Desert

Desert
Online at http://desert.ut.ac.ir

Desert 19-2 (2014) 131-135

Modelling the formation of Ozone in the air by using Adaptive Neuro-Fuzzy Inference System (ANFIS) (Case study: city of Yazd, Iran)

L. Rafati[a], M.H. Ehrampoush[a], A. Talebi[b], M. Mokhtari[a*], Z. Kherad Pisheh[a], H.R. Dehghan[a]

[a] *Department of Environmental Health, Shahid Sadoughi University of Medical Sciences, Yazd, Iran*
[b] *Faculty of Natural Resources, Yazd University, Yazd, Iran.*

Received: 25 May 2014; Received in revised form: 15 October 2014; Accepted: 27 October 2014

Abstract

The impact of air pollution and environmental issues on public health is one of the main topics studied in many cities around the world. Ozone is a greenhouse gas that contributes to global climate. This study was conducted to predict and model ozone of Yazd in the lower atmosphere by an adaptive neuro-fuzzy inference system (ANFIS). All the data were extracted from 721 samples collected daily over two successive years, from April 2012 to 29 March 2014. The concentration of pollutants and meteorological variables including NO_X, temperature, wind speed and wind direction were considered as input and ozone (O_3) as the output of model. The results showed that among five membership functions used in the model, the Gaussian membership function with R^2 equal to 0.949, RMSE equal to 2.430 and correlation coefficient equal to 0.974 was obtained as the best model to predict the concentration of ozone in the lower atmosphere. This study showed that predicting and modelling ozone using an adaptive neuro-fuzzy inference system (ANFIS) is appropriate and, due to the expansion of the city of Yazd in the not too distant future, it is necessary to pay more attention to the permissible threshold values of pollutants such as ozone.

Keywords: Modelling; Ozone concentration; Adaptive Neuro-Fuzzy Inference System (ANFIS); Yazd

1. Introduction

Ozone is a colourless gas with a slightly sweet taste and nasty smell that is regarded as a pollutant and life protector in the stratosphere. Troposphere ozone has high oxidizing power, high permeability coefficient similar to CO_2 (easy penetration into plant tissues), solubility in water (ten times the solubility of CO_2) and tends to react with water in the environment (Banan *et al*., 2013; Fuhrer *et al*., 1997). Ozone is an important greenhouse gas whose high concentration causes multiple diseases such as respiratory dysfunctions in humans, and has a negative effect on plants and objects (Huang *et al*., 2012; Neidell and Kinney, 2010). Ozone in the stratosphere layer is vitally important for living organisms on Earth, and by absorption of UV radiation hinders it from reaching the Earth's surface (Sousa *et al*., 2007). Ground-level ozone (troposphere) is produced through photochemical reactions related to volatile organic compounds and NO_X, in the presence of sunlight and hydroxyl radicals (Chandra *et al*., 2004). In recent years, the damaging effects of ozone on plants, animals, human health and food have gained the attention of the world as to the amount of surface ozone (De Laat *et al*., 2005).

Predictive models in the past included only simple correlations of empirical data, but today a vast amount of information has provided the possibility of the inclusion of the prediction models in simulations of the air pollution (Jang and Sun, 1993). Standard statistical methods may not sufficiently predict the complex and nonlinear behaviour of the model, but neural

* Corresponding author. Tel.: +98 913 3559789,
Fax: +98 351 6238555.
E-mail address: mhimokhtari@gmail.com

networks are used extensively in simulations due to their great ability in modelling nonlinear and unreliable data based on previous assumed equations. Another investigation method is fuzzy logic: to use this method the data should be sorted by clustering algorithm based on iterative optimization (Morabito and Versaci, 2003).

Sugeno and Yasukawa expanded the fuzzy-adaptive inference system with one output to an inference system (fuzzy-adaptive) with multiple outputs (Sugeno and Yasukawa, 1993). Jang also developed a self-structure inference network (neuro-fuzzy) (Jang, 1993).

Research shows that the clustering system of MATLAB software enables us to use the first grade Sugeno model (Ashrafi *et al*., 2012).

There have been various studies in predicting and modelling ozone concentrations by means of multiple linear regression (MLR), artificial neural networks (ANNs) and multi-gene genetic programming (MGP) (Sousa *et al*., 2007; Al-Alawi *et al*., 2008; Ozdemir *et al*., 2008).

One of the main issues in the development of statistical models is the selection of the model's input, which aims to minimize the number of inputs in order to achieve the best results. In other words, among all possible inputs for a model those that have the greatest effect on the results should be chosen. However, some of the input data may be correlated, and applying all of them to the model simply results in increasing the size and execution time of the model (Yildirim and Bayramoglu, 2006).

