AIR POLLUTION RELATED OBJECTIVE CLASSIFICATION OF AIR MASS TYPES FOR SZEGED, HUNGARY

László Makra[1], János Mika[2], Rita Béczi[1], Emőke Borsos[1], Zoltán Sümeghy[1], András Gál[3] and Béla Vitányi[3]

[1]Department of Climatology and Landscape Ecology, University of Szeged; Hungary;
[2]Hungarian Meteorological Service, Budapest, Hungary; e-mail: mika.j@met.hu;
[3]István Bocskai Secondary School and Economic Technical School, Szerencs, Hungary

Abstract

The paper determines characteristic air mass types over the Carpathian Basin for the winter (December, January, and February) and summer ((June, July and August) months with the levels of the main air pollutants. Based on the ECMWF data set, daily sea level pressure fields measured at 00 UTC were prepared for each air mass type (cluster) in order to relate sea level pressure patterns with the levels of air pollutants in Szeged. The data basis comprises daily values of twelve meteorological and eight pollutant parameters for the period 1997-2001. Objective definition of the characteristic air mass types occurred by using the methods of Factor Analysis and Cluster Analysis.

1. Introduction

Air quality, concentration of air pollutants are influenced not only by physical and chemical processes but also by meteorological processes as well as geographical and social factors. Some weather conditions, like moderate wind condition or calm air condition with temperature inversion, caused by an anticyclone, can contribute to extreme accumulations of air pollutants.

The aim of the study is to determine objectively defined characteristic air mass types for Szeged in the summer and winter months using multivariate statistical methods. Afterwards, for each air mass type having homogenous temperature and humidity parameters, concentrations of the main air pollutants are estimated. Then, mean sea level pressure fields of each air mass type for the North-Atlantic – European region are constructed in order to connect spatial distribution of the sea level pressure fields and air pollution of the Szeged region. Up to now, few studies have been published to analyse relations of the current weather and concentrations of air pollutants for Szeged. The study represents an objective weather classification, which might be the base of an air pollution observing/forecasting system, in order to avoid developing severe air pollution episodes in Szeged.

2. Data basis

The data basis consists of a 30-minute data set from the five-year period between 1997 – 2001 for the winter (December, January and February) and summer (June, July and August) months. The elements considered are, on the one hand, the average mass concentrations of the main air pollutants [CO, NO, NO_2, NO_x, SO_2, O_3 and TSP ($\mu g\ m^{-3}$)]; on the other hand, the daily values of the main climatic elements (temperature, humidity, air pressure, global radiation, wind direction, wind speed).The 12 meteorological parameters used are: mean temperature (T_{mean}, °C), maximum temperature (T_{max}, °C), minimum temperature (T_{min}, °C), daily temperature range ($\Delta T = T_{max} - T_{min}$, °C), wind speed (WS, $m\ s^{-1}$), relative humidity (RH, %), irradiance (I, $MJ\ m^{-2}\ day^{-1}$), saturation vapour pressure (E, mm Hg), water vapour pressure (VP, mm Hg), potential evaporation (PE, mm), dew point temperature (T_d, °C) and atmospheric pressure (P, mm Hg). The 8 pollution parameters considered are the average diurnal mass concentrations of the following pollutants: CO ($\mu g\ m^{-3}$); NO ($\mu g\ m^{-3}$), NO_2 ($\mu g\ m^{-3}$), SO_2 ($\mu g\ m^{-3}$), O_3 ($\mu g\ m^{-3}$) and TSP ($\mu g\ m^{-3}$) as well as the daily ratios of NO_2/NO and the daily maximum concentrations of O_3 ($\mu g\ m^{-3}$).

3. Methods

In order to reduce the dimensionality of the above meteorological data sets and thus to explain the relations among the 12 meteorological variables, the multivariate statistical method of factor analysis is used. Factor analysis was applied on the data set consisting of 12 columns (12 meteorological variables) and 450 rows (days) both for the winter and summer months.

The most commonly used technique for searching natural structure among observations is cluster analysis. This method groups objects into clusters so that objects in the same cluster are more similar to one another than they are to objects in other clusters. The aim is to maximize the homogeneity of objects within the clusters and also to maximize the heterogeneity between the clusters. Hence, the resulting clusters show then high internal (within cluster) homogeneity and, at the same time, high external (between cluster) heterogeneity.

When determining the synoptic types, only meteorological parameters are taken into account, excluding pollution data. Hence, the differences of the calculated mean pollution levels among the resultant synoptic types need a further statistical evaluation. This is performed by the method of one-way Analysis of Variance (ANOVA) for each pollutant. The Tukey's honestly significant difference test is applied in order to quantitatively compare the mean pollution levels between each pair of synoptic type (pairwise multiple comparisons) (McGregor and Bamzelis, 1995; Sindosi et al., 2003).

All statistical computations were performed with SPSS (version 9.0) software.

