J. Automation & Systems Engineering 3-2 (2009): 76-84

Regular paper

Journal of Automation & Systems Engineering

Recurrent Neural Network for Multi-steps ahead prediction of PM10 concentration

Sabri Ghazi, Mohamed Tarek Khadir

This paper describes the development of a Multi Layer Perceptron (MLP) recurrent neural network based model to perform a multi-steps ahead prediction of pollutant concentration. Receiving the latest k measurements of pollutant concentration and the meteorological parameters the model is able to predict the next k hourly concentration. The model has been applied to predict the PM10 (Particulate Matter with an aerodynamic diameter of 10 micrometer) concentration in Annaba city, northeast of Algeria (north of Africa). Compared with a single step prediction MLP model the recurrent model perform best for short-term prediction and give interesting performances for long-term prediction

Keywords: Artificial Neural Network, Recurrent Neural Network, Air quality prediction, NARX, PM10.

1. INTRODUCTION

Air pollution is a major concern in many cities in the world, many factors influences in air pollutants concentrations the most important are: metrological conditions, topology and population density. This makes air pollution difficult to model. Many air pollution prediction models have been studied (mathematical emission models, linear models, Artificial Neural Networks based models) in order to design an air quality prediction system to control the air pollution, and limit the influence of peaks periods by informing community, and taking the necessary precaution. PM10 is the sum of all solid and liquid particles suspended in air many of which are hazardous. This complex mixture includes both organic and inorganic particles, such as dust, pollen, soot, smoke, and liquid droplets. These particles vary greatly in size, composition, and origin. Particles in air are either: directly emitted, for instance when fuel is burnt and when dust is carried by wind, or indirectly formed, by a photochemical interaction between gaseous pollutants in the air space. The paper is organized as follows. In the next section, a brief regional description is presented; the section 3 describes the use of ANN for air pollution prediction. In particular, Section 4.1 presents the NARX model that is able to give multi-steps ahead prediction is presented, and in Section 4, some experimental results are shown, comparing the performance of the proposed NARX model with single step prediction MLP model that uses the identical inputs. Finally, Section 5 presents some conclusions.

2. REGIONAL DESCRIPTION

Annaba is situated in northeastern of Algeria (Figure 1), it is one of the biggest Algerian cities, it is known as an industry pole. With concentrated industrial activities, such as steel industry and petrochemical stations, the air quality becomes very bad especially in summer where certain meteorological conditions affect the PM10 concentration. The population near the industrial zones suffers very alarming air pollution rates. An air pollution control

LabGED, Department of computer science, University Badji Mokhtar Annaba, Algeria. ghazi@labged.net, khadir@labged.net

system will be very useful to keep the population informed especially those suffering a chronic disease such as Asthma.

Fig. 1. Geographic map of Annaba, northeast of Algeria (Google Earth ©).

3. ANN FOR PREDICTION OF AIR POLLUTANT CONCENTRATION

Artificial Naturals Networks (ANN) are an alternative to classical statistics methods used for prediction in air quality monitoring stations. It will be very efficient to use the large amount of available data from monitoring stations in order to design efficient predictions models such as ANN based models. Relationship between meteorological parameter and air pollutant concentration is very difficult to model, Malkar and Boznar [1] presents an MLP based model for SO2 concentration prediction, it receives the meteorological and emission parameters as inputs, the model shows an efficient predictions and outperform the ARMA (Autoregressive Moving Average) based model. Malkar et al. [2] presents a pattern recognition inspired approach to select the parameter for the air pollution prediction model, in order to optimize the training time and increases model performance.

