



OPEN

Artificial intelligence accuracy assessment in NO₂ concentration forecasting of metropolises air

Seyedeh Reyhaneh Shams¹, Ali Jahani^{2✉}, Saba Kalantary³, Mazaher Moeinaddini⁴ & Nematollah Khorasani⁴

Air quality has been the main concern worldwide and Nitrous oxide (NO₂) is one of the pollutants that have a significant effect on human health and environment. This study was conducted to compare the regression analysis and neural network model for predicting NO₂ pollutants in the air of Tehran metropolis. Data has been collected during a year in the urban area of Tehran and was analyzed using multi-linear regression (MLR) and multilayer perceptron (MLP) neural networks. Meteorological parameters, urban traffic data, urban green space information, and time parameters are applied as input to forecast the daily concentration of NO₂ in the air. The results demonstrate that artificial neural network modeling ($R^2 = 0.89$, RMSE = 0.32) results in more accurate predictions than MLR analysis ($R^2 = 0.81$, RMSE = 13.151). According to the result of sensitivity analysis of the model, the value of park area, the average of green space area and one-day time delay are the crucial parameters influencing NO₂ concentration of air. Artificial neural network models could be a powerful, effective and suitable tool for analysis and modeling complex and non-linear relation of environmental variables such as ability in forecasting air pollution. Green spaces establishment has a significant role in NO₂ reduction even more than traffic volume.

High levels of air pollution is presently one of the significant environmental issues and problems in many countries. Pollutants are a mixture of natural and human-made sources and many of them pass over the standard limit^{1,2}. Air pollutants emission has had a considerable influence on the increase in mortality rates, so healthy air is one of the main issues for human and environment^{3,4}. Nitrogen dioxide (NO₂) is one of the important families of air-polluting chemical compounds worldwide. NO₂ gas is emitted from indoor sources (gas stoves, vented gas heaters, and smoking) and outdoor sources (high-temperature combustion process, mobile sources such as vehicles and traffic, and industrial processes)^{5,6}. Moreover, a mainly part of NO₂ concentration linked to secondary production of NO₂ in air through photochemical processes. Most nitrogen oxides are NO, but this gas is rapidly oxidized to NO₂ in the presence of O₃ ozone⁷. NO₂ could lead to many unfavorable impacts on human health, environment and biological ecosystems such as acid rain, ozone layer reduction, and global warming⁸. Epidemiological studies show that high-level exposure to NO₂ in the air as a pollutant lead to increasing approximately 5–7% of the lung cancers among people (ex-smokers and non-smokers)⁹. Moreover, these studies reveal that a positive correlation between air pollutants such as NO₂ and adverse effects on human health such as changes in the kidney, liver, heart, and decrease resistance immunity to infectious diseases^{10–12}. World Organization Health (WHO) reported that NO₂ associated with adverse effects such as increases in respiratory symptoms, asthma prevalence and incidence, cancer incidence, adverse birth outcomes and mortality¹³. Air pollution is a huge challenge in metropolitans such as Tehran city as the capital of Iran. The overcrowded population of Tehran and the increasing number of vehicles as well as the concentration of industries are causes of air pollution, in the past two decades¹⁴. It is estimated that, between March of 2016 and 2017, 5% of days are classified as a clean air day, 71% as moderate clean air day, 22% unhealthy air for the sensitive group, and 2% unhealthy air for all people¹⁵. There are increasing levels of NO₂ pollution in many countries, especially in metropolises such as Tehran. Therefore forecast, controlling and counteracting NO₂ is a vital issue in urban management⁴.

Artificial neural networks (ANNs) are mathematical models that can be used for complex and non-linear processes^{16,17}. ANNs can simulate the behavior of the human brain^{18,19}. In recent years, ANNs have been

¹Department of Environmental Pollution, Faculty of Environment, College of Environment, Karaj, Iran. ²Research Center of Environment and Sustainable Development and College of Environment, Tehran, Iran. ³Department of Occupational Health Engineering, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran. ⁴Department of Environment, Faculty of Natural Resources, Tehran University, Karaj, Iran. ✉email: ajahani@ut.ac.ir

successfully used for predicting and modeling the concentration of ambient air pollution. Furthermore, ANNs are widely used in short- and long-term applications for forecasting pollutants²⁰. Various studies applied the ANNs for pollution prediction without considering a variety of environmental factors such as green space and parks, which is recognized as a shortcoming^{4,21–23}. Regression analysis is a traditional statistical technique for model generation. Multi-linear regression (MLR) analysis is an approach to evaluate the relationship between independent and dependent factors. The goal of this investigation is the prediction of the concentration levels of NO₂ as a factor determining the atmospheric pollution in Tehran city. For this purpose prediction accuracy, MLR and ANN models are selected to be compared. Many environmental variables were used to support model outputs, and finally, we identified the most accurate model for prediction of the concentration levels of NO₂ in Tehran. The result of model sensitivity analysis was illustrated to prioritize model variables.

Material and methods

Data collection. To perform this research, data from the Tehran Air Quality Control Company, Tehran Meteorological Organization, Tehran Transportation and Traffic Organization and urban green space information were collected. The parameters measured in these organizations were included five groups: the concentration of NO₂, urban traffic parameters (using Tehran traffic cameras data), urban green space information, meteorological data, and time parameters. The urban traffic parameters were including the length of the north–south (y) and east–west (x) cross streets (km), the average number of vehicles at the street intersections, and total number of vehicles at the intersections. The urban green space information were the number of parks, the total area of parks in each district, an average of the distance between each park and NO₂ monitoring station and park index adjacent to the NO₂ monitoring station (i.e. the area of the nearest park to the NO₂ monitoring station divided by the distance of the nearest park form NO₂ monitoring station) and the area of green spaces in district. Furthermore, meteorological data such as air temperature, rainfall, wind speed and direction, humidity, air pressure, and the length of sunshine per day, as well as time parameters such as 1-day time delay (NO₂ concentration in the last day), 2-day delay time (NO₂ concentration in the past two days), the day of the year (1–365), the desired month (1–30), the season (1: Spring–2: Summer–3: Autumn–4: Winter) and the warm and cold seasons (1: hot–2: cold) were considered as main variables affecting the air pollution of the Tehran. Data were collected for 1 year (2015). The stations of NO₂ monitoring and meteorological stations were located close to each other and in one area, as well as the traffic information was gathered on the streets near the stations.

