

Effect of Environmental Factors on Air Ion Concentration

Andris Skromulis, Juris Breidaks

*Rezeknes Augstskola, Faculty of Engineering,
Address: Atbrivosanas aleja 115, Rezekne, LV-4601, Latvia*

Abstract. The paper discusses and analyzes the effect of environmental factors on ion concentrations in urban air. Statistically mathematical method was used to analyze the measurements of air ion concentration collected in a period of several months, with ion size ranging from 0.75 to 36.6 nm. As environmental factors related to air ions, the following chemical and physical parameters of the atmospheric air were analyzed: CO, NO, NO₂, NO_x, SO₂, O₃, PM10, PM2.5, temperature, relative humidity, wind speed and direction. When analyzing data in this combination, there is often a problem of multicollinearity between air chemical and physical parameters. This paper addresses this problem by using the component regression. Regression equations were elaborated to understand the dependence of concentrations of various classes of positive and negative air ions on chemical and physical parameters of the air.

Keywords: air ions, air pollution, multicollinearity, multivariate regression, Principal Component Regression (PCR).

I INTRODUCTION

Research on chemical and mechanical air pollution nowadays is one of the most topical problems in the environmental science. Pollution of the atmosphere is very difficult to localize or eliminate, it is able to spread not only on regional but also on global scale. A special attention is paid to a regular monitoring of the quality of urban air environment, constantly looking for new ways to improve the existing monitoring system. One of possibilities is the inclusion of natural air ionization parameters into the range of parameters to be monitored [1].

In the atmospheric physics, the term "air ion" signifies all airborne particles that are electrically charged and serve as a basis for air conductivity [2, 3]. There are several air ion classifications depending on their size, however the classification that is becoming more popular is based on the physical composition of air ions [4]. Cluster ions consist only of ionized and polarized gas molecules that form clusters, but the heaviest ions are aerosol ions consisting not only of ionized and polarized gas molecules, but also of atmospheric aerosol particles as ion condensation nuclei ion.

Almost for a century, various scientific studies have been revealing and describing the impact of air ions on human beings and all living organisms [5, 6]. Ionized molecules of oxygen and other air components have a stronger effect on the human body. Air ions may be either healthy or harmful to the human health depending on their concentration levels

in the air and on the proportion of the positive and negative ions.

Over the past two decades, outdoor studies of air ions have become topical aimed at understanding of interactions of air ions and other natural biotic and abiotic components, which affect not only ecosystems, but also the global climate [7]. There is still a need for relevant and credible information about the influence of anthropogenic pollution on air ions both in urban and rural environment.

The mass spectrometer has allowed to identify different substances in the ambient air, such as NH₃, HNO₃, H₂SO₄, which are essential for the formation of atmospheric aerosols and heavy air ions (for example [8, 9, 10]). The spectrometry of cluster ion mobility is widely used in detection of various impurities in the atmospheric air [11, 12]; the most important substance classes of these impurities are the following: exhaust gases, poisonous chemical gases, military and explosive gases. The total air pollution also affects the mobility and size of air ions [13, 14].

Balance of air ions in the atmosphere depends on various and rather complex ion generation and ion loss physical mechanisms, therefore outdoor studies of air ions is a difficult, but necessary step to acquire new knowledge on atmospheric pollution and its interaction with the ecosystem. The results are often difficult to interpret.

Serious problem for analysis of air pollution and interaction of air ions is caused by multicollinearity of parameters (see the Figure 1). It is particularly important when dealing with the urban environment

that is characterized by a high level of anthropogenic pollution. Many atmospheric chemical and physical contamination components have a common source such as industrial emissions, transport, energy, etc., therefore there is a strong correlation among them, which causes problems in the classical regression analysis. The study aims to determine and apply a mathematical model of regression analysis that would minimize the negative impact of the multicollinearity on interpretation of the results and help to explain the concentration of air ions by changes of environmental factors.

II MATERIALS AND METHODS

In order to assess the impact of the anthropogenic air pollution on air ion concentrations, measurements of chemical air pollution, meteorological parameters and air ion concentrations are needed. Further analysis was based on data derived from the air monitoring station in Estonia (Tartu, Karlova residential district). Concentration of naturally charged air ions according to their mobility was measured using the Neutral Cluster and Air Ions Spectrometer "NAIS" [20]. Devices for measuring meteorological and chemical parameters are certified.