To get the most adaptable model for air pollution prediction Jorquera et al. [3] presents a comparison between three models: ANN based model, linear model and Fuzzy logic based model. These models have been applied to predict Ozone concentration in Santiago Chile, the Fuzzy and ANN model has shown an efficient prediction and outperform the linear model. Gardner and Droling [4] presents NO2 prediction model based on MLP topology, compared with linear regressor using the some input data and parameter, the MLP based model show better accuracy. In Gardner and Droling [5], the long-term tendency of Ozone concentration in London has been studied using ANN based model. The prediction of PM2.5 (Particulate matter with diagonal smaller the 2.5 micrometer) concentration has been well studied in Perez et al. [6], using ANN based model which is able to predict short-term concentration of PM2.5. In Rob et al. [8] a DVS (Delay Vector) method has been applied to detect the nonlinearity of the Ozone time series, and after detecting the nonlinearity two models have been used: recurrent NARMA based model, and MLP based model, where the two models have presented the very similar performances. The influence of metrological parameter on the precision of the prediction has been well studied in Hooyberghs Jef et al.

[9], where for each parameter an MLP based model is designed to receive the parameter and the lagged input. The comparison of the performances of these networks has shown that the BLH (Boundary Layer Height) is the most important parameter that affect the accuracy of the model. Another method is proposed by Corani [10] consisting to eliminate the parameter which has the lowest weight in an MLP network by the application of the OBD method (Optimal Brain Damage) [11].The obtained neural model has been compared with Lazy Learning based model in order to predict the Ozone and PM10 concentration in Milan, the lazy learning model has shown a better performance and outperform the neural model. Many air pollution prediction models generate the value of the studied pollutant concentration, in Prez et al. [12], the proposed MLP based model must give a decision concerning the class of air quality (A: good, B: bad, C: critical), the model has been compared with equivalent linear model, and has shown better performance. A Cyclo-stationary model for NO2 concentration prediction has been proposed in Bianchini et al. [13] , N MLP networks are designed, where N is the cyclo-stationary order of time series, the Cyclo-stationary algorithm show good performance especially for the prediction of the peak period. Several ANN of different topologies (MLP, RBF) have been compared in Ordieres [14] to predict of PM2.5 concentration, the RBF topology has shown the better performance. An Elman neural network based model is proposed in Brunelli et al. [15] for a three steps ahead prediction of SO2 concentration, another study is presented in Nikhil and al, [16] where an ANN prediction models is presented, in order to control and to predict the H_2 concentration in station of bio-hydrogen production, the model shows a good performance and the H_2 production process has been more optimized.

In this work we use a recurrent neural network model of the Nonlinear Autoregressive with eXogenous inputs (NARX) for the multi-steps ahead prediction of air pollutant concentration. The NARX model is very interesting for efficiently study the air pollution phenomena, by giving a long prediction horizon lets the user know the development of the pollutant concentration in the next k hours.

4. MODEL DEVELOPMENT

Many prediction models provide a One-step-ahead prediction, by estimating only the next value of time series using a fixed prediction horizon of 1 (one), this mean that the output of the model t+h depend on the lagged input passed in t.

It will be very useful for the user to get the h next values, where h is the horizon of the prediction. Providing the [t+1,.., t+h] value estimation will help the user in studying efficiently the behaviors of the phenomena like the air pollution.

A. Data preprocessing

For the proposed of the work we used data collected by the Algerian air quality monitoring network called SMASAFIA. SMASAFIA provides a real-time air pollution monitoring system composed by mobile and fixed stations. The available data is an hourly measurements of pollutants concentration of two years, form 1/1/2003 to 31/12/2003 and form 1/1/2004 to 31/12/2004, in Annaba exactly the station number 4.The table -I- presents the statistical properties of the data. The data set of 2003 was taking for training and those of 2004 of validation; this will help us to efficiently test and get the performance of the model. The pollutant concentration measurements are in microgram/ m3. To adapt the data with our model, we have applied normalization using equation (1), this will make the input ranges between [0 1], and in the output we will obtain a percentage indicating the degree of the peak alarm.

$$V_p = \frac{v_p}{\max(v_p) - \min(v_p)} \quad (1)$$

Where V_p is a vector of parameter, min and max are function that returns the maximum and minimum of the vector.

