The 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications
21-23 September, 2017, Bucharest, Romania

Models of Particulate Matter Concentration Forecasting Based on Artificial Neural Networks

Mihaela Oprea[1*], Marian Popescu[1], Elia Georgiana Dragomir[2], Sanda Florentina Mihalache[1]
[1]Automatic Control, Computers and Electronics Department
[2]Informatics, Information Technology, Mathematics and Physics Department
Petroleum-Gas University of Ploiesti
Bd. Bucuresti, 39, Ploiesti, Romania
mihaela@upg-ploiesti.ro

Abstract—**The paper presents some models based on artificial neural networks for particulate matter concentration forecasting. A methodology framework is proposed for selecting the best forecasting model from a set of neural networks models. First, two artificial neural network types (feed forward and radial basis) are analyzed for forecasting the particulate matter with diameter less than 10 µm concentration, based on the proposed methodology. Also, a forecasting model for particulate matter with diameter less than 2.5 µm is developed, and then tested on real time data provided by two air quality monitoring microstations built within the ROKIDAIR project. In both cases, statistical indicators are calculated in order to assess the performance of the forecasting models.**

Keywords—forecasting models; particulate matter air pollution; artificial neural networks

I. INTRODUCTION

Real time air pollution accurate forecasting is a difficult problem due to time series data which are usually, incomplete, complex and non-linear. Few air quality monitoring networks have included real time air pollutants forecasting modules [1], the majority of the national ones having online warning systems showing the air quality index (AQI) and some of them providing alerts (as colored codes) related to the exceedances of several air pollutants concentration, as well as some general information regarding the possible human health impact of more severe air pollution episodes [2]. Artificial neural networks proved to be good forecasters (see. e.g. [3]-[9]), better than the classical methods (e.g. linear regression, ARIMA [10]), with a real time response, and therefore, we have chosen such a method for particulate matter (PM) real time forecasting in urban regions. We focus on PM as among the main air pollutants, it has a greater impact on children health, especially the respirable PM fractions: PM_{10} and $PM_{2.5}$, i.e. PM with a diameter smaller than 10 µm and 2.5 µm, respectively.

The research work reported in this paper was performed under the ROKIDAIR project (http://www.rokidair.ro/), which joined under the coordination of Valahia University of Targoviste, the following partners: Norwegian Institute for Air Research (NILU), Petroleum Gas University of Ploiesti, and Politehnica University of Bucharest. Under this research project, it was developed ROKIDAIR DSS, an intelligent decision support system for $PM_{2.5}$ air pollution monitoring, analysis, forecasting and early warning in two pilot cities, Ploiesti and Targoviste, with the main goal of protecting children health during severe $PM_{2.5}$ air pollution episodes that can occur in the two cities [11]. The ROKIDAIR $PM_{2.5}$ continuous monitoring network is composed of microstations that measure apart from the $PM_{2.5}$ concentration, several meteorological parameters such as air temperature, relative humidity and atmospheric pressure [12]. The microstations were developed by Politehnica University of Bucharest.

The purpose of the research described in this paper was to identify an appropriate real time $PM_{2.5}$ forecasting model for ROKIDAIR DSS.

The paper proposes a methodology framework for selecting the best forecasting model from a set of neural networks models. First, two ANN types were used: feed forward and radial basis, and a comparative study between them was performed for PM_{10} short term forecasting in the Ploiesti city. Next, a forecasting model for $PM_{2.5}$ concentration is built using data from Ploiesti city and then tested on real time data from the above mentioned microstations.

II. THE ANN FORECASTING MODEL DEVELOPMENT METHODOLOGY

Artificial neural networks are nonlinear functions universal approximators, inspired from biology that can solve forecasting problems in various domains [13]. An ANN is composed of interconnected nonlinear processing units, named artificial neurons, which are organized under a certain topology with several layers between the input and the output layer. The links between artificial neurons have associated weights, which quantify the degree of the connection. Each artificial neuron has an activation function (e.g. linear, sigmoid). Several types of ANNs, recurrent and non-recurrent were developed so far, multilayer perceptron, Hopfield, Kohonen etc. According to the research work reported in the literature (see e.g. some recent reviews in [1], [5] [14]), the most suitable ANN models for air pollution forecasting based on time series are feed forward ANN (FF-ANN) and radial basis function ANNs (RBF-ANN). A feed forward ANN has an input layer, an output layer and zero, one or more hidden layers. The neurons outputs of each layer are connected with the inputs of the neurons from the next layer, except the outputs of the neurons from the output layer. FF-ANN training is usually performed with a backpropagation algorithm. Fig. 1 shows the structure of a feed forward ANN.

