Optimization of Neural Predictor for Air Pollution

Raihane Mechgoug, A. Taleb Ahmed, and Lakhmissi Cherroun

***Abstract*– This paper presents a neural network modeling applied to predict monthly ozone and carbon dioxide concentration time series. The performances of neural network are critically depending on learning parameters, on its structure and control parameters. Even when a suitable setting of parameters (weights) can be found, the ability of the resulting network to generalize the data after learning may be far from optimal. For these reasons it is well suitable and attractive to apply an optimization method using genetic algorithms. In this paper, we show how Genetic algorithms may provide a powerful tool for automating the design of neural networks. This pattern presents a better fit of non linearity concerning the air pollution data. Simulation results show the effectiveness of the proposed optimization method. The results are compared and discussed to demonstrate the better quality of the forecasting model.**

***Index Terms*—Air pollution, Neural network, Genetic Algorithm, Prediction, Time Series.**

I. INTRODUCTION

OZONE and carbon dioxide are a pollutant resulting from photochemical reactions of a variety natural and anthropogenic precursors. Under favorable meteorological conditions, ozone and dioxide carbon may accumulate in the atmosphere and reach such a high concentration level that can impose adverse effects on human health and ecosystem. Therefore, Air pollution can cause the human health, plants growth and daily mortality in numerous studies over the past decade. Therefore, forecasting and analysis of air quality are important topics of research today [1][2].

A time series is a finished suite of data indexed by time $(x_1, ..., x_n)$. The time indication can be a minute, hour, day, week, year… etc. An important class of the stochastic model used for the prediction describes the relation between the future value of time series and the previous ones. The choice of the model depends on the nature of series (stationary, non stationary, linear and non linear) [3][4].

Manuscript received October 18, 2011; revised December 16, 2011.

Raihane Mechgoug is with the Electrical Engineering Department, University of Biskra, Algeria, e-mail: raihane_2005@ yahoo.fr

A. Taleb Ahmed is with LAMIH FRE CNRS 3304 UVHC Le Mont Houy, Valenciennes, France, e-mail: taleb@univ-valenciennes.fr.

Lakhmissi Cherroun is with laboratory of Micro-waves Devices and Materials for Renewable Energies (DIMMER), University of Djelfa, Algeria, e-mail: cherroun_lakh@yahoo.fr.

The stationary time series Box and Jenkins (1970) suggested using the autocorrelation function and empirical partial autocorrelation functions. Their methodology based on the theoretical autocorrelation function of model mobile averages $MA(q)$ nullifies from the rank $q+1$ and the partial autocorrelation function of an autoregressive mod $AR(p)$ nullifies from the rank $p+1$. An alternative for the selection of model was introduced by Cleveland (1972). Its method uses the inverse autocorrelation function $\rho_i(h),(h \in Z)$, defined by the inverse of the spectral density. Several researches were afterward consecrated to the study of the inverse autocorrelation functions and inverse partial autocorrelation. In the case of the linear processes, Bhansali (1980) was the first one published a complete study concerning the asymptotic distribution of the inverse autocorrelation function estimated by both methods evoked previously. After these works, Bhansali (1983) developed a new approach based on the properties of the Hilbert spaces for defining the inverse partial autocorrelation function. The series not stationary undergo a pretreatment before using the model ARIMA which transform them to stationary series by using an operator according to the nature of stationeries (average not stationary, variance not stationary and seasonality), for non linear series; the models ARCH and GARCH and EAR were used [5][6]. The problem is forecasts are built based on transformed series which can lose their optimal characteristic when we express them in the same way as the initial data. These models remain very difficult to apply in most of the real temporal series.

In recent years, some methods have been proposed for handling forecasting air pollution. Jure & Rahela [7] presented a Q^2 learning to the analysis and prediction of ozone concentration. An Elman Neural Network for hourly SO_2 ground concentration is proposed by Brunelli and Piazza [8]. YU and Wenfang [9] employed the multi artificial neural network, where a model optimizing the back-propagation algorithm by using genetic algorithms to forecast ozone concentration and a support vector machine was used to accurately classify the data into their corresponding categories. Biljana and Mile [10] proposed the prediction of data for ozone concentration in ambient air using a modeling technique of support machine and radial neural networks.

