On using Adaptive Hybrid Intelligent Systems in PM10 Prediction

S. A. Asklany, Khaled Elhelow, M .Abd El-Wahab

Abstract *A comparative study based on producing two intelligence systems applied to PM_{10} prediction was presented in this work. Adaptive Network –based Fuzzy Inference System (ANFIS) used in build a system has three weather elements as input variables (Wind Speed, Wind Direction and Temperature) and the PM_{10} as output variable for PM_{10} nowcast model. Another technique used ANFIS in prediction of chaotic time series to get 6 hours forecast for PM_{10} from the present data. For developing the models, thirteen years hourly data for Mansoria station coordinates 29° 300′ 0″ N, 45° 45′ 0″ E from 1995to 2007 has been used. Different models employing a different training and testing data sets had been studied. The criteria of performance evaluation are calculated for estimating and comparing the performances of the two techniques. The results show that the two presented models success tools in PM_{10} prediction with acceptable root mean square error (RMSE); the model built on using ANFIS for chaotic time series prediction has smaller error compared with the adaptive network fuzzy inference system.*

Keywords: *Air quality, artificial intelligence, pollution, ANFIS, soft computing*

I. INTRODUCTION

Many studies about health effects of particulate matter were recently published, since long exposure to PM10 (particles) with diameter smaller than 10 mm) causes respiratory and cardiovascular diseases [1], [2]. Neural computing can model air pollution with more advantages compared with ordinary statistical methods because air dynamics comprise: multiple seasonality and long memory [3], [4]. Use flexible neural tree give good result in Time-series forecasting [5]. An integrated artificial neural network model can forecast the maxima of 24 h average of PM_{10} concentrations [6]. The neural networks can use to develop a mathematical model for predicting daily concentrations of air pollution caused by the traffic in urban areas with high performance [7]. Two environmental prediction systems for two types of applications; one in the area of environmental protection for air pollution prediction and the other in the area of hydrology for flood prediction [8]. An intelligent rule based fuzzy system can use to predict rainfall events [9]. In this study, we propose a hybrid intelligent approach (ANFIS) for predicting PM_{10} concentration. In a hybrid intelligent approach,

Revised Version Manuscript Received on August 11, 2016.

Dr. S. A. Asklany, Department of Computer Science, Modern Academy Maadi, Cairo, Egypt.

Dr. Khaled Elhelow, Department of Mathematics, Elmagmaa University, Elmagmaa, Saudi Arabia.

Prof. M.Abd El-Wahab, Astronomy and Meteorology department, Cairo University, Cairo, Egypt.

Not only combine the learning capabilities of a neural network but also incorporate reasoning by using fuzzy inference with supervised learning capability, Thereby enhancing the capability of the system for prediction, as compared to using a single methodology alone. Previous studies have shown that ANFIS is a better predictor for a chaotic time series as compared to using a single technique like neural network [10]. Moreover, nonlinearity, uncertainty and complexity are the three predominant characteristics of PM_{10} phenomena which cause the process challenge for scientific evaluation. Our models merge both the fuzzy set and neural network theories in order to improve the objectiveness and to produce Now casting models for predicting air pollutant concentrations which frequently used in the “what-if” scenarios. These types of hypothetical scenarios can be help in justifying the benefits by reducing the concentrations of the targeted ambient air pollutants

II. METHODS

Roger Jang 1993, suggested Adaptive Neuro Fuzzy Inference system (ANFIS). It can serve as a basis for constructing a set of fuzzy ‘if-then’ rules with appropriate membership functions to get the stipulated input-output pairs. Here, the membership functions are tuned to the input-output data and excellent results are obtained [11].

In the ANFIS, input series are converted to fuzzy inputs by constructing membership functions for each input. The membership function pattern used for the input series is of the Trinangular shape. The fuzzy inputs with their associated membership functions form the inputs to the neural network. These fuzzy inputs are processed through a network of transfer functions at the nodes of different layers of the network to obtain fuzzy outputs with linear membership functions that are combined to give a single output [12]. Principally, ANFIS is about taking an initial fuzzy inference (FIS) system and adjust it with a back propagation algorithm based on the pairs of input-output data. The basic structure of a fuzzy inference system consists of three components: A rule base, which contains a selection of fuzzy rules; a database, which defines the membership functions used in the fuzzy rules; and an inference mechanism, which performs the inference procedure upon the rules and the given facts to derive a proper output or conclusion. In ANFIS, neural networks recognize patterns that help adaptation to environments. Fuzzy inference systems incorporate human expertise knowledge and perform interfacing and decision-making.

