AN ANALYSIS OF $PM_{2.5}$ RELATED AIR POLLUTION IN PLOIESTI CITY

Mihaela OPREA[1], Elia G. DRAGOMIR[2], C. IANACHE[1], R. IANACHE[1]

ABSTRACT. – An analysis of $PM_{2.5}$ related air pollution in Ploiesti city. The paper presents an analysis of $PM_{2.5}$ related air pollution in Ploiesti city, based on datasets with measurements of $PM_{2.5}$ concentration and some meteorological variables at PH-2 monitoring station, taken from the RNMCA public site (www.calitateaer.ro) and AirBase, and the dataset with in situ measurements of $PM_{2.5}$ concentration and some meteorological variables acquired through the ROKIDAIR project monitoring campaigns run during the period August 2014 – November 2016, in the Ploiesti city. $PM_{2.5}$ most correlated meteorological variables are highlighted and the Ploiesti city map with seasonal $PM_{2.5}$ air pollution distribution is provided. The paper experiments highlight that the $PM_{2.5}$ concentration evolution analysis must be done taking into consideration the other atmospheric and seasonal parameters, in order to obtain the most appropriate results.

Keywords: air pollution analysis, $PM_{2.5}$ air pollution episodes, meteorological parameters correlation.

1. INTRODUCTION

Particulate matter (PM) with the aerodynamic diameter less than 2.5 μm (i.e. $PM_{2.5}$) is one of the air pollutants of special concern due to its potential negative effects on human health of sensitive people (such as children and elderly people) (see e.g. a recent study presented in (Iordache et al., 2014), (Prieto-Parra et al., 2017), (Calderón-Garcidueñas et al., 2015), (Zeng et al., 2016), (Silva da Silva et al., 2015)). The analysis of $PM_{2.5}$ related air pollution is a first step toward a better protection of people living in cities, performed through more informed forecasting and early warning of $PM_{2.5}$ air pollution episodes. It requires real time monitoring of $PM_{2.5}$ concentration from most polluted sites of urban areas. The National Monitoring Network of Air Quality (RNMCA) in Romania measures two fractions of PM, PM_{10} and $PM_{2.5}$ at certain automated monitoring stations of some cities. For example, in Ploiesti city more RNMCA monitoring stations, from the total of six ones that exist at present, measures PM_{10}, while only one monitoring station, PH-2 (RNMCA), measures $PM_{2.5}$ and, moreover, due to technical problems it does not provide continuous time series with $PM_{2.5}$ concentrations.

[1] Petroleum-Gas University of Ploiesti, Faculty of Mechanical and Electrical Engineering, Department of Automatic Control, Computers and Electronics, 100680 Ploiesti, Romania, e-mail: mihaela@upg-ploiesti.ro, cornelianache@yahoo.com, ianacheteodor@yahoo.com

[2] Petroleum-Gas University of Ploiesti, Faculty of Letters and Sciences, Department of Informatics, Information Technology, Mathematics and Physics, 100680 Ploiesti, Romania, e-mail: elia.dragomir@yahoo.com

A possible solution is to develop continuous $PM_{2.5}$ monitoring networks in the air polluted cities. Under the ROKIDAIR project (www.rokidair.ro/en) it is developed such a monitoring network composed of $PM_{2.5}$ micro-stations for two Romanian pilot cities, Ploiesti and Targoviste, as part of the ROKIDAIR system cyberinfrastructure, introduced in (Iordache et al. 2015) and extended with artificial intelligence based modules for air pollution analysis, forecasting and early warning as shown in (Oprea et al., 2015). As a prerequisite, several $PM_{2.5}$ monitoring campaigns were run in different zones of the two cities, identifying the critical zones, where the micro-stations will be placed.

The paper presents statistical experimental results to discover $PM_{2.5}$ most correlated meteorological variables and provides the Ploiesti city map with the distribution of $PM_{2.5}$ average values measured during autumn season.

2. STATISTICAL ANALYSIS OF $PM_{2.5}$ AIR POLLUTION

2.1 Data Set

For the statistical analysis of $PM_{2.5}$ air pollution, we have used a data set that contains $PM_{2.5}$ hourly records measured at the PH-2 national monitoring station. This station is part of RNMCA and is located in the Ploiesti city center. The monitored meteorological parameters at PH-2 (RNMCA) are: atmospheric pressure, precipitations, relative humidity, solar radiation, temperature, wind direction, wind speed. The database contains 12001 records hourly measured from March 2015 until July 2016 (when $PM_{2.5}$ has the least missing records value). In order to determine the most important atmospheric parameters that may influence the $PM_{2.5}$ evolution and furthermore to be monitored in the campaigns, there were conducted different statistical methods as a preliminary analysis.

