Iran J Public Health, Vol. 51, No.1, Jan 2022, pp.196-204

Original Article

Spatio-Temporal Modeling of Ozone Distribution in Tehran, Iran Based on Neural Network and Geographical Information System

*Leila Sherafati [1], *Hossein Aghamohammadi Zanjirabad [1], Saeed Behzadi [2]*

1. *Department of Remote Sensing and GIS, Faculty of Natural resources and Environment, Science and Research Branch, Islamic Azad University, Tehran, Iran*
2. *Department of Surveying Engineering, Faculty of Civil Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran*

***Corresponding Author:** Email: aghamohammadi@srbiau.ac.ir

(Received 12 Dec 2020; accepted 15 Feb 2021)

Abstract

Background: Air pollution is one of the most important causes of respiratory diseases that people face in big cities today. Suspended particulates, carbon monoxide, sulfur dioxide, ozone, and nitrogen dioxide are the five major pollutants of air that pose many problems to human health. We aimed to provide an approach for modeling and analyzing the spatiotemporal model of ozone distribution based on Geographical Information System (GIS).

Methods: In the first step, by considering the accuracy of different interpolation methods, the Inverse distance weighted (IDW) method was selected as the best interpolation method for mapping the concentration of ozone in Tehran, Iran. In the next step, according to the daily data of Ozone pollutants, the daily, monthly, and annual mean concentrations maps were prepared for the years 2015, 2016, and 2017.

Results: Spatial and temporal analysis of the distribution of ozone pollutants in Tehran was performed. The highest concentrations of O_3 are found in the southwest and parts of the central part of the city. Finally, a neural network was developed to predict the amount of ozone pollutants according to meteorological parameters.

Conclusion: The results show that meteorological parameters such as temperature, velocity and direction of the wind, and precipitation are influential on O_3 concentration.

Keywords: Spatial analysis; Neural networks, Computer; Geographic information systems (GIS); Ozone

Introduction

Industrial and technological evolution, urban development, population growth, motor vehicle incensement, increasing the consumption of petroleum products, and in some cases specific regional and geographical conditions, all of caused air pollution to increase (1). Ozone is a secondary-level pollutant, and it is a very strong oxidizer produced by photochemical interactions of nitrogen oxides and hydrocarbons in the presence of sunlight (2). There is a direct relationship between heart and lung disease and exposure to pollutants. Therefore, increasing these pollutants has become a major challenge for managers and planners especially in metropolitan areas such as

Tehran (3). Air quality is a particularly important environmental issue, especially in metropolises, so reliable modeling is essential for investigating its behavior (4). Tehran's O_3 pollutants are in favorable condition between 2015 and 2017, and only a few days of which were in unhealthy conditions for sensitive groups. The number of "clean days" has decreased in recent years, on the other hand, the number of "healthy days" has also increased. The amount of ozone pollutants was in a clean state on most days of spring, fall, and winter. As O_3 is produced in the presence of sunlight, the concentration of this pollutant is increased in the warm months of the year, and the highest Air Quality Index (AQI) values are seen in the warm days of the year (5).

In recent years, many studies have been conducted to discuss air pollution. Some of these researches are discussed as follows.

Masoudi et al (6) calculated the relationships between the concentration of pollutants and meteorological parameters were expressed by multiple linear and nonlinear regression equations for both annual and seasonal conditions using SPSS software. Wang et al (7) used a series of monthly hybrid spatiotemporal land-use regression (LUR) (Two-stage) models consisting of two submodels based on the multiple linear regression (MLR) algorithm, general additive mixed models (GAMMs), and land use random forest (LURF) models.

On the other hand, Zhang et al (8) explored the sensitivity of ozone predictions from photochemical grid point simulations to small meteorological initial perturbations that are realistic in structure and evolution. Through both meteorological and photochemical ensemble forecasts with the Penn State/NCAR mesoscale model MM5 and the EPA Community Multiscale Air Quality (CMAQ) Model-3, the 24-hour ensemble means of meteorological conditions and the ozone concentrations compared fairly well against the observations for a high-ozone event that occurred on 30 Aug during the Texas Air Quality Study of 2000 (TexAQS2000). Yi et al (9) developed a neural network model for forecasting daily maximum ozone levels. They then compared the neural network's performance with those of two traditional statistical models, regression, and Box-Jenkins ARIMA. The neural network model for forecasting daily maximum ozone levels is different from the two statistical models because it employs a pattern recognition approach.