4. Results

4.2. Winter months

The application of Factor Analysis to the time series of the meteorological elements yielded 4 Factors explaining 86.51 % of the total variance. Table 1 displays the factor loadings after orthogonal rotation. Values higher than $|0.10|$ are statistically significant at the 5 % level; however, Table 1 shows only those exceeding $|0.60|$. This means that at least 36 % of the total variance of a parameter can be explained by a single Factor.

Table 1. Factor loadings for the winter months (December, January and February). Values higher than $|0.60|$ are only presented.

Meteorological parameters	Factor 1	Factor 2	Factor 3	Factor 4
Mean temperature, T_{mean}	0.94			
Maximum temperature, T_{max}	0.81			
Minimum temperature, T_{min}	0.84			
Daily temperature range, $\Delta T = T_{max} - T_{min}$			0.87	
Wind speed, WS				
Relative humidity, RH		-0.89		
Irradiance, I			0.74	
Saturation vapour pressure, E	0.93			
Water vapour pressure, VP	0.97			
Potential evaporation, PE		0.74		
Dew point temperature, T_d	0.97			
Atmospheric pressure, P				0.96

Factor 1 explains 50.86 % of the total variance and includes the three main air temperature variables (mean, maximum and minimum temperatures) and three important humidity parameters (saturation vapour pressure, water vapour pressure and dew point temperature).

Factor 2 (19.85 % of the total variance) includes only relative humidity with a negative sign and potential evaporation. The high loadings of opposite sign indicate an inverse relationship between the two variables. Namely, high potential evaporation is associated with low relative humidity and vice versa.

Factor 3 explains 8.72 % of the total variance and comprises daily temperature range and irradiance.

Factor 4 is slightly weaker than *Factor 3* and explains 7.08 % of the total variance. It comprises atmospheric pressure only.

Next, Cluster Analysis was applied to the four factor score time series. The resultant five dominant clusters as air mass types and the corresponding pressure patterns with the associated pollution levels are described as follows.

Cluster 1 This can be named "anticyclone over the Carpathian Basin". This pressure pattern is characterized by a high pressure system over Central Europe. This cluster (air mass type) amounts to 12.5 % of the total number of days. During such weather conditions primary pollutants (CO, NO_2, SO_2 and TSP, except NO) are highly concentrated in the city, as a consequence of poor ventilation and temperature inversions formed during the night (Horváth et al., 2002). Under the prevalence of this air mass type, in accordance with the high amounts of the total irradiance, concentration of the secondary pollutants (O_3 and O_{3max}) is relatively high.

Cluster 2 This type is named "anticyclone over the Mediterranean". This is the most frequent cluster with 30.0 % of the total number of days. This pressure pattern forms an anti-cyclonic ridge over the Carpathian Basin with calm or slight breezes. Irradiance reduced to the half and the temperature parameters are significantly higher during this air mass type, comparing to those values in *Cluster 1*. The ozone levels are lower because of the lower irradiance. The reason of the lower concentration in primary pollutants might be the higher wind speed.

Cluster 3 Anticyclone stretches out from the region of Azores. This situation is only characteristic in February. During prevalence of this type an anticyclone can reach Central and even Eastern Europe causing undisturbed irradiance with calm weather. This situation causes high temperatures and high winds. Lower concentrations of CO, SO_2 and TSP comparing to those in *Cluster 1* can be attributed to the difference of these mean cluster fields in the prevailing winds. Since average concentration of NO during the *Clusters 1* and *3* are the same, higher enrichment of the ozone parameters in *Cluster 3* relating to that in *Cluster 1* can be explained with the higher irradiance.

Cluster 4 Anticyclone over Southern Europe and North Africa. This cluster does not differ significantly from *Cluster 3*. In this situation the high pressure system from SW Europe extends over the Eastern part of the Mediterranean. Because of the very low wind speeds levels concentrations of primary pollutants (except SO_2) are extremely high. At the same time, irradiance is also high. However, concentration of the ozone parameters is not high due to the highest levels of NO, implying the lowest levels of NO_2/NO, which inhibit the formation of O_3 via the destruction process: $NO + O_3 \longrightarrow NO_2 + O_2$.

Cluster 5 Intense zonal current over Europe. This air mass type amounts to 20.1 % of the total number of days and is mostly frequent in December. During this situation Szeged is characterised by strong winds. Its pressure pattern corresponds to a zonal current over the Carpathian Basin, which involves fairly low levels of the primary pollutants especially those of SO_2 and TSP with their lowest concentrations. On the other hand, the lowest irradiance (as in *Cluster 2*) with a medium level of NO result in the lowest concentrations of the ozone parameters.