TABLE I: STATISTICAL PROPERTIES OF THE DATA

Parameter	2003 MEAN	2004 MEAN	2003 STD	2004 STD
PM10	51.70	27.76	51.66	26.38
Wind speed m/s	2.65	2.12	1.78	1.27
Humidity	63.52	71.92	16.50	14.33
Temperature	18.69	16.82	7.76	6.30

B. One-step ahead prediction

The single step prediction based models are widely used in air pollution prediction, receiving in training step, a past measures of concentration and the meteorological parameters as inputs, and the t+h as forced output, the model can efficiently perform a one-step-ahead prediction, the model can be formalized by :

$$PM10(t+1) = f\left(PM10(t), TEM(t), Hu(t), WS(t)\right) \quad (2)$$

Where:

t: the time in hour,

PM10: measured of PM10 concentration value,

TEM: measured of humidity,

WS: measured of wind speed.

After several experiments we obtained the best architecture of the single step ahead model given by a (4-25-1), with 25 neurons in the hidden layer, and 4 neurons in the inputs layer.

We used the log sigmoid as activation function. Table -II- shows the performances in terms of MSE (Mean Squared Error) of the tested networks architecture.

TABLE II : PERFORMANCES OF TESTED MLP MODELS

Architecture of network	MSE error using the validation set	MSE error using the training set
4-10-1	2.113	0.061
4-12-1	3.119	0.043
4-15-1	4.1111	0.040
4-20-1	2.1114	0.370
4-25-1	1.7006	0.046

C. Multi-steps ahead prediction:

A recurrent neural network model (NARX), and it can be formalized by:

$$y(t) = f\left(u(t-2), u(t-1), .., u(t), y(t-k), .., y(t-1)\right) \quad (3)$$

Where $u(t)$ and $y(t)$ are the input and output of the model at time t, k is the number of lagged inputs, the function f can be mapped by an MLP network. Figure -2- illustrates a simple NARX model. Recurrent neural network are capable of representing arbitrary nonlinear dynamical mappings Cybenko [15], such as those commonly found in nonlinear time series prediction. They have been applied for solving a variety of prediction problems, such as forecast the electric load Costa et al. [17]. The NARX model receives as input the metrological parameters and the k last hourly measures of the PM10 and N2 concentration, and gives in output the k peak of the next hourly PM10 and N2 concentration, the prediction horizon is equal to k.