Therefore, in dealing with these data it is better to choose the most effective ones and other data correlated with that should be removed from the input (Ashrafi *et al*., 2012).

Due to geographic and climatic conditions, population growth, urbanization and orientation to industry in Yazd, it is likely that the city is dealing with environmental problems. The aim of this study is modelling and predicting the formation of the ozone in Yazd city using ANFIS by means of MATLAB 2011a software.

2. Material and methods

2.1. Study area

Yazd province, with an area of about 131,517 km^2, is located in the central Iranian plateau. This province in located at a 29º 35ʹ to 35º 7’ range of northern latitude and an eastern longitude of 52º 50ʹ to 58º 16’. The region of study is Yazd city. Yazd is generally considered to be a dry province because it is located in a global dry belt of straw (25- 40 degrees of north latitude) and is away from the sea. Yazd is located on a yellow radiation belt and has a high potential for receiving solar energy. For simulation of network inference systems (fuzzy-adaptive) a data set is required consisting four variables: NO_x, temperature, wind speed and wind direction as network input and the ozone concentration as output of the network. All of the data were extracted from 721 samples collected daily during two successive years, from April 2012 to 29 March 2014. Table 1 shows the parameters used in modelling.

Table 1. Model Parameters

Parameter	Unit
O_3	$\mu g/m^3$
NO_x	$\mu g/m^3$
Temperature	°C
Wind speed	m/s
Wind direction	0°–360°

2.2. Criteria for model evaluation

Two common statistical indicators including correlation coefficient (R^2) and root mean square error (RMSE) were used to evaluate the performance of the models. These parameters are calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (SPI_i - SPI_j)^2}{\sum_{i=1}^{n} (S\hat{P}I_i - \overline{SPI_i})^2} \quad (1)$$

$$RMSE = \frac{\sqrt{\sum_{i=1}^{n} (SPI_i - S\hat{P}I_i)^2}}{n} \quad (2)$$

In these equations $S\hat{P}I_i$, SPI_i, $\overline{SPI_i}$ and n are observed data, calculated data, the mean number of observed data and the number of data, respectively. It is clear that the closeness of R^2 to 1 and the low value of RMSE indicate the accuracy of the method.

2.3. Adaptive neuro-fuzzy inference system (ANFIS)

In this study, a fuzzy-neuro inference system (ANFIS) that is one of neuro-fuzzy systems was used. The most common type of fuzzy inference system that has the capability of being placed on an adaptive network is the Sugeno fuzzy system (Sugeno and Yasvkava, 1993).

In this study, the forward selection method was used for selecting model inputs. As mentioned above, in this method, first, the variable that has the highest correlation coefficient with the output variable was considered as the first input. Then, other variables enter the model one by one and if they can improve results of the model or reduce RMSE they are chosen, and otherwise they are excluded (Chen *et al*., 2004; Wang *et al*., 2006; Khan *et al*., 2007).

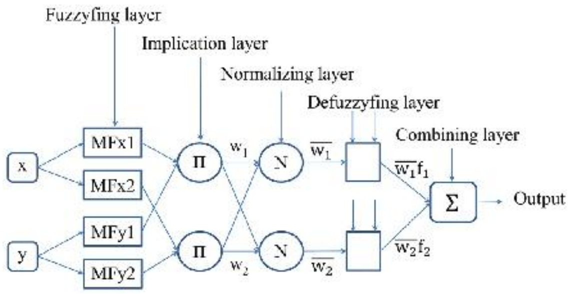

Fig. 1. Structure of ANFIS model equal to inference system discussed (Jang, 1993)

Afterwards, the overall structure of the models consisting of model inputs, output and functions is set. Then, a number of meteorological data are introduced to the system.

3. Results

Table 2 shows the results of the model. Figure 2 shows the regression obtained and Figure 3 shows the time series of pollution observed and predicted.

