Research Article

Comparison of RNN and ANFIS in Concentrations of Carbon Monoxide and Fine Particles Forecasting in Tehran

Razeghi,A.[1 *], Khademi,M.[2], Moadel,M.[3]

[1]Department of Mechanical Engineering, Email: St_a_razeghi@azad.ac.ir
[2] Department of Applied Mathematics, Email: Khademi@azad.ac.ir
[3]Department of Energy Systems Engineering, Email: masoud.moadel@gmail.com
Islamic Azad University South Tehran Branch, Tehran, Iran

Received date: 15 September 2013; Accepted date: 20 September 2014

Academic Editor: Farshid Keynia

Abstract

One of the most signification current discussions in environmental science is the prediction of pollutant's concentration. The aim of this study is to predict concentration of CO and PM10, RNN and adaptive neuro fuzzy inference system (ANFIS) and compare this two well-known methods in prediction of air pollutants. In this regard, meteorological data of Aghdasiyeh synoptic station between 2008 to 2011 have been utilized for training ANFIS and RNN. Six models with different input parameters were evaluated. Results of this study show that neural network techniques are very good for predicting the CO-AQI and PM10-AQI. In general, RNN model has a better result in CO prediction and model 3 with RMSE =0.060 and R^2 =93.1 is the best one. Whereas ANFIS model is better in PM10 forecasting that indicate these data have fuzzy feature.

Keywords: CO-AQI, PM10-AQI, ANFIS, RNN.

1 Introduction

In recent years, the problem of air pollution in Tehran has been caused environmental problems and various respiratory, heart and skin diseases. So on some days pollution crisis have forced the government to prorogate the city. Therefore research in the field of air pollution has been important and specialists in various disciplines seek to reduce, control, and predict the amount of pollutants in different parts of Tehran. The earliest and most important problem in process of planning and controlling air pollution is predicting the concentration of pollution in different areas. The key problem with this explanation is that existed models are not able to forecast the concentration of air pollution with acceptable precision. Many studies have been carried out in this field and several methods have been used.

Many of them try to establish the mathematical relationship between the concentration of pollutants and climatic characteristics using statistical and numerical techniques. The performance of these models is not acceptable for various reasons and has not gained a major concern. Another computer technique to predict different values is artificial intelligence technique (Chelani, A.B.; Chalapati R,C.V.; Phadke,K.M.; Hasan, M.Z., 2002). Recurrent Neural Network (RNN) and Adaptive Neuro Fuzzy Inference System (ANFIS) models are so and use data analysis. In recent years by development of artificial intelligence techniques, these methods are used to simulate, prediction and behavior of a system. These methods are very suitable to approximate functions that have not a known mathematical relation between effective parameters (Gardner, M.W.; Dorling, S.R., 1998). Various studies have demonstrated the effectiveness of these networks in modeling nonlinear systems. To create an appropriate model of the system, all input parameters and their impacts on output parameters must be considered simultaneously that often is very difficult and practically impossible, while the techniques of artificial intelligence can establish a logical relationship between the desired parameters. In the researches, tow general approaches are discussed for modeling and forecasting air pollution. The first approach is modeling by atmospheric diffusion model. To this end, the equations describing the physical and chemical changes in concentration of pollutants and meteorological data are required. Collet and oduyemi (1997) have discussed these models in a review paper (Collet, R.S., Oduyemi, K., 1997).

The second approach is a statistical model that aims to determine relationships between input and output data. Regression models are the most famous statistical models that are used for predicting and modeling the different concentrations. One of the major limitations to these models is the lake of proper learning in nonlinear systems. While artificial intelligence techniques (ANN, ANFIS, GA, etc.) are very suitable for modeling the nonlinear system. Agirre et. al (2006), showed that the concentration of sulfur dioxide can be predicted by neural network more accurately (Agirre-Basurko, E., Ibarra-Berastegi, G., Madariaga, I., 2006). Hooybergs et al. (2005), applied multi-layer neural network (MLP) to predict PM10 concentration and demonstrated the superiority of MLP than regression models (Hooyberghs, J.; Mensink, C.; Dumont, G.; Fierens, F.; Brasseur, O., 2005). Comrie (1997) used neural networks and regression models to predict ozone concentrations and showed neural network model is a better method to predict the concentration of pollutants (Comrie, 1997).

