Global J. Environ. Sci. Manage., 3(3): 333-340, Summer 2017
DOI: 10.22034/gjesm.2017.03.03.010

SHORT COMMUNICATION

Artificial neural network forecast application for fine particulate matter concentration using meteorological data

M. Memarianfard, A.M. Hatami, M. Memarianfard*

Department of Civil Engineering, K.N. Toosi University of Technology, Tehran, Iran

Received 24 September 2016; revised 20 December 2016; accepted 28 December 2016; available online 1 June 2017

ABSTRACT: Most parts of the urban areas are faced with the problem of floating fine particulate matter. Therefore, it is crucial to estimate the amounts of fine particulate matter concentrations through the urban atmosphere. In this research, an artificial neural network technique was utilized to model the $PM_{2.5}$ dispersion in Tehran City. Factors which are influencing the predicted value consist of weather-related and air pollution-related data, i.e. wind speed, humidity, temperature, SO_2, CO, NO_2, and $PM_{2.5}$ as target values. These factors have been considered in 19 measuring stations (zones) over urban area across Tehran City during four years, from March 2011 to March 2015. The results indicate that the network with hidden layer including six neurons at training epoch 113, has the best performance with the lowest error value (MSE=0.049438) on considering $PM_{2.5}$ concentrations across metropolitan areas in Tehran. Furthermore, the "R" value for regression analysis of training, validation, test, and all data are 0.65898, 0.6419, 0.54027, and 0.62331, respectively. This study also represents the artificial neural networks have satisfactory implemented for resolving complex patterns in the field of air pollution.

KEYWORDS: *Air pollution; Artificial neural network (ANN); Meteorological data; $PM_{2.5}$ concentration; Tehran City*

INTRODUCTION

Different kinds of airborne microbes suffer different effects because of meteorological parameters as well as influence their abundance (Fierer *et al.,* 2008; Oliveira *et al.,* 2009; Pyrri and Kapsanaki-Gotsi 2011; Li *et al.,* 2011; Raisi *et al.,* 2013). In reality, any airborne microbe has got certain effects that can be specifically considered in a special solution which concerns that microbe. In this research, to ease the solution all airborne microbes are assumed together an average effect which must be obtained. There are three main meteorological parameters effects as temperature (*t*), wind speed (*w)* and humidity (*h)*, which have been defined and measured well and assessed in the paper. There are great concerns regarding visibility reduction, adverse health effects, and climate change due to atmospheric fine particle as $PM_{2.5}$ (with an aerodynamic diameter of 2.5 microns or less) and some other effects (Pope *et al.*, 2002; Molina and Molina, 2004; Forster *et al.,* 2007; Wang *et al.*, 2016). $PM_{2.5}$ has high potential to adsorb or condense more toxic air pollutants such as organic compounds, metals, etc. and pose greater health risks, when its surface areas are increased (Oberdorster *et al.*, 2005). Rapid process of industrializing, urbanization, energy utilization and the associated population growth in urban areas has caused serious particulate matter pollution in many cities worldwide, from Asian mega cities to even modern cities in Europe and America. (Han *et al.*, 2014; Hu *et al.*, 2014; Huang *et al.*, 2014; Lary *et al.*, 2014; Lin *et al.*, 2014; Liu *et al.*, 2013; Song *et al.,* 2015; Zhang *et al.,* 2012; Trizio *et al.,* 2016). Based on the results of about 1600 cities of 91

*Corresponding Author Email: *amir.m.hatami@gmail.com*
Tel.: .:+98 21 8820 1431 Fax: +98 21 2205 2573
Note: Discussion period for this manuscript open until September 1, 2017 on GJESM website at the "Show Article".

countries, the world's average $PM_{2.5}$ concentration is 28.4–56.8 μg/m^3 during 2008 to 2013, ranging from 26 to 208 μg/m^3 (WHO, 2014; WHO, 2015).