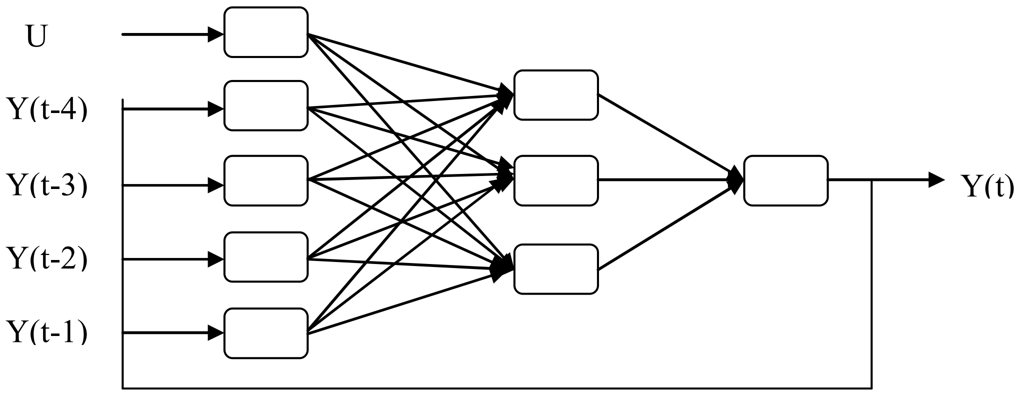

Fig. 2. Simple NARX model, with ARX order k=4

TABLE III : THE PERFORMANCE OF THE TESTED NARX MODELS

K the Delay order	IA Validation set	IA training set	MSE error validation	MSE error training
2	0.7662	0.9468	1. 18	0.015
3	**0.8947**	**0.9520**	**1. 24**	**0.028**
4	**0.8830**	**0.9432**	**1. 36**	**0.017**
5	0.8243	0.9283	1. 38	0.021
6	0.8372	0.9198	1. 40	0.017
7	0.7438	0.9011	1. 45	0.020
8	0.8376	0.8485	2. 18	0.020
9	0.7950	0.8554	2. 45	0.023
10	0.7734	0.8270	2. 49	0.025
11	0.8000	0.8300	2. 86	0.021
12	0.7800	0.7885	3. 58	0.028
MLP (single step)	**0.8625**	**0.96253**	**1. 38**	**0.040**

V. SIMULATION RESULTS

In order to test the performances of the proposed models, we used a measured data of the PM10 time series. We used the 2004 hourly measures of meteorological parameters and the PM10 concentrations values as inputs, for each vector the model calculates the next k hourly concentrations. This can clearly show the behaviors and performance of the model. We have changed the delay order of lagged input (k) form 2 to 12, and have obtained 11 different networks configurations. The index of agreement cited in Gardner and Dorling [4], and formalized by equation (4) to calculate the performance of the model.

$$IA = 1 - \frac{\overline{\left(c_p - c_0\right)}^2}{\left|c_p - \bar{c}_0\right|^2 + \left|c_0 - \bar{c}_0\right|^2} \qquad (4)$$

Where:

Cp is the predicted value of the concentration, and C_0 is the measured value. Table -III- presents the performance of the tested networks. Figure 4 presents a multi-steps prediction (3 values) of the k=h=3.

Figure 5, 6 and 7 show the performances of the NARX model with k=h=4, 5, 6. Respectively figure 8 shows a large window of prediction of the NARX with k=h=4 hours predictions.

The NARX model shows a good performance in a short-term prediction, and outperforms the single step ahead model. As shown in the Table III, for 3 and 4 steps prediction horizon the NARX model gives a value of IA of 89.47 and 88.30 of successful prediction and the single step MLP predictor gives 86.25 of successful prediction.

In a long-terms prediction 8 hours ahead (k=8) the model shows interesting performance, Figure 9 presents a long-terms prediction 8 hours in advance.

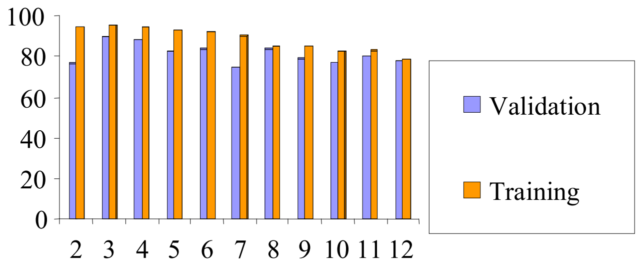

Fig 3. Performance of NARX models.

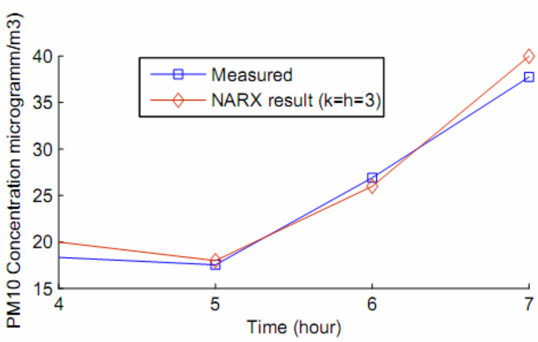

Fig 4. Results of NARX with 3 ARX order.

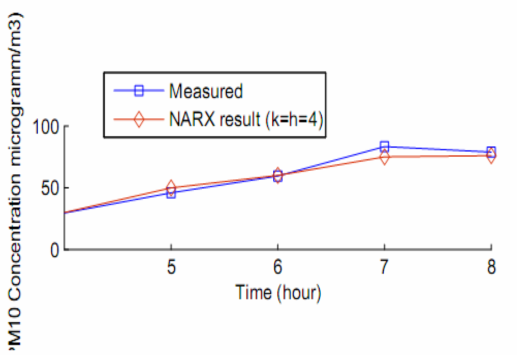

Fig. 5. The Results of NARX with 4 ARX order.

