An Urban Air Pollution Early Warning System Based on $PM_{2.5}$ Prediction Applied in Ploiesti City

MIHAELA OPREA*, MARIUS OLTEANU, RADU TEODOR IANACHE
Petroleum-Gas University of Ploiesti, 39 Bucuresti Blvd., 100680, Ploiesti, Romania

Fine particulate matter with a diameter less than 2.5 μm (i.e. $PM_{2.5}$) is an air pollutant of special concern for urban areas due to its potential significant negative effects on human health, especially on children and elderly people. In order to reduce these effects, new tools based on $PM_{2.5}$ monitoring infrastructures tailored to specific urban regions are needed by the local and regional environmental management systems for the provision of an expert support to decision makers in air quality planning for cities and also, to inform in real time the vulnerable population when $PM_{2.5}$ related air pollution episodes occur. The paper focuses on urban air pollution early warning based on $PM_{2.5}$ prediction. It describes the methodology used, the prediction approach, and the experimental system developed under the ROKIDAIR project for the analysis of $PM_{2.5}$ air pollution level, health impact assessment and early warning of sensitive people in the Ploiesti city. The $PM_{2.5}$ concentration evolution prediction is correlated with $PM_{2.5}$ air pollution and health effects analysis, and the final result is processed by the ROKIDAIR Early Warning System (EWS) and sent as a message to the affected population via email or SMS. ROKIDAIR EWS is included in the ROKIDAIR decision support system.

Keywords—decision support system, early warning system, air pollution, $PM_{2.5}$ prediction, $PM_{2.5}$ monitoring.

Air pollution in cities has a direct impact on the population health as revealed by majority of the epidemiological studies reported recently in the literature (e.g., [1, 2, 16, 17]). One of the air pollutants that is of special concern for urban areas is $PM_{2.5}$ (fine particulate matter with an aerodynamic diameter of less than 2.5 ∝m), due to its potential significant negative effects on the health of sensitive people, such as children and elderly people. In order to minimize these effects, new tools are needed by the local and regional environmental management systems for the provision of the decision makers' expert support in planning the air quality in cities and informing in real time the vulnerable population about the $PM_{2.5}$-related air pollution episodes that may occur in a certain urban region [3, 4]. Such tools need to use data collected from continuous $PM_{2.5}$ monitoring stations placed in the city's air pollution critical zones [5]. Different solutions were recommended so far in the literature for the implementation of modern environmental support tools. For example, a decision support system (DSS) is described in [6], combinations of a GIS-based system with DSS and/or an artificial intelligence (AI) approach are introduced in [7-9]. The integration of an air pollution prediction module in the environmental management system is an important step in performing air quality analysis and human health impact assessment in cases of air pollution episodes. In this sense, we are developing an integrated solution, which is based on DSS, GIS and AI, under the ROKIDAIR research project (http://www.rokidair.ro/en). The overall cyber-infrastructure is described in [18].

The aim of the ROKIDAIR environmental DSS tool is to perform the analysis of $PM_{2.5}$ air pollution levels, health impact assessment and early warning of vulnerable people in two pilot cities, Ploiesti and Targoviste, in order to provide a better protection of children against air pollution threats in urban areas from Romania. The project joins research efforts from three Romanian universities: Valahia University of Targoviste (coordinator), Petroleum-Gas University of Ploiesti and Politehnica University of Bucharest, and a Norwegian partner i.e., Norwegian Institute for Air Research (NILU). In this paper, we focus on a first version of the air pollution early warning experimental system, which is based on $PM_{2.5}$ short-term prediction, and was tested for Ploiesti city (Romania).