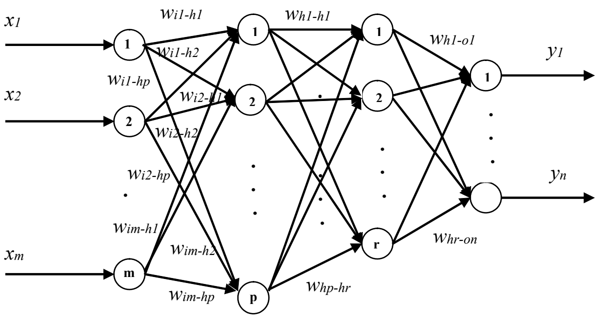

Figure 1. The structure of a feed forward artificial neural network (m×p×r×n).

The radial basis function ANN is similar with FF-ANN, only that every artificial neuron from the hidden layers has a “radial basis function” (e.g. a Gaussian function) with two parameters: “center” and “width”. An RBF-ANN has a number of clusters and a clustering seed value which is the starting value for the “center” of the clusters.

The design of an ANN (FF-ANN or RBF-ANN) takes into account three fundamental aspects: setting the ANN architecture (i.e. number of layers, number of neurons in each layer, number of clusters), choosing the right activation function for the neurons of each ANN layer, and determining the weights of the ANN connections with a learning algorithm, during ANN training.

The architecture of the best PM artificial neural network based forecasting model is chosen by experiment according to the methodology framework proposed in this section. The proper model is selected from a set of models built in two steps: (1) it is determined the most appropriate time window for the ANN forecasting model based solely on PM past measurements and, (2) it is found the most influent atmospheric parameter on the PM concentration, in the analyzed area, selected from the site specific air pollutants and meteorological parameters, which is added to the time window found in the first step.

The database with time series of hourly air pollutants concentrations and meteorological parameters measurements is provided as the methodology input, while the output is the next hour PM concentration value. Under the methodology framework, several ANN architectures are tested in order to choose the best one. Some steps of the methodology are detailed with information related to the use of a data mining software tool in the implementation phase.

A. *Methodology – PM short term forecasting ANN model development*

Inputs: PM and other atmospheric parameters measurements - hourly time series data bases

Output: next hour PM (predicted value) and the best ANN forecasting model

Steps:

1. *Data preprocessing*. In this step, the data are processed in order to eliminate irrelevant information or noisy data, to fill the incomplete records, to normalize or generalize the validated values. The database can be rearranged to be recognizable by the data mining software tool.
2. *Selection of relevant parameters*. The atmospheric parameters relevant to PM concentration value prediction can be empirically selected using the chemical processes experts’ knowledge. They can be determined also by experiments using the model’s statistical parameters values or some specific data mining techniques, such as Principal Components Analysis (PCA), which identifies the relevant atmospheric parameters based on the inputs and the predicted variables correlations.
3. *Applying the overfitting avoiding method*. A common solution is to divide the database into three sets: the training set, the validation set and the testing set. Another option is to use cross-validation with 10 folds.
4. *Setting the ANN architecture*. This step involves setting the number of nodes in the input layer, *no_input_variables*, (i.e. setting the optimal time window for the next hour PM prediction), the number of nodes in the hidden layer, *no_nodes_hidden_layer* (usually, one hidden layer is enough for a forecasting ANN model), the suitable activation function for the neurons of each layer etc.
5. *Training parameters adjustment*. The parameters of the learning algorithm (i.e. backpropagation) are experimentally determined for each type of ANN forecasting model: the number of training epochs (*Tep*), the learning rate (*LR*), the momentum (*m*) for the FF-ANN model, and the number of clusters, the clustering seed value (starting value for the "center" of the clusters), for the RBF-ANN model, in order to avoid an overtrained or undertrained ANN.
6. *ANN training* with the parameters set in steps 4 and 5.
7. *Validation of the obtained ANN architecture*.
8. *Testing the ANN forecasting model*.
9. *ANN forecasting model performance analysis*. At this stage, the statistical parameters that can be used are the correlation coefficient between variables (R^2), the mean absolute error (MAE), the root mean square error (RMSE), the training error (Terr) etc. The parameters can be compared with the limits set in the literature by human experts or with those obtained by other models designed to predict atmospheric parameters value.
10. *Selection of the best ANN forecasting model*.

The best ANN forecasting model is chosen from the various forecasting models designed, trained, validated, tested and analyzed during steps 4 to 10 of the proposed methodology. The forecasting performance of a model is measured using a set of statistical parameters.