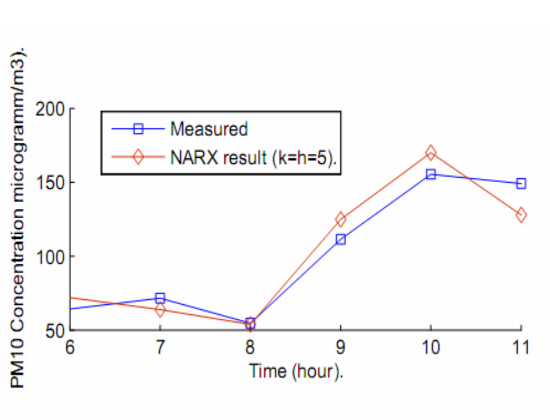

Fig. 6. Results of NARX with 5 ARX order.

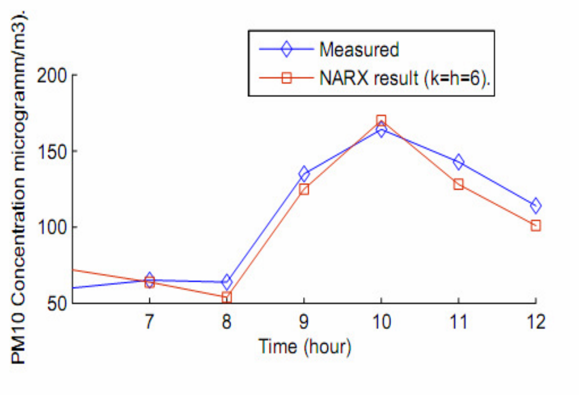

Fig. 7. Results of NARX with 6 ARX.

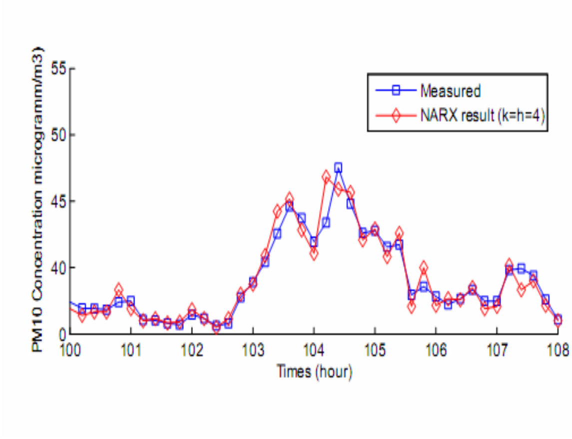

Fig. 8. Result of NARX with 4 ARX order.

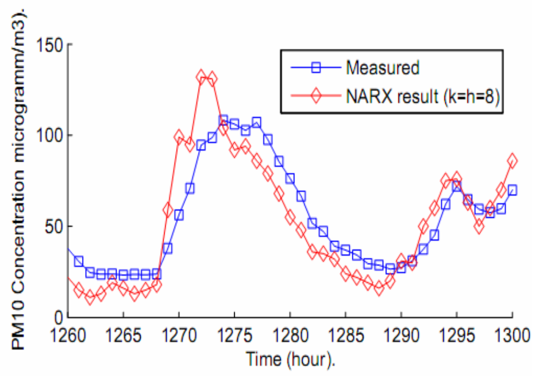

Fig. 9. Result of NARX with 8 ARX order.