Tehran as the capital of Iran with approximately 8.5 million inhabitants is plagued by severe air pollution (Lotfabadi, 2014) In the past few years, due to urbanization, industrialization and population growth in Tehran, the issue of air pollution especially the PM has become extremely crucial. In recent years, the artificial neural technics (ANN) were started to be utilized for forecasting particulate matter concentrations (Perez and Reyes, 2002; Kukkonen *et al.*, 2003; Lu *et al.*, 2003; Ordieres *et al.*, 2005; Zhou *et al.*, 2014; Feng *et al.*, 2015; Mohammadizadeh *et al.*, 2016). The ability of ANN to change easily to suit the different situations has directed, to their use in the majority of scientific fields. Some applications of ANN in the atmospheric sciences during the 1990s were indicated in the studies that Gardner and Dorling had done (Gardner and Dorling, 1998). A variety of $PM_{2.5}$ sources such as power stations, transportation, natural disasters and heating systems in the residences have brought about considerable difficulties for $PM_{2.5}$ evaluation. In this research, the back-propagation learning algorithm for modeling and prediction was utilized. An error estimation technique has been conducted analysis upon indicating "R" and "IA", the number of neurons on the hidden layer and epochs have been estimated. The study has been carried out in Tehran City capital of Iran and the data has been employed and performed in period of 2011 to 2015.

MATERIALS AND METHODS

The daily recorded data set was provided from the urban air recording stations in Tehran. The Geographical map and air pollution monitoring station locations in Tehran are shown in Fig. 1.The stations in Tehran districts are constantly monitoring air and report the daily data, which parts of recorded data are used in this paper. The recorded data set is composed of nitrogen dioxide, sulfur dioxide, carbon monoxide, $PM_{2.5}$, wind speed, temperature, and relative humidity (Qin *et al.,* 2014).

According to the data, during 1461 days from 21 March 2011 to 20 March 2015, as presented in Table 1, including descriptive statistics of $PM_{2.5}$, SO_2, NO_2, CO, temperature, humidity, and wind speed. Moreover, the corresponding data histograms are shown in Fig. 2.

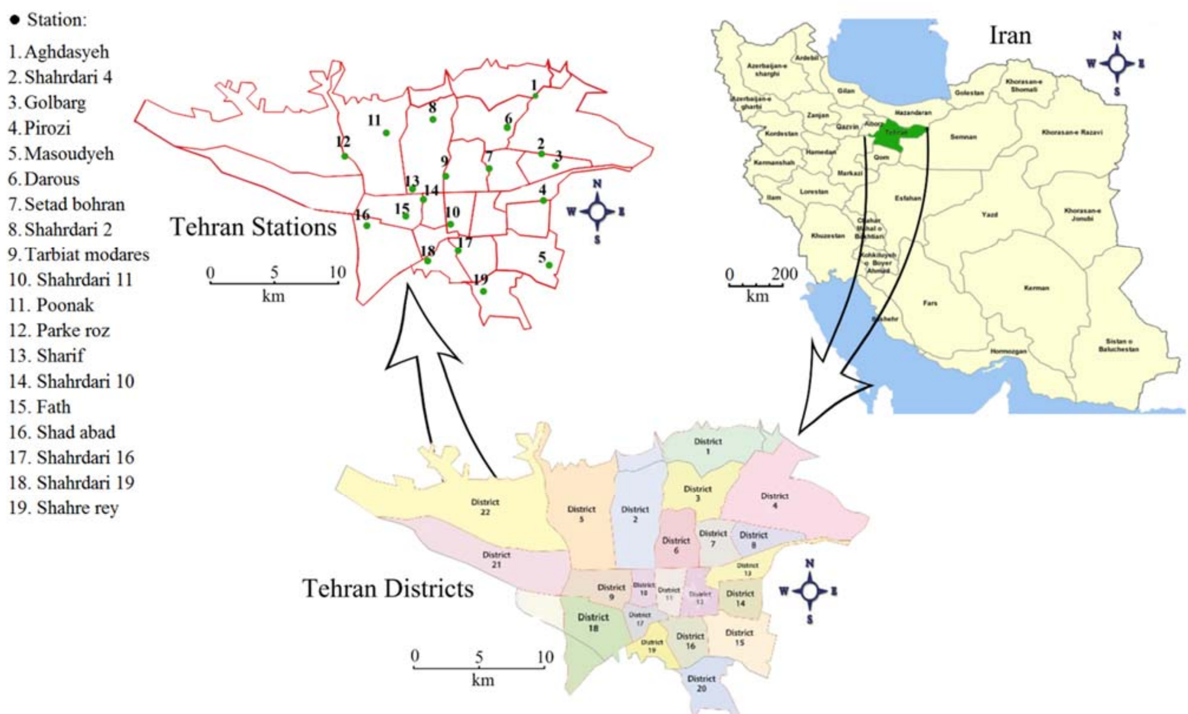

Fig. 1: The study area of research performance and air pollution monitoring stations in Tehran City