The application of the methodology is realized as a two steps algorithm following the two cases of a forecasting problem formulation [15] which are given below.

The forecasting problem can be solved either as a time series problem for one parameter (case I), i.e. only the PM concentrations time series, or as a time series problem for more parameters (case II) including PM concentration, other PM related air pollutants concentrations and meteorological parameters. Suppose that x represents the PM concentration that needs to be forecasted in a $t+k$ time window. Problem formulation for case I is given by (1), while for case II is given by (2).

$$x(t+k) = f1(x(t), x(t-1), x(t-2), \ldots, x(t-r-1), x(t-r)) \quad (1)$$

$$x(t+k) = f2(p_1, p_2, \ldots, p_m; x(t), x(t-1), x(t-2), \ldots, x(t-r-1), x(t-r)) \quad (2)$$

where: t is time, $p_1, p_2, \ldots, p_m$ – are parameters that influence the evolution of x, t-r is time until are considered the previous values, and $f1$ and $f2$ are the specific forecasting functions.

As the short term (e.g. next hours) evolution of PM concentration can be influenced, apart from the current and past values of PM concentrations, by other atmospheric factors such as meteorological [16] (e.g. air temperature, wind speed, relative humidity) and other air pollutants (e.g. CO, SO_2, NO_x), we have extended the application of the methodology framework with a second step in which the most influent atmospheric parameter is detected and added in the time window of the ANN forecasting model. Thus, the two forecasting cases formalized by (1) and (2) are integrated in the methodology framework, providing a larger set of PM forecasting ANN models from which the model with the best forecasting performance is chosen.

B. Methodology framework

This methodology provides the PM ANN forecasting model with the best forecasting performance)

- *step I.* Apply the methodology for case I (given by (1)) – finding the optimal time window setting for the ANN forecasting model dependent only on PM time series; provides the best ANN forecasting model – I;
- *step II.* Apply the methodology for case II (given by (2)) – simplified case (one additional atmospheric parameter): finding the most influent atmospheric parameter (from other air pollutants and meteorological parameters) related to PM and extend the time window with this one; provides the best ANN forecasting model – II;
- **return** the best PM ANN forecasting model from the two models found in steps I and II.

Based on the described methodology we have developed several PM forecasting models from which the best ones were chosen to be used in the ROKIDAIR DSS system.

III. PM_{10} FORECASTING MODELS

We have applied the proposed methodology framework to the FF-ANN and RBF-ANN forecasting models, for PM_{10} next hour forecasting in the Ploiești city at PH-2 monitoring station from the Romania National Air Quality Monitoring Network (RNMCA). The implementation of the forecasting models was done with the WEKA Data Mining software tool (http://www.cs.waikato.ac.nz/ml/weka/). In the following, details related to the data set and experiments that were performed are given.

A. The data set

The data set used in the experiments contains time series of hourly recorded values for the major atmospheric parameters from January 2009 to December 2009 and partially from 2011, taken from the National Air Quality Monitoring Network web site (http://www.calitateaer.ro). The data were gathered at PH-2 monitoring station, located in the center of the Ploiesti city. Apart from the air pollutants concentrations that are monitored at PH-2 station, the following meteorological parameters are measured: wind direction (degree), wind speed (m/s), temperature (°C), relative humidity (%), atmospheric pressure (mbar), solar radiation (W/m^2), precipitations (mm). The database had 7641 records, before preprocessing.

B. The FF-ANN PM_{10} forecasting model

In this experiment we have followed the proposed methodology in order to build an FF-ANN forecasting model to predict the next hour PM_{10} concentration at the PH-2 monitoring station, based on the previously prepared data set. The proper value for the PM_{10} time window and the most influent atmospheric parameter related to PM_{10} next hour forecast were determined.

In the first stage different FF-ANNs were built in order to determine the optimal architecture for next hour PM_{10} concentration forecasting. Using a time window up to 10 hours ago, we have designed several FF-ANNs architectures in order to find the optimal number of input layer nodes (a number of 10 nodes in the input layer reflects that for the forecast are used the past 10 hours recorded values of PM_{10}). Thirteen FF-ANNs with different number of nodes in the input layer and in the hidden layers were built, trained, cross validated and tested using the same set of training data and the WEKA software tool. Some details on the relevant ANN models built during step I of the methodology framework are given in Table I. The parameters that were analyzed are: *Ph* – number of past hours, *Hn* - number of nodes in the hidden layer, *LR* - learning rate, m - momentum, *Tep* - training epochs, *Terr* - training error, *R* - correlation coefficient, *RMSE* – root mean squared error, *TBM* - time taken to build the ANN forecasting model.