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

The neural networks have particular's properties which make them a prime target and suitable for solving a wide family of problems. However, they also have some limitations and disadvantages which limit their individual application in some cases [11]. There are a various forms for combining the neural network and genetic algorithms. The study of neural network and genetic algorithms combination can be summarize in three general aspects: learn genetically the neural network, genetic optimization of network topology and optimization of its parameters [11]. The combination of neural network and genetic algorithms allows merging advantages of both approaches and having more capacity to surmount difficulties and the limitations which characterize each approach [3].

The objective of this paper is to use a real genetic algorithm to automate de design of a neural network predictor; where a multi-layer feedforward neural network is used with back-propagation learning algorithm. Its topology and control parameters are selected by an optimization process. The designed predictor will be employed for predicting the monthly ozone and carbon dioxide concentration time series in Arosa, Switzerland.

This paper is organized as follows. In section 2, we present briefly a review of the prediction and neural network theory. In section 3, the optimization of the neural predictor using the real genetic algorithms is explained and discussed. Section 4 gives the simulation results of application of this method to predict the monthly ozone, carbon dioxide concentration time series in Arosa Switzerland. A comparison study with other models will be presented in section 5. Section 6 Conclude the paper.

II. PREDICTION AND NEURAL NETWORK (NN)

A neural network (NN) contains a number of relatively simple processing units called neurons. The neurons are interconnected through synaptic weights, and are grouped in layers using synaptic links for connecting neurons in adjacent layers. Three different layers types can be distinguished: input layer, hidden and output layers [12]. The NN architecture considered here is a multilayer perceptron and the training phase is employed using Back-propagation algorithm (BP).

The topology of neural predictor requires the determination of various parameters: number of hidden layers (NHL), number of nodes (Nhl) for each layer *l*, activation functions (f^{l}, f^{out}) for each layer, Presence or absence of the bias (b), learning rate (η), Momentum coefficient (α), input number (N), times that separate two successive input (δ). The GA has been used to determine the appropriate set of the parameter listed above.

III. OPTIMIZING NEURAL PREDICTOR STRUCTURE

Neural predictor (NP) is defined by genetic encoding in which the chromosome composed of the different characteristics of NP. We consider a specific representation of the parameters related to the neural predictor structure; the chromosome is constituted of 17 genes containing the neural model (Number of hidden layers (NHL), numbers of neurons in each layer l (Nhl)), and other genes representing the activation function type for the hidden layers and the output layer (f^{l}, f^{out}), number of neurons in the input layer (N), presence or absence of bias (b), sampling lag (δ), learning rate (η), momentum coefficient (α), variation interval of the synaptic weights for the 2 hidden layer and the output layer.

A. Generation of the Initial Population

The chromosome in the initial population is produced by a stochastic generation as represented in figure 1. Each gene is defined in a subset S_i, where $i = 1...7$ reported in the second row of the Table 1.

TABLE.I
THE SUBSET OF EACH GENE

	NHL	η	α	b	N	Nh1	Nh2	δ	f^1	f^2	f^{out}
set	S_1	S_2	S_2	S_3	S_4	S_5	S_5	S_6	S_7	S_7	S_7

	$high_1$	low_1	$high_2$	low_2	$high_3$	low_3
set	S_2	S_2	S_2	S_2	S_2	S_2

Generally, in most applications of NN; it is sufficient that the network has few hidden layers. The value of (NHL) has been allowed to vary from 1 to 2 while the maximum value of (Nhl) is calculates using the following equation:

$$MAX_{Nhl} = \frac{(N+1)}{2} + \sqrt{ND} \quad (1)$$

Where ND : number of training data. All genes are defined in the following subsets:

$$S_1 = \{1,2\}, S_2 = [0,1], S_3 = \{0,1\}, S_4 = \{2,3,..,30\}, \quad (2)$$
$$S_5 = \{3,4,..,40\}, S_6 = \{2,3..,10\}, S_7 = \{1,2,3\}.$$

For S_7 the value $1 \equiv$ the activation function is sigmoid, $2 \equiv$ the activation function is hyperbolic tangent, $3 \equiv$ the activation function is pure linear. For S_3: $0 \equiv$ absence of the bias, $1 \equiv$ presence of the bias. $[low_i, high_i]$: present the interval variation of the synaptic weights in the two hidden layers and output layer.