Table 1: The descriptive statistics of daily air pollutants and meteorological parameters as input data (from 21 March 2011 to 20 March 2015)

Statistic parameters	Temperature	Wind speed	Humidity	SO_2	NO_2	CO	$PM_{2.5}$
N	1461	1461	1461	1461	1461	1461	1461
Mean	18.419	3.065	34.631	30.42	56.72	38.49	97.80
Std. Error of Mean	0.25994	0.047403	0.469062	0.204	0.303	0.242	0.666
Median	18.950	2.750	30.250	30.00	57.00	37.00	96.00
Std. Deviation	9.9359	1.8118	17.9289	7.799	11.578	9.258	25.455
Variance	98.722	3.283	321.448	60.818	134.055	85.706	647.975
Skewness	-0.141	9.711	0.932	2.080	-0.015	0.540	0.449
Std. Error of Skewness	0.064	0.064	0.064	0.064	0.064	0.064	0.064
Kurtosis	-1.249	215.742	0.202	9.782	0.291	0.098	0.565
Std. Error of Kurtosis	0.128	0.128	0.128	0.128	0.128	0.128	0.128
Minimum	-2.025	0.500	9.00	16	20	19	26
Maximum	36.625	46.000	93.625	88	97	77	204

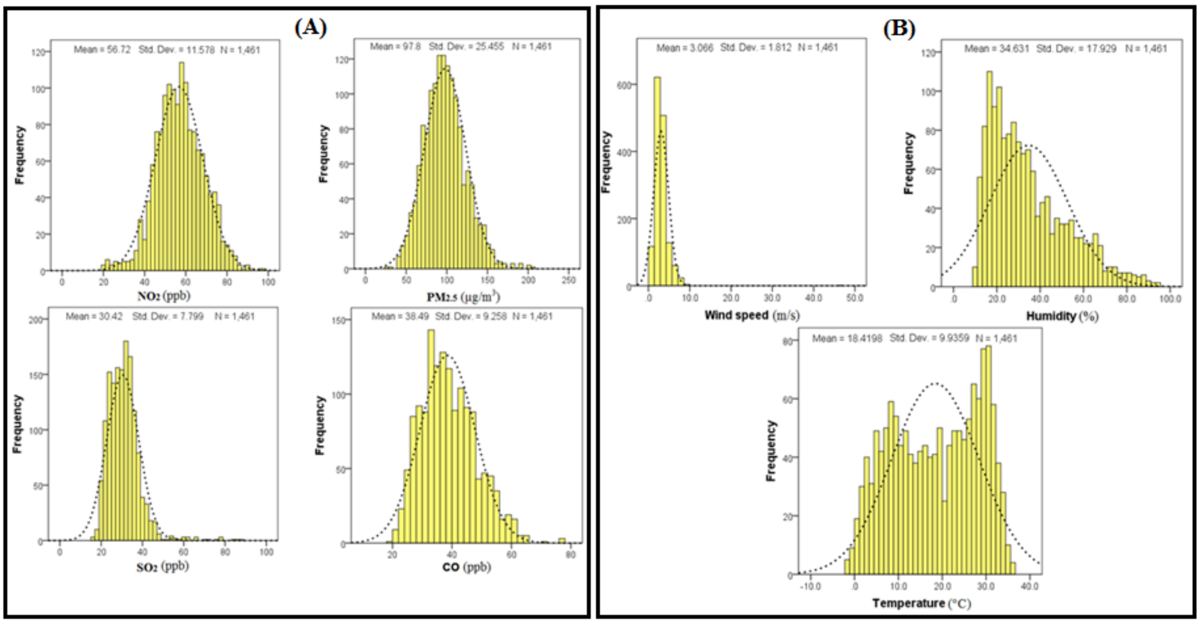

Fig. 2: (A) The histograms of measured $PM_{2.5}$, NO_2, SO_2 and CO (24h. averages).
(B) The histograms of measured humidity, wind speed, and temperature (24h. averages)

Neural network approach

The most ANN-based studies in the field of air pollution have suggested the back-propagation learning algorithm for modeling and prediction. In this algorithm, the data split into three parts:

1-assessing data set: which forms the bulk of the data that can be used for the training purposes?

2-examining data set: this data can be used to examine the performance quality of the trained model.

3-validating data set: this data can be used to validate the model.