Table 1 presents some preliminary statistics that characterize the database used in this analysis. $PM_{2.5}$ has a target role, it is the dependent variable analyzed based on the input (independent) parameters. The minimum, maximum, mean and standard deviation is calculated for each variable as well as the non missing value.

Table 1. ***The database variables general description***

Variable	Role	Mean	Deviation	Non Missing	Min	Max	Skewness	Kurtosis
Atmospheric pressure	INPUT	997.91	7.74	11305	974.64	1021.19	0.23	0.08
Precipitations	INPUT	0.30	0.65	11319	0.00	8.11	6.12	56.44
Relative Humidity	INPUT	67.82	19.06	11310	13.01	100.00	-0.35	-0.72
Solar Radiation	INPUT	63.21	108.90	11306	1.00	453.34	2.32	4.81
Temperature	INPUT	14.87	9.57	11307	-10.75	37.70	0.05	-0.57
Wind Direction	INPUT	97.81	122.05	11317	0.00	360.00	0.65	-1.21
Wind Speed	INPUT	0.29	0.26	11312	0.00	1.20	2.05	3.76
PM_2_5	TARGET	17.40	17.89	3554	0.00	377.82	10.46	140.50

The Skewness and Kurtosis statistics are also presented. A greater than 0 Skewness value indicates an inclination towards the left of the time series distribution, representing more extreme values to the right. A value lower than 0 presents an inclination to right. From Table 1 it can be noticed that the time series of $PM_{2.5}$, precipitations, solar radiation and wind speed present more extreme values to the right, while the rest of time series distribution are symmetric around the mean. The Kurtosis statistic reflects the high probability for extreme value (a value greater than 3) such as for precipitations, solar radiation and $PM_{2.5}$ time series. A Kurtosis value less than 3 characterizes a flatter distribution that has values spread out over a longer interval around the mean.

2.2 Experiments and discussion

Several statistical analyses were performed with SAS software package. We have studied the meteorological parameters' possible influence on the $PM_{2.5}$ evolution. In order to highlight the most correlated meteo variables to the $PM_{2.5}$ concentration we have used SAS Enterprise Guide and SAS Enterprise Miner 7.1 Workstation 13.2.

The main steps that were followed are: (1) calculate the Pearson and Spearman correlation coefficients; (2) determine the crosscorrelations of $PM_{2.5}$ versus the inputs; (3) determine the similarity measure of $PM_{2.5}$ and the meteorological parameters.

The Pearson and Spearman coefficients are calculated between $PM_{2.5}$ and the meteorological parameters in order to determine a possible correlation among them. Both coefficients may determine a possible relantionship. They have values between -1 and 1 (as closest are to -1 or 1 as stronger the relantionship is), but a meaningful relationship can exist even if the correlation coefficients are 0. The Pearson coefficient measures only the linear relantionships, as the Spearman coefficient reflects only the monotonic relantionships.

Table 2 presents Pearson correlation values around 0 for all the meteo parameters thus a linear relantionship may not exists. The Spearman correlation values are slightly higher which may reflect a weak monotonic relantionship between $PM_{2.5}$ and temperature or solar radiation (the values are negative so as higher temperature or solar radiation values are as lower $PM_{2.5}$ concentrations). Positive values are determined for relative humidity and atmospheric pressure.

Table 2. ***The correlation coefficients of the $PM_{2.5}$ and the database variables***

Input	Pearson Correlation	Spearman Correlation
REP_Relative_Humidity	0.092151	0.31244
REP_Atmospheric_pressure	0.079556	0.2489
REP_Wind_Direction	0.016782	0.09942
REP_Precipitations	0.014403	0.09137
REP_Wind_Speed	-0.027331	0.03816
REP_Solar_Radiation	-0.066968	-0.23327
REP_Temperature	-0.080977	-0.35044

Next, we have performed crosscorrelations analysis. SAS Enterprise Miner 13.2 provides crosscorrelation analysis for timestamped data. Crosscorrelation analysis is a similarity measure of two time series as a function of time lag, which indicates the time points where the two series are correlated.