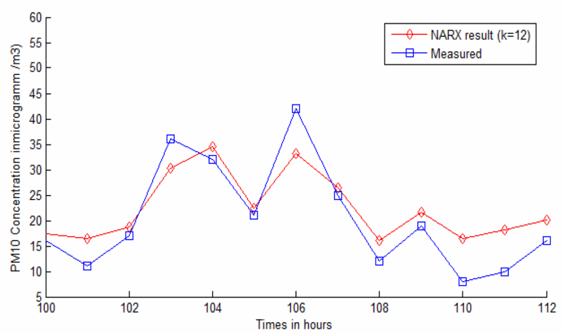

Fig. 10. Result of NARX with 12 ARX order.

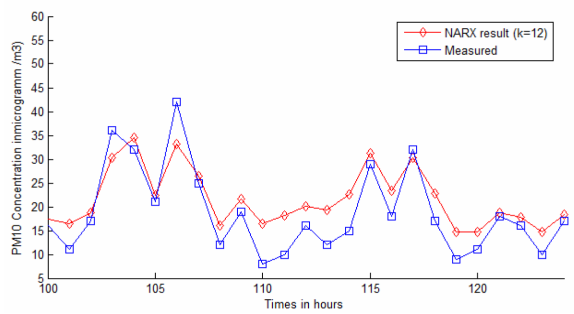

Fig. 11. 24 hours, result of NARX with 12 ARX order.

We can see clearly that the NARX model with K=12, can give encouraging results. Figure 10, illustrates 12 hours estimation of the PM10 concentration, figure 11 present the next day prediction of the PM10 concentration. In the case of long-terms prediction the NARX model can well predict the peaks periods, and do not overestimate theses values.

Figure 12, illustrate the performance of the NARX model with k=12 on a large data window (15 days), we can observe that the model is able to give interesting performance in the different position of the time series.

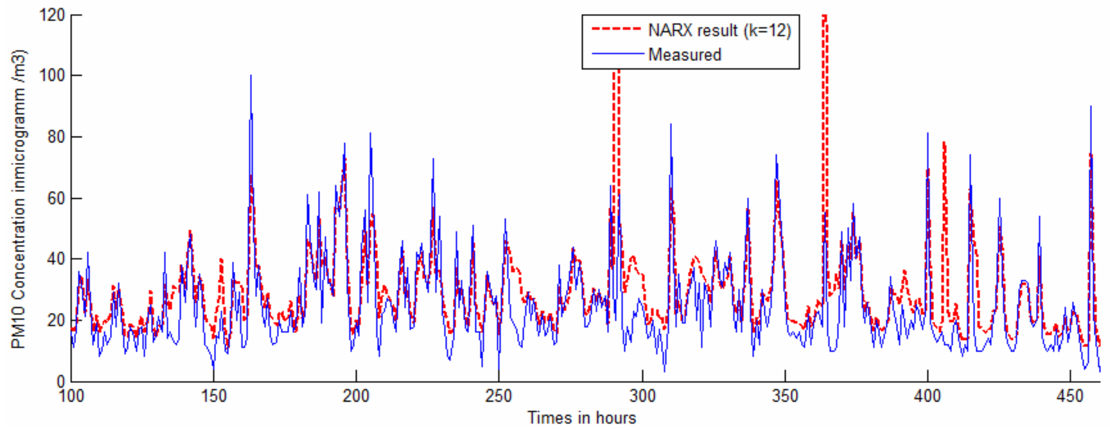

Fig. 12. 15 Days result of NARX with 12 ARX order.

6. CONCLUSION

The NARX based models are able to give efficient short-term (h= 3, 4 hours) prediction of air pollutant concentration, and outperform equivalent single steps ahead based model. For long-term prediction (horizon = 5 to 12 hours) the NARX model gives interesting performances and may be useful for predicting of the maximum or average of the daily pollutant concentration. Adding emission parameter may also enhance the accuracy of the prediction results.

ACKNOWLEDGMENT

This work was supported by the TASSILI program n 07 MDU 714.

REFERENCES

[1] P. Mlakar, and M. Boinar, Short-term air pollution prediction on the basis of artificial neural network. 2nd International Conference on Air pollution, Barcelona, Computer Simulation Southampton, Boston, Computational Mechanics Publ, pages 545-552, 1994.

