the previous period. Optimal model 3 at the 48th hour prediction case has values $R^2 = 94,021\%$ and $MSE = 0,000376697$ and uses the data from the referent hour and its previous 35 hours. It can be observed that in both prediction cases, with the increase of number of previous hours from which the prediction data are used, accuracy increases sharply by increasing R^2 values to a certain value, then increases or even decreases slightly using more data from the previous period and by decreasing MSE values sharply at first and then decreases or even increases slightly.

This suggests that during the prediction, an optimal number of previous hours must be found to obtain satisfactory performance of the model with a simpler structure.

The optimal prediction model for the 24th hour prediction case can predict the concentration of pollutants for all hours from the current hour to the 24th hour ahead of the current hour by proper selection of referent hour. This means that the referent hour must be away from the hour for which the prediction is made for 24 hours. The same analogy can be applied to the optimal model for the 48th hour prediction case.

With these two optimal models, the air pollutant concentration can be predicted for each hour over the next 48 hours. For the prediction of air pollutant concentrations up to 24th hours ahead, the optimal model for 24th hour prediction case is used because it

has better values of R^2 and MSE than the optimal model for 48th hour prediction case. For the hours from 24th to 48th hours ahead, the model for 48th hour prediction case is used.

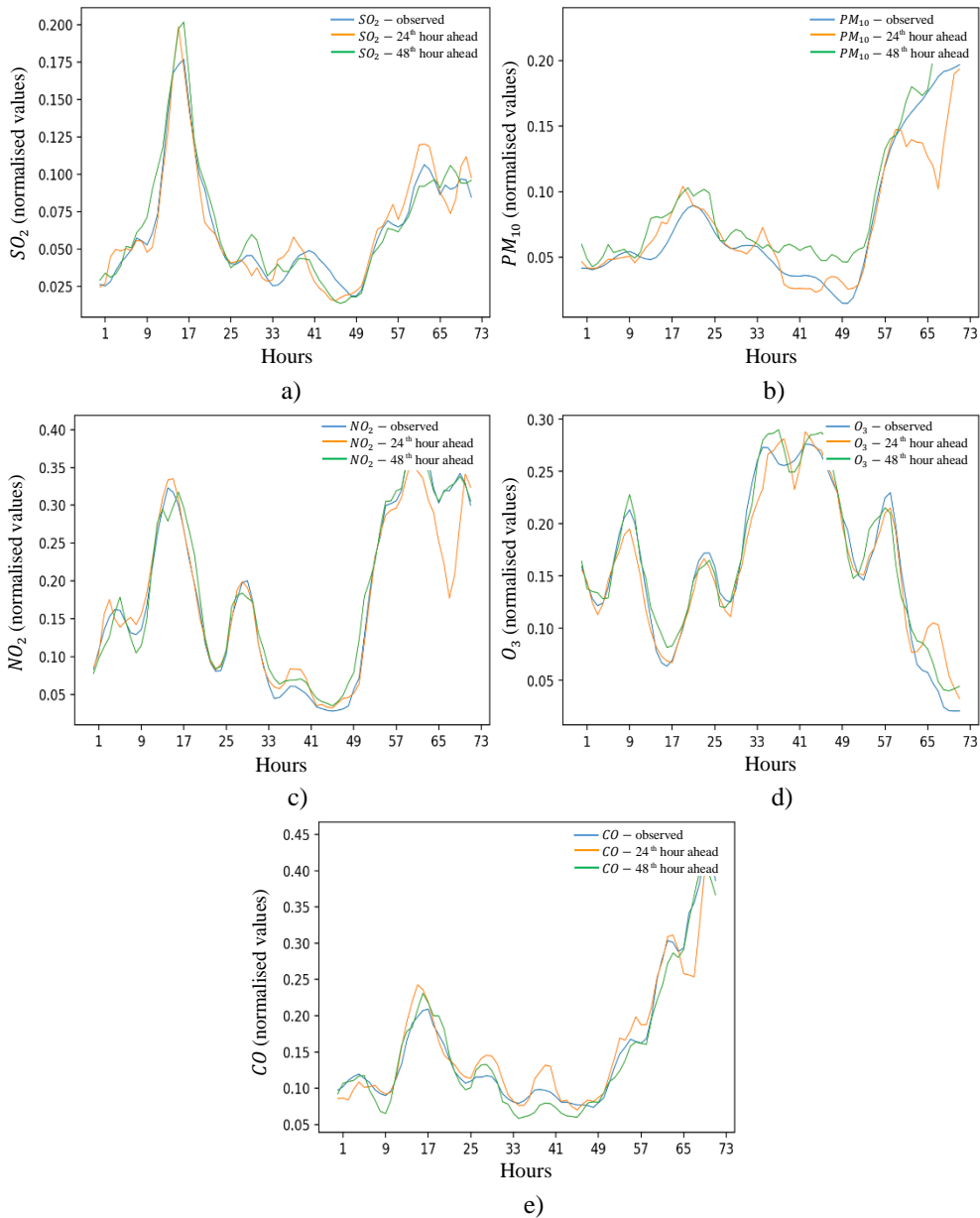


Figure 7. Graphical presentation of the predicted and observed values for selected period of 72 hours for 24th and 48th hour ahead a) SO_2 , b) PM_{10} , c) NO_2 , d) O_3 and e) CO