The main steps of the methodology followed for the development of the ROKIDAIR EWS system are divided in pre-requisite steps (steps 1÷4) and development steps (5÷7), as follows:

Pre-requisite steps:

- Analyze the urban region from the viewpoint of $PM_{2.5}$ air pollution and select a set of in-situ monitoring points;
- perform several $PM_{2.5}$ monitoring campaigns during weekdays and storing the data collected in a database;
- analyze the data from the monitoring database and identify the $PM_{2.5}$ critical zones (with higher levels of $PM_{2.5}$), where the $PM_{2.5}$ monitoring stations will be installed;
- develop a $PM_{2.5}$ prediction module based on a prediction model, selected from several $PM_{2.5}$ prediction models (e.g., statistical, AI-based such as artificial neural network etc); The prediction module will provide for each monitoring station separately, the $PM_{2.5}$ predicted values in the selected time window (e.g., next hour, next day), values that will be stored in the ROKIDAIR databases.

ROkidAIR EWS development steps:

Develop the $PM_{2.5}$ air pollution early warning system, that will analyze the $PM_{2.5}$ concentration level, the predicted value, the $PM_{2.5}$ air quality index – AQI, and will provide the expert message (with information and recommendation – measures to reduce the possible negative effects of $PM_{2.5}$ air pollution episodes on children health) that will be sent to vulnerable people.

Testing the developed ROKIDAIR EWS on various data (collected from the ROkidAIR databases or taken from the Romanian National Air Quality Monitoring Network - RNMCA site *www.calitateaer.ro*, for the PH-2 station, which monitors $PM_{2.5}$ in the Ploiesti city).

* *email: mihaela@upg-ploiesti.ro*

Running the experimental ROkidAIR EWS.

Since our main purpose is to reduce the $PM_{2.5}$ effects on children health, we are using for air quality assessments, the Air Quality Index (AQI), which quantifies the daily air quality report. This indicator shows how clean or unhealthy the air is, and what associated health effects might be of concern for certain groups of people. The correlation between the $PM_{2.5}$ concentration, AQI and the air quality level (derived from the EU standard: http://ec.europa.eu/environment/air/quality/standards.htm) is quantified by specific rules that are included in the ROkidAIR knowledge base. The rules are used by the analysis and decision module. Some examples of rules are given below:

Rule I-1 // analyze ConcPM2_5 the predicted concentration of PM2.5

IF Conc_PM2_5 d" 12 **THEN**

AQ_level = Good; //AQI = 1; AQI_ColorCode = Green;

Rule II-3 // AQI = 3; AQI_ColorCode = Orange

IF Conc_PM2_5≥35.5 **AND** Conc_PM2_5 d" 55.4 **THEN**

AQ_level = Unhealthy_for_Sensitive_People

Rule III-1 // informing message type

IF AQ_level = Good **THEN**

Health_Impact = No_impact_on_health;

Rule III-3 // early warning message type

IF AQ_level = Unhealthy_for_Sensitive_People **THEN**

Health_Impact = Impact_on_Sensitive_People_health;

Depending on the type of health impact, the decision module of the ROKIDAIR DSS system will choose a certain message type: *informing, early warning* or *alerting*, that will be sent as an email or SMS to the vulnerable people affected by the $PM_{2.5}$ air pollution episode.

Summarizing, the early warning system will receive the $PM_{2.5}$ short-term prediction (next hour or next day) for each $PM_{2.5}$ monitoring station from the pilot city, and will perform an analysis of the $PM_{2.5}$ air quality level and human health impact (by using the rules from ROkidAIR knowledge base) in each critical zone of the city. It will provide a corresponding message (i.e., *inform, warning, alerting* type message) with some recommendations, that will be sent via email and SMS to sensitive people, possible affected by the $PM_{2.5}$ air pollution episodes and to other registered users of the ROkidAIR geoportal.

Experimental part

The prediction problem

Air quality modeling can be realized with two types of mathematical models: deterministic (based on the analysis of atmospheric physical and chemical processes) and statistic (based on time series). Another class of models is given by the artificial intelligence-based models. The air quality in a certain geographic region can be determined by analyzing a set of parameters that includes: measurements of the area specific air pollutants concentrations and meteorological parameters (air temperature, atmospheric pressure, wind speed and direction, relative humidity etc.). The estimation of a certain air pollutant concentration evolution in a given time window is performed by applying a prediction method. As follows, the prediction problem formulation and some criteria used to evaluate the predictors' performance are briefly presented.