[2] P. Malkar and S. Josef, Determination of Features for Air Pollution Forecasting Models, Proceedings of 3rd ed. International Conference on Intelligent Information Systems (IIS ’97) IEEE, 1997.

[3] H. Jorquera, R. Prez, A. Cipriano, A. Espejo, M.V. Letelier and G. Acua, Forecasting ozone daily maximum levels at Santiago, Chile, Atmospheric Environment, Volume 32, Issue 20, Pages 3415-3424, 25 September 1998.

[4] M.W. Gardner and S.R. Dorling, Neural network modeling and prediction of hourly NOx and NO2 concentration in urban in London, Atmospheric Environment, vol. 33, pages 709-719, 1999.

[5] M.W. Gardner, and S.R. Dorling, Meteorologically adjusted trends in UK daily maximum ozone concentrations, Atmospheric Environment, vol. 34, Issues 2, pages 171-176, 1999.

[6] P. Perez, A. Trier, J. Reyes, Prediction of PM2.5 concentrations several hours in advance using neural networks in Santiago, Chile., Atmospheric Environment 34, pages 1189-1196, 2000.

[7] P. Perez, J. Reyes, Prediction of maximum of 24-h average of PM10 concentrations 30 h in advance in Santiago, Chile, Atmospheric Environment vol. 36, pages 4555-4561, 2002.

[8] R. Foxall, I. Krcmar, G. Cawley, S. Dorling and D.P. Mandic, Nonlinear modeling of air pollution times series, IEEE 0-7803-7041,2001

[9] J. Hooyberghs, C. Mensink, G. Dumont, F. Fierens and O. Brasseur, A neural network forecast for daily average PM10 concentrations in Belgium, Atmospheric Environment, Volume 39, Issue 18, Pages 3279-3289 , June 2005.

[10] G. Corani, Air quality prediction in Milan: feed-forward neural networks, pruned neural networks and lazy learning, Ecological Modelling, Volume 185, Issues 2-4 , Pages 513-529, 10 July 2005.

[11] D. Stork, and B. Hassibi, Second order derivatives for network pruning: Optimal Brain Surgeon, In T. J. Sejnowski G. E. Hinton and D. S. Touretzky, editors Advances in Neural Information Processing Systems (NIPS), 5, pages 164-171, 1993.

[12] M. Bianchini, E. Di Iorio, M. Maggini, C. Mocenni and A. Pucci, A Cyclostationary Neural Network Model for the Prediction of the NO2 Concentration, ESANN'2006 proceedings – European Symposium on Artificial Neural Networks Bruges (Belgium), d-side publi., ISBN 2-930307-06-4, 2006.

[13] J. B. Ordieres, E. P. Vergaraa, R. S. Capuzb and R. E. Salazarc, Neural network prediction model for fine particulate matter (PM2.5) on the US-Mexico border in ElPaso, Environmental Modeling and Software, Volume 20, Issue 5, Pages 547-559 , May 2005.

[14] U. Beunelli, V. Piazza, L. Pignato, F. Sorbello and S. Vitabile, Two-days ahead prediction of daily maximum concentrations of SO2,O3, PM10, NO2, CO in the urban area of Palermo, Italy , Environmental Atmospheric environment, vol. 41, no 14, pages 2967-2995, 2007

[15] G. Cybenko, .Aproximation by superposition of sigmoidal function, Math. Cont. Sig. syst., vol. 2, pages 303-314, 1989.

[16] N.B. Özkaya, A. Visa, C.Y. Lin, J.A. Puhakka, and O. Yli-Harja, An Artificial Neural Network Based Model for Predicting H2 Production Rates in a Sucrose-Based Bioreactor System, proceedings of World Academy of Science, Engineering and Technology, volume 27, pages 1307-1312, 2008,

[17] M. Costa, E. Pasero, F. Piglione, and D. Radasanu, Short term load forecasting using a synchronously operated recurrent neural network, Proceedings of the International Joint Conference on Neural Networks, 1999.