Figure 1 sketches the $PM_{2.5}$ crosscorrelations versus the input series. The plot shows both positive and negative time lags, but is not necessarily symmetric about lag 0. The graphics shape highlights clearly that the correlation between $PM_{2.5}$ concentration evolution and any meteo parameter presents some time points where the time series are correlated (e.g the peaks over 0.2 between lags -18 and 0, or over -0.2 between -7 and 0).

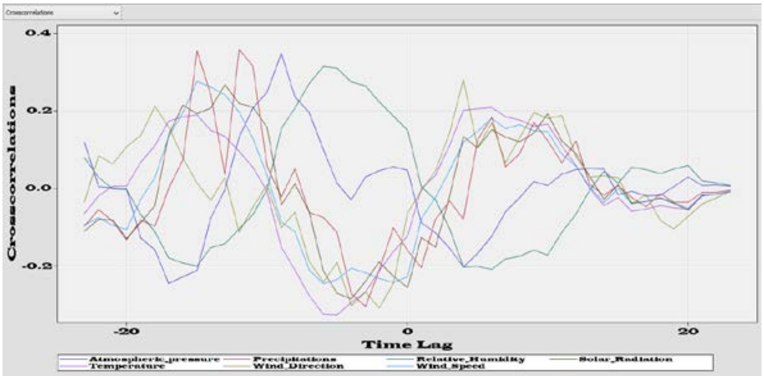

Fig. 1. ***The crosscorrelations of target PM2.5 versus input series***

Furthermore, in order to determine possible $PM_{2.5}$ relantionships with meteorological variables, the similarity measure is computed. This is a metric that measures the distance between the input and the target sequence. Figure 2 is a bar chart comparing the similarity measure of each input time series to the target $PM_{2.5}$. The conclusion is that the atmospheric pressure and the relative humidity are most similar to the $PM_{2.5}$ time series followed by temperature and wind direction.

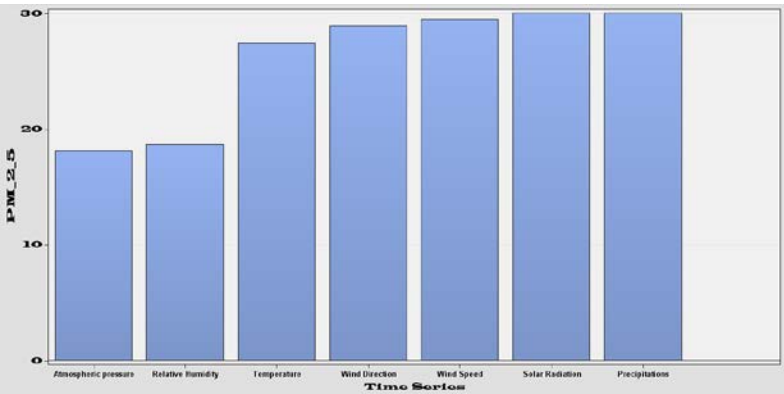

Fig. 2. ***The Similarity measure of $PM_{2.5}$ versus Input***

All three methods (*Pearson and Spearman coefficients*, *crosscorelation* and *similarity measure* with the distance map) that were applied on the current database in order to determine which meteorological time series are correlated with $PM_{2.5}$, have discovered that atmospheric pressure and relative humidity are most similar to particulate matter evolution. However, the indirect relationship, i.e. when a time series values increase the other one's decrease (the negative correlations coefficients values), are hardly discovered. Thus, in the $PM_{2.5}$ monitoring campaigns it was taken into consideration that atmospheric pressure and relative humidity influence the $PM_{2.5}$ concentration evolution.

3. $PM_{2.5}$ MONITORING CAMPAIGNS

$PM_{2.5}$ monitoring campaigns were made in Ploiesti city, with portable measurement systems that use either an infrared beam or a laser beam to detect the concentration of the particulate matter. These campaigns were made during the time period of August 2014 – November 2016.

In the preliminary study stage, from August 2014 to March 2015, for $PM_{2.5}$ measurements it was used the Casella Microdust Pro system that uses an infrared beam. Regarding the data acquisition, the Casella Microdust Pro system automatically calculates the average concentration and the maximum concentration that was recorded during the entire analysis of the time series collected.