Fig. 3 shows the architectural plan of the neural network proposed model with 6 predictor variables. The proposed model has been developed with the aim of presenting minimum possible error through predicting forecast computations, upon 6 entering parameters as input, 6 middle layers (hidden) parameters which are lead to one output parameter (no. 13) indicated as 6:6:1 scheme. It should be noted that the combination of SO_2, NO_2 and CO (in gaseous state) are named as precursors (input data) which are assessed upon archive correlations in ANN motor to create the amount of new particles as secondary $PM_{2.5}$ output.

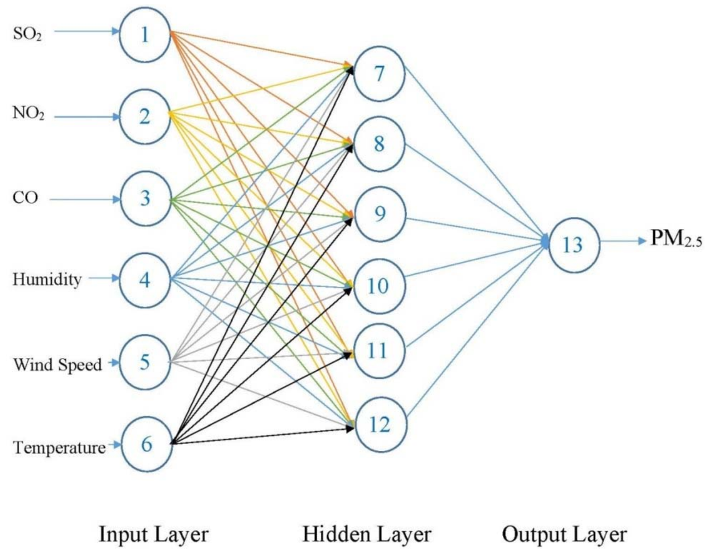

Fig. 3: Architecture of the proposed 6:6:1 ANN $PM_{2.5}$ model

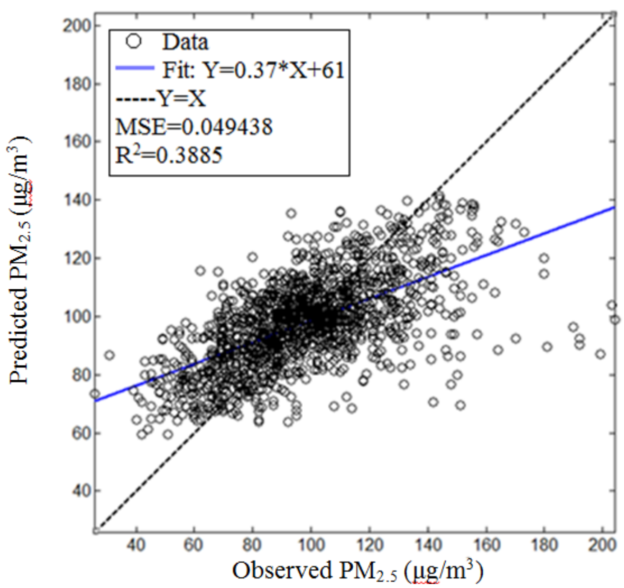

Fig. 4: Scattered plot of the observed and predicted results of $PM_{2.5}$

Error and indicator 'MSE' & 'R^2'

For the performance of the ANN-based models to assess the accuracy of the estimation, there are several parameters, i.e. error and indictor: 'MSE' & 'R^2', which are calculated according to Eqs. 1 and 2.

$$(The\ mean\ square\ error) MSE = \frac{1}{N}\sum_{i=1}^{N}[P_i - O_i]^2 \quad (1)$$

$$(The\ correlation\ factor) R^2 = \frac{\sum_{i=1}^{N}[P_i - \bar{O}]^2}{\sum_{i=1}^{N}[O_i - \bar{O}]^2} \quad (2)$$

Where N, O_i, P_i and $\bar{O}$ are the number of observations, observed value, predicted value, and average value, respectively.

RESULTS AND DISCUSSION

In this paper, the ANN approach was utilized for modeling and the data was divided into three groups for training the network (50 % of data), validating (25% of data), and testing (25% of data) the network. The function optimization technique used is the scaled conjugate gradient algorithm. The accomplishment of the model has been estimated with calculation of the mean square error (MSE) as the statistical criteria and "R". Scattered plot of the observed and predicted concentrations of $PM_{2.5}$ using the data from 21 March 2011 to 20 March 2015 have been shown in Fig. 4. R^2 and MSE values were found to be 0.3885 and 0.049438, respectively.