There are two categories of prediction problems, depending on the datasets that are used (as shown in [10]):

- prediction based only on the predicted parameter time series;
- prediction based on the predicted parameters time series and other related parameters.

Suppose we have the parameter x whose value will be predicted in the $t+k$ time window. For the first category (case (a)), the problem formulation is given by relation (1).

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

where:

t is time;

t-r is the time until past values of the parameter are used by the prediction model;

f_a is the prediction function specific to problems from the first category.

For the second category (case (b)), the general formulation is given by relation (2).

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

where:

t is time;

$p_1(.), p_2(.), \ldots, p_m(.)$ – the parameters that would influence the evolution of the x parameter;

t-r is the moment of time until past values are considered;

f_b is the prediction function specific to problems from the second category.

The two functions f_a and f_b represent the x parameter predictors.

Prediction Methods

As we have used in our experiments only time series with $PM_{2.5}$ concentrations, we have solved the first type of prediction problem, (a), and we have implemented three types of prediction methods: a simple one, named exponential smoothing, and two AI-based methods: artificial neural networks (ANN) and adaptive neuro-fuzzy inference system (ANFIS). The last two methods were described and compared in [11]. In the next section, we present the exponential smoothing method that was used in the experiments performed by the experimental ROkidAIR EWS.

Prediction Model for $PM_{2.5}$ Time Series Using Exponential Smoothing

Exponential smoothing represents a set of techniques used for filtering, distinguishing trends in time series and also for forecasting. Time series forecasting implies computing at current time t, a value which is k steps ahead ($t+k$). This forecasting model presumes a constant process. For such a model, first order exponential smoothing is described by relation (3) as given in [12]:

$$\tilde{y}_t = \lambda y_t + (1-\lambda)\tilde{y}_{t-1} \quad (3)$$

where:

$\tilde{y}_t$ = current smoothed value;

y_t = the value of the last observation;

$\tilde{y}_{t-1}$ = previous smoothed value;

λ = tuning parameter, $\lambda \in (0, 1)$.

The tuning parameter or discount factor λ influences directly how close the smoothed series follows the actual values of the time series. It is practically the weight put on the last observation, while 1-λ is the weight associated with the previous smoothed value.

Because of the constant level associated with the first-order exponential smoother, the one step-ahead value is simply the output y_t given by relation (4):

$$\hat{y}_{t+1} = \tilde{y}_t \quad (4)$$

where $\hat{y}_t$ = predicted value.

As new observations (i.e., measurements) are made (at $t+1$), a prediction error can be computed with relation (5), as the difference between the predicted value and the current observation [12]:

$$e_t(1) = y_{t+1} - \hat{y}_{t+1}(t) \quad (5)$$

This prediction error can be used as a correction factor when computing the next value with relation (6):

$$\hat{y}_{t+1}(T) = \hat{y}_t(1) + \lambda e_t(1) \quad (6)$$

The model prediction accuracy is given by two indicators that are usually used: the Sum of Squared Errors (*SSE*) and Root Mean Square Error (*RMSE*), computed as in [12], and given by relations (7) and (8), respectively:

$$SSE(\lambda) = \sum e_{t-1}^2(1) \quad (7)$$

$$RMSE(\lambda) = \sqrt{\sum \frac{e_{t-1}^2}{n}} \quad (8)$$

where: n = total number of considered values.