Starting with April 2015 it was used DustTrak DRX measurement system. This system gives the possibility to measure simultaneously the following particulate matter fractions: PM_{10}, PM_4, $PM_{2.5}$, PM_1 and Total fraction. Being protected by a weatherproof protective enclosure and having heating elements for air samples at the inlet system, the DustTrak DRX measurement system is designed for continuous outdoor measurements. The collected data are registered in a file, where the system records the minimum, the maximum and the average value of the monitored parameters.

The selection of the monitoring points, took into account the following aspects (Dunea et al., 2016):

- Quasi-radial placing of the monitoring points on the map;
- The location in Ploiesti city of the automated stations of RNMCA;
- The proximity of the monitoring points to the pediatric hospital of Ploiesti, kindergartens, and schools.

Considering these aspects, the monitoring campaigns were performed in 12 sampling points of Ploiesti city. These points are illustrated in figure 3.

For the preliminary study stage, the measurement period, on each of the 12 monitoring points, was 15 minutes. The monitoring campaigns were carried in the time slot 08:00-20:00 and at least two days after a rainfall. In the second study stage, from April 2015 to November 2016, the measurement time was raised at 1 hour and the number of monitoring points was decreased to 4 points. These points are PH-3, PH-4, PH-6 and PH-8, under the ROKIDAIR network (Fig. 3).

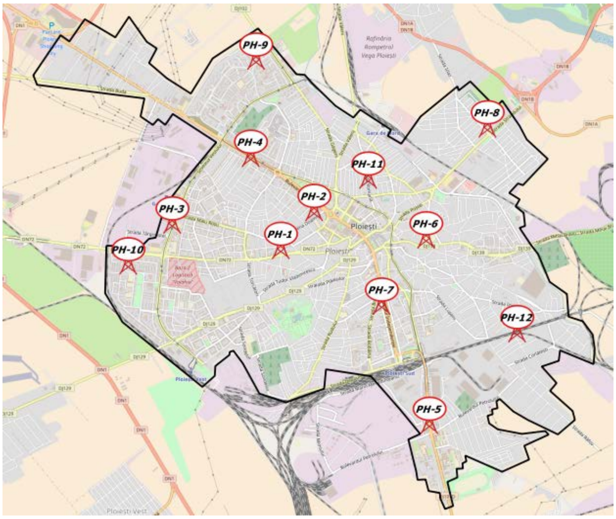

Fig. 3. ***The location of the monitoring points***

4. ANALYSIS OF $PM_{2.5}$ MONITORING CAMPAIGNS DATA

In the analysis of data obtained from monitoring campaigns, it was taken into consideration only $PM_{2.5}$ particles between April 2015 and November 2016 without the winter months. The 20 series of measurement data (approximately 1 or 2 series per month) obtained during the monitoring campaigns with Dusttrak DRX system, were separated in 3 categories: data measurements from *autumn months* (September, October and November), data measurements from *spring months* (April, May) and data measurements from *summer months* (June, July and August).

Figure 4 illustrates the difference, in each of the 12 points, between the average value of $PM_{2.5}$ concentration in autumn months, the average value of $PM_{2.5}$ in spring months and the average value of $PM_{2.5}$ in summer months.

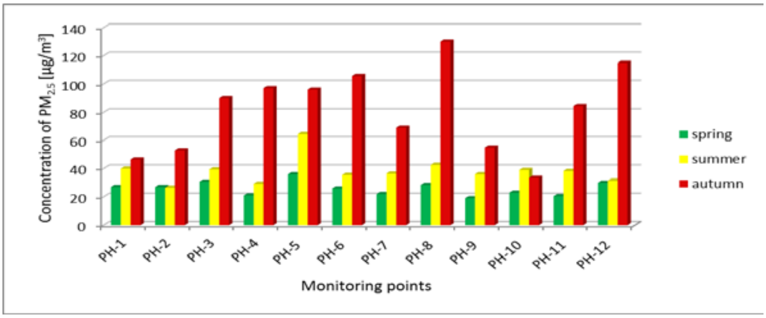

Fig. 4. ***The average value of $PM_{2.5}$ concentration in spring, summer and autumn***

As it can be seen, the concentration levels of $PM_{2.5}$ in autumn months are higher than the ones in spring and summer months. The reason behind these results is that in autumn months people use heating equipment (like central heating, wood stove etc.), that emit a huge quantity of particulate matter in the atmosphere.