TABLE I. THE FF-ANN EXPERIMENTAL RESULTS – STEP I (*TIME WINDOW* – PAST HOURS)

Model	Ph	Hn	Terr	R	RMSE	TBM[s]
FFANN1	10	6	0.0162	0.8515	13.2817	0.47
FFANN2	9	6	0.0175	0.8518	13.3623	0.43
FFANN3	8	6	0.0177	0.8661	12.5851	0.41
FFANN4	7	6	0.0218	0.8686	12.4465	0.37
FFANN5	6	6	0.0215	0.882	11.8267	0.36
FFANN6	5	6	0.0257	0.8857	11.5666	0.34
FFANN7	5	3	0.0300	0.8976	10.9507	0.19
FFANN8	5	4	0.0244	0.8939	11.1448	0.23
FFANN9	5	2	0.0352	0.8956	11.0482	0.16
FFANN10	4	3	0.0298	0.8795	11.8915	0.17
FFANN11	4	2	0.0506	0.8901	11.3408	0.14
FFANN12	3	2	0.0411	0.848	13.3842	1.82
FFANN13	2	1	0.0416	0.8867	11.4946	0.09

The highest correlation of the FF-ANNs architectures was provided by FFANN7 (R=0.8976).

Different tests were performed in order to adjust the FF-ANN training parameters for each of the thirteen FF-ANNs. Fig. 2 shows the best FF-ANN architecture identified during step I of the proposed methodology.

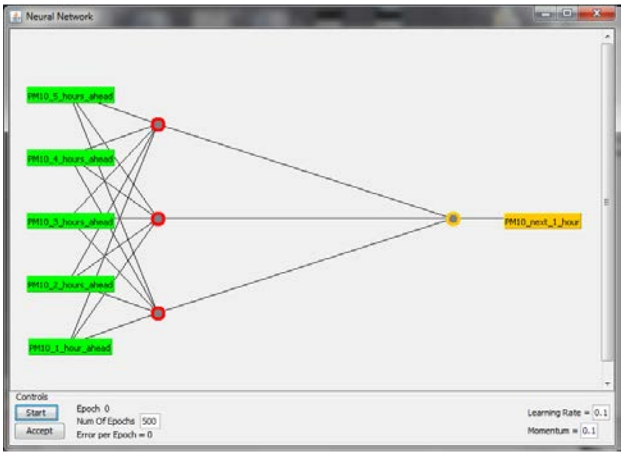

Figure 2. The best FF-ANN architecture – step I (*time window* = 5).

In the second step, in addition to the five input nodes, another input node was added, representing the last 5 hours mean of various atmospheric parameters. The purpose of this step was to identify the atmospheric parameter that is the most influent to next hour PM_{10} concentration forecast along with the past 5 hours recorded values for PM_{10} concentration. Thus, the proper FF-ANN architecture after the second step was set to six input nodes and one output node. After various tests it was determined the proper number of nodes in the hidden layer as well as the parameters of the backpropagation learning algorithm (*LR*, *m* and *Tep*). The sixth node on the input layer represented successively the values recorded for different atmospheric parameters, such as: sulphur dioxide, nitrogen oxides, carbon monoxide, and meteorological parameters (temperature, wind speed, relative humidity etc). The experiments performed on the PH-2 data set revealed that SO_2 was the most influent air pollutant and air temperature, relative humidity and wind speed were the most influent meteorological parameters on PM_{10} concentration. A selection of experimental results is given in Table II.

TABLE II. THE FF-ANN EXPERIMENTAL RESULTS – STEP II (MOST INFLUENT ATMOSPHERIC PARAMETER)

Parameter	Hn	Terr	R	RMSE	TBM [s]
SO_2	5	0.0224	0.7034	11.6683	0.22
SO_2	3	0,0223	0.6054	13.2098	0.17
SO_2	4	0.0220	0.7026	11.6678	0.2
NO_x	5	0.0254	0.6698	12.2275	0.19
CO	5	0.0254	0.6698	12.2275	0,19
Temp	5	0.0250	0.6775	12,0879	0.17
Rel. Humid.	5	0.0255	0.6434	12.6409	0.19
Wind Speed	5	0.0251	0.6708	12.1825	0,19

C. *The RBF-ANN PM_{10} forecasting model*

The second experiment had the same purpose as the first one performed with the FF-ANN forecasting model, i.e. the next hour PM_{10} concentration value prediction. In this case we have built several RBF-ANN architectures that use radial basis function as activation functions.

