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

MIHANEA 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 µm), 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.rockidaire.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 (România).

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:

- Perform the analysis 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 model 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

858 http://www.revistadechimie.ro REV.CHIM.(Bucharest) ● 68 ● No. 4 ● 2017
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 PM$_{2.5}$

**Rule II-3 // AQI = 3; AQI ColorCode = Orange**

**Rule III-1 // informing message type**

- IF Conc PM2.5 < 35.5 THEN AQ level = Unhealthy_for_Sensitive_People
- IF Conc PM2.5 = Good THEN Health_Impact = No_impact_on_health
- IF Conc PM2.5 >35.5 AND Conc PM2.5 < 55.4 THEN AQ level = Unhealthy_for_Sensitive_People

**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]:

$$\hat{x}_{t+k} = \lambda \hat{x}_t + (1-\lambda)\hat{x}_{t-1}$$

where:
- $\lambda$ – tuning parameter, $\lambda \in (0,1)$. The tuning parameter or discount factor $\lambda$ 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 $1-\lambda$ is the weight associated with the previous smoothed value.
- $\hat{x}_t$ = current smoothed value;
- $\hat{x}_{t-1}$ = the value of the last observation;
- $\hat{x}_{t-2}$ = previous smoothed value.

Because of the constant level associated with the first-order exponential smoother, the one step-ahead value is simply the output $\hat{y}_t$ given by relation (4):

$$\hat{y}_{t+1} = \hat{y}_t$$

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]:

\[
e_{t+1}(t) = y_{t+1} - \hat{y}_{t+1}(t) \tag{5}
\]

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

\[
\hat{y}_{t+1}(t) = \hat{y}_{t}(t) - \lambda e_{t+1}(t) \tag{6}
\]

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:

\[
SSE(\lambda) = \sum e_{t+1}^2(\lambda) \tag{7}
\]

\[
RMSE(\lambda) = \sqrt{\frac{\sum e_{t+1}^2(\lambda)}{n}} \tag{8}
\]

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 [\( \mu g/m^3 \)], and a continuous time series of 744 values was used.

![Fig. 1. Measured values of PM\(_{2.5}\) from the TS-Ploiesti time series](image)

By running the algorithm with different values for \( \lambda \), 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.

<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>SSE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>915.50</td>
<td>1.1092</td>
</tr>
<tr>
<td>0.2</td>
<td>421.01</td>
<td>0.7522</td>
</tr>
<tr>
<td>0.3</td>
<td>186.17</td>
<td>0.5002</td>
</tr>
<tr>
<td>0.4</td>
<td>78.38</td>
<td>0.3119</td>
</tr>
<tr>
<td>0.5</td>
<td>57.77</td>
<td>0.2786</td>
</tr>
<tr>
<td>0.6</td>
<td>165.96</td>
<td>0.4723</td>
</tr>
<tr>
<td>0.7</td>
<td>489.73</td>
<td>0.8113</td>
</tr>
<tr>
<td>0.8</td>
<td>1322.34</td>
<td>1.3331</td>
</tr>
<tr>
<td>0.9</td>
<td>3962.56</td>
<td>2.3078</td>
</tr>
</tbody>
</table>

![Table 1: Statistical indicators for the TS-Ploiesti sample series with different \( \nu \) values](image)

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 PM\(_{2.5}\) 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.
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 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-
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., SCHIEPP, W., HEYES C., AMANN, M., Modelling PM$_{2.5}$ impact indicators in Europe: Health effects and legal compliance, Environmental Modelling & Software, 2015, http://dx.doi.org/10.1016/j.envsoft.2015.02.022.


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


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).


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


Manuscript received: 10.01.2017