Multiple linear regression model (MLR). Multi-linear regression analysis (MLR) was implemented in the form of stepwise. This model was used to examine the relationship between the daily concentration of NO₂ of air as dependent data and influential variables as independent data. The most accurate MLR equation was obtained based upon statistical parameters including the highest correlation coefficient, the lowest mean square error root (RMSE) and the number of descriptors in the model (n), and the greater value of the F statistic^{2,24,25}. The R statistic parameter shows the accuracy of the regression line and the greater value of R can better fit between observed and predicted values^{26,27}.

Neural network model. In this study, the daily concentration of NO₂ of air was predicted using an artificial neural network model. The multilayer perceptron (MLP) architecture has been used successfully to model some difficult and diverse problems and nonlinear functions such as air quality forecasting. It is composed of a system of layered and interconnected neurons or nodes, namely, an input, one or more hidden layers, and output layers^{15,28,29}. In the current study, logarithmic sigmoid and linear activation functions were examined to optimize the network. The back-propagation (BP) training algorithm is found to be the most common and powerful nonlinear statistical technique in MLP networks^{29,30}. All computations were developed in MATLAB R2016b software. In learning process to detection the performance of designed the neural network model, the following statistical indicators such as correlation coefficient (R²) (Eq. 1), mean absolute error (MAE) (Eq. 2), mean square error (MSE) (Eq. 3), and root means square error (RMSE) (Eq. 4) were calculated^{31,32}:

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

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

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}, \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad (4)$$

where y_i and \hat{y}_i are the targets and network outputs, \bar{y}_i is the mean of target values, and n is the number of samples, respectively.

R ²	RMSE	F	Sig
0.816	13.151	605.079	0.000

Table 1. Multivariate regression model results in forecasting NO₂ concentration in Tehran city.

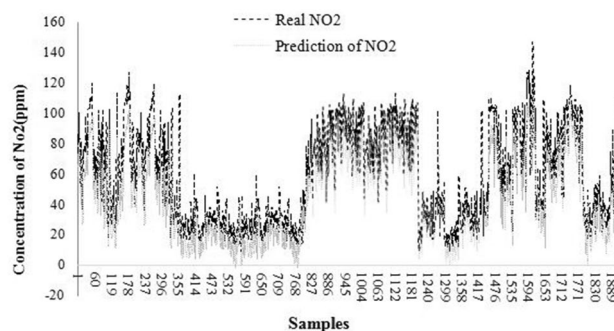


Figure 1. The result of stepwise-multi linear regression in forecasting the NO₂ concentration of air.

To determine the most affecting factors on the model output, a sensitivity analysis was performed. In this method, each factor was changed in the range of the standard deviation, while other factors were equal to their value of the average. The standard deviation outputs of the model for each factor change were calculated as the sensitivity of the model finding for that factor.

Results

The study area. Tehran metropolis, the capital of Iran, with a population of over 8,000,000 individuals, is located in the south of the Alborz mountains and the northern margin of Iran's central desert. Its geographical longitude is 51° 2' E to 51° 36' E with an approximate length of 50 km. Its geographical latitude is from 35° 35' N to 35° 50' N with an approximate width of 30 km. The altitude of this city at the northern point is about 1800 m and at the southernmost point is about 950 m above sea level. This city has been extended to 730 km^{24,33}.

MLR model. According to the result of the stepwise-multi linear regression analysis, the desired month is from the desired year, length of the north–south (y) and east–west (x) street, wind speed and direction, rainfall, the air temperature, the humidity, the length of sunshine per day, the one-day time delay and the two-day time delay have a significant impact on the predicted daily NO₂ concentrations in Tehran.

There is a relation between the rate of change of affecting factors and the air NO₂ concentration in Tehran. The best result for performance stepwise-multi linear regression in predicting the NO₂ concentration of air set out in Table 1 and Eq. (5).

$$NO_2 = 0.210X_1 - 0.947X_2 - 0.409X_3 - 0.05X_4 - 0.851X_5 - 2.202X_6 - 0.144X_7 - 0.072X_8 + 0.294X_9 + 0.691X_{10} + 0.212X_{11} + 10.031. \quad (5)$$

where X_1 is the desired month is from the desired year, X_2 is the length of the east–west (x) street, X_3 is the length of the north–south (y) street, X_4 is wind direction, X_5 is the wind speed, X_6 is the rainfall, X_7 is the air temperature, X_8 is the humidity, X_9 is the length of sunshine per day, X_{10} is the 1-day time delay, and X_{11} is the 2-day time delay, respectively.

The findings obtained from stepwise-multi linear regression in forecasting the NO₂ concentration of air are shown in Fig. 1.

ANN modeling. Before neural network training, the factors influencing air pollution, as input data need to be normalized so that the data were converted to numbers between 1 and – 1. Data were randomly divided into three subsets: 20% samples selected for validation, 20% as test sets and 60% as training tests and then input data were weighed in the first layer and moved to the middle layer. Next, the outputs were weighted through connections between the middle layer and the output layer. Finally, the findings were presented in the output layer. The most accurate structure of neural networks optimized using 27 neurons and one hidden layer. The optimal structural characteristics of the neural network are presented in Table 2.