The purpose of this research is to develop a model that communicates AQI (CO and PM10) to meteorological parameters (maximum temperature, minimum temperature, relative humidity, solar radiation, wind speed and wind direction) and forecast them with a reasonable accuracy. In this way, RNN and ANFIS are used to modeling CO-AQI and PM10-AQI and the effectiveness of input parameters is investigated in forecast accuracy.

2 Material and Methods

2.1 ANFIS mathematical basis

Neural network (NN) was presented in 40th decade of twenty century by Warren McCulloch and Walter Pitts. They proved NN can calculate any arithmetic and logic function.

Theory of fuzzy released in 1965 by Asgarzadeh, Iranian scholar and professor at UC Berkeley's of America. In complex systems that are difficult to understand or issues related to argument, decision-making and human inference, the fuzzy logic is considered as an effective tool.

Each fuzzy systems and neural networks have advantages and disadvantages. The fuzzy system is able to use human language and can human experience can be used, while not capable of learning. In other words, using the observed data fuzzy systems can't be trained. But, neural network using data sets are capable of learning. Neural networks are Non-committal and not able to use human language.

First time in 1993, Zhang could use the capabilities of neural networks and fuzzy systems and provide a system base on fuzzy logics and adaptive neural networks. These systems are known as Adaptive Neuron Fuzzy Inference System (ANFIS).

ANFIS uses a Takagi-Sugeno fuzzy system with a progressive networks structure. If the output of each layer is shown by O_k^l (output of Kth node in lth layer), ANFIS is composed of tow law and tow inputs (x and y) which describe by two membership functions.

In the layer 1, the output of each node is:

$$O_{1,i} = \mu_{A_i}(x) \qquad for\ i = 1,2$$
$$O_{1,i} = \mu_{B_{i-2}}(y) \qquad for\ i = 3,4$$

So, $O_{1,i}(x)$ is essentially the membership grade for 'x' and 'y'. The membership functions could be anything but for this purpose we will use the bell shaped function given by:

$$\mu_A(x) = \frac{1}{1 + \left|\frac{x - c_i}{a_i}\right|^{2b_i}}$$

Where a_i, b_i, c_i *are parameters to be learnt. These are the premise parameters.*

In layer 2 every node is fixed. This is where the t-norm is used to 'AND' the membership grades. For example:

$$O_{2,i} = w_i = \mu_{A_i}(x)\mu_{B_i}(y), \quad i = 1,2$$

Layer 3 contains fixed nodes which calculate the ratio of the firing strengths of the rules:

$$O_{3,i} = \overline{w}_i = \frac{w_i}{w_1 + w_2}$$

Also in layer 4 the nodes in this layer are adaptive and perform the consequent of the rules:

$$O_{4,i} = \overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i)$$

The parameters in this layer (p_i, q_i, r_i*) are to be determined and are referred to as the consequent parameters.*

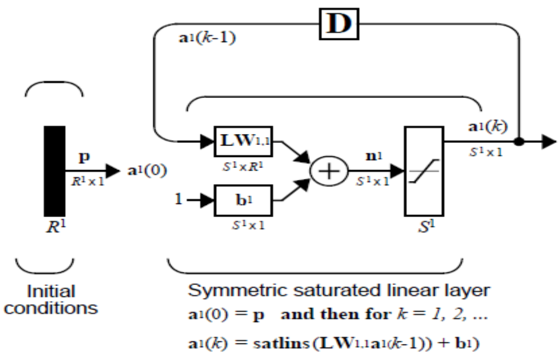
Eventually, in the layer 5, there is a single node here that computes the overall output:

$$O_{5,i} = \sum_i \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}$$

The learning of these systems is the concept of determination of non linear parameters in first layer (related to fuzzy membership functions) and linear parameters in fourth layer in which for an arbitrary input, the desired output is achieved. Hybrid learning method is one of the most effective ways of learning fuzzy inference system based on adaptive neural network. The method of training the first layer is back-propagation and least squares error estimation is used in the fourth layer. For more details refer to [14].