Ezimand and Kakroodi (10) also used various Meteorological data, the concentration of other pollutants, traffic information, industrial agglomeration, and remote sensing data as study parameters on ozone concentrations. They initially analyzed the data to find meaningful relationships between the study parameters and ozone concentrations and Ozone concentration was then predicted using multivariate linear regression and spatio-temporal changes of ozone concentration were finally investigated. Ghazali et al (11) examined the transformation of nitrogen dioxide (NO2) into ozone (O3) in the urban environment using a time series plot. Data on the concentration of environmental pollutants and meteorological variables were employed to predict the concentration of O3 in the atmosphere. The possibility of employing multiple linear regression models as a tool for the prediction of O3 concentration was tested.

To reduce the level of ozone pollutants, and their negative impacts on the people of Tehran, it is better to present appropriate management plans. Nowadays, Geographical Information System (GIS) methods are used in various fields of air pollution modeling such as: creating databases to store spatial information of pollutants, locating people, modeling air pollutants, estimating levels of pollution effects on individuals, and identifying relevant patterns (12). In this study, we first examined the amount of ozone distribution in the years 2015, 2016, and 2017. This study aimed to predict the ozone pollution behavior based on previous years with artificial intelligence. This model can be used as a suitable platform for managing health alerts. In this study, we proposed a model using a neural network to predict ozone contaminants using meteorological parameters such as temperature, precipitation, humidity, etc. which have a significant influence on the distribution of ozone.

Methods

The study area is Tehran metropolitan, the capital of Iran, with longitude 51° 2´ E to 51° 36´ E, and latitude 25° 34´ N to 35° 50´ N, with an area of approximately 800 km^2. The height of the city is 1200 m in the south and 2000 m in the north. The study area is shown in Fig. 1. There is also a general slope from north to the south, but there is a lot of topographical roughness in the city. Tehran, with a population of about 9 million, is the capital and most important city of Iran. The increasing number of vehicles, the presence of numerous factories in the west of the metropolis, along with the surrounding Alborz Mountains, are major contributors to air pollution. In this study, for analyzing and modeling ozone pollutants in Tehran, daily data of air quality stations were used between 2015 and 2017. The daily data of the year 2018 and 2019 were also used to validate the presented models for O3 emission distribution. This information was obtained from the air quality control website at www.air.tehran.ir. Meteorological and geographical conditions are factors that have a great impact on the distribution of pollutants, especially ozone. For this purpose, meteorological maps such as 24-hour precipitation, wind speed, wind direction, minimum, maximum, and average temperature, minimum, maximum, and average humidity were used.