In order to avoid over-training problem in this study, two indicators of the network were utilized, which are optimum choice of the hidden neuron numbers and error goal. According to input data, the neural network system that was required had 6 nodes on the input layer and one node on the output layer as well. Upon the error estimation method through analysis, the number of neurons on the hidden layer and epochs were estimated. The numbers of neurons were available in the hidden layer varied from 2 to 30. The MSE index was used to reach the optimum number of neurons in the hidden layer. As a result, the network layer with 6 neurons in the input layer, as a single hidden layer, including 6 hidden neurons plus a single neuron in the output layer will be able to represent best forecast.

MSE values for training, validation and test have been shown in Fig. 5(a and b) shows Schematic showing of the predicted and observed concentrations of $PM_{2.5}$ data during the four years, from 21 March 2011 to 20 March 2015.

R values for training, validation, test, and all data regression analysis have been shown in Fig. 6. The "R" value for regression analysis of training, validation, test and all data are 0.65898, 0.6419, 0.54027, and 0.62331, respectively. The correlation factor values of near 0.6 are normal with regards to random climate changing data. The correlation factor value corresponding to test may be considered as 0.54 so poor. Therefore, the ANN proposed model capability to handle such random variation can be accepted as fair, so acceptable. Actually in climatic modeling, generally, such expectation of 0.95 correlation factor may be impossible.

CONCLUSION

Generally, modeling of climate pollution, including several irregular trends, does not obey certain mathematical theory. The use of ANN, as a tool for predicting future climatic conditions upon the trends taken from previous measured data, can be the best method to be utilized for modeling. Furthermore, employing certain algorithm or constitutive relations, may be quite helpful in reducing the errors.

The present work indicates that the predicted power of artificial neural network models depends on several vital parameters, namely choice of six inputs data, six parameters in hidden layers, learning algorithm, and types of stopping criteria. One of the prime issues in developing optimal ANN is over-training. It arises when ever network learns the noisy details

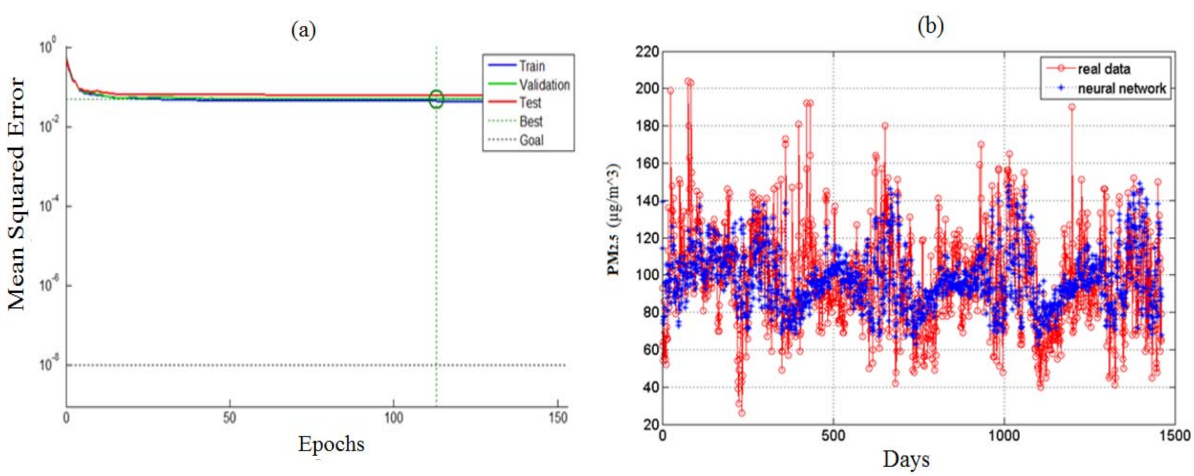

Fig. 5: (a) MSE values for training, validation and test
(b) Schematic showing of the predicted and observed concentrations of $PM_{2.5}$ data