The maximum value of R² as well as the minimum amount of MSE in the test set and train set considered (Table 3). Indeed it shows a very high level of neural network accuracy in forecasting the NO₂ concentration of air in the city of Tehran.

The scatter plot represents a correlation between data that show the accuracy of a neural network model for forecasting the concentration of NO₂ in Tehran³⁴. The scatter plot of the neural network (MLP) output versus target values of the forecasting the NO₂ concentration of air for training, validation, test, and all data are shown

Network structural features	Hidden layer	Output layer
Type of network	Multilayer perceptron	Multilayer perceptron
Transfer function	Sigmoid tangent	Linear
Optimal algorithm	Conjugate gradient	Conjugate gradient
Learning period	300	300
Momentum coefficient	0.9	0.9
Number of neurons	26	1
Normalization	- 1 to 1	- 1 to 1

Table 2. The optimal structure of an artificial neural network for the forecast of NO₂ concentration.

Function			Test set				Train set			
activation	Training	Structure	R ²	MSE	RMSE	MAE	R ²	MSE	RMSE	MAE
Logsig-Logsig-Purelin	LM	23-26-26-1	0.741	0.235	0.485	0.355	0.879	0.125	0.354	0.252
Tansig-Tansig-Purelin	GDM	23-22-22-1	0.562	0.386	0.621	0.468	0.564	1.631	1.277	0.521
Tansig-Purelin	BQN	23-26-1	0.763	0.208	0.456	0.327	0.836	0.163	0.403	0.278
Tangsig-Purelin	CGB	23-26-1	0.742	0.232	0.481	0.334	0.877	0.122	0.35	0.25
Tansig-Tansig-Purelin	GDA	23-19-19-1	0.722	0.246	0.496	0.341	0.821	0.181	0.425	0.294
Logsig-Logsig-Purelin	CGP	23-24-24-1	0.711	0.254	0.504	0.353	0.829	0.172	0.414	0.288

Table 3. The performance of the best artificial neural networks for the forecast of NO₂ concentration in the air. *LM* Levenberg–Marquardt, *GDM* gradient descent with momentum, *BQN* BFGS quasi newton, *CGB* conjugate gradient with Powell/Beale, *GDA* gradient descent with adaptive learning rate, *CGP* conjugate gradient with Polak/Ribiere.

in Fig. 2. As can be seen from Fig. 2, the value of the coefficient (R^2) proves the relatively high correlation between output and target values.

Sensitivity analysis of NO₂ concentration. Sensitivity analysis aims to evaluate the most important input parameters affecting on the model output. As can from Fig. 3, the value of the park area, an average of green space and one-day time delay are detected as the most influential inputs that influence the NO₂ concentration in the air.

Figure 4 shows the effect of varying most important parameters including park area, area of green space, and one-day time delay on the level of NO₂ concentration.

Discussion

The objective of this study was to forecast the air NO₂ concentration in Tehran metropolis by applying MLR and MLP models. The comparison between MLR and artificial neural network modeling demonstrates that the neural network model ($R^2 = 0.89$, $RMSE = 0.32$) performs more accurately than multiple regression analysis ($R^2 = 0.81$, $RMSE = 13.151$). The MLR analysis had a lower R^2 value than the MLP model. There is no doubt that air pollution is influenced by several factors including meteorological parameters, pollutant sources, green spaces quality, and quantity, etc. which require a non-linear computing tool as alternatives to the traditional approach. Artificial Neural Network (ANN) models could be used as interpolation methods for complex and non-linear problems such as predicting and modeling air pollution^{17,35,36}. Several research exhibits the performance of ANNs and traditional regression models to predict air pollutant concentrations^{19,37–39}. Dragomir et al. compared the multiple linear regression and multilayer perceptron for the forecast of the NO₂ concentration in Romania. This study focuses on the dependence between meteorological factors including temperature, pressure, wind speed, wind direction, solar radiation, rainfall, and relative humidity and their influence on measured NO₂ concentration. The results indicated that MLP has a higher correlation coefficient than MLR in the forecasting of air quality and meteorological factors have an impact upon the NO₂ concentration³⁸. Rahimi used MLP and MLR for prediction of the NO₂ and NO_x concentrations according to meteorological variables. The best structure for the MLR model RMSE and R^2 were 3.6 and 0.42 respectively in comparison to the MLP model with $RMSE = 0.0046$ and $R^2 = 0.82$. The findings in this study shows that performance of MLP is superior in comparison to MLR for prediction NO₂ concentrations in urban environments³⁹. Cabaneros et al. illustrated that MLP models can accurately forecast NO₂ concentrations ($R^2 = 0.9$, $RMSE = 23.45$) based on air pollution and meteorological data¹⁹. Cakir and Sita (2020) developed a non-linear method (MLP) and MLR model utilizing air pollution and meteorological data to predict concentration values of NO₂. The results demonstrate that there are no significant differences between two methods. They recommended that more studies be done in this area⁴⁰. Although there are several studies in this area, however, they did not formulate comprehensive influential variables with green space and traffic data. To determine the daily averaged concentration of PM10, and PM2.5 in the Adriatic coast of Italy, Biancofiore et al. compared ANN and MLR models. This research was used, as input, daily values of air temperature, humidity,