2.2 RNN

Neural networks are divided into two categories, base on the transmission pass of data, feed-forward NNs and recurrent NNs (RNN). Unlike the feed-forward NN, RNNs don't transform data in only one direction. An example of this network is shown in figure 1. The applications of RNN are optimization, parallel processing and pattern recognition. Hopfield, Elman, Jordan and Elman-Jordan are of these types.



an neural network

nt neural network is composed of three layers and t layer, called a context unit. These kinds of neural networks have a memory that can detect and generate time-varying patterns. Elman network generally use sigmoid function in hidden layer and linear function in output layer. This NN can estimate every function by adjusting the number of neurons.

Jordan network is the same as Elman with the difference that the context feeds from output layer instead of hidden layer. A combination of Elman NN and Jordan NN (Elman-Jordan) which has further capabilities is used in this paper.

2.3 Data

The data are related to Aghdasiyeh synoptic station, located in Tehran with the longitude of N 37/51, Latitude E 47/35 and the height above sea level, 1548 meters, which were measured and registered by automatic devices with 1-minute intervals. As respects the target to predict the average of pollutants concentration in a day, the daily average of parameters were used. Air pollution data are in the format of Air Quality Index (AQI) which is an index for reporting daily air quality.

TABLE 2. Data specification

Std	Min	Max	Mean	
110.52	1	365	193.03	Day num
19.83	11.06	99.9	42.85	humidity
8.19	-5.9	27.2	10.18	MIN temp
10.01	-1.5	40.1	20.61	MAX temp
8.58	0.57	35.50	17.19	radiation
0.62	0.23	5.97	1.519	wind speed
51.03	11.8	296.58	119.49	wind dir.

In this research, maximum temperature, minimum temperature, relative humidity, number of days, wind speed and wind direction, Between December 22, 2008 and December 25, 2011 were used. By the way selection of the number of days as input parameter is for modeling the season and also solar radiation have been used to investigate the impacts of air pollution on incoming solar radiation. As regards there are gap in registered data, initially validation test were done for them. So they were compared with the registered data of before days and next days and also other years, then a decision was made by using statistical methods. If the missed data for a day wasn't too much, they were estimated by using AQI of before days and next days and utilization of weight estimation method, otherwise that day was ignored. Finally, the data of 534 days was approved and used in the process of analysis.

Both the ANFIS and RNN, neural network methods are applied for estimating the CO-AQI and PM10-AQI in Tehran city based on the parameters which are mentioned above.

TABLE 1. Models based on different combinations of input parameters.

Model	Input Parameters

1	maximum temperature, minimum temperature, relative humidity, solar radiation, wind speed, wind direction, number of days
2	maximum temperature, minimum temperature, relative humidity, number of days, solar radiation
3	maximum temperature, minimum temperature, solar radiation, wind speed, wind direction, number of days
4	maximum temperature, minimum temperature, relative humidity, solar radiation, wind speed, wind direction
5	relative humidity, solar radiation, wind speed, wind direction, number of days
6	maximum temperature, minimum temperature, relative humidity, number of days

The procedure utilized in the development of the ANFIS and RNN models starts with input data normalization (i.e. target values) in the range of 0 to 1 followed by the dataset matrix size.

After being normalized, sub-datasets are created and prepared for training and testing including the data of 373 days for training, 80 days for validation or checking and the data of 81 days for testing. The output values are generated from test data using two models and finally, the performance of the discussed models is verified by comparison of output and target values. All these steps are carried out by MATLAB software. One of the advantages of these models is to obtain good results with 534 data.