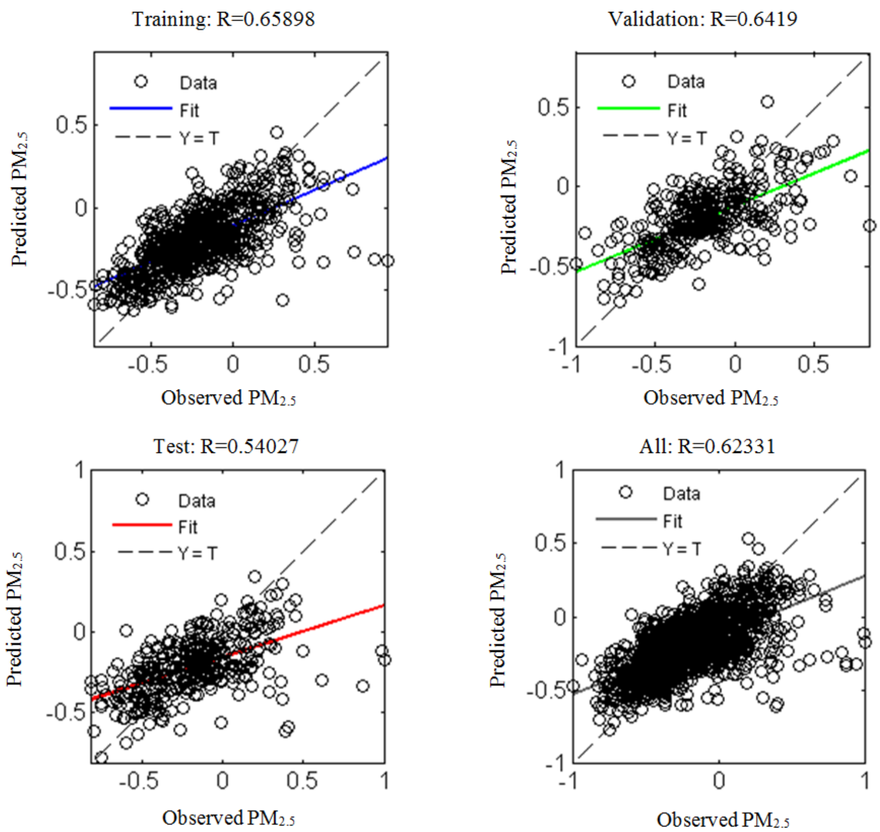

Fig. 6: "R" values for training, validation, test, and all data for regression analysis

in assessing the data. The obtained results indicate that the network with 6 hidden neurons at training epoch 113 has the best performance with the lowest MSE value (MSE=0.049438). This explains that this network prediction is closely matching with actual observation. Further, the "R" value for regression analysis of training, validation, test and all data are 0.65898, 0.6419, 0.54027, and 0.62331, respectively. Accordingly, the ANN-based approach is capable of producing accurate evaluation data set in the field of air pollution.

ACKNOWLEDGEMENT

The authors would like to acknowledge Professor S.A. Sadrnejad, the faculty member of K.N. Toosi University for his valuable supports throughout this research performance.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

ABBREVIATIONS

ANN	Artificial neural network
CO	Carbon monoxide
IA	Information Assurance
MSE	Mean square error
N	Number of observations
NO_2	Nitrogen dioxide
O_i	Observed value
$\bar{O}$	Average value
P_i	Predicted value
PM	Particulate matter
$PM_{2.5}$	Particles less than or equal to 2.5 micrometers in diameter
PM_{10}	Particles less than or equal to 10 micrometers in diameter
SO_2	Sulfur dioxide

REFERENCES

Feng, X.; Li, Q.; Zhu, Y.; Hou, 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, Atmos. Environ., 107: 118-128 (**11 Pages**)

Fierer, N.; Liu, Z.; Rodriguez-Herna´ndez, M.; Knight, R.; Henn, M.; Hernandez, M. T., (2008). Short-term temporal variability in airborne bacterial and fungal populations. Appl. Environ. Microbiol., 74(1): 200–207 (**8 Pages**).

Forster, P.; Ramaswamy, V.; Artaxo, P.; Berntsen, T.; Betts, R.; Fahey, D.W.; Haywood, J.; Lean, J.; Lowe, D.C.; Myhre, G.; Nganga, J.; Prinn, R.; Raga, G.; Schulz, M.; Dorland, R.V., (2007). Changes in atmospheric constituents and in radiative forcing. (Climate Change 2007: Physical Science Basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Gardner, M.W.; Dorling, S.R., (1998). Artificial neural networks (the multilayer perceptron)-a review of applications in the atmospheric sciences. Atmos. Environ. 32: 2627–2636 (**10 Pages**).