Measurements were carried out in different seasons in a period of less than a year. Average half-hour values of the measured parameters were analyzed. In order to obtain an overall picture of the effects of atmospheric pollution on basic air ion classes, all air ion sub-classes measured with the NAIS were summed up in the following ion classes with the following ion size ranges:

1. Cluster ions (0.75 – 1.54 nm)
2. Middle ions (1.54 – 7.50 nm)
3. Heavy ions (7.50 – 36.6 nm)
4. Total range of ions (0.75 – 36.6 nm).

Overall, 13756 records were subjected to statistical analysis.

Due to the strongly expressed multicollinearity among the measured parameters, it was not possible to apply the classical regression method of analysis. As an adequate solution to this problem the method of Multivariate Regression was chosen. The purpose of the principal component regression (PCR) is to estimate the values of a response variable at the basis of selected principal components (PCs) of the explanatory variables. There are two main reasons for regressing the response variable on the PCs rather than directly on the explanatory variables. Firstly, the explanatory variables are often highly correlated (multicollinearity) which may cause inaccurate estimations of the least squares (LS) regression coefficients. This can be avoided by using the PCs in place of the original variables since the PCs are uncorrelated. Secondly, the dimensionality of the regressors is reduced by taking only a subset of PCs for prediction.

Multivariate regression methods, such as principal component regression (PCR), enjoy are very popular in a wide range of fields, including the natural sciences. The main reason is that they have been designed to confront the situation that there are many correlated, predictor variables, and relatively few samples - a situation that is common, especially in chemistry where developments in spectroscopy since the seventies have revolutionised chemical analysis.

In the context of usual multiple linear regression (MLR), the least-squares solution for

$$Y = XB + \varepsilon \tag{1}$$

is given by

$$B = (X^T X)^{-1} X^T Y \tag{2}$$

Often the problem is that $X^T X$ is singular, either because the number of variables (columns) in X exceeds the number of objects (rows), or because of collinearities. PCR circumvent this by decomposing X into orthogonal scores T and loadings P

$$X = TP \tag{3}$$

and regressing Y not on X itself but on the first columns of the scores T . In PCR, the scores are given by the left singular vectors of X , multiplied with the corresponding singular values, and the loadings are the right singular vectors of X . This, however, only takes into account information about X , and therefore may be suboptimal for prediction purposes.

In PCR, we approximate the X matrix by the first a principal components (PCs), usually obtained from the singular value decomposition (SVD):

$$\begin{aligned} X &= \tilde{X}_{(a)} + \varepsilon_X = (U_{(a)} D_{(a)}) V_{(a)}^T + \varepsilon_X = \\ &= T_{(a)} P_{(a)}^T + \varepsilon_X \end{aligned} \tag{4}$$

Then we regress Y on the scores, which leads to regression coefficients

$$B = P(T^T T)^{-1} T^T Y = V D^{-1} U^T Y \tag{5}$$

where the subscripts a have been dropped for clarity.

Choosing the number of extracted factors, one subset (the training set) is used to fit the model, and the other subset (the test set) is offered to measure how well the model fits. The number of factor chosen is usually the one which shall minimize the square root of the mean squared error of prediction (RMSEP) [15, 16, 17].

$$e_i = y_i - \hat{y}_i \tag{6}$$

$$RMSEP = \sqrt{\frac{1}{n} \sum_{i=1}^n y_i^2} \tag{7}$$

The R^2 values returned by "R2" are calculated as

$$R^2 = 1 - \frac{SSE}{SST}, \tag{8}$$

where SST is the (corrected) total sum of squares of the response, and SSE is the sum of squared errors for either the fitted values (i.e., the residual sum of squares), test set predictions or cross-validated predictions [17, 18, 19].