5. Conclusion

Air quality is the world's growing problem because it has a negative impact on human health. The problem is complex and requires many measures to be taken to improve air quality. Air quality is determined by the concentration of pollutants in the air. Sarajevo faces poor air quality during winter. The prediction of concentration of pollutants is imposed as a necessity in order to protect human health.

In this paper, neural network models are developed to predict the concentration of air pollutants for the 24th and 48th hour ahead, which can predict the concentration of air pollutants for each hour ahead up to an hour for which the model is intended, including that hour.

For the development of the prediction models, NARX-SP neural network type is used, where output values from the previous period and exogenous inputs for the same period are used.

In both prediction cases, concentration of the same pollutants are predicted, SO_2 , PM_{10} , NO_2 , O_3 and CO , for a different hour ahead. As inputs for the development of the models, according to the NARX-SP principles, concentration of these pollutants from the selected previous period are used with the values of meteorological parameters, month of the year, day of the week and hour in the day from the same previous period and meteorological parameters, month of the year, day of the week and hour in the day for

hour for which the prediction is being made.

For both prediction cases, 5 models are developed that differ according to the selected previous hours from which the prediction data are used. For both cases, optimal prediction models are selected.

Optimal models for both prediction cases have high values of R^2 and low values of MSE , so it can be concluded that models have high prediction accuracy. By using these models, it is possible to predict the concentration of pollutants for the 24th and 48th hour ahead with great accuracy.

The graphical presentation of predicted and observed values of the pollutant concentrations showed very small deviations of predicted from observed values, which confirmed the high accuracy of models' predictions.

It can also be concluded that NARX-SP is a powerful neural network type that can achieve high prediction accuracy for this type of data. Analyzing the prediction accuracy of the developed models for both prediction cases, it can be concluded that the optimal models can predict the concentration of pollutants for a certain hour ahead with great accuracy, but also for all hours ahead up to the hour for which the model is intended with the correct choice of the referent hour.

The analysis of the selection of optimal models shows that optimal number of previous hours from which the data is used, must be determined, to achieve satisfactory performance and a simpler model structure.

References:

- Abrahamsen, E. B., Brastein, O. M., & Lie, B. (2018). Machine learning in Python for weather forecast based on freely available weather data. In L.E. Øi, T. Komulainen, R.T. Bye & L.O. Nord (Eds.), *Proceedings of the 59th Conference on Simulation and Modelling (SIMS 59)* (pp. 169-176). Oslo, Norway: Metropolitan University.
- Bai, Y., Li, Y., Wang, X., Xie, J., & Li, C. (2016). Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions. *Atmospheric Pollution Research*, 7(3), 557-566. doi: 10.1016/j.apr.2016.01.004