Case Study of $PM_{2.5}$ Prediction for Ploiesti City

We have applied in our experimental EWS system the exponential smoothing prediction method on the data collected by the Romanian National Air Quality Monitoring Network at the monitoring stations from the Ploiesti city, which are publicly available on the www.calitateaer.ro site. In this case study, we have used the values of $PM_{2.5}$ concentration measured at PH-2 monitoring station located in Ploiesti, during the period 28.05.2015 – 29.06.2015. The representation of the time series data (TS-Ploiesti) is shown in figure 1. The values are given in [μg/m^3], and a continuous time series of 744 values was used.

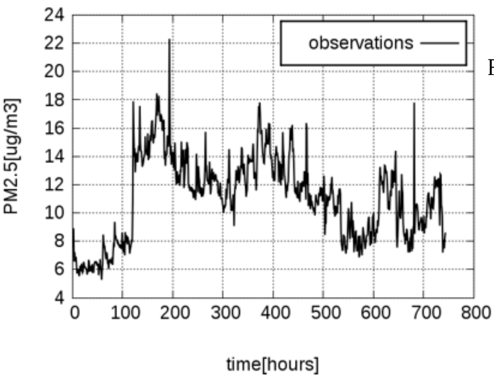

Fig. 1. Measured values of $PM_{2.5}$ from the TS-Ploiesti time series

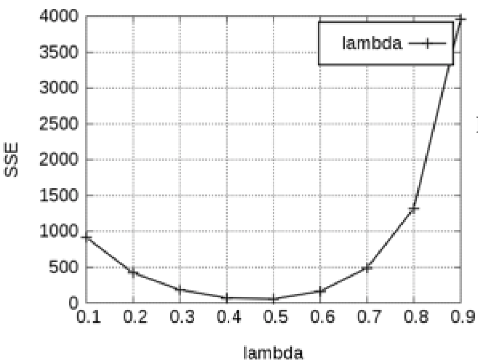

Fig. 2. Plot of the errors obtained for different λ

By running the algorithm with different values for l, the following results were obtained.

The values presented in (table 1) indicate an optimal value around 0.5, as can be seen also from figure 2.

By further refining a local optimum search near 0.5 value, an error of SSE=51.1027 has been obtained for λ= 0.4648.

Figure 3 presents the predicted and actual values for an interval of 24 h (from hour 530 to 554) in order to better visualize the two curves.

Table 1
STATISTICAL INDICATORS FOR THE TS-PLOIESTI SAMPLE SERIES WITH DIFFERENTν VALUES

λ	SSE	RMSE
0.1	915.50	1.1092
0.2	421.01	0.7522
0.3	186.17	0.5002
0.4	78.38	0.3119
0.5	57.77	0.2786
0.6	165.96	0.4723
0.7	489.73	0.8113
0.8	1322.34	1.3331
0.9	3962.56	2.3078

As it can be seen from the figure, a good prediction is obtained by this simple method. Further investigations should draw a conclusion regarding the time window for which the computation will take into consideration, seasonal or trend influences having negative effects on the accuracy of exponential smoothing methods.

The PM2.5 Air Pollution Early Warning Experimental System for Ploiesti City

Our experimental $PM_{2.5}$ air pollution early warning system for Ploiesti is based on the architecture of the ROkidAIR intelligent system, described in [13].

$PM_{2.5}$ Monitoring in the Critical Zones of Ploiesti City

According to the methodology used in the ROkidAIR project, the first step was to identify the critical zones in the Ploiesti city, which required monitoring and analysis from the viewpoint of the $PM_{2.5}$ air pollutant. Figure 4 shows the map with the critical zones that were identified during the $PM_{2.5}$ monitoring campaigns run in 2014 and 2015.

The Experimental Early Warning System ROkidAIR EWS

The proposed air pollution early warning system has the following modules: a Database module, a GIS module and a Messaging module that include the analysis sub-module. The structure of this system is detailed in [14] and is shown in figure 5.