III RESULTS AND DISCUSSION

In general, the air ion concentration correlates with air pollution and meteorological conditions (See Fig. 1. Correlation coefficients above 0.3 are colored light gray). For different air ion classes, this correlation is manifested in different ways. For example, correlation of cluster ions with the chemical air pollution is relatively weak, while correlation with the temperature is moderately strong. In the area of

middle air ions this correlation is even weaker. It would be necessary to mention the nature itself of the middle air ions, i.e. they constitute a transition from cluster ions to heavy ions or aerosol ions, and their concentration, under natural conditions, is usually very low, perhaps only a few or a few dozen ions per 1cm³ of air. This is the fact which leads to a relatively weak correlation. In the area of heavy ions (especially positive ones), there is a strong correlation with the chemical and physical (PM 10) air pollution. Heavy ions contain aerosol particles as condensation nuclei. Not only air ions can stick to neutral aerosol particles, but air-polluting gas molecules can also absorb on them.

	CO	NO	NO	NOx	O3	PM10	PM2,5	SO2	Rel. Hum.	Temp.	Wind Dir.	Wind Sp.	Cluster "-"	Cluster "+"	Middle "-"	Middle "+"	Heavy "-"	Heavy "+"	Total "-"	Total "+"	
CO, mg/m3	1																				
NO, ug/m3	0,75	1																			
NO2, ug/m3	0,83	0,66	1																		
Nox, ug/m3	0,84	0,96	0,84	1																	
O3, ug/m3	-0,52	-0,38	-0,64	-0,50	1																
PM10, ug/m3	0,48	0,35	0,48	0,43	-0,24	1															
PM2,5, ug/m3	0,15	0,09	0,15	0,12	-0,12	0,20	1														
SO2, ug/m3	0,41	0,31	0,40	0,37	-0,21	0,29	0,09	1													
Rel. Humidity, %	0,05	-0,01	0,05	0,01	-0,55	-0,20	-0,01	-0,07	1												
Temperature, °C	-0,44	-0,33	-0,45	-0,40	0,55	-0,12	-0,12	-0,39	-0,36	1											
Wind Dir., deg.	-0,15	-0,12	-0,15	-0,14	0,12	-0,21	-0,07	-0,11	0,06	0,05	1										
Wind Speed, m/s	-0,24	-0,18	-0,33	-0,25	0,32	-0,23	-0,02	-0,07	-0,08	0,01	0,03	1									
Cluster ions "-"	-0,13	-0,07	-0,16	-0,10	0,11	-0,02	-0,03	-0,04	-0,10	0,32	0,02	-0,03	1								
Cluster ions "+"	-0,11	0,01	-0,17	-0,05	0,10	0,00	-0,03	-0,11	-0,10	0,49	0,01	-0,09	0,80	1							
Middle ions "-"	-0,08	-0,02	-0,07	-0,04	0,15	-0,03	-0,03	-0,02	-0,20	0,09	0,03	0,12	0,18	0,05	1						
Middle ions "+"	-0,02	0,04	-0,02	0,02	0,12	0,02	-0,02	0,00	-0,23	0,08	0,03	0,09	0,11	0,08	0,41	1					
Heavy ions "-"	0,16	0,19	0,18	0,20	-0,09	0,15	0,03	0,07	-0,06	0,02	-0,06	-0,08	0,08	0,10	0,14	0,10	1				
Heavy ions "+"	0,48	0,49	0,49	0,53	-0,27	0,33	0,08	0,25	-0,09	-0,21	-0,09	-0,15	0,13	0,09	0,11	0,21	0,41	1			
Ions (Total) "-"	0,15	0,18	0,17	0,19	-0,08	0,14	0,02	0,07	-0,08	0,03	-0,05	-0,08	0,13	0,13	0,20	0,12	1,00	0,41	1		
Ions (Total) "+"	0,45	0,48	0,46	0,51	-0,24	0,32	0,08	0,23	-0,11	-0,16	-0,08	-0,15	0,19	0,17	0,15	0,31	0,41	0,99	0,42	1	

Fig. 1. Matrix of correlation of data used in calculations

As shown in Fig. 1, among different air ion classes there are different correlations. The correlation matrix demonstrates the above-mentioned problem of multicollinearity: correlation of the components of chemical and physical air pollution is strong or moderately strong.

The principal component regression was used to solve the problem of multicollinearity without excluding any of parameters from the analysis.

For further analysis we will use a number of principal components where R² reaches its highest level, but the RMSEP is at the lowest level (See table 2 and table 3).