These intelligent systems adapt themselves and learn to do better in changing environments. Figure 1 show the basic ANFIS architecture

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

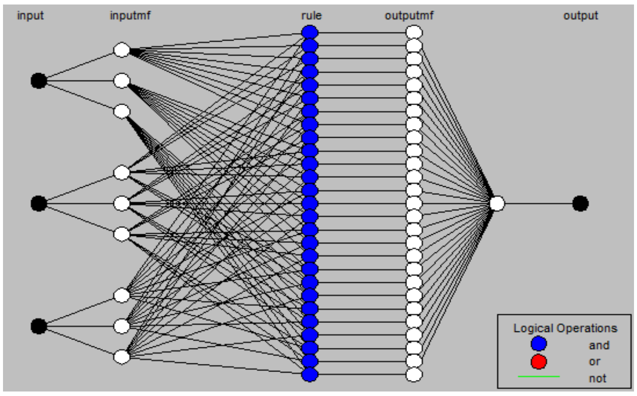

Fig.1. Basic ANFIS Architecture

III. EXPERIMENTS

A. ANFIS as Now casting model for PM_{10} prediction.

We used Neuro-fuzzy to build the model depend on thirteen years data divided into training, checking and testing data sets. First step is handling the data to be normalized between 0 and 1 (Pejman Tahmasebi, Ardeshir Hezarkhani, 2010) i.e. [0 1], by using the following formula:

$$X_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (1)$$

Where x is the data which should be normalized, the x_{max} and x_{min} are maximum and minimum of the original data respectively and X_{norm} is the normalized data after calculated.

Since Normalization, is reasonable for many practical problems, variables have different importance and make different contributions to the output(s). Such a method can be used for feature selection, or for reducing the size of input vectors through keeping the most important variables [13]. This is especially applicable to a neural networks or neuro- fuzzy models [14]. The dataset is then partitioned into a training set and a checking set.

The function **exhsrch= exhsrch(1, trn_data, chk_data, input_name);** performs an elaborate search within the available inputs to select from them the most influence the PM_{10} concentration. The first parameter to the function specifies the number of input combinations to tri during the search, **exhsrch** builds an ANFIS model for each combination and trains it for one epoch and reports the performance achieved. The most important weather parameters we got are: wind speed, wind direction and temperature, selected as three inputs with PM_{10} as single output for the system as in the figure 2.

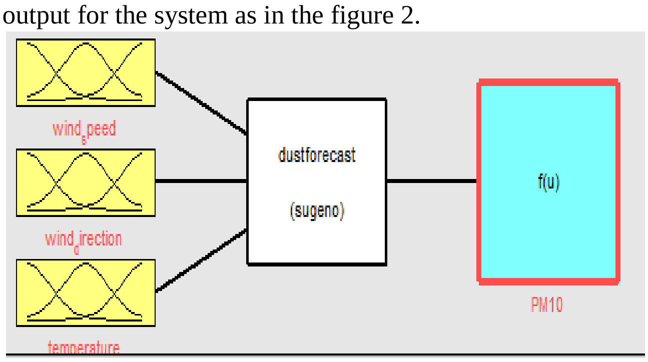

Fig.2. ANFIS system based on the Takagi–Sugeno inference

The process of edit the ANFIS by loading training and checking data sets then generate FIS (Fuzzy Inference System) assign three triangular membership functions to each input parameter. The triangular curve is a function of a vector, x, and depends on three scalar parameters a, b, and c, as given by

$$f(x,a,b,c) = max\left(min\left(\frac{x-a}{b-a},\frac{c-x}{c-b}\right),0\right) \qquad (2)$$

The surface viewer shows the relationship between the inputs and the output which clearly manifest that the concentration of the PM_{10} increase with increasing the temperature and wind speed concurrence with the northerly wind direction as in figure 3.