In figure 5 it is mapped the distribution of $PM_{2.5}$ average values measured during autumn season (which was the season with $PM_{2.5}$ concentration higher than spring and summer) in the years 2015 and 2016, in Ploiesti city. This map was generated using Inverse Distance Weighting algorithm to determine the value of particulate matter concentration between the measuring points. As it can be seen, the East side of the Ploiesti city is more polluted than the West side.

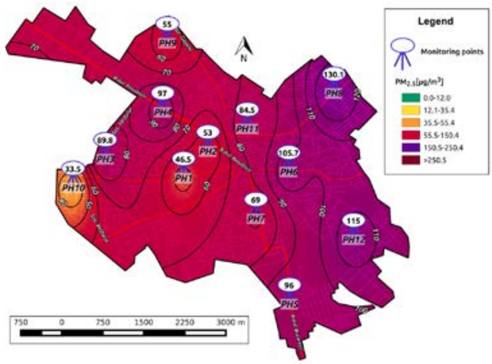

Fig. 5. ***Distribution of $PM_{2.5}$ average values measured during autumn season 2015-2016***

5. CONCLUSIONS

An efficient air pollution analysis in urban areas need to be based on concentrations measurements of particulate matter fractions, as e.g. PM_{10}, $PM_{2.5}$, $PM_{1.0}$, apart from those of the common air pollutants, NO_x, CO, SO_2, NO_2 etc, due to their major potential effects on human health of sensitive people such as children and elderly. In this paper there were conducted some preliminary experiments in order to identify different correlations between $PM_{2.5}$ concentration and other atmospheric parameters.

The results of the statistical analysis of $PM_{2.5}$ related air pollution and of the $PM_{2.5}$ monitoring campaigns data analysis, that were performed for the Ploiesti city, are very useful for the selection of $PM_{2.5}$ most correlated meteorological variables that will be used at a certain monitoring site, by the forecasting module of the ROKIDAIR system.

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.

REFERENCES

1. Calderón-Garcidueñas L., Franco-Lira M., D'Angiulli A., Rodríguez-Díaz J., Blaurock-Busch E., Busch Y., Chao C., Thompson C., Mukherjee P., Torres-Jardón R., Perry G. (2015), *Mexico City normal weight children exposed to high concentrations of ambient PM2.5 show high blood leptin and endothelin-1, vitamin D deficiency, and food reward hormone dysregulation versus low pollution controls. Relevance for obesity and Alzheimer disease*, Environmental Research, Volume 140, 579-592.
2. Dunea D., Iordache Şt, Radulescu C, Pohoata A, Dulama I.D. (2016), *Multidimensional Approach To The Influence of Wind on the Variations of Particulate Matter and Associated Heavy Metals in Ploiesti City, Romania*, ROMANIAN JOURNAL OF PHYSICS, Volume 61, Number *7-8*, 1354-1368.
3. Iordache Șt., Dunea D., Bøhler T. (2014), *Current status of citizens' protection against the risk of air pollution in urban areas*, (in Romanian), Șt. Iordache and D. Dunea (eds.), Methods to assess the effects of air pollution with particulate matter on children's health (in Romanian), MatrixRom, 1-44.
4. Iordache Șt., Dunea D., Lungu E., Predescu L., Dumitru D., Ianache C., Ianache R. (2015), *A Cyberinfrastructure for Air Quality Monitoring and Early Warnings to Predict Children with Respiratory Disorders*, Proceedings of the International Conference CSCS20, Bucharest, 789-796.
5. Oprea M., Ianache C., Mihalache S.F., Dragomir E.G., Dunea D., Iordache Șt., Savu T. (2015), *On the Development of an Intelligent System for Particulate Matter Air Pollution Monitoring, Analysis and Forecasting in Urban Regions*, Proceedings of the 19th Int. Conf. ICSTCC, Cheile Grădiștei, 711-716.
6. *** (2016), *SAS Enterprise Miner 13.2 Software and Reference Help*.
7. Silva da Silva C., Rossato J.M., Rocha J.A.V., Vargas V. M. V. (2015), *Characterization of an area of reference for inhalable particulate matter (PM2.5) associated with genetic biomonitoring in children*, Mutation Research/Genetic Toxicology and Environmental Mutagenesis, Volume 778, 44-55.
8. Zeng X., Xu X., Zheng X., Reponen T., Chen A., Huo X. (2016), *Heavy metals in PM2.5 and in blood, and children's respiratory symptoms and asthma from an e-waste recycling area*, Environmental Pollution, Volume 210, 346-353.