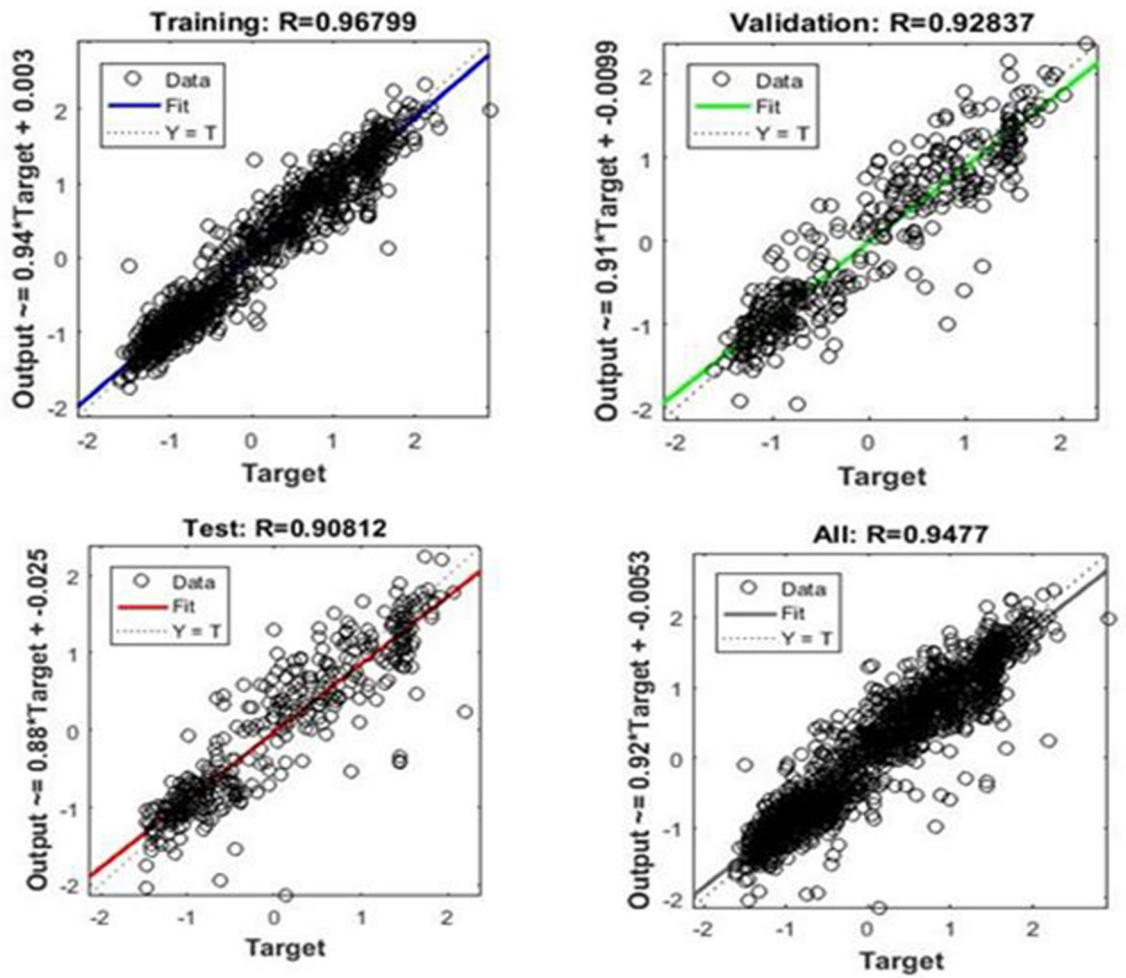


Figure 2. Scatter plots related to data modeling results. Training data, validation data, test data, and total data.

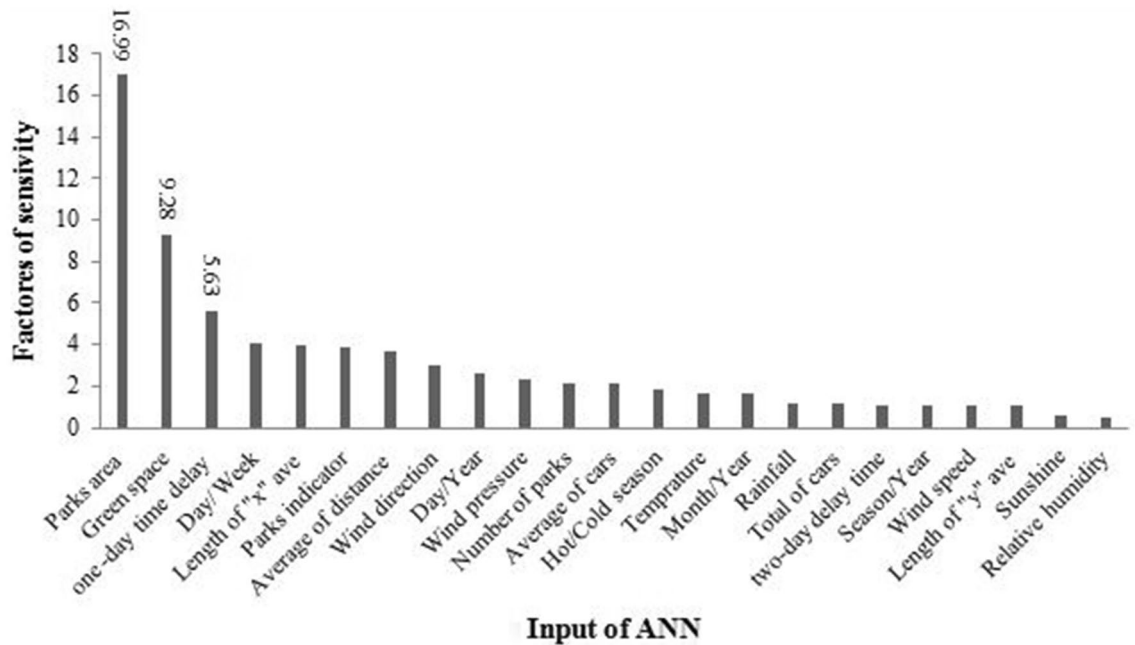


Figure 3. Sensitivity analysis of affecting parameters on NO₂ prediction.