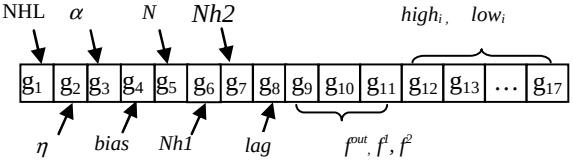

Fig. 1. The chromosome representation.

The subsets S_i where $i = 1:7$ are the same for both ozone and dioxide carbon predictions.

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

B. Fitness Function

To evaluate the goodness of an individual, the neural network is trained with a fixed number of alleles and then will be evaluated according to the obtained parameters. The parameters which seem better to describe the goodness and the effectiveness of the neural configuration that optimize the mean square error *MSE* at the end of training phase. This function is calculated by the following equation:

$$f(x) = \frac{\sum_{i=1}^{N} \left(T_i - \hat{T}_i\right)^2}{N} \quad (3)$$

T_i: Measured value, $\hat{T}_i$: predicted value and N: the number of observations.

C. Genetics Operators

1. Selection: is the process of selecting a pair of individuals from the population. In our work, we used *remainder selection* that assigns parents deterministically from the integer part of each individual's scaled value and then uses roulette selection on the remaining fractional part [13].

2. Crossover: is the reproduction process, in which two chromosomes exchanges some of their corresponding genes [13]. The crossover used is *heuristic crossover* that returns a child that lies on the line containing the two parents, a small distance away from the parent with the better fitness value in the direction away from the parent with the worse fitness value. We can specify how far the child is from the better parent by the parameter Ratio. The default value of Ratio is 1 or 2. If $parent_1$ and $parent_2$ are the parents, and $parent_1$ has the better fitness value, the function returns the child is:

$$child = parent_2 + R \times (parent_1 - parent_2) \quad (4)$$

3. Mutation: The mutation used is *adapt feasible Mutation* which randomly generates directions that are adaptive with respect to the last successful or unsuccessful generation. The feasible region is bounded by the constraints and inequality constraints. A step length is chosen along each direction so that linear constraints and bounds are satisfied.

IV. SIMULATION RESULTS

In order to test the effectiveness of the proposed method, we apply the optimized neural network for the air pollution time series. Monthly Ozone and dioxide carbon concentration over a period of 10 years ranging from 1965 to 1996 were available from Hipel and Mcleod [14]. The objective is to optimize the neural network structure, which could predict the ozone and dioxide concentration over the next time step of one month based on previous values of time series, as explained in the following sections. The data have been divided into two sets: the first contains a data of 25 years for training the network when the data of the last 7 years were reserved for testing the resulting predictor.

A. Ozone Forecast

In this section, the optimization process of the neural network is developed for forecasting the monthly ozone concentration. The data used is partitioned in two parts, in which 365 data are used for the training process and 36 are used for the test phase. The inputs of the neural predictor are the past values of ozone concentration and the output is the ozone concentration corresponding to the next month.

Table 2 shows the genetic algorithms parameters used to build the neural predictor. In table 3, we present the parameters values of the resulting neuronal structure generated by the genetic algorithms optimization. Fig. 2 shows the mean square error (MSE) of the fitness function evolution during the genetic optimization. In Fig. 3 and Fig. 5, we can see how the neural network predicts rather well the monthly ozone concentration time series for the training and test phases. The optimized neural predictor can predict effectively this concentration. Figures 4 and 6, present the prediction errors for the two phases; training and test. The mean square error of the neural network training is depicted in Fig. 7. As seems, the error decreases during the time.