3 Result and Discussion

In this study, ANFIS and RNN is used in order to predict CO-AQI and PM10-AQI by meteorological and air pollution data of Aghdasiyeh synoptic station. In this regard, at first, available data was evaluated and valid data was selected. Then they were normalized in the range of 0 and 1 for better network training. Recurrent neural network with three hidden layers with varying number of neurons were trained and the best network with the lowest error was selected. ANFIS models generated by fuzzy c-means clustering (FCM) and hybrid method (combination of back-propagation and least square)

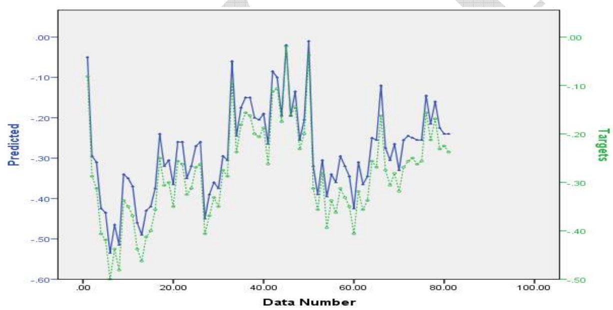
TABLE 2. ANFIS and RNN models performance and structure in prediction of CO-AQI

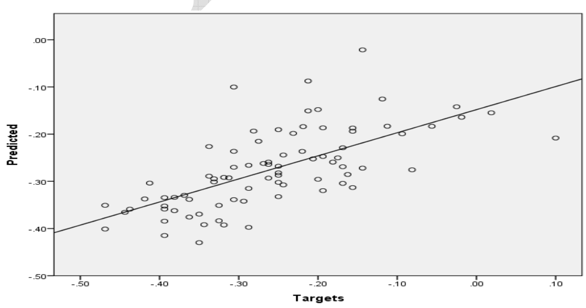
Model	ANFIS	RNN

	NC	R^2	RMSE	MBE	NS	R^2	RMSE	MBE
1	10	93.7	0.066	0.0002	8-7-1	94.4	0.064	-0.0004
2	34	92.4	0.067	-0.0002	13-8-1	93.0	0.064	-0.0004
3	18	90.9	0.074	-0.0003	12-8-1	93.1	0.060	-0.0003
4	39	90.1	0.077	-0.0004	12-5-1	91.8	0.070	-0.0005
5	32	91.3	0.069	0.0006	15-9-1	93.7	0.064	-0.0005
6	34	92.3	0.065	0.0002	10-4-1	93.0	0.065	-0.0002

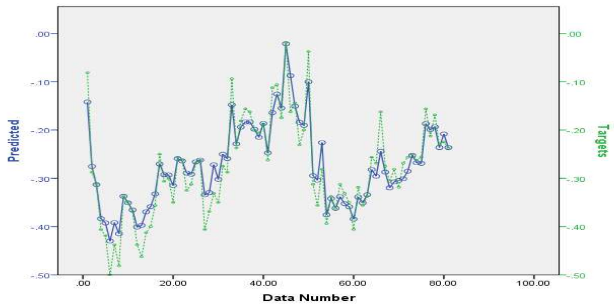
Six models with different input parameters were examined to determine the effect of each parameter on accuracy of prediction. Table 1 shows the input parameters for each of the models. As you see in Table 2 and 3, neural network method has a very good ability to predict the concentrations of air pollutants. Of the two methods used in AQI prediction, RNN model has a better result with RMSE =0.060 and R^2 =93.1 in prediction of CO-AQI.