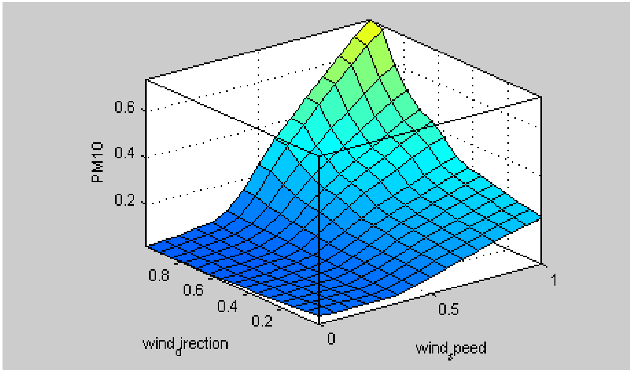

Fig.3. FIS surface viewer wind direction and wind speed with the PM_{10}

A. The model results and validation

After determination the ANFIS structure the parameters associated with the membership functions adjusted through the learning process. The computation of these parameters is facilitated by a gradient vector. This vector gives a measure of how well the fuzzy inference system is modeling the input/output data to provide a set of parameters. When the gradient vector is obtained, several optimization routines can be applied in order to adjust the parameters to reduce the error measure. This error measure is usually defined by the covariance between actual and desired outputs. ANFIS uses a hybrid learning algorithm consisting of backpropagation and least-squares estimation. A common problem in ANFIS model building is the so-called “over fitting”, which occurs when we fit the fuzzy system to the training data so well that it no longer does a very good job of fitting the checking data. The result is a loss of generality. However, optimal number of epochs can only be found through experiments [15].To find the optimal number of epochs to our experiment, the error curve is plotted. We found that the Root Mean Squared Error (RMSE) decrease till reached training epoch equals to twenty.

B. Performance Evaluation

Evaluate the performance of the system has been done by two methods, first calculate the ANFIS root mean square error and then compare it against the linear regression root mean square error.

RMSE (Root Mean Square Error) is used,

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

This defined as: $RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\hat{\theta}_j - \theta_j\right)^2}$(3)

Where N is the total number of prediction, $\hat{\theta}_j$is the predicted output and $\theta_j$is the observed data.

C. Liner Regression

The ANFIS prediction can be compared against a linear regression model by comparing their respective RMSE (Root mean square) values against checking data. Regression analysis is used to predict the value of one variable on the basis of other variables; here linear regression is used to predict the PM_{10} concentration on the basis of wind speed, wind direction and temperature. The ANFIS information is given in the table 1show that RMSE against checking data: ANFIS: 0.115 against Linear Regressions RMSE: 0.417ANFIS model outperforms the linear regression model.

Table 1. Final ANFIS information for the first model.

Number of Nodes:	78
Number of linear parameters	27
Number of nonlinear parameters	27
Total number of parameters	54
Number of training data pairs	63557
Number of checking data pairs	15404
Number of fuzzy rules:	27
Training RMSE	0.07
Checking error	0.115
Linear regression RMSE	0.417

D. The Model Testing

To test the model we used two days data which unseeded during the training process, the result we obtained can present in the figure 4and 5, they show a very good forecasting capability.

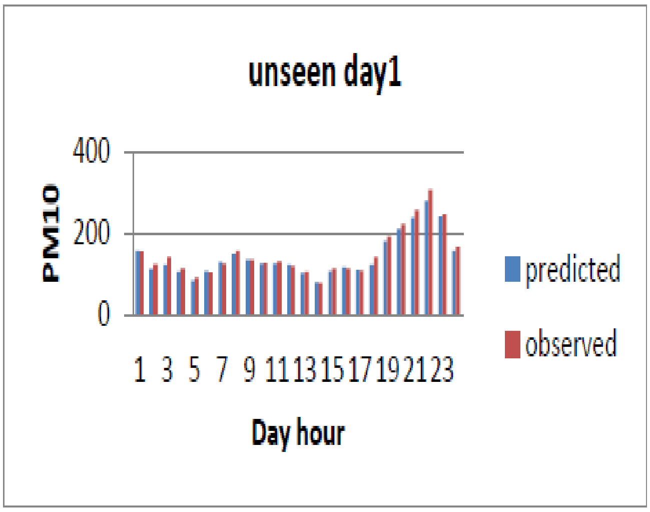

Fig.4. FIS evaluation for test data day 1.