TABLE. II
THE CONTROL PARAMETERS OF THE GA

N° of generation	50
Size of population	50
Size of chromosome	17
Probability of crossover	1
Probability of Mutation	0.1

TABLE. III
OPTIMUM STRUCTURE OF NP (OZONE FORECAST)

N° of hidden layers	1
learning coefficient	0.34112
Momentum coefficient	0.53289
the presence/ absence of t he bias	Presence (1)
Number of the input	20
Lag time	1
Number of neurons in the hidden layer 1	8
Number of neurons in the hidden layer 2	1
the activation function in the hidden layer1	Pure linear 3
the activation function in the hidden layer2	hyperbolic 2
the activation function in the output layer	Pure linear 3
int_{out}	[0.16, 0.4559]
int_1	[0.1404, 0.771]
Int_2	[0.1572, 0.7688]

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

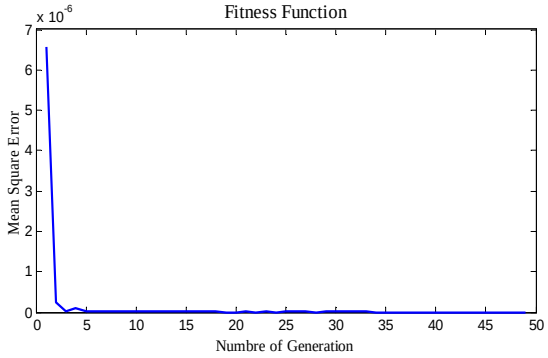

Fig. 2. Evolution of fitness function (ozone forecast).

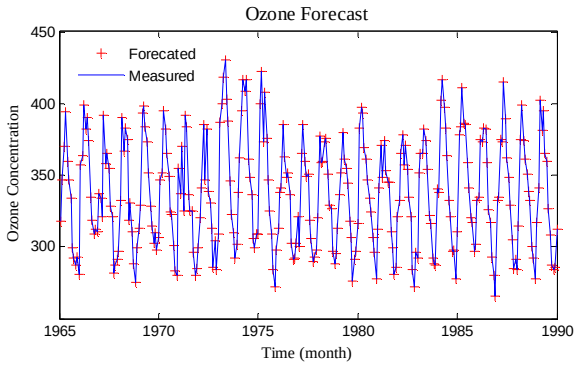

Fig. 3. Forecasted and measured ozone concentration (training stage)

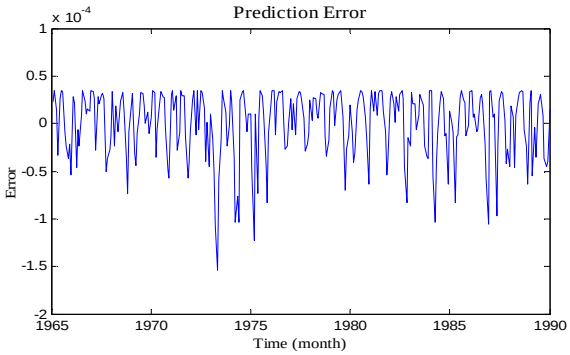

Fig. 4. Ozone prediction error (training stage)

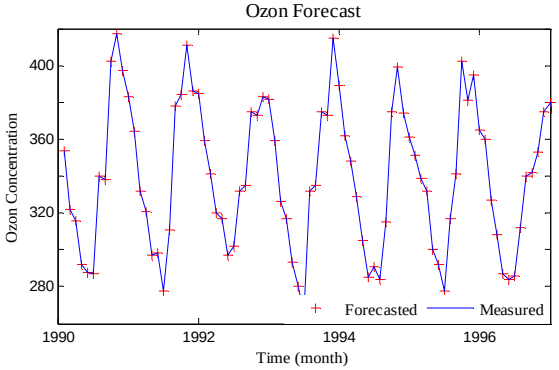

Fig. 5. Forecasted and measured ozone concentration (test stage)

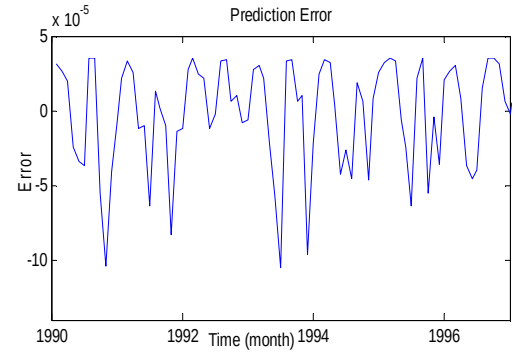

Fig. 6. Ozone prediction error (test training stage)