The same methodology was followed. Therefore, in the first stage, it was determined the optimal time window for the past values of PM_{10} concentration. The same approach was used: several RBF-ANN architectures with different numbers of input nodes were designed, trained, cross validated and tested on the same data set as in the first experiment (FF-ANN). The number of input layer nodes varied between 2 and 10. The parameters of the RBF-ANN forecasting model are the number of clusters and the clustering seed value. For our data set, these parameters were experimentally set. The experimental results obtained for the relevant RBF-ANN architectures built during the step I are given in Table III.

TABLE III. THE RBF-ANN EXPERIMENTS RESULTS – STEP I (*TIME WINDOW* – PAST HOURS)

Model	Ph	R	RMSE	TBM [s]
RBFANN1	10	0.7574	16.5786	0.02
RBFANN2	9	0.7589	16.1592	0.02
RBFANN3	8	0.7599	16.1307	0.03
RBFANN4	7	0.7592	16.1593	0.04
RBFANN5	6	0.7618	16.0745	0.03
RBFANN6	5	0.7623	16.0605	0.08
RBFANN7	4	0.7801	15.5257	0.02
RBFANN8	3	0.7903	15.2048	0.06
RBFANN9	2	0.8018	14.8322	0.02

Comparing the values obtained for the statistical parameters, the best RBF-ANN architecture was RBFANN9, which predicts the next hour PM_{10} concentration using the time window of 2 past hours (i.e. the values of PM_{10} concentration measured one hour ago and two hours ago), with the correlation coefficient of 0.8018, and RMSE (14.8322) and the time taken to build the model of 0.02 seconds.

We have performed similar experiments during step II with those for the FF-ANN PM_{10} forecasting model that revealed the same atmospheric parameters influence on PM_{10}.

D. *Comparative analysis*

We were interested to compare the experimental results obtained by the two models, feed forward ANN and radial basis function ANN. Table IV synthesized the experimental results obtained by the ANN architectures with the best performance: FFANN7, FFANN7+ (FFANN7 + SO_2), RBFANN9, RBFANN9+ (RBFANN9 + NOx / CO). The comparative analysis of these forecasting models highlights that the FF-ANN forecasting model performed better than the RBF-ANN forecasting model. The best value for the correlation coefficient (0.8976) was obtained by the FFFANN7 model, a feed forward network that uses in the forecast process only the past values of PM_{10} concentration. The same model, FFANN7 obtained the minimum value of RMSE (10.9507).

The main conclusion of the comparative analysis is that the best PM_{10} next hour forecasting results are obtained by the FFANN7 model, which uses only the past 5 hours PM_{10} concentration values.

TABLE IV. THE COMPARATIVE PERFORMANCE ANALYSIS

Model	PM_{10} time window	Air parameter added	R	RMSE	TBM [s]
FFANN7	past 5 hours	-	0.8976	10.9507	0.19
FFANN7+	past 5 hours	SO_2	0.7034	11.6683	0.22
RBFANN9	past 2 hours	-	0.8018	14.8322	0.02
RBFANN9+	past 2 hours	NO_x / CO	0.6292	12.684	0.02

At the end of the comparative analysis discussion we want to emphasize that our methodological framework proposes the selection of the relevant parameters for short term PM forecasting with a principal components analysis or with a two steps algorithm for obtaining the best ANN forecasting model (step I – best time window determination; step II – extending the time window with the most influent PM related atmospheric parameter) which is characterized by optimal statistical parameters (as shown by the results of the two experiments, described in the previous section and the comparative analysis between the best PM_{10} ANN forecasting models). This two steps algorithm represents the main contribution of our methodology. These types of solutions can be integrated in the proposed methodology (for step II – the relevant parameters selection).

IV. THE $PM_{2.5}$ FORECASTING MODEL

The most influent parameters on $PM_{2.5}$ were identified using SAS Enterprise Miner 7.1 Workstation 13.2 software package. Thus, the best correlations with $PM_{2.5}$ concentration were given by relative humidity, temperature and atmospheric pressure.

The data set used in the following experiment is from an air quality monitoring station from Ploiesti city, Romania, and contains around 6000 samples of hourly data for each of the parameters: $PM_{2.5}$ concentration, temperature, relative humidity, and atmospheric pressure. The measured ranges for these parameters are:

- $PM_{2.5}$: [0.07…51.40] µg/m3;
- Temperature: [-13.23…37.24] ^{0}C;
- Relative humidity: [11.17…100] %;
- Atmospheric pressure: [733.66…766.57] mmHg.