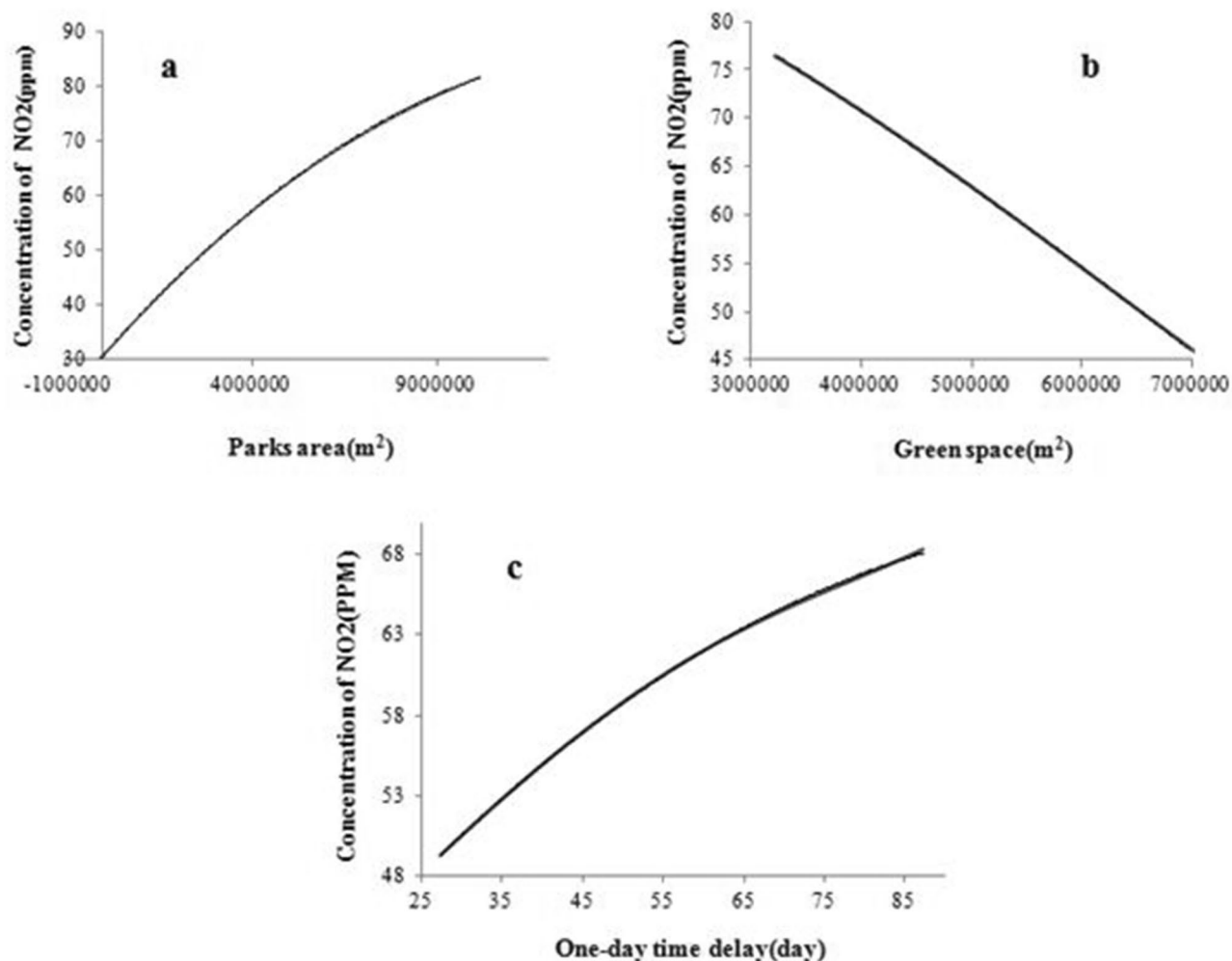


Figure 4. The trend of NO₂ prediction output changes with varying (a) parks area (b) green space. (c) One-day time delay.

air pressure, wind speed, and wind direction. The neural network model has the best correlation coefficient ($R=0.88$) and performance compared to the MLR model ($R=0.86$) in forecasting air pollution⁴¹. Alimissis et al. worked on an air pollution prediction model based on artificial neural networks and multiple linear regression in the greater area of metropolitan Athens in Greece. The selected air pollutants were NO₂, NO, O₃, CO, and SO₂. The findings for each five air pollutants highlighted that neural network models to be significantly superior compared to MLR owing to their ability to effectively predict the complex and nonlinear issues such as air pollution⁴². Although implementation of regression methods is simple, however, the results of various studies also exhibit that regression methods may not offer precise predictions in some areas such as air pollution in comparison to ANNs models; and ANNs have generally superior performance^{39,43,44}. The comprehensive data collection including green space and traffic data is the most advantages of our study. As a result, our ANN model is more accurate than other developed models in research. It is concluded that green space and traffic variables, especially in metropolitans, should be considered to achieve more accurate models.