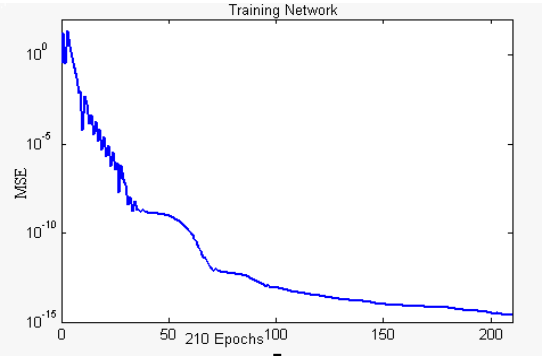

Fig. 7. MSE for the training neural network algorithms (Ozone Forecast)

B. Dioxide Forecasts

In this section, we present the Monthly dioxide prediction. The data used is partitioned in two parts: 300 data are used for the training and 84 for the test phase. The parameters of genetic algorithms are the same used in the previous section for ozone forecast. The optimal parameters for obtaining the minimum prediction error after using the genetic algorithms optimization are shown in table 4.

TABLE. IV
OPTIMUM STRUCTURE OF NP (DIOXIDE FORECAST)

N° of hidden layers	2
learning coefficient	0.8594
Momentum coefficient	0.3911
the presence/ absence of t he bias	Presence (1)
Number of the input	24
Lag time	4
Number of neurons in the hidden layer 1	10
Number of neurons in the hidden layer 2	1
the activation function in the hidden layer1	Pure linear 3
the activation function in the hidden layer2	Pure linear 3
the activation function in the output layer	Pure linear 3
int_{out}	[0.2781, 0.4554]
int_1	[0.5467, 0.98]
Int_2	[0.4264, 0.82]

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

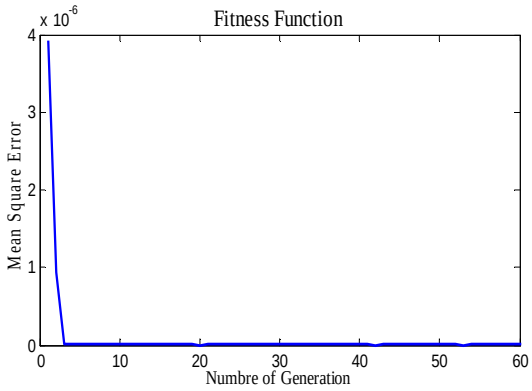

Fig. 8. Evolution of the fitness function (dioxide forecast)

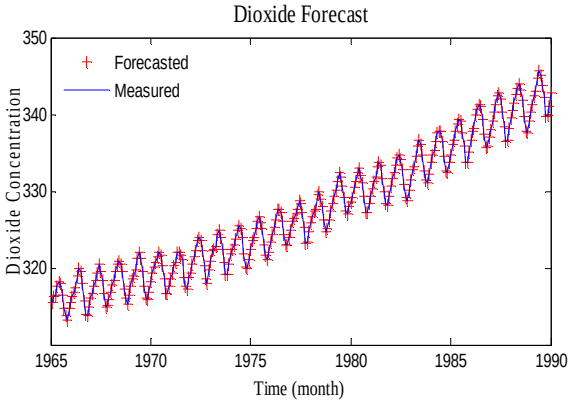

Fig. 9. Forecasted and measured dioxide concentration (training stage)

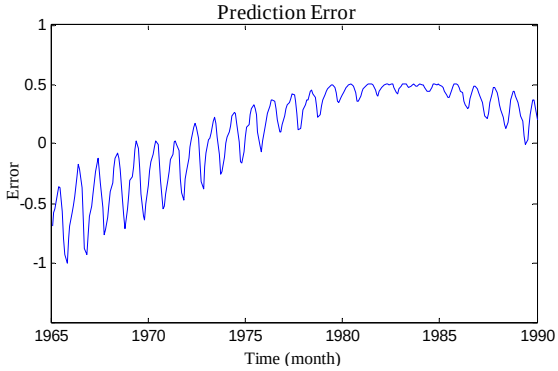

Fig. 10. Prediction error (training stage)

Fig. 8 shows the evolution of the fitness function during the optimization phase using the genetic algorithm. In Fig. 9 and Fig. 11, we can see also as monthly dioxide concentration forecast how the neural network can predict sufficiently the dioxide concentration time series for the training and test phases. It is able to predict the dioxide concentration with a minimum error. Figures 10 and 12 show the prediction error for the training and test phases. In Fig. 13, the MSE for the Gradient descent algorithm is decreases during the time indicating the goodness of the training process.