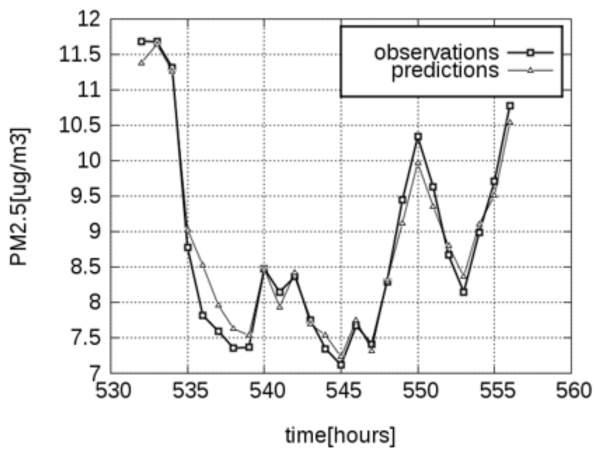

Fig. 3. Measured and predicted values for λ=0.4648

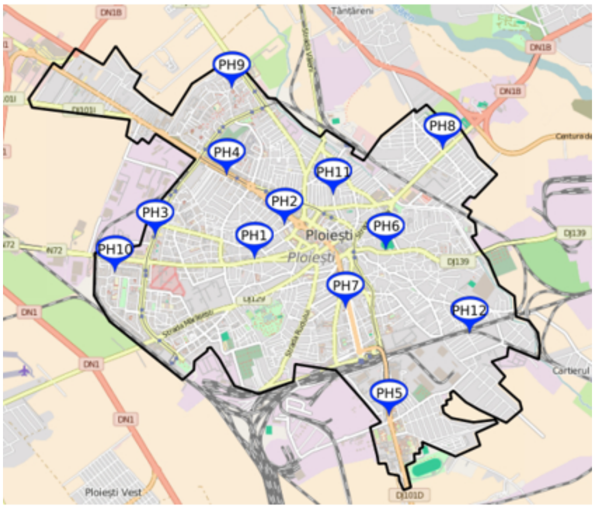

Fig. 4. The $PM_{2.5}$ air pollution monitoring points in Ploiesti city used in the Rokidair monitoring plan

The *Database module* has two databases: (1) a database that contains the measured data and the predicted data; and (b) a database which has the geospatial data (buildings, roads etc.) that will be used by the GIS module. The *GIS module* (Geographic Information System) contains a map server, named *Geoserver*, and a JavaScript library called Openlayers. With those two components, the GIS module takes the data from the Database module to generate maps that will be displayed in a web page. The $PM_{2.5}$ prediction module is integrated on the web page of the ROkidAIR geoportal [20]. The ROkidAIR user can choose from the main menu what data will be displayed on the map: measured data or predicted data. In addition, the user can select the layers that will be shown on the map, like base layer, buildings layer, education entities layer etc.

In order to send the messages to the ROkidAIR EWS users, the *Messaging module* will make an analysis on the data taken

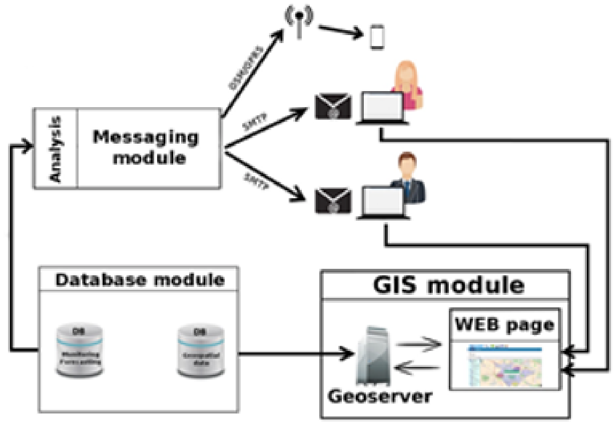

Fig. 5. The structure of the air pollution EWS [19]

from the Database module. We divided the users in two categories: sensitive groups and general public. This module will send email messages to all the users and SMS messages to the users that belong to the sensitive groups.