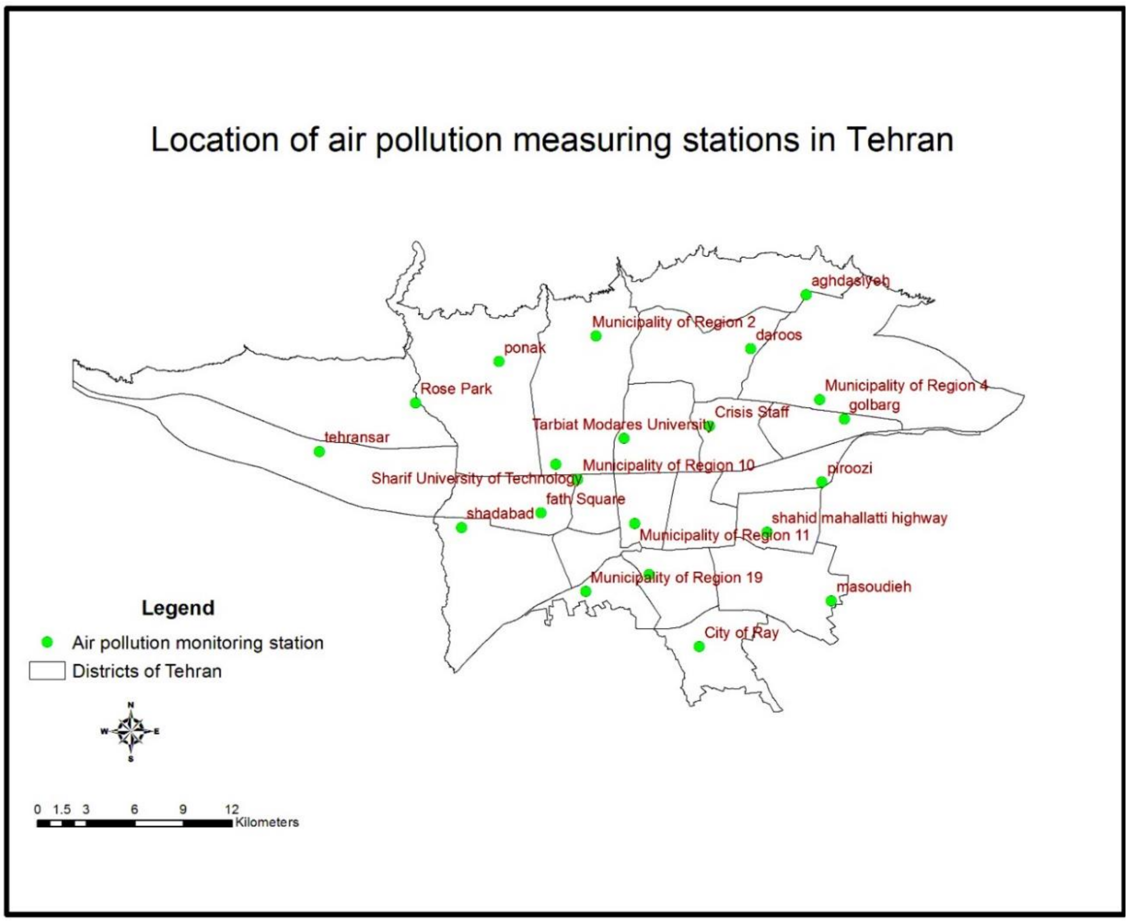


Fig.1: Location of the air quality control stations in Tehran, Iran (Owned by Air Quality Control Company)

After data preparation, it is necessary to select the best interpolation method to crate the continuous map of O_3 pollutant concentration. Among the various interpolation methods, three methods of IDW, Spline, and Kriging, which are more common than other methods, were selected to interpolate O3 pollutant concentration information. To interpolate, each of the interpolation methods was applied to the contaminant concentration data of the most polluted day in summer, then their RMSE error was determined. Radial Basis Functions (RBF) are used in the Spline interpolation method. Moreover, among the different methods of kriging, the ordinary kriging method is selected for interpolation. To obtain the best accuracy in the Kriging method, various semi-variogram models were tested on the data. The RMSE value for these three methods is shown in Table 1; the IDW method is the best one to map O_3 pollutants.

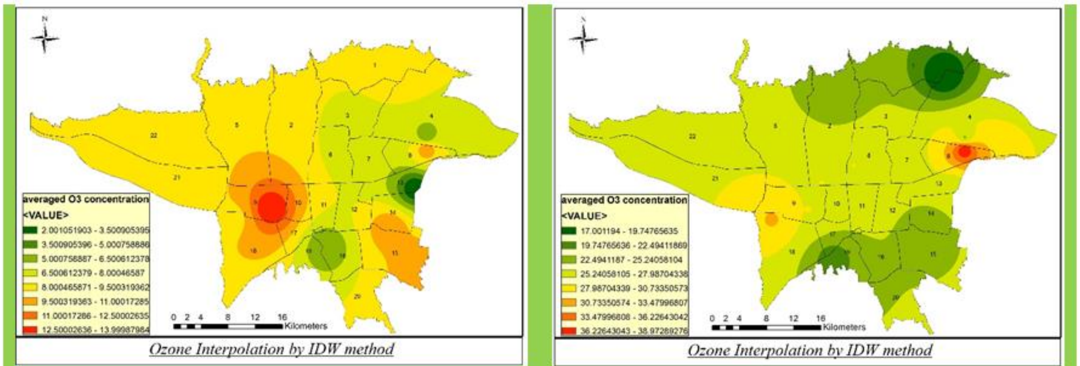
Table 1: RMSE values of different methods in O3 pollutant interpolation