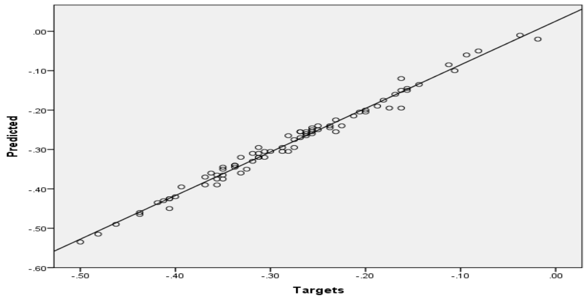
It can be seen that model 3 with maximum temperature, minimum temperature, solar predicted the CO-AQI with acceptable accuracy (RMSE =0.065 and R^2 =92.3). Figure 1-a and 1-b shows measured versus predicted CO-AQI for third model of RNN (model 3). Overall, the results show that the maximum error occurred in areas where the concentration of pollutants are significantly higher than warning range (AQI>100). Table 3 contains the results of the PM10-AQI forecasting. Contrary to the results





(A)





(B)

Figure 2. Comparison between measured and predicted PM10-AQI. (A) ANFIS model, (B) RNN model.

for CO, ANFIS model provides better results.

According to Table 1 and 3, relative humidity, solar radiation, wind speed, wind direction and number of days are the best combination to use as input parameters. Model 5 gives the best result by using these parameters (RMSE =0.043 and R^2 =95.9). Also figure 2 shows the results of PM10-AQI prediction.

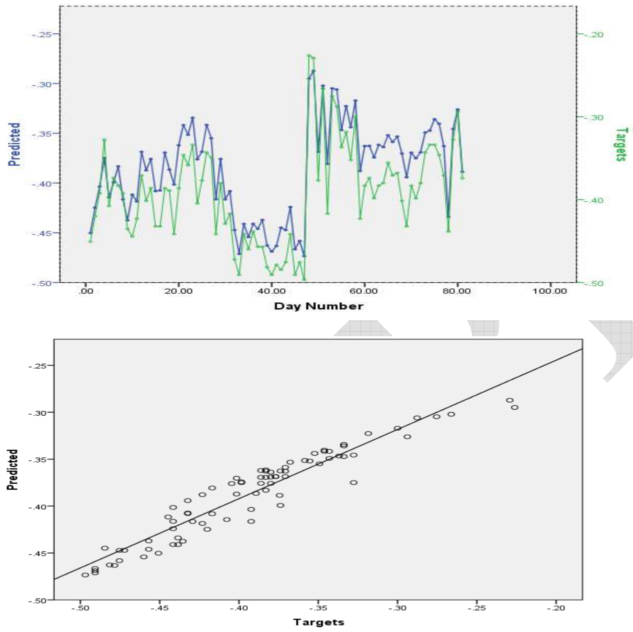
TABLE 3. ANFIS and RNN models performance and structure in prediction of PM10-AQI

Model	ANFIS				RNN			
	NC	R^2	RMSE	MBE	NS	R^2	RMSE	MBE
1	10	95.4	0.054	0.0001	8-7-1	92.8	0.055	0.024

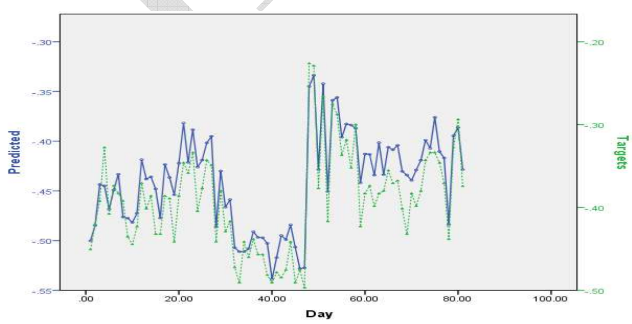
2	34	94.8	0.057	-0.0001	13-8-1	94.5	0.055	0.0001
3	18	92.6	0.050	0.0003	12-8-1	95.1	0.052	0.0000
4	39	91.1	0.077	-0.0004	12-5-1	94.9	0.070	0.0005
5	32	95.9	0.043	-0.0000	15-9-1	95.3	0.046	0.0001
6	34	95.3	0.049	0.0003	10-4-1	96.4	0.050	0.0001