As can from Fig. 3, green spaces and parks have a key role in the prediction of NO₂ air. Figure 4A presents that there is a positive correlation between the concentration of NO₂ and the parks area. A reason for this result could be the fact that the vegetation cover of parks especially in metropolises are so limited to grasses and bushes and lack of trees is one of the disadvantages. Hence, more human activities around urban parks cause more air pollution. On the other hand, urban green spaces (vegetation in parks, squares, street tree lines, etc.) play a significant role in NO₂ reduction in the air. This observation is in agreement with those of the previous studies^{45–50}. Slemi et al. provided that tree cover and the level of air pollutant concentrations were two key parameters in removing pollutants⁴⁷. Janhall showed that differently designed vegetation and vegetation cover dispersion impact on air quality. Moreover, trees and vegetation should be high and porous enough to allow air to pass through them because the air that passes not through vegetation is not filtered; and vegetation should be as close to sources of contaminants as possible⁴⁵. Vos et al. indicated that trees could result in a higher concentration of NO₂. This result may be explained by the fact that vegetation cover may lead to obstructing the wind flow. This means that the aerodynamic effect decreases ventilation and filtering capacity of vegetation⁵¹. As one can see Fig. 4B, there is a negative correlation between the concentration of NO₂ and area green space, so the concentration of NO₂

decreases with the increased area green space. Selmi et al. demonstrate the contribution of trees in public green spaces in air pollution removal in Strasbourg city, France. The results showed that trees removed about 88 tons of pollutants during a year. The amount of air pollution removed by trees was approximately 14 tons for NO₂, 1 ton for CO, 56 tons for O₃, 12 tons for PM₁₀, 5 tons for PM_{2.5}, and 1 ton for SO₂⁴⁸. In the United States, urban vegetation is estimated to occupy 3.5% of the land, which could lead to the absorption of 0.711 ton of pollutant. This amount is equivalent to 3.8 billion US dollars. Urban green spaces are often regarded as a substantial air purification service⁵². Although, the tree cover and the level of air pollutant concentrations were two key parameters in removing pollutants⁴⁷. Moreover, we concluded that green spaces has a significant role in NO₂ reduction even more than traffic variables.

Conclusion

Atmospheric pollution is now recognized as a permanent concern and a major environmental issue at the global level with adverse health worldwide. Because air pollution can effect on human health, quality of life and environment, the predicting of air pollution values has received much attention number of researchers due to closely relate it health of people and environment. The results showed an accurate prediction approach from the field of data processing techniques and artificial neural network model could be a powerful, effective and suitable tool for analysis and modeling complex and non-linear relation of environmental variables such as ability in forecasting air pollution. Green spaces establishment has a significant role in NO₂ reduction even more than traffic volume. It can thus be suggested developing different ANNs techniques survey for ambient air pollutants prediction and forecasting. It's worthwhile in the aspect of environmental, health, economic aims and useful for governmental authorities and decision-makers.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 8 September 2020; Accepted: 7 January 2021

Published online: 19 January 2021

References

- Pawul, M. & Śliwka, M. Application of artificial neural networks for prediction of air pollution levels in environmental monitoring. *J. Ecol. Eng.* <https://doi.org/10.12911/22998993/64828> (2016).
- Djebbari, N., Rouainia, M. Artificial neural networks based air pollution monitoring in industrial sites. In *2017 International Conference on Engineering and Technology (ICET)* 1–5. (IEEE, 2017).
- Ghadi, M. E., Qaderi, F. & Babanezhad, E. Prediction of mortality resulted from NO₂ concentration in Tehran by Air Q+ software and artificial neural network. *Int. J. Environ. Sci. Technol.* **16**(3), 1351–1368. <https://doi.org/10.1007/s13762-018-1818-4> (2019).
- Beigzadeh, Z. et al. Evaluation of pulmonary function of construction workers in Tehran city in 2017. *Health Saf. Work.* **8**(4), 323–332 (2019).
- Cyrys J, Heinrich J, Richter K, Wölke G, Wichmann HE, Group IS, for Environment GN. Sources and concentrations of indoor nitrogen dioxide in Hamburg (west Germany) and Erfurt (east Germany). *Sci. Total Environ.* **250**(1–3), 51–62. [https://doi.org/10.1016/S0048-9697\(00\)00361-2](https://doi.org/10.1016/S0048-9697(00)00361-2). (2000)
- Mohammadi, M. J. et al. Dispersion modeling of nitrogen dioxide in ambient air of Ahvaz city. *Health Scope* <https://doi.org/10.17795/jhealthscope-32540> (2016).
- Mavroidis, I. & Chaloulakou, A. Long-term trends of primary and secondary NO₂ production in the Athens area. Variation of the NO₂/NO_x ratio. *Atmos. Environ.* **45**(38), 6872–6879. <https://doi.org/10.1016/j.atmosenv.2010.11.006> (2011).
- Boningari, T. & Smirniotis, P. G. Impact of nitrogen oxides on the environment and human health: Mn-based materials for the NO_x abatement. *Curr. Opin. Chem. Eng.* **13**, 133–141. <https://doi.org/10.1016/j.coche.2016.09.004> (2016).
- Khaniabadi, Y. O. et al. Exposure to PM₁₀, NO₂, and O₃ and impacts on human health. *Environ. Sci. Pollut. Res.* **24**(3), 2781–2789. <https://doi.org/10.1007/s11356-016-8038-6> (2017).
- Hadei, M. et al. Estimation of mortality and hospital admissions attributed to criteria air pollutants in Tehran Metropolis, Iran (2013–2016). *Aerosol. Air Qual. Res.* **17**, 2474–2481. <https://doi.org/10.4209/aaqr.2017.04.0128> (2017).
- Nasiri, E. F., Kebria, D. Y., Qaderi, F. An experimental study on the simultaneous phenol and chromium removal from water using titanium dioxide photocatalyst. *Civ. Eng. J.* <https://doi.org/10.28991/cej-0309117>. (2018)
- Nordic Expert Meeting (1995: Oslo, Norway), Clench-Aas, J, Krzyzanowski, M. & World Health Organization. Regional Office for Europe. Quantification of health effects related to SO₂, NO₂, O₃ and particulated matter exposure: Report from the Nordic Expert Meeting, Oslo, 15–17 October 1995. (WHO Regional Office for Europe, Copenhagen, 1995).
- WHO (2013) Review of evidence on health aspects of air pollution-REVIHAAP project: final technical report. World Health Organization Regional Office for Europe. <http://www.euro.who.int/en/health-topics/environment-andhealth/airquality/publications/2013/review-of-evidence-on-health-aspects-of-air-pollution-revihaap-project-final-technical-report>. Accessed Feb 2015.
- Golbaz, S., Farzadkia, M. & Kermani, M. Determination of Tehran air quality with emphasis on air quality index (AQI); 2008–2009. *Iran Occup. Health.* **6**(4), 62–68 (2010).
- Hassanpour Matikolaei, S. A., Jamshidi, H. & Samimi, A. Characterizing the effect of traffic density on ambient CO, NO₂, and PM_{2.5} in Tehran, Iran: An hourly land-use regression model. *Transp. Lett.* **11**(8), 436–446. <https://doi.org/10.1080/19427867.2017.1385201> (2019).
- Jahani, A. & Saffariha, M. Human activities impact prediction in vegetation diversity of lar national park in Iran using artificial neural network model. *J. Int. Env. Ass. Man.* **17**(1), 42–52. <https://doi.org/10.1002/ieam.4349> (2021).
- Mishra, D. & Goyal, P. Development of artificial intelligence based NO₂ forecasting models at Taj Mahal. *Agra. Atmos. Pollut. Res.* **6**(1), 99–106. <https://doi.org/10.5094/APR.2015.012> (2015).
- Jahani, A., Goshtasb, H. & Saffariha, M. Tourism impact assessment modeling in vegetation density of protected areas using data mining techniques. *Lan. Deg. & Dev.* **31**(12), 1502–1519. <https://doi.org/10.1002/ldr.3549> (2020).
- Cabaneros, S. M., Calautit, J. K. & Hughes, B. R. Hybrid artificial neural network models for effective prediction and mitigation of urban roadside NO₂ pollution. *Energy Proc.* **142**, 3524–3530. <https://doi.org/10.1016/j.egypro.2017.12.240> (2017).
- Cabaneros, S. M., Calautit, J. K. & Hughes, B. R. A review of artificial neural network models for ambient air pollution prediction. *Environ. Model Softw.* **119**, 285–304 (2019). <https://doi.org/10.1016/j.envsoft.2019.06.014> (2019).

