BULETINUL Universităţii Petrol – Gaze din Ploieşti	Vol. LXI No. 3/2009	389 - 394	Seria Tehnică

Artificial Intelligence applied in Air Pollution Forecast

Elia Georgiana Petre

Petroleum – Gas University of Ploieşti, 39 Bucureşti Blvd., Ploieşti, ROMÂNIA
e-mail: elia_petre@yahoo.com

Abstract

This paper presents three different systems used to predict hourly or daily values of some greenhouse gases. These gases emissions are the main factor that influences the climate change that lately has become one of the greatest environment issues. Therefore, the monitoring and control of these dangerous gases, like their prediction, represent important tasks for environmental and health authorities. Having set some comparison aspects, this study has been performed on three artificial intelligence based systems. AirPolTool, developed in Istanbul, Turkey can predict SO_2, CO_2 and PM_{10} for three days ahead. The second system, BISTAPOF was designed in Bilbao, Spain to forecast 8 hours ahead the SO_2, CO, NO_2, NO and O_3 emissions. Airthess is a model built to predict the next day's ozone concentration in Thessaloniki, Greece.

Key words: *artificial intelligence, air pollution, neural networks, genetic algorithms.*

Introduction

One of the most important and interesting area developed in the last years is computer technology and Artificial Intelligence. In the last decade, this domain has known an explosive development and has been used in various scientific fields, like: economy, industry, process control, medicine, human resources, banking system etc.

Among artificial intelligence techniques are:

Artificial neural networks: "Non-linear predictive models that learn through training and resemble biological neural networks in structure."(Actual biological neural networks are incomparably more complex.)

Genetic algorithms: Optimization techniques similar with the genetic process that uses as methods the genetic combination, mutation, and natural selection. It also includes the "chromosome" in which all the necessary information is laid for building a model (the next generation) until the best is found.

Principal component analysis (PCA) is a method that may support the investigation of "hidden" relationships between the variables examined.

Multiple linear regression attempts "to model the relationship between two or more explanatory variables and a response variable by fitting a linear equation to observed data [6]".

This paper presents comparatively some application of the artificial intelligence techniques in the air pollution domain. Lately, there was an increase of reports about our activity negative impact on the ecosystems around us. Scientific researches show that climate change represents

one of the greatest environmental, social and economic threats facing the planet. The warming of the climate system is now evident from observations of increases in global average air and ocean temperatures, widespread melting of snow and ice, and rising global mean sea level [7].

The explosive emissions of carbon dioxide (CO_2) represent the main gas responsible for climate change, as well as of other 'greenhouse' gases. Therefore, all these gases must be intensively monitoring. Scientists have developed different projects to predict these greenhouse gases values in order to reduce them significantly.

Brief overview of Air Pollution Forecast Systems

Next, there are presented three air pollution application designed to predict future values of pollutants. All three models are based on an artificial intelligence technique: artificial neural networks and try to forecast the values of the most important pollution parameters: sulfur dioxide (SO_2), carbon dioxide (CO_2), carbon monoxide (CO), particular matter (PM_{10}), nitrogen dioxide (NO_2), ozone (O_3), nitrogen dioxide (NO_2), nitrogen monoxide (NO).

AirPolTool

This system is a user friendly website which predicts the values of some pollutants for the next three days in Greater Area of Istanbul, Turkey. Based on the artificial neural networks model, it publishes the forecast values for SO_2, CO_2, PM_{10}.

The article ”An online air pollution forecasting system using neural networks”, published by Atakan Kurt, Betul Gulbagci, Ferhat Karaca and Omar Alagha from Department of Computer Engineering, Fatih University specifies that there is enough one neural network to obtain acceptable accuracy for these pollution parameters [3].

BISTAPOF (Bilbao Short-Term Air Pollution Forecast)

It is a software product built in 2005 by Gabriel Ibarra-Berastegi from University of the Basque Country, Bilbao, Spain. Developed in MS Visual Basic 5.0, this software predicts the values of five air pollution parameters (SO_2, CO, NO_2, NO and O_3) at rush hours for six locations in Bilbao, Spain.

Airthess

It is a project supported by the Organization of the Master Plan of Thessaloniki and the Prefecture of Central Macedonia, Greece. It tries to simulate the ozone formation phenomenon based on the “hidden” relationship between ozone (O3), nitrogen dioxide (NO2) and meteorological variables.

Comparative Description

Although, all three systems are using as inputs almost the same air pollution parameters as well as the meteorological variables, there are completely different input values for those parameters: one of it collects data registered around Istanbul, the second one extracts the needed data from a wide spread network sensors in Bilbao and the third gathers the required data around Thessaloniki. The input variables used by these three systems are presented by comparison in table 1.