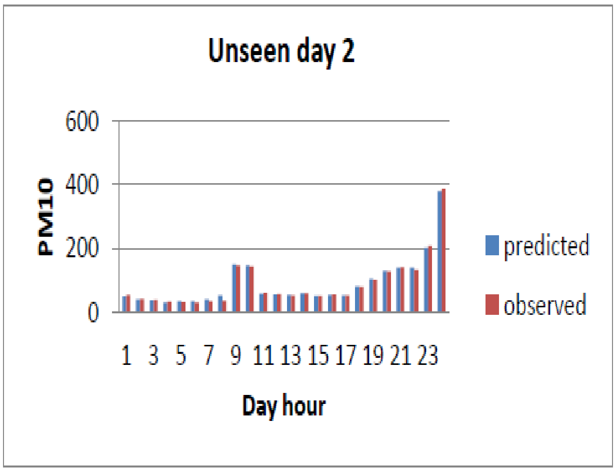

Fig.5. FIS evaluation for test data day 2

E. Chaotic Time Series Prediction (model 2)

Time series is widely observed in many aspects of our lives. The prediction of future values based on past or/and present information therefore is very useful. Many useful methods in time series prediction have been established. Daily temperature, stock market and so forth are examples of time series. Basically, there are two main goals in time series analysis: First, identifying the nature of the phenomenon represented by the sequence of observations and secondly prediction future values [16] .To forecast future of PM_{10} based on its characteristics, Chaotic Time Series can be used. In the last two decades, chaotic time series analysis has attracted a lot of research attention. The interests of these researches have been lately focused on the techniques of chaotic time series modeling and on prediction of future time series values Chaotic behavior can be described as bounded fluctuations of the output of a non-linear system which is very sensitive to initial conditions [17], [18].

One of the mathematical descriptions of chaotic behavior is the delay differential equations (DDEs). A time-dependent solution of a system of DDE is not only determined by its initial state at a given moment; but also the solution profile on an interval with length equal to the maximal delay has to be given. That is, we need to define an infinite dimensional set of initial conditions between t=-$\tau$and t=0. So, DDEs are infinite-dimensional problem, even if we have only a single DDE. One of the most commonly used DDE is the Mackey-Glass model [19] which was proposed as a model for the production of white blood cells but now is quite frequently used in other time series domains [20].

The model built on the following equation:

$$\frac{dx}{dt} = Ax_\tau \frac{\theta^n}{\theta^n + {x_\tau}^n} - Bx \qquad (4)$$

Where A, $\theta$and n are parameters. For τ>17, this equation is known to exhibit chaos.

In this experiment we select the observed PM_{10} data for year (2003) from our data, filter successive 1200 hourly data, 700 records for training and 500 one for check the validation of the model.

In time series prediction the past values of PM_{10} up to time‘t’ are used to predict the value at some point in the future ‘ $t + p$ ’. The standard method for this type of

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

prediction is to create a mapping from D points of the time series spaced 'Δ' apart; that is [*x(t-(D-1)Δ*]….*x(t-Δ),x(t)* to predict a future value $x(t+p)$, where $D = 4$ and $\Delta = p = 6$ are used. The number of times the entire data set is used to check and validate the prediction is called the epoch number. ANFIS is used for the entire process of training and evaluation of FIS .

The first step of model building is to construct a set of initial membership by ANFIS which uses a hybrid learning algorithm to identify parameters of Sugeno type fuzzy inference mechanism, whose output MFs are only linear or constant [21].