21. Zallaghi, E., Goudarzi, G., Haddad, M. N., Moosavian, S. M. & Mohammadi, M. J. Assessing the Effects of nitrogen dioxide in urban air on health of west and southwest cities of Iran. *Jundishapur J. Health Sci.* <https://doi.org/10.5812/jjhs.23469> (2014).
22. Johansson, C. *et al.* Impacts on air pollution and health by changing commuting from car to bicycle. *Sci. Total Environ.* **584**, 55–63. <https://doi.org/10.1016/j.scitotenv.2017.01.145> (2017).
23. Pourbabaki, R., Karimi, A. & Yazdanirad, S. Modeling the consequences and analyzing the dangers of carbon disulfide emissions using ALOHA software in an oil refinery. *J. Health Field.* **6**(3), 24199 (2019).
24. Akan, R. & Keskin, S. N. The effect of data size of ANFIS and MLR models on prediction of unconfined compression strength of clayey soils. *SN Appl. Sci.* **1**(8), 843. <https://doi.org/10.1007/s42452-019-0883-8> (2019).
25. Jahani, A. Forest landscape aesthetic quality model (FLAQM): A comparative study on landscape modelling using regression analysis and artificial neural networks. *J. For. Sci.* **65**(2), 61–69. <https://doi.org/10.17221/86/2018-JFS> (2019).
26. Jahani, A. & Mohammadi, F. A. Aesthetic quality modeling of landscape in urban green space using artificial neural network. *J. Nat. Environ. (Iran. J. Nat. Resour.)* **69**(4), 951–963 (2017).
27. Kalantary, S., Jahani, A., Pourbabaki, R. & Beigzadeh, Z. Application of ANN modeling techniques in the prediction of the diameter of PCL/gelatin nanofibers in environmental and medical studies. *RSC Adv.* **9**(43), 24858–24874. <https://doi.org/10.1039/C9RA04927D> (2019).
28. Saffariha, M., Jahani, A. & Potter, D. Seed germination prediction of *Salvia limbata* under ecological stresses in protected areas: An artificial intelligence modeling approach. *BMC Eco.* **20**, 48–58. <https://doi.org/10.1186/s12898-020-00316-4> (2020).
29. Akbarifard, S., Sharifi, M. R. & Qaderi, K. Data on optimization of the Karun-4 hydropower reservoir operation using evolutionary algorithms. *Data Brief.* <https://doi.org/10.1016/j.dib.2019.105048> (2020).
30. Qaderi, K., Akbarifard, S., Madadi, M. R., Bakhtiari, B. Optimal operation of multi-reservoirs by water cycle algorithm. In *Proceedings of the Institution of Civil Engineers-Water Management*, **171**(4), 179–190. (Thomas Telford Ltd, 2018). <https://doi.org/10.1680/jwama.16.00034>.
31. Samiei, S. & Pourbabaki, R. Risk factors of low back pain using adaptive neuro-fuzzy. *Arch. Occup. Health.* **3**(2), 339–345 (2019).
32. Alefi, M., Sadeghi Yarandi, M. & Karimi, A. Modeling of occupational risk factors in the development of musculoskeletal disorders in nurses. *Arch. Occup. Health.* **4**(1), 474–479 (2020).
33. Jahani, A. & Rayegani, B. Forest landscape visual quality evaluation using artificial intelligence techniques as a decision support system. *Stoc. Env. Res. Ris. Ass.* **34**(10), 1473–1486. <https://doi.org/10.1007/s00477-020-01832-x> (2020).
34. Jahani, A. & Saffariha, M. Aesthetic preference and mental restoration prediction in urban parks: An application of environmental modeling approach. *Urb. For. Urb. Green* **54**, 126775. <https://doi.org/10.1016/j.ufug.2020.126775> (2020).
35. Meng, X. *et al.* A land use regression model for estimating the NO₂ concentration in Shanghai, China. *Environ. Res.* **137**, 308–315. <https://doi.org/10.1016/j.envres.2015.01.003> (2015).
36. Shams, S. R., Jahani, A., Moeinaddini, M. & Khorasani, N. Air carbon monoxide forecasting using an artificial neural network in comparison with multiple regression. *Model. Earth Syst. Environ.* **6**(3), 1–9 (2020).
37. Wang, P., Liu, Y., Qin, Z. & Zhang, G. A novel hybrid forecasting model for PM10 and SO₂ daily concentrations. *Sci. Total Environ.* **505**, 1202–1212. <https://doi.org/10.1016/j.scitotenv.2014.10.078> (2015).
38. Dragomir, C. M., Voiculescu, M., Constantin, D. E. & Georgescu, L. P. Prediction of the NO₂ concentration data in an urban area using multiple regression and neuronal networks. *AIP Conf. Proc.* **1694**(1), 040003 (2015).