Table 1. Input air pollution parameters

	O_3	SO_2	NO	NO_2	CO	PM_{10}	NMHC	CH_4
AirPolTool	✓	✓	✓	✓	✓	✓	✓	✓
BISTAPOF	✓	✓	✓	✓	✓	-	-	-
Airthess	✓	-	-	✓	-	-	-	-

Input Data Collection

In order to forecast the value of the air pollution parameters, AirPolTool System collects the pollutant levels from related websites and stored daily in a MySQL database by a PHP script.

The collected data are:

- Meteorological data (humidity, day and night temperature, pressure, wind speed, wind direction, daily condition values), which are available from http://bbc.co.uk;
- Air pollution measurement data (SO_2, PM_{10}, CO, NO, NO_2, non-methane hydrocarbons NMHC, O_3, methane CH_4) which are published on the website http://www.ibb.gov.tr by Istanbul Metropolitan Municipality. These data are retrieved and place in the database.

For the models built by Bilbao Short-Term Air Pollution Forecast are used records regarding to meteorological phenomena (measured at six sensors as is shown in figure 1), traffic (registered at 181 locations throughout the central area of Bilbao) and air pollution from the years 2000 and 2001.

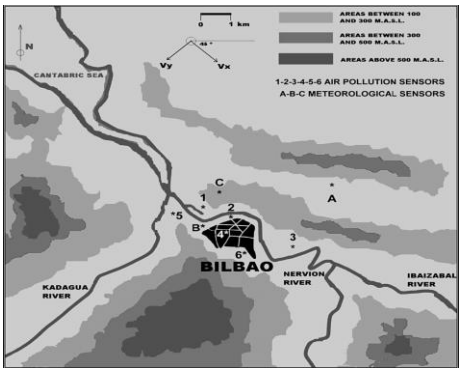

Fig. 1. Sensors distribution in Bilbao [1]**.**

Similar to the other two systems, Airthess analyses the values registered for O_3, NO_2 and meteorological parameters like: air temperature, wind direction, wind speed and humidity for the time period of 2001 and 2004.

Table 2. Input meteorological parameters

	Air temperature	Humidity	Wind direction	Wind speed	Radiation	Atmospheric pressure
AirPolTool	✓	✓	✓	✓	-	✓
BISTAPOF	✓	✓	-	✓	✓	✓
Airthess	✓	✓	✓	✓	-	-

Table 2 was planned to highlight the meteorological parameters used to build the prediction model by each of these systems. Air temperature, humidity and wind speed are considered important in all three built model.

On the other hand, beside these parameters, wind direction and wind speed, as well as atmospheric pressure are included in the neural net's input set of AirPolTool. Along these, BISTAPOF uses in the artificial neural networks' input layer the recorded data for atmospheric pressure and radiation. In the Airthess's model, atmospheric pressure is considered another important factor in prediction as is shown in table 2.

Prediction model

AirPolTool is "a real time system and it has been working since 01.08.2005 [3]". Meteorological air conditions and air pollutant levels are used for building a forecasting model with artificial neuronal network. It uses a feed forward back propagation neural network implemented in MATLAB. The model has seven input nodes; ten nodes in the hidden layer and three nodes in the output layer as is shown in figure 2.

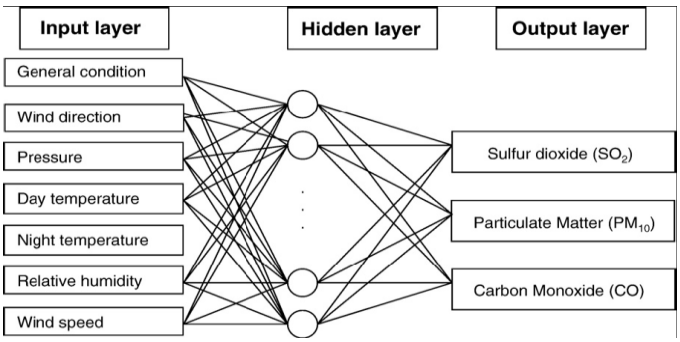

Fig. 2. Artificial neural network used in AirPolTool [3].

AirPolTool is using a single neural network to predict the values of three parameters (SO_2, CO and PM_{10}) for the next 3 days.

On the other hand, BISTAPOF uses a lot more neural nets to forecast the values of SO_2, CO, NO_2, NO and O_3 with only 8 hours ahead. In order to estimate pollutants hourly values, there were built statistical prognostic models (216, as a result, for each location, pollutant, and number of hours ahead), with the different types of neural networks and multiple linear regression. In the training process are used data sets from 2000, while data belonging to the year 2001 are reserved for testing and selecting the best model. To avoid overfitting of the networks, genetic algorithms were applied first to select the best training and validation sets and secondly to identify the most relevant inputs to use in the model.

To sum up, we present the artificial intelligence techniques used in air pollution forecast examples in table 3.

Table 3. Artificial intelligence techniques

	Artificial neural networks	Genetic algorithms	Multi linear regression	Principal Component Analysis
AirPolTool	✓	-	-	-
BISTAPOF	✓	✓	✓	-
Airthess	✓	-	-	✓

According to table 3, Airthess uses a statistical method, named Principal Component Analysis, to reduce the variables number stored in databases. Then, the forecast is made using a neural network with one hidden layer: the network consisted of one input layer, one hidden layer with 10 nodes and the output layer [2]. This model can predict the parameters values 24 h in advance.