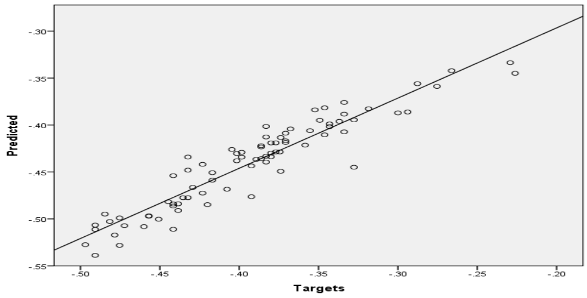
As the results show, the concentration of pollutants is not just a function of car's fuel and other polluting factors, but may also be a function of weather conditions and temperature.

As you can see in table 1 and table 2, MBE for both CO and PM10 show ANFIS sometimes predict AQI more than real amount and sometimes fewer measured amount, while the RNN is usually overestimated.



(A)





(B)

Figure 3: Comparison between measured and predicted PM10-AQI. (A) ANFIS model, (B) RNN model.

4 Conclusion

This research set out to determine the capability of ANFIS and RNN and also investi-gate the effect of each meteorological parameter in prediction of pollutants. The results show that artificial intelligence models (RNN and ANFIS) are able to forecast AQI with acceptable precision. The most interesting results are that RNN is better in CO prediction, while ANFIS has a better result in PM10 forecasting. This finding further support this idea that PM10 data have fuzzy feature. So likely fuzzy systems such as ANFIS, Local Linear Model Tree (LOLIMOT) and etc. have a better ability for modeling PM10.

Thus, artificial intelligence models can forecast concentration of pollutants in different days of the year by using data for meteorological stations and takes necessary plan to reduce and control emission rate. Also, the authorities in Tehran Air Quality Control Agency can review and analyze the distribution of each of the pollutants by using these results.

REFERENCES

[1] A. Mellit, S.A. Kalogirou, S. Shaari, H. Salhi, A. Hadj Arab, 2008. Methodology for predicting sequences of mean monthly clearness index and daily solar radiation data in remote areas: Application for sizing a stand-alone PV system. Renewable Energy 33, 1570–1590.

[2] Agirre-Basurko, E., Ibarra-Berastegi, G., Madariaga, I., 2006. Regression and multilayer perceptron-based models to forecast hourly O3 and NO2 levels in the Bilbao area. Environmental Modelling& Software, 21, 430–446.

[3] (2013)[Online]Available at: http://en.wikipedia.org/wiki/Adaptive_neuro_fuzzy_inference_system

[4] Chelani, A.B., Chalapati R,C.V., Phadke,K.M., Hasan, M.Z., 2002. Prediction of sulphur dioxide concentration using artificial neural networks. Environmental Modelling & Software, 17, 161–168.

[5] Collet, R.S., Oduyemi, K., 1997. Air quality modelling: a technical review of mathematical approaches. Meteorological Applications, 4, 235-246.

[6] Comrie, A., 1997. Comparing neural network and regression models for ozone forecasting. Journal of Air and Waste Management Association, 47, 653–663.

[7] Gardner, M.W., Dorling, S.R., 1998. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmospheric Environment, 32(14-15), 2627-2536.

[8] Hooyberghs, J., Mensink, C., Dumont, G.; Fierens, F.; Brasseur, O., 2005. A neural network forecast for daily average PM10 concentrations in Belgium. Atmospheric Environment, 39, 3279–3289.

[9] Mandic. D., Chambers. J., 2001. Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability. s.l. Willy.

[10] T.R. Sumithira, A. Nirmal Kumar, R. Ramesh Kumar, 2012. An adaptive neuro-fuzzy inference system (ANFIS) based Prediction of Solar Radiation: A Case study. Journal of Applied Sciences Research, 8(1), pp. 346-351.