Interpolation method	***RMSE***
IDW	12.256 ($\mu g/m^3$)
SPLINE	16.797 ($\mu g/m^3$)
ORDINARY KRIGING	16.771 ($\mu g/m^3$)

By using the IDW interpolation method, pollutant maps are generated for all days of years 2015, 2016, and 2017. These maps can be used to investigate the relationship of O_3 dispersion with wind velocity, relative humidity, precipitation, and temperature. Moreover, based on the results of daily interpolation, monthly average maps of these pollutants are produced throughout the city of Tehran for three years. The amount of pollutants that changes during different months is also monitored. Besides, the average, maximum and minimum levels of pollutants in different regions of Tehran are also calculated. Furthermore, the annual changes of O3 pollutants are carried out in Tehran as well as in different regions of the city during the mentioned period. In the next step, the concentration of O3 pollutants was analyzed using a neural network. Meteorological parameters are one of the most important factors in increasing or decreasing the concentration of ozone in Tehran. Therefore, in this study, a neural network was developed to investigate the relationship and prediction of O3 contaminants concentration with climatic conditions. A neural network can be defined as an adaptive system that incorporates several simple processing elements modeled based on the human brain (13). The processing elements, which are the neurons, come together to complete a processing path. These processing elements usually fall into layers with regular plates. The input layer acts as a processor, which provides the results to the network after processing. The middle layer is called the hidden layer, which is a computational neural layer. The last layer is the output layer which shows the network output in response to a specified input. There are different models to show how this layer is working, one of known as a Feedforward Neural Network. There is no horizontal relation of the layers in this network. In this network, the transmission is defined in input and output through the external network, which monitors the adjustment of the parameters.

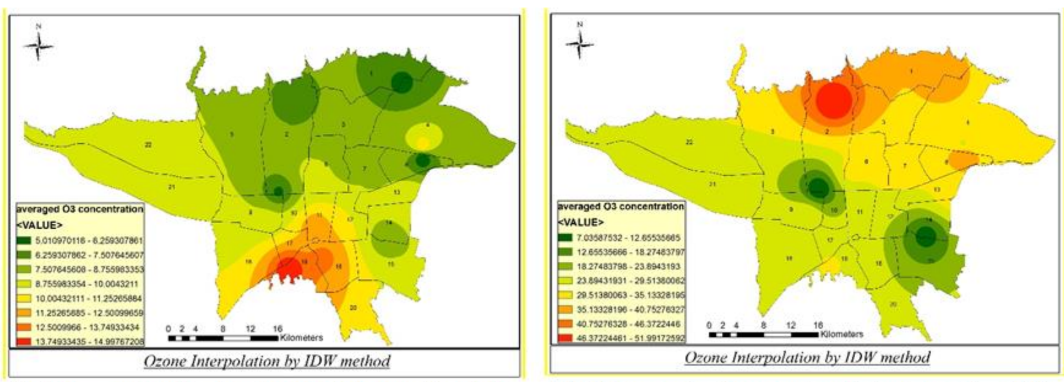
Results

To carry out a systematic analysis of the changes for O_3 pollutants in Tehran on a daily, monthly, and yearly basis, the IDW continuous interpolation map of these pollutants was created for all days. Subsequently, monthly and annual maps were prepared using the combination of all these daily interpolation maps. Figures 2, 3, and 4 show the maps of the lowest and highest O3 pollution months in 2015-2017. For the O_3 pollutants, the lowest and highest monthly averages were in Nov and Aug respectively for the year 2015 (Fig. 2). For the year 2016, the monthly lowest average concentration was for Jan and the highest was in Jul (Fig. 3). For the year 2017, the monthly lowest average concentration was also in Jan and the highest was in Jul (Fig. 4). The highest average accumulation of O_3 was in southwestern of Tehran. Concentration increases in the suburbs relative to the city center since the concentration of O_3 in the suburbs is higher than in the central areas, the concentration of these pollutants is lower in the central areas.