Table 2. Membership type, RMSE, COREL and R^2 for different ANFIS models

MEMtype	RMSE	Crr.	R2
Gbellmf	3.942	0.151	0.722
'gauss2mf	3.670	0.154	0.720
gaussmf *	2.430	0.974	0.949
Trapmf	3.391	0.781	0.679
Trimf	3.718	0.856	0.727

*the best combination

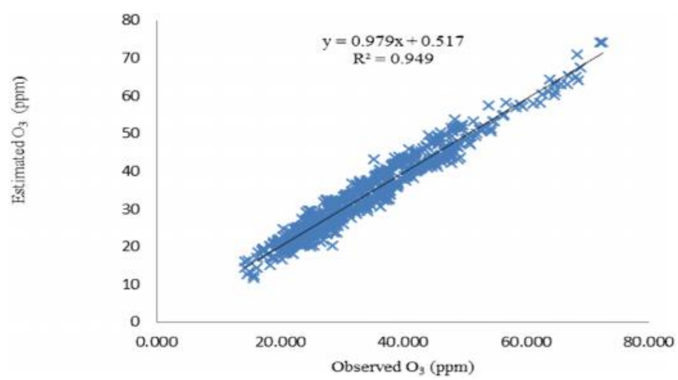

Fig. 2. The regression between observed and simulated data for the best ANFIS model for O_3 parameter

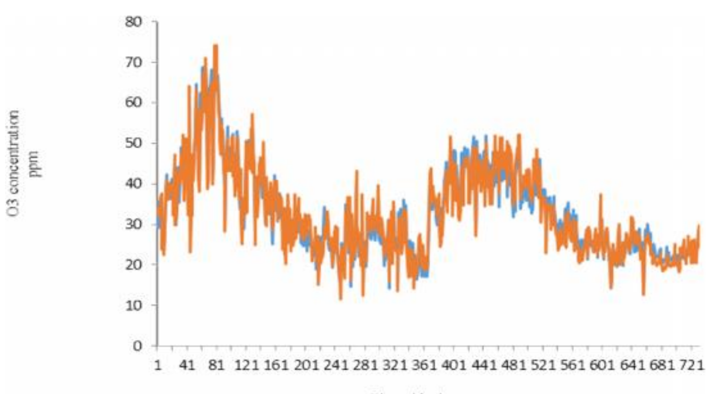

Fig. 3. Time series of observed and predicted pollution by the ANFIS model

4. Discussion

The purpose of this study was to predict the ozone generated in Yazd by pre-reactants and other factors affecting the formation of ozone. Given that modelling ozone production is very complex and existing theories cannot answer the complexities, a hybrid system called ANFIS system in MATLAB toolbox was used.

A study was conducted in 2013 on modelling and predicting the formation of ozone in the air of Mashhad city using a neural fuzzy network based on inference fuzzy-adaptive systems. It was observed that the error obtained for the designed network for the training dataset was 1.8%, and for the pilot data this was 2.4%. In the back-propagation algorithm, in comparison to combinative patterns, more appropriate answers with a lower percentage error were observed (Seghat Al-Eslami *et al*., 2013).

Ozdemir *et al*. performed a prediction of stratospheric ozone concentrations by the ANN. The results showed that the ozone concentrations estimated were compatible with the observed values, and this introduced ANN as one of the compromising methods in the estimation of multiple air pollutants (Ozdemir *et al*., 2008).

In a study conducted in 2002 by Abdul-Wahab et al., the assessment and prediction of ozone concentration using ANN was performed. This study showed that the concentrations of NO, SO_2, relative humidity, non-methane hydrocarbons and NO_2 had a huge effect on the prediction of ozone. In addition, temperature has an important role when the amount of solar radiation is low. They considered ANN to be a promising approach for the modelling of air pollutants (Abdul-Wahab and Al-Alawii, 2002).

A study by Johanyak and Kovacs, on the prediction of surface ozone concentration based on a fuzzy model, showed the model applied using LESFRI to have the best results with a low number of values (Johanyak and Kovacs, 2011).

A study by Savic *et al*., in predicting and modelling SO_2 by ANFIS, satisfactory results were not obtained. This might be due to the low number of input and output data (two months) in these studies (Savic *et al*., 2013). In our study this issue was taken into consideration and the number of data entered into the model was 721 for two successive years, so we obtained successful results.

According to the results of various studies mentioned above, these studies considered the modelling and prediction of concentrations of air pollutants by different models to be suitable.

5. Conclusion

Based on the results of this study, ANFIS seems suitable for modelling and predicting ozone levels in Yazd. Considering the city’s development and expansion, it is necessary to note that the more is inputs and output, the greater the accuracy of predictions obtained from model simulations. It will be necessary in the near future to implement traffic control projects, in order to be able to lower pollutants such as ozone value to their thresholds and permissible limits by controlling determinant factors in the formation of ozone.

Acknowledgement

The writers of this paper thank the Department of the Environment (DOE) Office of Yazd Province for providing air pollution data, and the statistics department of the Meteorological Centre of Yazd for providing meteorological data.