The messaging module will send the SMS messages using a GSM/GPRS modem that is connected to the computer. The communication with the GSM/GPRS modem is assisted by an SMS gateway which is based on a software application, named Gammu. The information contained by the warning messages must be clear and easy to understand by the users of the early warning system, because some of the users are not accustomed with the technical language. From the different instruments that can be used to send the warning messages, we have chosen the SMS messages and the email messages.

Description of Databases

The software application has two databases: (1) a database called **main DB** that will save the measured and predicted values of the $PM_{2.5}$ pollutant, data regarding the users of the software application, data about the monitoring stations and other relevant data; and (2) a database called **geospatial DB** in which we will save data about roads, buildings etc. Figure 6 shows the relationships between the tables of the databases.

Case Study of Running the Experimental EWS for Two $PM_{2.5}$ Air Pollution Scenarios

Figure 7 shows the ROkidAIR system interface web site (a), and a graph (b) with mean values of the $PM_{2.5}$ concentration registered in Ploiesti, during a $PM_{2.5}$ air pollution episode on August 28, 2015 – *scenario-1*.

Figure 8 shows an example of the ROkidAIR EWS run, in the case of *scenario-2*, with an early warning sent by email on March 6, 2016 at 17:00 hour, as an expert message (fig. 8 (a)) and as SMS (fig. 8 (b)), informing the vulnerable people about the possibility of occurring a $PM_{2.5}$ air pollution episode in the PH8 urban area, during the next hour (i.e. around 18:00 h), when the estimated value of the AQI is 171. The re-

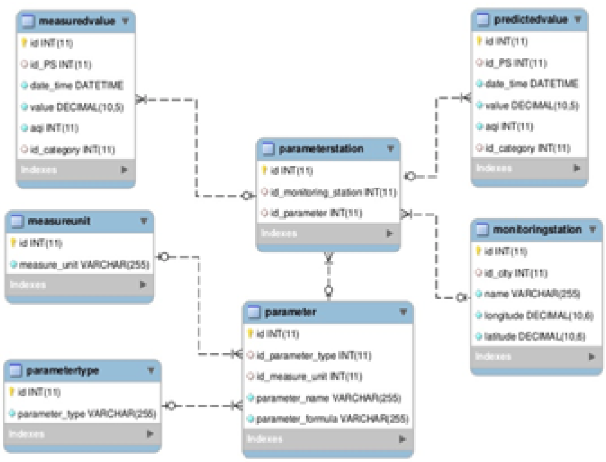

Fig. 6. Relationships between the tables of the ROKIDAIR EWS application databases

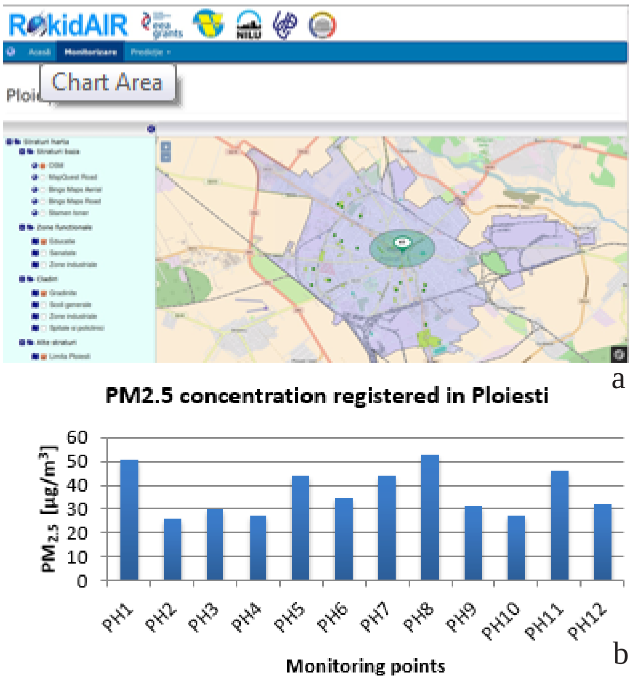

Fig. 7. The ROkidAIR system web interface for the Ploiesti city (a) and the $PM_{2.5}$ air pollution Ploiesti graph during a $PM_{2.5}$ air pollution episode (registered on August 28, 2015) – *scenario-1* (b)