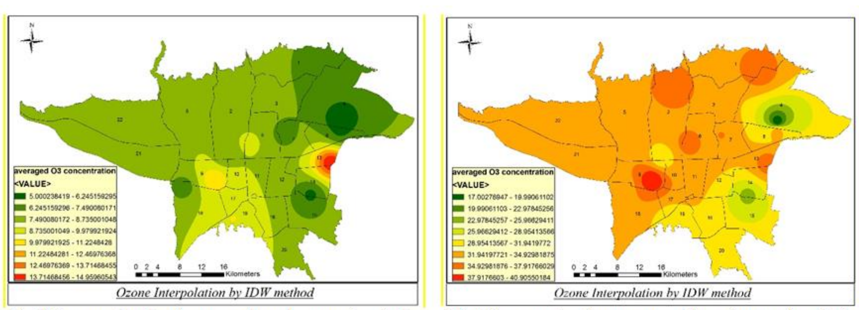
(a) O3 map in the least polluted month of the year (November) (b) O3 map in the most polluted month of the year (August)

Fig. 2: The averaged O3 concentration 2015, Tehran, Iran



(a) O3 map in the least polluted month of the year (January) (b) O3 map in the most polluted month of the year (July)

Fig.3: The average O3 concentration in July 2016, Tehran, Iran



(a) O3 map in the least polluted month of the year (January) (b) O3 map in the most polluted month of the year (July)

Fig. 4: The average O3 concentration in July 2017, Tehran, Iran

In the year 2015, the highest and the lowest annual average concentration of O3 pollutants belong to regions 8 and 1, respectively. Moreover, the most and the least polluted area in the most polluted month of the year belongs to zones 8 and 21 respectively. The most and the least polluted area in the cleanest month of the year belongs to zones 19 and 4 respectively. In the year 2016, the highest and lowest annual mean O_3 concentrations belong to zone 1 and 15 respectively. In addition, the most and the least polluted area in the most polluted month of the year belongs to zones 4 and 15 respectively. The most and the least polluted area in the cleanest month of the year belongs to zones 1 and 14 respectively. In the year 2017, the highest and lowest annual mean O_3 concentrations belong to Region 9 and 19 respectively. Moreover, the most and the least polluted area in the most polluted month of the year belongs to zones 13 and 19, respectively. The most and the least polluted area in the cleanest month of the year belongs to zones 6 and 19 respectively. A multilayer feed forward neural network with a back propagation algorithm was also designed to predict O_3 contaminant concentration. In Fig. 5, the structure of this network can be seen.

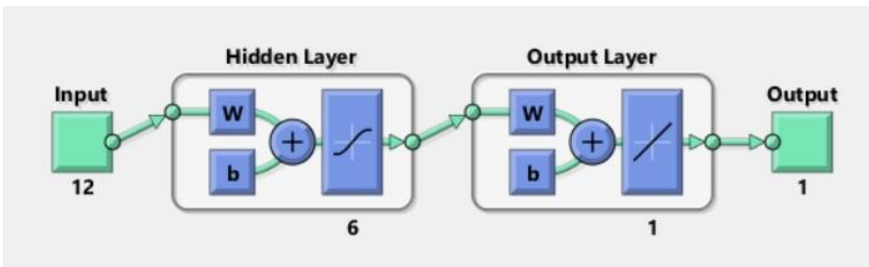


Fig. 5: The schema of developed neural network

As it is clear from the network structure, this network has an output that represents the concentration of pollutants daily. The network has twelve inputs, including wind speed, wind direction, temperature, and rainfall data for each day and the day before that. The data used for network training include 240 d of years 2015, 2016, and 2017. At the next step, the neural network was designed and trained. Seventy percent of samples were considered as training data, and 15% as validation data, and 15% as test data. Figure 6 shows the correlation coefficient obtained for each data set.