Regression equation coefficients obtained with a help of the principal component regression were

summarized in the table 1. The sign and value of the coefficient show the influence of each variable on the concentration of air ions of each specific class.

Regression equation highlights the profound impact of CO on the air ion concentration which increases when air ions become bigger. In the area of cluster ions, the impact is lower, while in the area of heavy ions there is the highest impact, which also influences the air ion concentration in general. Both in the area of cluster ions and heavy ions, the presence of CO reduces the concentration of negative ions and increases the concentration of positive ones. This effect could explain the lack of negative air ions which is often observed in a polluted atmospheric environment.

TABLE 1
MULTIVARIATE REGRESSION COEFFICIENTS (ROUNDED)

Parameter / Ion Class	Cluster Ions		Middle Ions		Heavy Ions		Ions (Total)	
	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive
CO	-2.67	6.64	-24.61	-22.74	-164.92	58.20	-192.20	42.10
NO	0.94	6.58	-2.68	-2.84	3.73	6.96	1.99	10.70
NO ₂	-0.29	2.06	-1.77	-2.02	5.87	8.38	3.81	8.42
NO _x	-0.45	-3.72	1.89	2.11	3.30	2.90	4.75	1.28
O ₃	-0.43	-1.01	-0.05	-0.23	-3.24	-2.14	-3.72	-3.38
PM10	0.03	0.10	-0.10	-0.07	3.46	3.83	3.40	3.86
PM2,5	0.01	0.03	-0.04	-0.04	-0.06	-0.05	-0.08	-0.06
SO ₂	3.31	4.04	-0.14	-0.61	10.71	11.06	13.88	14.48
Rel. Humidity	-0.14	-0.14	-0.86	-1.29	-3.16	-5.82	-4.16	-7.25
Temp.	2.08	4.61	0.06	0.12	17.32	-0.49	19.46	4.23
Wind Direction	0.01	0.01	0.03	0.04	-0.23	0.11	-0.19	0.16
Wind Speed	-1.57	-3.68	9.45	9.33	-10.64	1.02	-2.76	6.67

TABLE 2
R² VALUES OF PRC REGRESSION FOR PRINCIPAL COMPONENTS

Number of principal components.	R ² , Cluster Ions		R ² , Middle Ions		R ² , Heavy Ions		R ² , Ions (Total)	
	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive
1	0.0005702	0.0002741	0.001516	0.0006379	0.005258	0.01706	0.004701	0.01502
2	0.0134813	0.0053189	0.004933	0.0006494	0.038530	0.27499	0.034487	0.25070
3	0.0136028	0.0054994	0.005181	0.0011777	0.040145	0.28353	0.036093	0.25925
4	0.0233351	0.0242611	0.036738	0.0414312	0.044579	0.29319	0.042324	0.27463
5	0.0235593	0.0242821	0.036755	0.0418616	0.047670	0.30272	0.045371	0.28382
6	0.0241145	0.0301941	0.044522	0.0549076	0.048119	0.30615	0.045983	0.28758
7	0.1154693	0.3157421	0.045134	0.0572499	0.054510	0.30637	0.054367	0.28854
8	0.1218842	0.3364217	0.047814	0.0611757	0.058017	0.30637	0.057809	0.28855
9	0.1359462	0.3497774	0.047884	0.0614623	0.058209	0.30688	0.058121	0.28939
10	0.1363843	0.3510621	0.055531	0.0671997	0.058244	0.30688	0.058125	0.28942
11	0.1364403	0.3512581	0.057873	0.0687347	0.058552	0.30697	0.058533	0.28947
12	0.1364405	0.3512681	0.057875	0.0687366	0.058552	0.30697	0.058533	0.28947