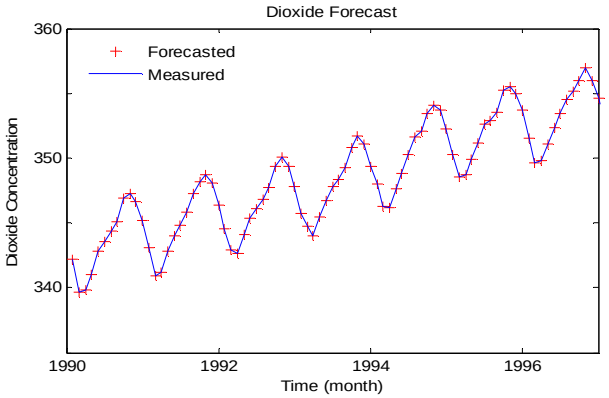

Fig. 11. Forecasted and measured dioxide concentration (test stage)

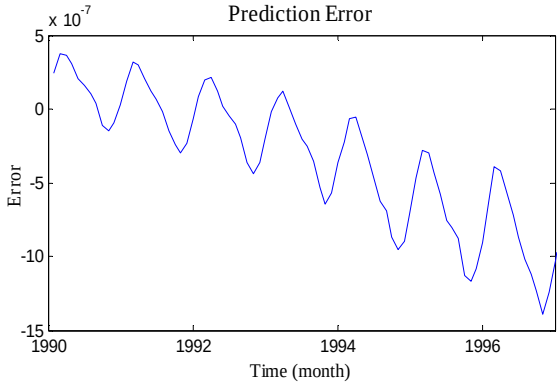

Fig. 12. Prediction error (test phase)

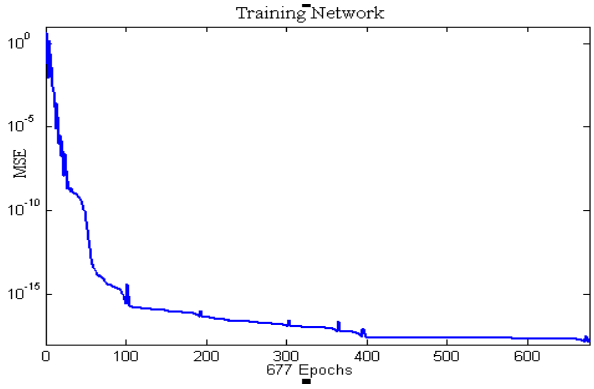

Fig. 13. MSE for the training neural network (dioxide forecast)

C. Comparison study

In order to compare the performances of the proposed method with the existing methods for forecasting the air pollution, we use the following criteria's as parameters of comparison: the mean square error (MSE), the mean bias error (MBE) and the mean absolute (MAE); where:

$$MSE = \frac{\sum_i (y_i - a_i)^2}{\sum_i (y_i - \bar{a}_i)^2} \quad (5)$$

$$MBE = \frac{1}{N}\sum_{i=1}^{N} (Forecasted_i - Measured_i) \quad (6)$$

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|Forecasted_i - Measured_i\right| \quad (7)$$

Table 5 shows the comparison results of the prediction performances between our method and other existing methods. The data in last rows in table 5 are taken from [8][9][10]. As depicted, the proposed optimization algorithm presents good results for air pollution prediction. These parameters of error prediction (MSE, MBE, MAE) are smaller using our method than the others works. We argue that our proposed design technique can produce the smallest amount of performance as we augment the number of the training data and choose a best initial population.