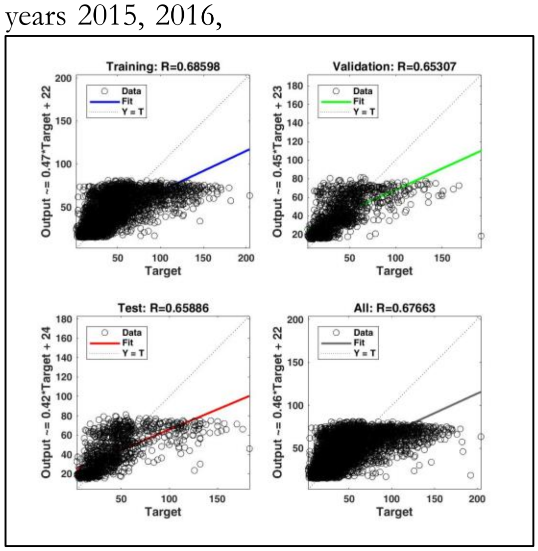


Fig. 6: Regression graphs for training, validation, and test data

Data of 2018 and 2019 were used to obtain the prediction accuracy of this network. Firstly, the average concentration of O_3 was selected for 85 d from different months of the year. Then the average concentration of O_3 pollutants for these 85 d was estimated using a neural network. The prediction accuracy of this network was estimated at 68% on hot days, and 77% for O_3 pollutants on cold days. The meteorological parameters of temperature, velocity, and direction of the wind, and the amount of precipitation affect the concentration of O_3 pollutants.

Discussion

Ozone is one of the major components of photochemical smog in the photochemical oxidant group of air, formed by the reaction of nitrogen oxides (NOx) and volatile organic compounds (VOCs) in the presence of sunlight (photochemical reaction). Therefore, most ozone is formed in sunny conditions. Therefore, appropriate management plans are needed to reduce the adverse effects of this pollutant on human health. In addition, the role of environmental factors and urban management must be investigated in the increase of this pollutant and its temporal and spatial distribution. According to this study, the highest concentrations of O_3 are from in the warm months of July and Aug and the lowest concentrations are formed in cold months of the year with less solar radiation. The highest average accumulation of O_3 was in southwestern Tehran. Concentration increases in the suburbs relative to the city center since the concentration of O_3 in the suburbs is higher than in the central areas, the concentration of these pollutants is lower in the central areas. Considering the accuracy of different interpolation methods in this study, the IDW method was selected as the best interpolation method for mapping O_3 pollutant concentration in Tehran, while the kriging method yielded the worst accuracy. One of the main causes of this problem can be related to the number and location of ground stations, which are not adequately and homogeneously distributed in the city. The distribution of ozone pollutants depends on various factors such as meteorological parameters, traffic, urban planning, land-use changes, etc. In this study, a neural network was developed to investigate the relationship and prediction of O3 accumulation rates with the meteorological data of the year 2015, 2016, and 2017. According to the data of ozone pollutants in the year 2018 and 2019, the accuracy of neural networks for hot and cold days of the year were about 68% and 77% respectively. Therefore, the meteorological parameters of temperature, wind speed and direction, and precipitation are significantly related to the concentration of O_3 pollutants.

In this study, a spatio-temporal model was expressed based on a neural network. In previous studies, a linear model such as a multivariate linear regression has been used (7, 10), (11).

The prediction of O_3 gas is made (7). In this paper, linear models such as Multiple Linear Regression (MLR) Algorithm, General Additive Mixed Models (GAMMs) and Land Use Random Forest (LURF) models were used. The best prediction accuracies are 0.466, 0.747 and 0.695, respectively for these three methods, while the accuracy obtained in the present study was 0.77. This indicates that the neural network method is more accurate than linear models. The multivariate linear regression method was used to model ozone in Tehran (10). Comparing its results with the present study, district 8 and 13, located in the east of Tehran are known as the most contaminated areas in both studies. Therefore, the east of Tehran Province is considered as a dangerous area in terms of ozone gas. The factors affecting O3 gas such as temperature, wind speed and humidity were investigated (11). The model used in this research was the same as the multivariate linear regression model. In that study, despite the spatial influencing factors, non-spatial analysis was used to examine O_3. The advantage of the present study over the mentioned research is its spatiality analyzes.

Other similar studies, such as (6) and (8), have used nonlinear models such as nonlinear regression equations. While in this research, it has been tried to use computational intelligence methods.

The ANN used in this paper is a model that combines both linear and nonlinear methods. This model has eliminated the disadvantages of the previous methods. It is more accurate than the previous ones.