TABLE 3
RMSEP VALUES OF PRC REGRESSION FOR PRINCIPAL COMPONENTS

Number of principal components.	RMSEP, Cluster Ions		RMSEP, Middle Ions		RMSEP, Heavy Ions		RMSEP, Ions (Total)	
	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive
1	50.80	67.77	72.45	82.74	1331	852.4	1350	883.3
2	50.47	67.60	72.33	82.74	1309	732.0	1329	770.4
3	50.47	67.59	72.32	82.72	1308	727.7	1328	766.0
4	50.22	66.95	71.16	81.04	1305	722.8	1324	758.0
5	50.21	66.95	71.16	81.02	1303	717.9	1322	753.2
6	50.20	66.74	70.87	80.46	1302	716.1	1321	751.2
7	47.79	56.06	70.85	80.37	1298	716.0	1316	750.7
8	47.62	55.21	70.75	80.20	1296	716.0	1313	750.7
9	47.24	54.65	70.75	80.19	1295	715.8	1313	750.3
10	47.22	54.60	70.46	79.94	1295	715.8	1313	750.3
11	47.22	54.59	70.38	79.87	1295	715.7	1313	750.2
12	47.22	54.59	70.38	79.87	1295	715.7	1313	750.2

In general, nitrogen oxides show a slightly weaker trend. They are cause a moderate increase of the number of heavy air ions, but it should be noted that NO and NO₂ slightly increases the concentration of positive cluster ions. NO_x is the sum of these nitrogen oxides, but the regression analysis shows that the total effect is ambiguous - NO_x slightly reduces the concentration of cluster ions.

The impact of ozone is not very strong, but it should be noted that in all air ion classes it manifests itself negatively. It is considered that O₃ as a strong

oxidant can quickly reduce the number of air ions, but this effect is weak in low ozone concentrations in the atmospheric boundary layer under normal conditions.

SO₂ is associated with a moderate increase in air ion concentrations, except middle-sized ions.

Impact of dustiness on air ions manifests itself mainly in the increase of the number of heavy or aerosol ions, or in increase in the number of aerosols, because aerosols form a part of large ions as nuclei. Contrary to expectations that aerosols would cause decrease of cluster ion concentrations because, when a

cluster ion collides with an aerosol, it gives its charge to this aerosol thus forming a large ion, relations of dustiness and air ions were found to be almost neutral.

Relative humidity of the air, which is basically the relative content of water vapor in the air, influences negatively the air ion concentration. Water vapor molecules are good condensation nuclei in the atmosphere, and air ions, when connecting with these molecules, give away their charge and cease to exist. Besides, the bigger air ion, the greater possibility for it to come into contact with a water molecule in the air.

Temperature slightly increases the concentration of all air ions, because, when the temperature rises, the intensity of the movement of molecules in the air increases as well, which contributes to mutual collisions of molecules and ions, resulting in possible recombination of ions with opposite signs or in formation of new ions due to such collisions.

Similarly, when the wind speed increases, it enhances the mixing of atmospheric layers and spread of pollution, which may result in decrease of concentration of individual ion groups. On the other hand, the wind also contributes to the raise of radioactive isotopes above the earth's surface at a height of several meters, where they serve as atmospheric ionizers.

Wind direction in the city is important, as it determines whether polluted air from industrial districts and transport nodes, or clean and fresh air from suburban areas, seaside, etc. shall be brought to the monitoring point. For precise interpretation of the impact of wind direction on air ions, location of pollution emission points in each specific city should be analyzed.

IV CONCLUSION

By using the principal component regression, a positive result has been achieved, i.e. the problem of multicollinearity has been solved, which was the main problem in analyzing such type of data. A significant advantage of this method is that all data can be analyzed without excluding any of them because of multicollinearity. When using the classical regression, the excluded data may turn out to be significant.

The linear regression applied to value R^2 is not very high, only in 2 cases it exceeds 0.3, the RMSEP is also relatively high. It means that the data contain errors, and the selected 12 variables are not sufficient to explain air ion concentration and its fluctuations.

In further studies, it would be necessary to expand the range of the measured parameters to find new factors that affect air ion concentration significantly. In further statistical analysis, it would be helpful to use non linear regression and method of cross-validation.

Basing on the obtained regression coefficients of the regression equation, the following conclusions can be made. Anthropogenic atmospheric pollution is one

of the factors influencing the air ion concentration, charge and size.

In general, there is a trend that air pollution contributes to increase of concentration of positively charged air ions; concentration of heavy air ions or aerosol ions raises as well. This effect creates environment unfavorable for human health with increased air pollution, because the human health needs sufficient concentration of negative cluster ions in the inhaled air [21].

Air ion concentration, charge, chemical composition and structure are important factors for assessing the air quality. Their role and interaction with other components of the atmospheric air has not been adequately studied yet and requires further researches in the future.