A representation of the evolution of three of the mentioned parameters time series is shown in Fig. 3-5.

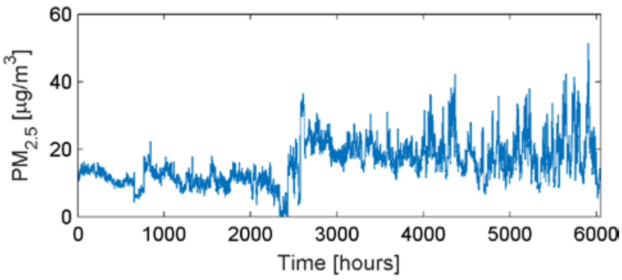

Figure 3. $PM_{2.5}$ concentration time series.

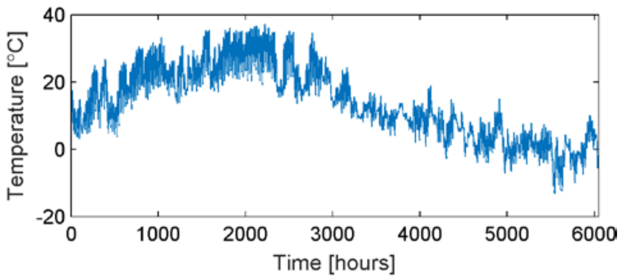

Figure 4. Temperature time series.

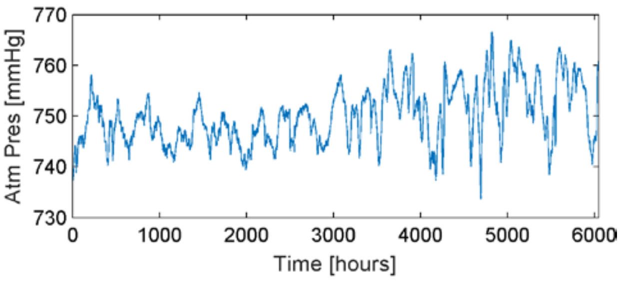

Figure 5. Atmospheric pressure time series.

The forecasting model is built using MATLAB®, the structure of the model being represented in Fig. 6.

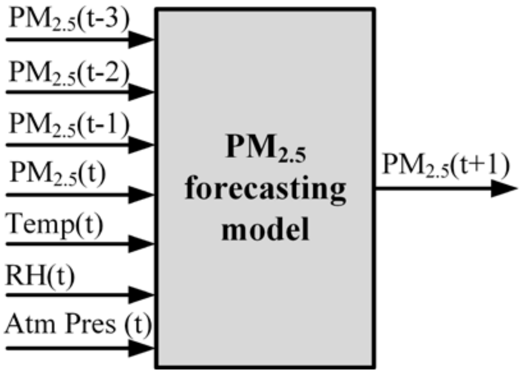

Figure 6. $PM_{2.5}$ forecasting model.

As it can be seen in Fig. 6, the forecasting model has seven inputs, namely: $PM_{2.5}$ concentration from current hour to three hours ago, current hour values for temperature, relative humidity and atmospheric pressure. The output of the model is the prediction of the next hour $PM_{2.5}$ concentration.

The neural network based model has the following structure: an input layer with seven neurons, one hidden layer, and an output layer with one neuron. The type of neural network used in this study is feed-forward back propagation, the training algorithm is Levenberg-Marquardt, and the adaptive learning function is the gradient descent with momentum weight and bias. The structure of the model was modified by changing the number of neurons in the hidden layer (from five to twelve).

The performance of the model is evaluated by calculating statistical indicators, such as: root mean square error (RMSE), index of agreement (IA), and correlation coefficient (R).

The results for this model are presented in Table V.

TABLE V. STATISTICAL INDICATORS FOR $PM_{2.5}$ FORECASTING MODEL

ANN structure	RMSE [μg/m^3]	IA	R
7 x 5 x 1	1.6351	0.9831	0.9668
7 x 6 x 1	1.6334	0.9831	0.9668
7 x 7 x 1	1.6028	0.9836	0.9680
7 x 8 x 1	1.6334	0.9830	0.9668
7 x 9 x 1	1.6070	0.9836	0.9679
7 x 10 x 1	1.5990	0.9837	0.9682
7 x 11 x 1	1.5895	0.9839	0.9686
7 x 12 x 1	1.6264	0.9832	0.9671

The best results are obtained for the structure with 11 neurons in the hidden layer, as this structure presents the smallest RMSE and the index of agreement and correlation coefficient closest to 1.

For this structure, a representation of a partial view of the comparison between testing data and the predicted values for the $PM_{2.5}$ concentration is shown in Fig. 7.