TABLE. V
COMPARISON RESULTS WITH OTHER METHODS

Method	MSE	MAE	MBE
Proposed method ozone forecast (Training phase)	$6.6772.10^{-15}$	$6.566.10^{-8}$	$-3.177.10^{-10}$
Proposed method ozone forecast (test phase)	$6.9057.10^{-15}$	$6.816.10^{-9}$	$-4.30.10^{-9}$
Proposed method dioxide forecast(Training phase)	$9.3307.10^{-18}$	$2.66.10^{-9}$	$6.59.10^{-10}$
Proposed method dioxide forecast (test phase)	$2.0916.10^{-17}$	$3.56.10^{-9}$	$-3.003.10^{-9}$
Brunelli & Piazza [8]	/	0.014424	/
Yu & Wenfang [9] BPNN (test stage)	5.52	4.2	- 0.01
Yu & Wenfang [9] SVM (test stage)	18.01	13.38	0.09
Biljana & Mile [10]	/	0.001	/

V. CONCLUSION

In this paper we have shown how Genetic Algorithm can be used in a successful way of optimizing the neural network in order to deal with the problem of prediction of air pollution; predict the ozone and dioxide concentration. The objective of the optimization task is to find the optimum structure and parameters of the neural predictor that can provide the best prediction of the ozone and dioxide concentration time series for the next month. The obtained results show the efficiency of the proposed method and the capacity of the designed predictor to achieve the desired task with a minimum error when it is compared with other models. The combination of the two techniques (neural and genetic) constitutes an effective methodology to exploit the power and the flexibility of each approach. The resulting neural network is an intelligent system automating by an optimization phase for the prediction task.

REFERENCES

[1] A.C. Comrie, "Comparing Neural Networks and Regression Models for ozone forecasting," *Journal of Air & Waste Manage. Assoc.*, 47, 653-663 (1997).

[2] G. Corani, "Air quality prediction in Milan: feed-forward neural networks, pruned neural networks and lazy learning," *Ecological Modelling*, 185, 513-529 (2005)

[3] O. Castilo and P. Melin "Hybrid Intelligent Systems for Time Series Prediction Using Neural Networks, Fuzzy Logic, and Fractal Theory," *IEEE trans. Neural Networks.* Vol .13, pp. 1395-1408, November 2002.

[4] M. Han and J. Xi and S. Xu and F.L. Yin "Prediction of chaotic time series based on the recurrent predictor neuronal network," *IEEE Trans. Signal Processing*, Vol. 52, pp. 3409-3416, 2004.

[5] M. Anuarg and M.C. Deo, "forecasting wind with neural networks", Marine Structures 16, pp. 35-49, 2003.

[6] Y. Radhika and M. Shashi, "Atmospheric Temperature Prediction using Support Vector Machines'', *International Journal of Computer Theory and Engineering*, Vol. 1, No. 1,pp 55-58, April 2009.

[7] Z. Jure; Z. Rahela and V. Daniel, ''Prediction of Ozone Concentrations'', Preprint submitted to Elsevier Science, 2005.

[8] U. Brunelli, V. Piazza and L. Pignato, '' Hourly Forecasting of so_2 Pollutant Concentration Using An Elman Neural Network", *Springer*-Verlag Berlin Heidelberg, LNCS 3931, pp,65-69,2006.

[9] F. YU, Z. Wenfang and S. Dezhi, "Ozone concentration forecast method based on genetic algorithm optimized back propagation neural networks and support machine data classification", Atmospheric Environment 45, *Elsevier*, pp, 1979-1985, 2011.

[10] M.B. Biljana and S. Mile, "Prediction of Missing Data of ozone concentration Using Support Vector machines and Radial Basis Neural Networks," *Informatica31*, pp. 425-430, 2007.

[11] P. Bonissone, T. Chen, K. Goebel, " Hybrid Soft Computing Systems: Industrial and Commercial Applications", *IEEE Proceedings, Special Issue on Computational Intelligence*, Vol. 87, No. 9, pp. 1641-1667, 1999.

[12] X-B Yan and al, "Time Series Forecasting with RBF Neural Network", *International conference on Machine Learning and cybernetics*, pp. 4680-4683, August 2005.

[13] L.W. Lee and L.H. Wang, ''Temperature Prediction and TAIFEX Forecasting Based on Fuzzy Logical Relationships and Genetic Algorithms'', expert systems with applications, pp. 1-12, *Elsevier* 2006 Ltd.

[14] R. Hyndman, collection of about 800 time series from many different fields. Available: http://www.robjhyndman.com/TSDL/

ISBN: 978-988-19252-1-3
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)