Han, L.J.; Zhou, W.Q.; Li, W.F.; Li, L., (2014). Impact of urbanization level on urban air quality: a case of fine particles ($PM_{2.5}$) in Chinese cities. Environ. Pollut., 194: 163–170 (**8 Pages**).

Hu, J.L.; Wang, Y.G.; Ying, Q.; Zhang, H.L., (2014). Spatial and temporal variability of $PM_{2.5}$ and PM_{10} over the North China plain and the Yangtze River Delta. China. Atmos. Environ. 95: 598–609 (**12 Pages**).

Huang, R.J.; Zhang, Y.; Bozzetti, C.; Ho, K.F.; Cao, J.J.; Han, Y.; Daellenbach, K.R.; Slowik, J.G.; Platt, S.M.; Canonaco, F.; Zotter, P.; Wolf, R.; Pieber, S.M.; Bruns, E.A.; Crippa, M.; Ciarelli, G.; Piazzalunga, A.; Schwikowski, M.; Abbaszade, G.; Schnelle-Kreis, J.; Zimmermann, R.; An, Z.; Szidat, S.; Baltensperger, U.; El Haddad, I.; Prevot, A.S., (2014). High secondary aerosol contribution to particulate pollution during haze events in China. Nature, 514: 218–222 (**5 Pages**).

Kukkonen, J.; Partanen, L.; Karppinen, A.; Ruuskanen, J.; Junninen, H.; Kolehmainen, M.; Niska, H.; Dorling, S.; Chatterton, T.; Foxall, R.; Cawley, G., (2003). Extensive evaluation of neural network models for the prediction of NO_2 and PM_{10} concentrations, compared with a deterministic modeling system and measurements in central Helsinki. Atmos. Environ., 37: 4539–4550 (**12 Pages**).

Lary, D.J.; Faruque, F.S.; Malakar, N.; Moore, A.; Roscoe, B.; Adams, Z.L.; Eggelston, Y., (2014). Estimating the global abundance of ground level presence of particulate matter ($PM_{2.5}$). Geospatial Health, 8: 611–630 (**20 Pages**).

Li, M.; Qi, J.; Zhang, H.; Huang, S.; Li, L.; Gao, D., (2011). Concentration and size distribution in an outdoor environment in the Qingdao coastal region. Sci. Total Environ., 409: 3812–3819 (**8 Pages**).

Lin, J.T.; Pan, D.; Davis, S.J.; Zhang, Q.; He, K.B.; Wang, C.; Streets, D.G.;Wuebbles, D.J.; Guan, D.B., (2014). China's international trade and air pollution in the United States. Proc. Natl. Acad. Sci. U. S. A. 111: 1736–1741 (**6 Pages**).

Liu, Z.; Guan, D.B.; Crawford-Brown, D.; Zhang, Q.; He, K.B.; Liu, J.G., (2013). A low-carbon road map for China. Nature, 500: 143–145 (**3 Pages**).

Lotfabadi, P., (2014). High-rise buildings and environmental factors. Renewable Sustainable Energy Rev., 38: 285–295 (**11 Pages**).

Lu, W.Z.; Wang, W.J.; Wang, X.K.; Xu, Z.B.; Leung, Y.T., (2003). Using improved neural network model to analyze RSP, NO_x and NO_2 levels in urban air in Mong Kok, Hong Kong. Environ. Monit. Assess., 87: 235–254 (**20 Pages**).

Mohammadizadeh, M.J.; Karbassi, A.R.; Nabi Bidhendi, Gh.R.; Abbaspour, M., (2016). Integrated environmental management model of air pollution control by hybrid model of DPSIR and FAHP. Global J. Environ. Sci. Manage., 2(4): 381-388 (**8 pages**).

Molina, M.J.; Molina, L.T., (2004). Megacities and atmospheric pollution. J. Air Waste Manag. Assoc. 54: 644-680 (**37 Pages**).

Oberdorster, G.; Oberdorster, E.; Oberdorster, J., (2005). Nano toxicology: an emerging discipline evolving from studies of ultrafine particles. Environ Health Perspectives 113: 823–839 (**17 Pages**).

Oliveira, M.; Ribeirio, H.; Delgado, J.L.; Abreu, I., (2009). The effects of meteorological factors on airborne fungal spore concentration in two areas differing in urbanization level. Int. J. Biometeorol., 53, 61–73 (**13 Pages**).