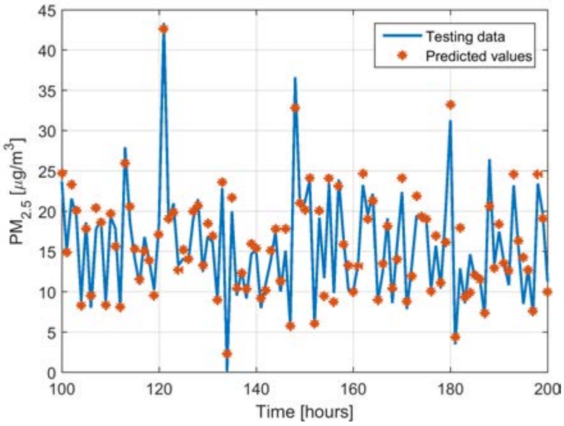

Figure 7. Partial view of the comparison between testing data and predicted values of $PM_{2.5}$.

Taking into account the results from Table V and Fig. 7 it can be considered that the proposed model is validated and can be used further in testing data from other air quality stations of the same type.

Thus, the previously validated forecasting model is applied on real time data provided by two air quality monitoring microstations in Ploiesti city, Romania. The microstations are developed within the ROKIDAIR research project and measure $PM_{2.5}$ concentration, temperature, relative humidity and atmospheric pressure.

At each hour, a computer program developed in MATLAB® automatically fetch data from the microstations databases. These data are necessary for testing the forecasting model and refer to the current hour, one hour ago, two hours ago, and three hours ago $PM_{2.5}$ concentrations, current hour temperature, relative humidity and atmospheric pressure. The same program applies the forecasting model on these data, thus generating the prediction for the next hour $PM_{2.5}$ concentration.

The results obtained by testing the forecasting model on data from the two microstations are presented in Fig. 8-9.

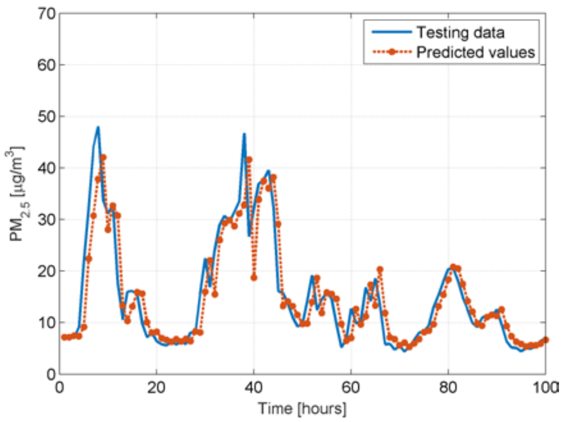

Figure 8. Comparison between real data from Microstation-1 and predicted data.

For Microstation-1 the statistical indicators were: RMSE=4.8873 μg/m^3, IA=0.9388, and R=0.8878.

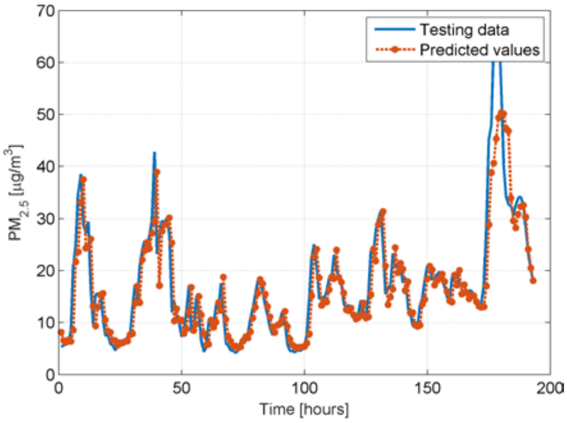

Figure 9. Comparison between real data from Microstation-2 and predicted data.

For Microstation-2 the statistical indicators were: RMSE=4.9870 μg/m^3, IA=0.9400, and R=0.8954.

The obtained results lead to the idea that the proposed $PM_{2.5}$ forecasting model performs well when tested with real time data, other than those used to build the model.

V. Conclusion

The paper presented some artificial neural network based models for real time PM_{10} and $PM_{2.5}$ concentration forecasting. A methodology framework is proposed for particulate matter short term forecasting based on ANN having as purpose the selection of the best forecasting model from a set of neural networks models. Two PM_{10} forecasting model based on two types of ANNs (feed forward and radial basis) are analyzed. Also, a $PM_{2.5}$ forecasting model was built and tested on real time data from two monitoring stations from Ploiesti city developed within the ROKIDAIR project. The most appropriate $PM_{2.5}$ forecasting model was integrated in the ROKIDAIR DSS and is currently performing real time forecasting in two Romanian cities, Ploiesti and Targoviste under the ROKIDAIR $PM_{2.5}$ continuous monitoring network with microstations developed at the Politehnica University of Bucharest.