V REFERENCES

- [1] Noviks G., Skromulis A. Gaisa vides kvalitātes monitoringa pilnveidošanas iespēju analīze. In: Environment. Technology. Resources: Proceedings of the 7th International Scientific and Practical Conference. Volume 1. Rezekne Higher Education Institution, Rezekne, 2009, pp. 273-283.
- [2] Dolezalek H., Reiter R., Kröling P. Basic comments on the physics, occurrence in the atmosphere, and possible biological effects of air ions. *Int. J. Biometeorol.*, No. 29.1985, pp. 207-242.
- [3] Tammet H. Air ions. In: *CRC Handbook of Chemistry and Physics*, 79th edition, No. 14, Boca Raton, Ann Arbor, London, Tokyo: CRC Press, 1998, pp. 32-34.
- [4] Hörrak U., Salm J., Tammet H. Statistical characterization of air ion mobility spectra at Tahkuse Observatory: classification of air ions. *J. Geophys. Res. Atmos.*, No. 105. 2000, pp. 9291-9302.
- [5] Чижевский А. Л. Аэроионификация в народном хозяйстве. Москва: Госпланиздат, 1969, 185 с.
- [6] Rim Y. Psychological Test Performance of Different Personality Types on Sharav Days in Artificial Air Ionisation. *Int. J. Biometeor.*, 1977, vol. 21, number 4, pp 337 – 340.
- [7] Hari P., Kulmala M. Station for Measuring Ecosystem – Atmosphere Relations (SMEAR II) Boreal Environment Research 10. 2005, pp. 315-322.
- [8] Tanner, D.J., Eisele, F.L. Ions in oceanic and continental air masses. *Journal of Geophysical Research* 96. 1991.
- [9] Eisele, F.L., and D.J. Tanner, Measurement of the gas phase concentration of H₂SO₄ and methane sulfonic acid and estimates of H₂SO₄ production and loss in the atmosphere, *J. Geophys. Res.*, 98, 1993, pp. 9001-9010.
- [10] Ziereis H., Arnold F. Gaseous ammonia and ammonium ions in the free troposphere. *Nature* 321. 1986, pp. 503 – 505.
- [11] Carr, T.W. *Plasma Chromatography*. Plenum, New York. 1984, p 259.
- [12] Eiceman, G.A., Karpas, Z. *Ion Mobility Spectrometry*. CRC Press, Boca Raton. 1994, p 288.
- [13] Hoppel, W.A. Measurement of the mobility distribution of tropospheric ions. *Pageoph*. 81. 1970, pp. 230-245.
- [14] Parts, T., Salm, J. The effect of pyridine and its homologues on mobility spectra of positive small air ions. *Acta et Comm. Univ. Tartuensis* 947. 1992, pp. 24-30.
- [15] Ying Li, A comparison study of principle component regression, partial least squares regression and ridge regression with application to FTIR data, Master thesis in statistics, Faculty of Social Sciences, Uppsala Universitet. Sweden, June 2010.
- [16] Jolliffe, I. T. *Principal component analysis*, Spring-Vlg, New-York, 1986.

- [17] Mevik, B.-H., Wehrens, R., Liland, K. H., Partial Least Squares and Principal Component regression; R package version 2.4-3, 2013., <http://cran.r-project.org/web/packages/pls/>
- [18] Mevik, B.-H., Wehrens, R., The pls Package: Principal Component and Partial Least Squares Regression R. *Journal of Statistical Software* 18(2). 2007, pp.1-24.
- [19] Mevik, B.-H., Cederkvist, H. R. Mean Squared Error of Prediction (MSEP) Estimates for Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR). *Journal of Chemometrics*, 18(9). 2004, pp. 422–429.
- [20] Mirme A., Tamm E., Mordas G., Vana M., Uin J., Mirme S., Bernotas T., Laakso L., Kirsikko A., Kulamala M. A wide-range multi-channel air ion spectrometer. *Boreal. Env. Res.* 12. 2007, pp.247 – 264.
- Krueger A. P., Struble A. E., Yost M. G., Reed E. I. Electric fields, small ions and biological effects. *Intern. Jour. Biomet.*, Vol. 22, No. 3. 1978., pp. 202-212.