Fig. 8.Example of messages sent by the experimental ROkidAIR EWS via email (a) and SMS (b) for a $PM_{2.5}$ air pollution episode from *scenario-2*

commendation for people with cardiovascular and respiratory problems as well as children and elders is to avoid outdoor activities or, at least, to reduce longer physical effort.

Conclusions

The paper presents an air pollution early warning system, ROkidAIR EWS, based on the prediction of $PM_{2.5}$ air pollutant that was developed for Ploiesti city. The main purpose of the system is to inform the population via email and/or SMS about $PM_{2.5}$ air pollution episodes, which can affect the health of sensitive people (e.g., children), providing also some recommendations to reduce negative effects by reducing the exposure time (e.g., the time when children are doing outdoor activities in kindergartens or schools).

The core of the ROKIDAIR EWS is the prediction module, which provides the forecasting of the $PM_{2.5}$ concentration evolution in the next hours or next day, for each $PM_{2.5}$ monitoring station situated in the critical zones of the Ploiesti city. The prediction model can be selected from a simple one (e.g., exponential smoothing) to a more complex one (e.g., a feed forward artificial neural network model). The system is based on GIS and is included in the ROKIDAIR decision support system.

As a future work, we shall investigate the adaptation of the exponential smoothing prediction model to different types of $PM_{2.5}$ time series in order to obtain an automated criterion for choosing the l parameter (e.g., by using a computational intelligence technique as in [15] or a heuristic rule).

Acknowledgement: 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.

References

1.KIESEWETTER, G., SCHOEPP, W., HEYES C., AMANN, M., Modelling PM2.5 impact indicators in Europe: Health effects and legal compliance, Environmental Modelling & Software, 2015, http://dx.doi.org/10.1016/ j.envsoft.2015.02.022.

2.SCHIAVON, M., ANTONACCI, G., RADA, E.C., RAGAZZI, M., ZARDI, D., Modelling Human Exposure to Air Pollutants in an Urban Area, Rev. CHim. (Bucharest), **65**, no. 1, 2014, p. 61.

3.LIU, H.-Y. ET AL., Facilitating knowledge transfer: decision support tools in environment and health, Environmental Health, 11:S17, 2012.

4.BONTOS, M.D., VASILIU, D., Short-term Health Impact Assessment of Air Polution in Targoviste City (Dambovita County), Rev. Chim. (Bucharest), **67**, no. 9, 2016, p. 1854.

5.BUCUR, E., DANET, A.F., Particulate Matter and Polycyclic Aromatic Hydrocarbon Air Pollution in Areas of Bucharest with Heavy Road Traffic,Rev. Chim. (Bucharest), **67**, no. 4, 2016, p. 621.

6.BOOTY, W.G., WONG I., LAM D., RESLER O., A decision support system for environmental effects monitoring, Environmental Modelling & Software, 24, pp. 889-900, 2009.

7.ELBIR, T., MANGIR, N., SIMSIR, S., EREN, T., OZDEMIR, S., Development of a GIS-based decision support system for urban air quality management in the city of Istanbul, Atmospheric Environment, 44, pp. 441-454, 2010.

8.MIRCEA, M., ZANINI, G., BRIGANTI, G., CAPPELLETTI, A., PEDERZOLI A., VITALI, L., PACE, G., MARRI, P., SILIBELLO C., FINARDI S., CALORI, G., Modelling air quality over Italy with MINNI atmospheric modelling system: from regional to local scale, in: Air pollution modelling and its application, Steyn, D.G., TriniCastelli, S. (Editors), 2010.