Acknowledgment

The research leading to these results has received funding from EEA Financial Mechanism 2009-2014 under the project ROKIDAIR "*Towards a better protection of children against air pollution threats in the urban areas of Romania*" contract no. 20SEE/30.06.2014. We thank to Prof. S. Iordache and Assoc. Prof. D. Dunea from Valahia University of Targoviste and T. Bohler from NILU, research project promoters.

References

[1] Y. Zhang, M. Bocquet, V. Maller, C. Seigneur, and A. Baklanov, "Real-Time air quality forecasting, part I: History, techniques, and current status", *Atmospheric Environment*, vol. 60, pp. 632-655, 2012.

[2] P. Hajek, and V. Olej, "Predicting Common Air Quality Index – The Case of Czech Microregions", *Aerosol and Air Quality Research*, vol. 15, pp. 544-555, 2015.

[3] H.J.S. Fernando, et al., "Forecasting PM_{10} in metropolitan areas: Efficacy of neural networks", *Environmental Pollution*, vol. 163, pp. 62-67, 2012.

[4] M. Fu, W. Wang, Z. Le, and M.S. Khorram, "Prediction of particulate matter concentrations by developed feed-forward neural network with rolling mechanism and gray model", *Neural Comput & Applic*, vol. 26, Is. 8, pp. 1789-1797, 2015.

[5] M. Oprea, E.G. Dragomir, S.F. Mihalache, and M. Popescu, "Prediction methods and techniques for $PM_{2.5}$ concentration in urban environment", In *Methods for the assessment of air pollution with particulate matter to children's health*, S. Iordache and D. Dunea (Editors), Matrix Rom, Bucharest, 2014 (in Romanian).

[6] M. Oprea, S.F. Mihalache, and M. Popescu, "Computational Intelligence-based $PM_{2.5}$ Air Pollution Forecasting", *International Journal Of Computers, Communications & Control*, vol. 12, Is. 3, pp. 365–380, 2017.

[7] P. Perez, and J. Reyes, "An integrated neural network model for PM_{10} forecasting", *Atmospheric Environment*, vol. 40, pp. 2845-2851, 2006.

[8] N.H.A. Rahman, M.H. Lee, Suhartono, M.T. Latif, "Artificial neural networks and fuzzy time series forecasting: an application to air quality", *Qual & Quant*, vol. 49, Is. 6, pp. 2633-2647, 2015.

[9] B.T. Ong, K. Sugiura, and K. Zettsu, "Dynamically pre-trained deep recurrent neural networks using environmental monitoring data for predicting $PM_{2.5}$", *Neural Computing & Applications*, vol. 27, pp. 1553-1566, 2016.

[10] L.A. Díaz-Robles, et al., "A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile", *Atmospheric Environment*, vol. 42, pp. 8331-8340, 2008.

[11] M. Oprea, et al., "On the development of an intelligent system for particulate matter air pollution monitoring, analysis and forecasting in urban regions", *Proceedings of ICSTCC 2015*, October 14-16 2015, pp. 711-716, 2015.

[12] D. Dunea, T. Savu, L. Marsteen, "Monitoring of size-segregated particulate matter fractions with optical instruments in urban areas", *Scientific Papers. Series E. Land Reclamation, Earth Observation & Surveying, Environmental Engineering*. Vol. V, pp. 99-104, 2016.

[13] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice-Hall, Englewood Cliffs, NJ, 1994.

[14] M.A. Elangasinghe, N. Singhal, K.N. Dirks, J.A. Salmond, and S. Samarasinghe, "Complex time series analysis of PM_{10} and $PM_{2.5}$ for a coastal site using neural network modelling and k-means clustering", *Atmospheric Environment*, vol. 94, pp. 106-116, 2014.

[15] M. Oprea, and E.G. Dragomir, "PM forecasting by using artificial neural networks", *ICEEM08 Book of Abstracts*, Iasi, September 2015.

[16] A.P.K. Tai, et al., "Meteorological modes of variability for fine particulate matter ($PM_{2.5}$) air quality in the United States: implications for $PM_{2.5}$ sensitivity to climate change", *Atmos. Chem. Phys.*, vol. 12, pp. 3131–3145, 2012.