ANFIS can serve as a basis for constructing a set of fuzzy if-then rules with appropriate membership functions to generate the stipulated input-output pairs. Once the initial MFs are found, it improve them through training. The most important advantage of using ANFIS is that all its parameters can be trained like a neural network, but with reasoning in a fuzzy logic system [22]. ANFIS applied a combination of the least-squares method and the back-propagation gradient descent method for training. In this study, we use GENFIS1, a commonly used function, to generate an initial single-output FIS matrix from training data. This FIS is used to provide initial conditions for ANFIS training. We find that MFs number "3" and membership function type "gbellmf" (generalized bell MFs) are the appropriate for used after testing many other types of membership functions with number varies from 2 to four. The bell membership function has some advantages such as being a little more flexible than the Gaussian membership functions. Therefore the parameters of network would be more adjusted by using the bell membership function [23].

The generalized bell function depends on three parameters *a*, *b*, and *c* as given by

$$f(x,\sigma,c) = e^{\frac{-(x-c)^2}{2\sigma^2}} \quad (5)$$

F. Performance Evaluation

RMSE (Root Mean Square Error) is used to evaluate the performance of the model. Figure 6 shows the original time series (dotted blue) and the one predicted by ANFIS (solid green).

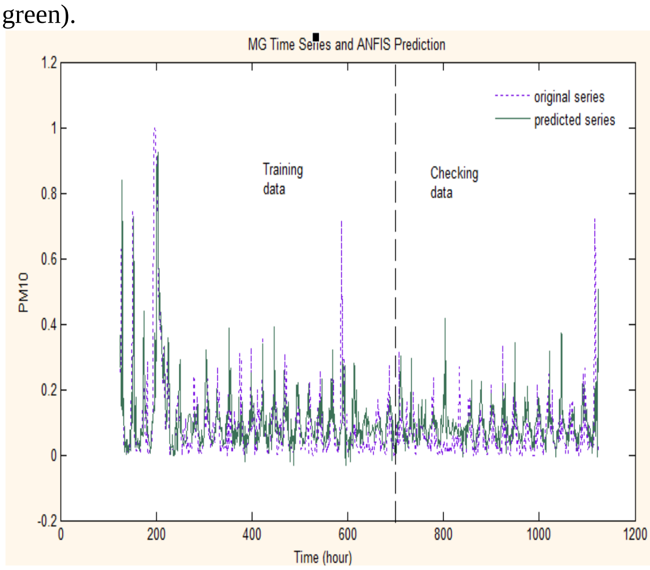

Fig. 6. ANFIS PM10 time series prediction.

The figure shows clearly that the model has very good prediction ability and RMSE of the model is 0.073. ANFIS information can be summarized in table 2

Table 2. Final ANFIS information for model 2.

Number of nodes:	405
Number of linear parameters	27
Number of nonlinear parameters	36
Total number of parameters	441
Number of training data pairs	500
Number of checking data pairs	500
Number of fuzzy rules:	81
Training RMSE	0.112
Checking error	0.072

G. ANFIS against Auto regression (AR)

In order to forecast the future values of a time series, a wide spectrum of methods is available. The traditional method of predicting time series data is auto regression method (AR), which expresses the current value of a time series by a finite linear aggregate of previous values. The definition the AR is as follows:

$$X_t = \sum_{i=1}^{N} a_i X_{t-i} + \varepsilon \qquad (6)$$

Where a_i are the auto regression coefficients, X_t is the series under investigation, and N is the order. The noise term or residual ε, is almost always assumed to be Gaussian white noise.

The aim of AR analysis is to integrate the best values for a_i for a given a series X (t). These methods assume that the series X (t) is linear and stationary. This may be quite advantageous in certain situations, particularly when it is difficult to identify the main components in a series and to construct suitable models for them. However, the very flexibility of AR is also its disadvantage. Unless one has some experience in time series analysis, such models may not yield sensible forecasts. The main problem of AR is that it performs piece-wise linear approximation, and it is difficult to model "volatile" time series. Such a model is likely to break down when it is used for forecasting.. To compare the performance, a model using the traditional autoregressive method is also built with the same data set. In this case, the autoregressive equation is as follows:

$$X(t+6) = a_0 + a_1 X(t) + a_2 X(t-6) + a_3 X(t-12) + a_4 X(t-18) + \varepsilon \quad (7)$$

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

After we program the autoregressive model the function (vgxdisp) can be used to Display multivariate time series parameters and standard errors in different formats after built our AR time series the following are the resulted regression coefficients are reported in the

table 3 From results in table 3 we can see that all parameters obtained are insignificant. This shows the problem of AR method: it is difficult to deal with the non-probabilistic type uncertainties in time-series data. Therefore, in this case it is meaningless to compare RMSE of the model built with the traditional autoregressive method and that of our model. Nevertheless, it shows that the dynamic adaptive neuron-fuzzy logic system proposed in Pm_{10} prediction is more powerful and flexible in pattern recognition when time-series data are used.