- Boussaada, Z., Curea, O., Remaci, A., Camblong, H., & Bellaaj, N.M. (2018). A Nonlinear Autoregressive Exogenous (NARX) Neural Network Model for the Prediction of the Daily Direct Solar Radiation. *Energies*, 11(3), 620-641. doi: 10.3390/en11030620
- British Lung Foundation. (2017). *Types of air pollution*. Retrieved from <https://www.blf.org.uk/support-for-you/air-pollution/type>
- De Gennaro, G., Trizio, L., Di Gilio, A., Pey, J., Pérez, N., Cusack, M., Alastuey, A., & Querol, X. (2013). Neural network model for the prediction of PM_{10} daily concentrations in two sites in the Western Mediterranean. *Science of the Total Environment*, 463-464C, 875-883. doi: 10.1016/j.scitotenv.2013.06.093
- Dorffner, G. (1996). Neural Networks for Time Series Processing. *Neural Network World*, 6, 447-468.
- Dursun, S., Kunt, F., & Taylan, O. (2015). Modelling sulphur dioxide levels of Konya city using artificial intelligent related to ozone, nitrogen dioxide and meteorological factors. *International Journal of Environmental Science and Technology*, 12, 3915-3928. doi: 10.1007/s13762-015-0821-2
- European Environment Agency (2019). *Air quality in Europe – 2019 report*. Retrieved from: <https://www.eea.europa.eu/publications/air-quality-in-europe-2019>
- Feng, X., Li, Q., Hu, J., Jin, L., & Wang J. (2015). Artificial neural networks forecasting of $PM_{2,5}$ pollution using air mass trajectory based geographic model and wavelet transformation. *Atmospheric Environment*, 107, 118-128. doi: 10.1016/j.atmosenv.2015.02.030
- He, J., Yu, Y., Liu, N., & Zhao, S. (2013). Numerical model-based relationship between meteorological conditions and air quality and its implication for urban air quality management. *International Journal of Environment and Pollution*, 53(3/4), 265-286. doi: 10.1504/ijep.2013.059921
- Horne, B. G., Siegelmann, H. T. & Giles, C. L. (2015). What NARX networks can compute. In M. Bartosek, J. Staudek & J. Wiedermann (Eds.), *SOFSEM '95: Proceedings of the 22nd Seminar on Current Trends in Theory and Practice of Informatics* (pp. 95-102). Berlin, Germany: Springer-Verlag.
- Kumar, A., & Goyal, P. (2013). Forecasting of air quality index in Delhi using neural network based on principal component analysis. *Pure and Applied Geophysics*, 170, 711-722. doi: 10.1007/s00024-012-0583-4
- Levine, D. M., Stephan, D. F., Krehbiel, T. C., & Berenson, M. L. (2008). *Statistics for Managers Using Microsoft Excel®*. London: Pearson Education Inc.
- Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C., & Chi, T. (2017). Long short-term memory neural network for air pollutant. *Environmental Pollution*, 231(1), 997-1004. doi: 10.1016/j.envpol.2017.08.114
- Lin, T., Horne, B.G., Tino, P., & Giles, C.L. (1996). Learning long-term dependencies in NARX recurrent neural networks. *IEEE Transactions on Neural Networks*, 7(6), 1329-1552. doi: 10.1109/72.548162
- Marsland, S. (2015). *Machine learning: an algorithmic perspective*. Boca Raton: CRC Press, Taylor & Francis Group.
- Mu, C., Qiu, B. Z., & Liu, X.H. (2017). A new method for figuring the number of hidden layer nodes in BP. *International Journal of Recent and Innovation Trends in Computing and Communication*, 5(9), 101-114.

- Ozel, G., & Cakmakyapan, S. (2015) A new approach to the prediction of PM_{10} concentrations in Central Anatolia Region, Turkey. *Atmospheric Pollution Research*, 6, 735-741. doi: 10.5094/apr.2015.082
- Pai, T. Y., Hanaki, K., & Chiou, R. J. (2013). Forecasting hourly roadside particulate matter in Taipei county of Taiwan based on first-order and one-variable grey model. *Clean-Soil Air Water*, 41, 737-742. doi: 10.1002/clen.201000402
- Pisoni, E., Farina, M., Carnevale, C., & Piroddi, L. (2009). Forecasting peak air pollution levels using NARX models. *Engineering Applications of Artificial Intelligence*, 22, 593-602. doi: 10.1016/j.engappai.2009.04.002
- Russo, A., Raischel, F., & Lind, P. G. (2013). Air quality prediction using optimal neural networks with stochastic variables. *Atmospheric Environment*, 79, 822-830. doi: doi.org/10.1016/j.atmosenv.2013.07.072
- Urbanski, S. P., Hao, W. M., & Nordgren, B. (2011). The wildland fire emissions inventory: western United States emission estimates and an evaluation of uncertainty. *Atmospheric Chemistry and Physics*, 11, 12973-13000. doi: 10.5194/acp-11-12973-2011
- Wang, J., Wang, Y., Liu, H., Yang, Y., Zhang, X., Li, Y., Zhang, Y., & Deng, G. (2013). Diagnostic identification of the impact of meteorological conditions on $PM_{2.5}$ concentrations in Beijing. *Atmospheric Environment*, 81, 158-165. doi: 10.1016/j.atmosenv.2013.08.033
- Wang, L., & Bai, Y. (2014). Research on prediction of air quality index based on NARX and SVM. *Applied Mechanics and Materials*, 602-605, 3580-3584. doi: 10.4028/www.scientific.net/AMM.602-605.3580
- Wen, C., Liu, S., Yao, X., Peng, L., Li, X., Hu, Y., & Chi, T. (2019). A novel spatiotemporal convolutional long short-term neural network for air pollution prediction. *Science of the Total Environment*, 654, 1091-1099. doi: 10.1016/j.scitotenv.2018.11.086
- Wu, S., Feng, Q., Du, Y., & Li, X. (2011). Artificial neural network models for daily PM_{10} air pollution index prediction in the urban area of Wuhan, China. *Environmental Engineering Science*, 28(5), 357-363. doi: 10.1089/ees.2010.0219

Mirza Pasic

Mechanical engineering
faculty Sarajevo, University
of Sarajevo
Sarajevo,
Bosnia and Herzegovina
mirza.pasic@mef.unsa.ba

Izet Bijelonja

Mechanical engineering
faculty Sarajevo, University
of Sarajevo
Sarajevo,
Bosnia and Herzegovina
bijelonja@mef.unsa.ba