Experimental forecast results

The AirPolTool model results are published in real time on the site http://airpol.fatih.edu.tr. The site provides forecast for ten districts in Greater Istanbul Area, for a certain period of time request by the user.

In BISTAPOF system, the best model is identified using a wide range of performance indicators, like "the Pearson correlation coefficient R and its square R2, which represents the proportion of the observed variance explained by the model" [1]. The results as well as a demo version of this software can be accessed on the website: http://www.ehu.es/eolo/software/bistapof_demo/index.html.

The performance of the Airthess model may be investigated with the aid of the index of agreement that presents an overall estimation of the agreement between modeled and actual data.

Table 4. Forecast results

	Predicted air pollution indicators							**Period of time (maximum)**
	SO_2	CO	CO_2	NO	NO_2	O_3	PM_{10}	
AirPolTool	✓	-	✓	-	-	-	✓	3 days
BISTAPOF	✓	✓	-	✓	✓	✓	-	8 hours
Airthess	-	-	-	-	-	✓	-	24 hours

For a better analyze of these three systems, we have centralized all the forecast pollutants as well as the prediction period of time in table 4.

According to table 4, AirPolTool can predict values for SO_2, CO_2 and PM_{10} with 3 days ahead while BISTAPOF calculates the SO_2, CO, NO, NO_2 and O_3 values but only with 8 hours earlier. Airthess forecast a single gas value: ozone. He succeeded to predict values with 24 hours ahead.

Conclusions

In this paper, we have presented comparatively three artificial intelligence based systems applied in the air pollution domain. With that intention, first we have set some comparison aspects, like meteorological and air pollutants data used as input in the models, the predicted air pollution indicators or the maximum period of time for prediction. The artificial intelligence technique on which the model is based on was another important issue in our study. Then, we have studied the systems descriptions as well as the software applications.

Although, all three application are using artificial neural networks there are some different aspects. Genetic algorithms are applied in BISTAPOF along neural nets and multiple linear regression as a method to reduce the number of inputs. On the other side, Airthess implements Principal Component Analysis to minimize this number.

Artificial neural networks are powerful techniques that can be used with success in estimation and prediction problems. Moreover, these technologies have been found to excel in environmental systems which are complex, uncertain and non-linear. This is an important aspect because problems of the environment pollution have become a great threat for the planet.

References

1. Ibarra-Berastegi, G., Elias, A., Barona, A., Saenz, J., Ezcurra, A., Diaz de Argona, J. – *From diagnosis to prognosis for forecasting air pollution using neural networks: Air pollution monitoring in Bilbao*. Environmental Modelling & Software 23, 2008.
2. Karatzas, K., Kaltsatos, S. – *Air pollution modeling with the aid of computational intelligence methods in Thessaloniki, Greece*. Simulation Modelling Practice and Theory, Vol. 15, Issue 10, November 2007.
3. Kurt, A., Gulbagci, B., Karaca, F., Alagha, O. – *An online air pollution forecasting system using neural networks*. Environ. Int., doi:10.1016/j.envint.2007.12.020, 2008.
4. Matei V. – *Interacția substanțelor chimice cu agenți de mediu*. Editura Universității din Ploiești, 2004.
5. Oprea M., Nicoară S. – *Inteligența Artificială*, Editura Universității din Ploiești, 2005.
6. * * * – *Multiple Linear Regression*. http://www.stat.yale.edu/Courses/1997-98/101/linmult.htm, accessed on 15 march 2009
7. * * * – *Climate change*. http://ec.europa.eu/environment/climat/home_en.htm, accessed on 20 march 2009

Inteligenţa artificială aplicată în predicţia poluării aerului

Rezumat

Acest articol prezintă trei diferite metode utilizate pentru predicţia din oră în oră sau zilnică a unor gaze cu efect de seră. Emisiile acestor gaze reprezintă un factor important ce influenţează schimbarea climei care actual a devenit una dintre cele mai mari probleme de mediu. Astfel, monitorizarea şi controlul acestor gaze periculoase, ca şi predicţia lor, reprezintă o activitate importantă pentru autorităţile de mediu cât şi pentru cele de sănătate.Folosind anumite aspecte pentru comparaţie, acest studiu a fost făcut pe trei sisteme bazate pe inteligenţă artificială. AirPolTool, dezvoltat în Istanbul, Turcia poate prezice valorile SO_2, CO_2 *şi* PM_{10} *cu trei zile înainte. Al doilea sistem, BISTAPOF a fost creat în Bilbao, Spania pentru a determina cu 8 ore înainte valorile pentru* SO_2, CO, NO_2, NO *şi* O_3. *Airthess este un model construit să prezică pentru următoarea zi concentraţia de ozon în Thessaloniki, Grecia.*