Table 3 Regression Coefficients

Parameter	Value	Std. Error	t-Statistic
a(1)	0.493374	0.009816	50.2612
a(2)	-0.58989	0.081595	-7.2295
AR(1)(1,1)	1.50361	0.017052	88.1806
(1,2)	-0.0139	0.002355	-5.90215
(2,1)	0.39209	0.141747	2.76612
(2,2)	0.778644	0.019573	39.7819
AR(2)(1,1)	-0.50332	0.032671	-15.4056
(1,2)	0.010961	0.003011	3.63987
(2,1)	-0.3727	0.271608	-1.37219
(2,2)	0.009429	0.02503	0.376714
AR(3)(1,1)	0.000143	0.031179	0.004589
(1,2)	2.00E-05	0.003019	0.006617
(2,1)	-0.00181	0.259198	-0.00697
(2,2)	-0.03029	0.025094	-1.20706
AR(4)(1,1)	-0.00044	0.013974	-0.03119
(1,2)	-7.29E-05	0.00237	-0.03075
(2,1)	-0.0176	0.116159	-0.15155
(2,2)	0.031845	0.0197	1.61647
Q(1,1)	0.00019		
Q(2,1)	0.00014		
Q(2,2)	0.013141		

IV. RESULTS AND DISCUSSION

The two models presented in this work proved very high performances in PM_{10} prediction figure 4 and figure 5 reflect an efficient predictability of the hybrid inelegant by showing a slight difference between the measured and calculated values for the unseen two days from the measured data with marginally error, where figure 3 Complies with climatic studies on the dust phenomenon the area where the phenomenon associated with the northern high wind speeds and high temperatures which is reflected by the shape clearly.

Figure 6 competently proves the ability of the model to predict the phenomena of non-linear nature like the PM10 by shows almost match between the calculated and expected values where table 3 reflect the difficulty that AR faced when dealing with non-probabilistic type in time series prediction.

V. CONCLUSIONS

In this study, an adaptive neuron-fuzzy inference system is used to build two intelligent models for forecasting the PM10 concentration. Even though FIS models trained usually have very good forecasting ability, their performance is not ideal when applying to predict variable like PM10 which has non probabilistic and chaotic behavior. This raises the possibility that fuzzy logic models could be further improved so they should not only be able to represent frequently occurring relationship but also be able to update itself in view of new data i.e. learning possibility. The needs to background knowledge that will allow model to reinterpret and/or combine concepts in the data into new concepts that can lead to more accurate and/or simpler patterns. That is the direction for the soft-computing intent which applied in construction of our models and approved very good performance.