References

Al-Alawi, S.M., S.A. Abdul-Wahab, C.S. Bakheit, 2008. Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone. Environmental Modeling & Software, 23; 396- 403.

Abdul-Wahab, S.A., S.M. Al-Alawii, 2002. Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks. Environmental Modeling & Software, 17; 219- 228.

Ashrafi, K., G.H. Hoshyaripour, B. Nadjar Araabi, H. Keshavarzi Shirazi, 2012. Prediction of daily carbon monoxide concentration using hybrid FS-ANFIS model based on atmospheric stability analysis; case study: city of Tehran. Journal of the Earth and Space Physics, 38; 183-201. (in Persian)

Banan, N., M.T. Latif, L. Juneng, F. Ahmad, 2013. Characteristics of surface ozone concentrations at stations with different backgrounds in the Malaysian Peninsula. Aerosol and Air Quality Research, 13; 1090- 1106.

Chandra, S., J.R. Ziemke, X. Tie, G. Brasseur, 2004. Elevated ozone in the troposphere over the Atlantic and Pacific oceans in the Northern Hemisphere. Geophysical Research Letters, 31; 1- 4.

Chen, S., S.A. Billings, W. Luo, 2004. Orthogonal least squares methods and their application to nonlinear system identification. International Journal of Control, 50; 1873-1896.

De Laat, A.T.J., I. Aben, G.J. Roelofs, 2005. A model perspective on total tropospheric O3 column variability and implications for satellite observations. Journal of Geophysical Research, 110; 1- 14.

Fuhrer, J., L. Skarby, M.R. Ashmore, 1997. Critical levels for ozone effects on vegetation in Europe. Environmental Pollution, 91; 97-106.

Jang, J.S.R, 1993. ANFIS: adaptive-network-based fuzzy inference systems, IEEE Transaction on Systems. Man and Cybernetics, 23; 665- 685.

Jang, J.S. R., C.T. Sun, 1995. Neuro-fuzzy modeling and control. Proceedings of the IEEE, 83; 378-406.

Johanyak, Z.C., J. Kovacs, 2011. Fuzzy Model based Prediction of Ground-Level Ozone Concentration. Acta Technica Jaurinensis, 4; 113-125.

Khan, J.A., S.V. Aelst, R.H. Zamar, 2007, Building a robust linear model with forward selection and stepwise procedures. Computational Statistics and Data Analysis, 52; 239-248.

Huang, H.L., M.G. Lee, J.H. Tai, 2012. Controlling indoor bioaerosols using a hybrid system of ozone and catalysts. Aerosol and Air Quality Research, 12; 73– 82.

Neidell, M., P.L. Kinney, 2010. Estimates of the association between ozone and asthma hospitalizations that account for behavioral responses to air quality information. Environmental Science & Policy, 13; 97– 103.

Morabito, F., M. Versaci, 2003, Fuzzy neural identification and forecasting techniques to process experimental urban air pollution data. Neural Networks, 16; 493- 506.

Ozdemir, H., G. Demir, G. Altay, S. Albayrak, C. Bayat, 2008. Prediction of Tropospheric Ozone Concentration by Employing Artificial Neural Networks. Environmental Engineering Science, 25; 1249- 1254.

Savic, M., I. Mihajlovic, Z. Zivkovic, 2013. An anfis – based air quality model for prediction of SO_2 concentration in urban area. Serbian Journal of Management, 8; 25 – 38.

Seghat al-eslami, N., M. Mosavi, M. Aalami, 2013. Modeling and predicting the formation of ozone air Mashhad city using a neural network-based fuzzy inference system and fuzzy-adaptive. Iranian Chemical Engineering Journal, 12; 12-20. (in Persian)

Sugeno, M., T. Yasukawa, 1993. A fuzzy logic based approach to qualitative modeling. IEEE Transactions on Fuzzy Systems, 1; 7-31.

Sousa, S.I.V., F.G. Mrtins, M.C.M. Alvim-Ferraz, M.C. Pereira, 2007. Multiple linear regression and artificial neural networks based on principle components to predict ozone concentrations. Environmental Modeling & Software, 22: 97–103.

Wang, X.X., S. Chen, D. Lowe, C.J. Harris, 2006. Sparse support vector regression based on orthogonal forward selection for the generalized kernel model, Neurocomputing, 70; 462- 474.

Yildirim, Y., M. Bayramoglu, 2006. Adaptive neuro-fuzzy based modeling for prediction of air pollution daily levels in city of Zonguldak. Chemosphere, 63; 1575-1582.