39. Rahimi, A. Short-term prediction of NO₂ and NO_x concentrations using multilayer perceptron neural network: A case study of Tabriz. *Iran. Ecol. Process.* **6**(1), 4. <https://doi.org/10.1186/s13717-016-0069-x> (2017).
40. Cakir, S. & Sita, M. Evaluating the performance of ANN in predicting the concentrations of ambient air pollutants in Nicosia. *Pollut. Res. Atmos.* <https://doi.org/10.1016/j.apr.2020.06.011> (2020).
41. Biancifiore, F. *et al.* Recursive neural network model for analysis and forecast of PM10 and PM2.5. *Atmos. Pollut. Res.* **8**(4), 652–659. <https://doi.org/10.1016/j.apr.2016.12.014> (2017).
42. Alimissis, A., Philippopoulos, K., Tzani, C. G. & Deligiorgi, D. Spatial estimation of urban air pollution with the use of artificial neural network models. *Atmos. Environ.* **191**, 205–213. <https://doi.org/10.1016/j.atmosenv.2018.07.058> (2018).
43. Unnikrishnan, R. & Madhu, G. Comparative study on the effects of meteorological and pollutant parameters on ANN modelling for prediction of SO₂. *SN Appl. Sci.* **1**, 1394. <https://doi.org/10.1007/s42452-019-1440-1> (2019).
44. Ul-Saufie, A. Z., Yahya, A. S., Ramli, N. A. & Hamid, H. A. Comparison between multiple linear regression and feed forward back propagation neural network models for predicting PM10 concentration level based on gaseous and meteorological parameters. *Int. J. Appl. Sci. Technol.* **1**(4), 42–49 (2011).
45. Janhäll, S. Review on urban vegetation and particle air pollution—Deposition and dispersion. *Atmos. Environ.* **105**, 130–137. <https://doi.org/10.1016/j.atmosenv.2015.01.052> (2015).
46. Rahmati, O. *et al.* Hybridized neural fuzzy ensembles for dust source modeling and prediction. *Atmos. Environ.* **224**, 117320. <https://doi.org/10.1016/j.atmosenv.2020.117320> (2020).
47. Sarmazeh, N. S., Jahani, A. & Goshtasb, H. Ecological impacts assessment of recreation on quality of soil and vegetation in protected areas (case study: Qhamishloo national park and wildlife refuge). *J. Nat. Environ.* **70**(4), 881–891 (2018).
48. Selmi, W. *et al.* Air pollution removal by trees in public green spaces in Strasbourg city, France. *Urban For Urban Gree.* **17**, 192–201. <https://doi.org/10.1016/j.ufug.2016.04.010> (2016).
49. Jahani, A., Khalegh panah, R., Goshtasb, H. & Khorasani, N. Assessing users' recreational demand in urban parks in Tehran with the help of the artificial neural network. *Soc. Stud. Tour.* **14**(7), 339–362 (2020).
50. Jahani, A. Risks assessment of forest project implementation in spatial density changes of forest under canopy vegetation using artificial neural network modeling approach. *J. Spat. Anal. Environ. Hazards.* **6**(2), 21–34 (2019).
51. Vos, P. E., Maiheu, B., Vankerkom, J. & Janssen, S. Improving local air quality in cities: to tree or not to tree?. *Environ. Pollut.* **183**, 113–122. <https://doi.org/10.1016/j.envpol.2012.10.021> (2013).
52. Xing, Y. & Brimblecombe, P. Role of vegetation in deposition and dispersion of air pollution in urban parks. *Atmos. Environ.* **201**, 73–83. <https://doi.org/10.1016/j.atmosenv.2018.12.027> (2019).

Acknowledgements

We would also like to thank the Tehran Air Quality Control Company, the Tehran City Transportation and Traffic Organization, and the Meteorological Organization for their cooperation in providing the information required for this research. May this smallest appreciation to some of their efforts.

Author contributions

A.J. designed the experiments, analyzed the data and S.K. wrote the paper. R.S. performed the experimental measurement. A.J., R.S. conceived the mathematical model and ANN Modeling Techniques A.J., M.M., and N.K. conceived the idea and designed the experiment. A.J., M.M., and N.K. supervised research. All authors discussed the results and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021