REFERENCES

1. Alvim-Ferraz, M.C., M.C. Pereira, J.M. Ferraz, A.M.C. Almeida e Mello and Martins, F.G. European Directives for Air Quality: Analysis of the New Limits in Comparison with Asthmatic Symptoms in Children Living in the Oporto Metropolitan Area. Portugal. Hum. Ecol. Risk Assess.11(3 .)pp. 607-616. (2005).
2. Kim, K.-H. , Kabir, E. , Kabir, S. A review on the human health impact of airborne particulate matter. Environ. Int. 74: 136–143. (2015)
3. Zhiguo Zhang and Ye San .Adaptive wavelet neural network for hourly NOX and NO2 Concentrations. Winter Simulation Conference (WSC'04) - Volume 2(2004).
4. Jose´ Luis Aznarte M., Jose´ Manuel Benı´tez Sa´nchez , Diego Nieto Lugilde , Concepcio´n de Linares Ferna´ndez, Consuelo Dı´az de la Guardia , Francisca Alba Sa´nchez Forecasting airborne pollen concentration time series with neural and neuro-fuzzy models, Expert Systems with Applications. 32: 1218–1225. (2007)
5. Yuehui Chen, Bo Yang and Ajith Abrahan.Time-series forecasting using flexible neural tree model. Information Sciences: an International Journal. 174, Issue 3-4 :219 - 235 (2005)
6. Patricio Perez and Jorge Reyes. An integrated neural network model for PM10 forecasting. Atmospheric Environment. 40: 2845–2851. (2006).
7. Bogdana VUJIĆ, Srđan VUKMIROVIĆ, Goran VUJIĆ, Nebojša JOVIČIĆ, Gordana JOVIČIĆ and Milun, BABIĆ. Experimental and Artificial Neural Network Approaches for forecasting of traffic air pollution in urban areas: the case of Subotica. Thermal Science. 14:1-7(2010).
8. Oprea, M. and Alexandra, M. Applying Artificial Neural Networks in Environmental Prediction Systems, Recent Advances in Electrical Engineering Proceedings of the 11th WSEAS international conference on Automation & information:110-115. (2010).
9. Somia A. Asklany a, Khaled Elhelow , Youssef I.K. and Abd El-wahab, M. Rainfall events prediction using rule-based fuzzy inference system. Atmospheric Research. 101 : 228–236 (2011)
10. Manish Kakar, Hakan Nystrom, Lasse Rye Aarup, Trine Jakobi Nøttrup and Dag Rune Olse. Respiratory motion prediction by using the adaptive neuro fuzzy inference system (ANFIS). Phys. Med. Biol. 50 :4721–4728(2005).

11. Ciji Pearl Kurian , George, V.I., Jayadev Bhat and Radhakrishna S Aithal. ANFIS Model for the time series Prediction of interior daylight illuminanc.AIML Journal. 6 (3): 35-40(2006).
12. Mehmet Tektaş Weather Forecasting Using ANFIS and ARIMA MODELS A Case Study for Istanbul, Environmental Research, Engineering and Management.1(51):5 – 10. (2010).
13. Song, Q., & Kasabov, N. Weighted data normalizations and feature selection for evolving connectionist systems proceedings. In Proceedings of the eighth Australian and New Zealand intelligence information systems conference : 285–290. (2003).
14. Qun Song, Nikola Kasabov . TWNFI—a transductive neuro-fuzzy inference system with weighted data normalization for personalized modeling, Neural Networks, Vol. 19 , Issue 10, pp. 1591-1596 (2010).
15. Weiping Liu. Forcasting exchange rate change between USD and JPY by using dynamic adaptive Neuro-Fuzzy logic system. Asia Pacific Journal of Finance and Banking Research. 2(4) pp. 1-12. (2008).
16. Jang, J.S.R. ANFIS: Adaptive-Network-Based Fuzzy Inference Systems, IEEE Transactions on Systems. 23(3): 665-685. (1993).
17. Casdagli, M. A Dynamical Systems Approach to Modeling Input-output Systems, in Nonlinear Modeling and Forecasting. SFI Studies in the Sciences of Complexity Process, Addison-Wesley, New York, 12: 265-281. (1992).
18. Maysam Behmanesh, Majid Mohammadi, Vahid Sattari Naeini. Chaotic Time Series Prediction using Improved ANFIS with Imperialist Competitive Learning Algorithm, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-4 Issue-4:25-33.(2014)
19. Glass, L. and Mackey, M.C. From Clocks to Chaos, the Rhythms of Life, Princeton University Press. (1988).
20. Yadav, R.N. , Kalra, P.K. and John, J. Time series prediction with single multiplicative neuron model. Applied Soft Computing. 7 :1157–1163. (2007).
21. Sugeno, M. Industrial Applications of Fuzzy Control. Elsevier Science Publication Company. (1985)
22. Kodogiannis, V. and Lolis, A. Forecasting Financial Time Series Using Neural Network and Fuzzy System-based Techniques. Neural Computing & Applications. 11: 90-102. (2002).
23. Pejman Tahmasebi, Ardeshir Hezarkhani. Application of Adaptive Neuro-Fuzzy Inference System for Grade Estimation; Case Study, Sarcheshmeh Porphyry Copper Deposit, Kerman, Iran. Australian Journal of Basic and Applied Sciences .4 (3):408-420. (2010).