9.VIENNEAU, D., HOOGH DE, K., BRIGGS D., A GIS-based method for modelling air pollution exposure across Europe, Science of the Total Environment, 408, pp. 255-266, 2009.

10.OPREA, M., DRAGOMIR, E.G., MIHALACHE, S.F., POPESCU, M., Methods and techniques for PM2.5 concentration prediction in urban areas, pages 387-428, chapter 11 in the book: Methods for the assessments of particulate matter air pollution effects on children health, Iordache S., Dunea D. (editors), Matrix Rom, Bucharest (in Romanian), 2014.

11.OPREA, M., MIHALACHE, S.F., POPESCU, M., A comparative study of computational intelligence techniques applied to PM2.5 air pollution forecasting, Proceedings of the ICCCC2016, 10-14 May 2016, Baile Felix, Oradea, Romania, pp. 103-108, 2016.

12.MONTGOMERY, D.C., JENNINGS, C.L., KULAHCI, M., Introduction to time series analysis and prediction, John Wiley & Sons, 2008.

13.OPREA, M., IANACHE, C., MIHALACHE, S.F., DRAGOMIR, E.G., DUNEA, D., IORDACHE, S., SAVU T., On the development of an intelligent system for particulate matter air pollution monitoring, analysis and forecasting in urban regions, Proceedings of 19th ICSTCC2015, 14-16 October 2015, Cheile Gradistei, Romania, pp. 711-716, 2015.
14.IANACHE R.T., Study concerning the development of an automated system for particulate matter air pollution early warning, MSc dissertation, Petroleum-Gas University of Ploiesti, 2016 (in Romanian).
15.KARATZAS, K.D., VOUKANTSIS, D., Studying and predicting quality of life atmospheric parameters with the aid of computational intelligence methods, iEMS 2008, Proceedings of the International Congress on Environmental Modelling and Software, 4th Biennial Meeting of iEMSs, Sanchez-Marre, M., Bejar, J., Comas, J., Rizzoli, A., and Guariso, G. (Editors), pp. 1133-1139, 2008.
16. DUNEA, D.; IORDACHE, S.; LIU, H.-Y.; BØHLER, T.; POHOATA, A.; RADULESCU, C.; Quantifying the impact of PM2.5 and associated heavy metals on respiratory health of children near metallurgical facilities. Environ. Sci. Pollut. Res. 2016, 23, 15395–15406.
17.RADULESCU C., IORDACHE S., DUNEA D., STIHI C., DULAMA I. (2015), Risks assessment of heavy metals on public health associated with atmospheric exposure to PM2.5 in urban area, Romanian Journal of Physics, 60(7-8) p. 1171-1182
18.IORDACHE, S.; DUNEA, D.; LUNGU, E.; PREDESCU, L.; DUMITRU, D.; IANACHE, C.; IANACHE, R.; A Cyberinfrastructure for Air Quality Monitoring and Early Warnings to protect Children with Respiratory Disorders, Proceedings of the 20th International Conference on Control Systems and Computer Science (CSCS20–2015), Bucharest, Romania, 27-29, 2015, 789-796.
19. IORDACHE St, DUNEA D, IANACHE C, PREDESCU L, PREDESCU M (2016) Developing a cyberinfrastructure for air quality management in an urban environment 16th International Multidisciplinary Scientific GeoConference SGEM 2016, SGEM2016 Conference Proceedings, ISBN 978-619-7105-64-3 / ISSN 1314-2704, June 28 - July 6, 2016, Book4 Vol. 2, 331-338 pp; DOI: 10.5593/SGEM2016/B42/S19.043
20. DUNEA D., LUNGU, E., POHOATA, A., An Environmental Mapping System for Airborne Particulate Matter Monitoring in Urban Areas, Proceedings of the 14th International Conference on Software Engineering, Parallel and Distributed Systems (SEPADS ’15), At Dubai, United Arab Emirates February 22-24, 85-94

Manuscript received: 10.01.2017