Ordieres, J.B.; Vergara, E.P.; Capuz, R.S.; Salazar, R.E., (2005). Neural network prediction model for fine particulate matter ($PM_{2.5}$) on the US–Mexico border in El Paso (Texas) and Ciudad Jua´ rez (Chihuahua). Environ. Model. Software, 20: 547–559 (**13 Pages**).

Perez, P.; Reyes, J., (2002). Prediction of maximum of 24-h average of PM_{10} concentrations 30 h in advance in Santiago, Chile. Atmos. Environ. 36: 4555–4561 (**7 Pages**).

Pope, C.A.; Burnett, R.T.; Thun, M.J.; Calle, E.E.; Krewski, D.; Ito, K.; Thurston, G.D., (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. J. Am. Med. Assoc. 287: 1132-1141 (**10 Pages**).

Pyrri, I.; Kapsanaki-Gotsi, E., (2011). Diversity and annual fluctuations of cultivable airborne fungi in Athens, Greece: A 4-year study. Aerobiologia. 10: 2-15 (**14 Pages**).

Qin, S.; Liu, F.; Wang, J.; Sun, B., (2014). Analysis and forecasting of the particulate matter (PM) concentration levels over four major cities of China using hybrid models. Atmos. Environ., 98: 665–675 (**11 Pages**).

Raisi, L.; Aleksandropoulou V.; Lazaridis M.; Katsivela E., (2013). Size distribution of viable, cultivable, airborne microbes and their relationship to particulate matter concentrations and meteorological conditions in a Mediterranean site, Aerobiologia, 29: 233–248 (**16 Pages**).

Song, C.; Pei, T.; Yao, L., (2015). Analysis of the characteristics and evolution modes of $PM_{2.5}$ pollution episodes in Beijing, China during 2013. Int. J. Environ. Res. Public Health. 12: 1099–1111 (**13 Pages**).

Trizio, L.; Angiuli, L.; Menegotto, M.; Fedele, F.; Giua, R.; Mazzone, F.; Carducci, A.G.C.; Bellotti, R.; Assennato, G., (2016). Effect of the Apulia air quality plan on PM_{10} and benzo(a) pyrene exceedances. Global J. Environ. Sci. Manage., 2(2): 95-104 (**10 pages**).

Wang, H.L.; Qiao, L.P.; Lou, S.R.; Zhou, M.; Ding, A.J.; Huang, H.Y.; Chen, J.M.; Wang, Q.; Tao, S.K.; Chen, C.H.; Li, L.; Huang, C., (2016). Chemical composition of $PM_{2.5}$ and meteorological impact among three years in urban Shanghai, China, J. Cleaner Product., 112: 1302-1311 (**10 Pages**).

WHO, (2014). WHO's ambient air pollution database: Update 2014. http://www.who.int/phe/health_topics/outdoorair/databases/AAP_database_methods_2014.pdf?

WHO, (2015). Ambient outdoor air pollution in cities database 2014. http://www.who.int/phe/health_topics/outdoorair/databases/cities/en/.

Zhang, Q.; He, K.B.; Huo, H., (2012). Cleaning China's air. Nature 484: 161–162 (**2 Pages**).

Zhou, Q.; Jiang, H.; Wang, J., Zhou, J., (2014). A hybrid model for $PM_{2.5}$ forecasting based on ensemble empirical mode decomposition and a general regression neural network. Sci. Total Environ., 496: 264-274 (**11 Pages**).

AUTHOR (S) BIOSKETCHES

Memarianfard, M., M.Sc., Department of Civil Engineering, K.N. Toosi University of Technology, Tehran, Iran.
Email: *memarian.m7@gmail.com*

Hatami, A.M., Ph.D. Candidate, Department of Civil Engineering, K.N. Toosi University of Technology, Tehran, Iran.
Email: *amir.m.hatami@gmail.com*

Memarianfard, M., Ph.D., Assistant Professor, Department of Civil Engineering, K.N. Toosi University of Technology, Tehran, Iran.
Email: *memarian@kntu.ac.ir*

COPYRIGHTS

Copyright for this article is retained by the author(s), with publication rights granted to the GJESM Journal.
This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/).

HOW TO CITE THIS ARTICLE

Memarianfard, M.; Hatami, A.M.; Memarianfard, M., (2017). Artificial neural network forecast model for fine particulate matter concentration using meteorological data. Global J. Environ. Sci. Manage., 3(3): 333-340.

DOI: 10.22034/gjesm.2017.03.03.010

url: http://gjesm.net/article_23079.html