Multiple linear and nonlinear regression equations have been used to study O_3 in Iran (6). The results of this study showed that summer, especially August, has the highest concentration of O_3. In the present study, summer (July and Aug) was identified as the most polluted season. By comparing these research results with research (6), we can confirm the very high concentration of O_3 is in summer, especially in the first two months. This indicates a direct relationship between the effects of sunlight on O_3 pollutant.

Neural networks were used to model ozone pollution based on non-spatial data, while the main advantage of the present study is spatial analysis (9).

Neural network and Box-Jenkins ARIMA models were compared to predict ozone pollution (9). In this study, the neural network method has better accuracy than the other two methods for predicting ozone pollutants. The basic neural network model has been used in the study (9). In recent years, more advanced neural network models such as Feedforward Neural Network, Radial basis function Neural Network, Kohonen Self Organizing Neural Network, Recurrent Neural Network (RNN), Convolutional Neural Network and Modular Neural Network have been proposed with much higher accuracy than the general neural network. In this study, Feedforward Neural Network (the most well-known type of neural network) was used, proven to be more accurate than the general neural network.

Conclusion

Current research has examined the spatial and temporal variations of ozone. In this study, changes in time (different seasons) indicate changes in solar radiation, while location changes indicate changes in spatial components of the study area. The warm months of the year have the most ozone pollution, which is a sign of solar radiation efficacy. On the other hand, specific locations of the study area were also identified as contaminated areas.

Prediction of ozone pollutants in the study area was another result of this research. This prediction can help health decision-makers know when and where intensive care is needed so that they can take action to prevent its negative effects.

Ethical considerations

Ethical issues (Including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Conflict of interest

The authors declare that there is no conflict of interests.

References

1. Yildirim Y, Bayramoglu M (2006). Adaptive neuro-fuzzy based modelling for prediction of air pollution daily levels in city of Zonguldak. *Chemosphere,* 63:1575-1582.
2. Pu X, Wang T, Huang X, et al (2017). Enhanced surface ozone during the heat wave of 2013 in Yangtze River Delta region, China. *Sci Total Environ,* 603-604:807-816.
3. Vafa-Arani H, Jahani S, Dashti H, et al (2014). A system dynamics modeling for urban air pollution: A case study of Tehran, Iran. *Transportation Research Part D: Transport and Environment,* 31:21-36.
4. Jafarian H, Behzadi S (2020). Evaluation of PM2. 5 emissions in Tehran by means of remote sensing and regression models. *Pollution,* 6:521-529.
5. Jaoui M, Kamens R (2003). Gaseous and Particulate Oxidation Products Analysis of a Mixture of [alpha]-pinene+[beta]-pinene/O3/Air in the Absence of Light and [alpha]-pinene+[beta]-pinene/NOx/Air in the Presence of Natural Sunlight. *Journal of Atmospheric Chemistry,* 44:259-297.

6. Et A, Masoudi M (2018). Status and prediction of ozone as an air pollutant in Ahvaz City, Iran. *Caspian J Environ Sci,* 16:203-213.
7. Wang J, Cohan DS, Xu H (2020). Spatiotemporal ozone pollution LUR models: Suitable statistical algorithms and time scales for a megacity scale. *Atmospheric Environment,* 237:117671.
8. Zhang F, Bei N, Nielsen-Gammon JW, et al (2007). Impacts of meteorological uncertainties on ozone pollution predictability estimated through meteorological and photochemical ensemble forecasts. *J Geophys Res,* 112: D04304.
9. Yi J, Prybutok VR (1996). A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area. *Environ Pollut,* 92:349-357.
10. Ezimand K, Kakroodi AA (2019). Prediction and spatio – Temporal analysis of ozone concentration in a metropolitan area. *Ecological Indicators,* 103:589-598.
11. Ghazali NA, Ramli NA, Yahaya AS, et al (2010). Transformation of nitrogen dioxide into ozone and prediction of ozone concentrations using multiple linear regression techniques. *Environ Monit Assess,* 165:475-489.
12. Stroh E (2011). Assessing exposure to airborne pollutants. *Lund University, Faculty of Medicine Doctoral Dissertation Series*:2.
13. Behzadi S, Jalilzadeh A (2020). Introducing a Novel Digital Train Model Using Artificial Neural Network Algorithm. *Civil Engineering Dimension,* 22:47-51.