In order to determine if significant inter – air mass type differences in pollutant levels occur, analyses of variance (ANOVA) were performed on the pollutant parameters. The results are shown in Table 2. It can be established that, except NO_2vsNO, all the pollutants present significant inter – air mass type differences in mean concentration values at the 0.01 % probability level. Considering that differences are found among the mean pollutant levels, Tukey's honestly significant tests were applied in order to receive a pairwise multiple assessment of the differences. The statistically significant differences are shown in Table 3 at the 0.01 and 0.05 % probability levels, respectively. It can be seen that the pairs of air mass types (clusters) 3-4 differ significantly for five pollutants, while the types 1-2, 1-5 and 2-3 differ substantially for four pollutants. Generally, clusters 3-4 are considered to be the mostly different, since levels of the most pollutant pairs show substantial difference for them. This can mainly be explained by the fact that these two types show almost the highest difference in wind speed. On the other hand, *Cluster 2* seems to be an intermediate cluster considering pollution, since it shows fewer pairwise differences than the others. An exception to this is NO_2 with 3 out of 4 pairs of *Cluster 2* being different. The pairwise multiple comparisons between *Clusters 2* and *5* did not detect significant difference in any pollutant.

Table 2. ANOVA statistics for inter – air mass comparison of pollutant concentrations in the winter months (December, January and February)

	CO	NO	NO_2	NO_2vsNO	O_3	O_{3max}	SO_2	TSP
Mean square between groups	1516531.41	8183.16	2361.19	305.70	2057.77	6255.12	332.56	4971.82
Mean square within groups	137957.12	585.10	257.83	212.15	186.97	464.40	65.23	534.98
F-Ratio	10.99	13.99	9.16	1.44	11.01	13.47	5.10	9.29
Level of significance	0.01	0.01	0.01	0.22	0.01	0.01	0.01	0.01

Table 3. Air mass type – pollution difference matrix. Each matrix cell represents the comparison between two air mass types. Pollutants appearing in the matrix cells indicate significant inter – air mass difference in concentrations according to Tukey's honestly significant difference test (**bold**: 0.01 level of significance, *italics*: 0.05 level of significance), winter months (December, January and February)

	1	2	3	4
2	NO_2 **O_{3max}** **SO_2** **TSP**			
3	**CO** SO_2 TSP	**CO** **O_{3max}** NO_2 **O_3**		
4	**NO** SO_2	**NO** **NO_2** **TSP**	**CO** **NO** **O_{3max}** **O_3** TSP	
5	O_{3max} **SO_2** O_3 **TSP**		**CO** **O_{3max}** **O_3**	**NO** NO_2 **TSP**

4.3. Summer months

The application of Factor Analysis to the time series of the meteorological parameters resulted in 4 main factors explaining 84.36 % of the total variance. Table 4 shows factor loadings of the summer months after orthogonal rotation. Values higher than $|0.10|$ are statistically significant at the 5 % level; however, Table 4 shows only those exceeding $|0.60|$. This means that at least 36 % of the total variance of a parameter can be explained by a single Factor.

Factor 1, with 47.35 % of the total variance, includes the same parameters as in the winter case. These are partly temperature variables (mean, maximum and minimum temperatures) and partly humidity variables (saturation vapour pressure, water vapour pressure and dew point temperature).

Factor 2 (19.44 % of the total variance) comprises irradiance and potential evaporation, both with positive sign and relative humidity with negative sign. Increasing irradiance involves an increase in potential evaporation and a parallel decrease of relative humidity.

Table 4. Factor loadings for the summer months (June, July and August). Values higher than $|0.60|$ are only presented.

Meteorological parameters	Factor 1	Factor 2	Factor 3	Factor 4
Mean temperature, T_{mean}	0.88			
Maximum temperature, T_{max}	0.75			
Minimum temperature, T_{min}	0.82			
Daily temperature range, $\Delta T = T_{max} - T_{min}$				
Wind speed, WS				0.98
Relative humidity, RH		-0.94		
Irradiance, I		0.71		
Saturation vapour pressure, E	0.87			
Water vapour pressure, VP	0.95			
Potential evaporation, PE		0.80		
Dew point temperature, T_d	0.95			
Atmospheric pressure, P			0.92	

Factor 3 (8.86 % of the total variance) consists of only atmospheric pressure, itself.

Factor 4 (8.22 % of the total variance) is slightly weaker than *Factor 3* and contains only wind speed.

In the following, Cluster Analysis was applied to the four factor score time series. The analysis revealed ten clusters of days (air mass types) (each with at least 3.7 % of the total number of days). The summer season is characterised by much more air mass types compared to the winter period. Description of the mean sea level pressure distribution over the North Atlantic – European region and the variation of the number of days within the summer season are presented as follows.

Cluster 1 It comprises 6.1 % of the total number of days. It is characterised by a high pressure system extending over Europe except Scandinavia and includes the Carpathian Basin, too. At the same time the thermal low of SW Asia is also developed. In this situation air temperature is the lowest of all the ten clusters. The reason of it is that most of these days belong to June, early summer. During this situation levels of both the primary pollutants (CO, NO, NO_2, NO_2vsNO and TSP except SO_2) and the secondary ones (O_3 and O_{3max}) are the lowest.

Cluster 2 Early summer. In this weather type (8.7 % of the total number of days) the above pressure pattern is less characteristic, since both the high and the low pressures withdraw and weaken. Wind speed is the lowest of the all clusters. Concentrations of the pollutants increase, except SO_2. Levels of NO reach highest values in this type.

Cluster 3 Typical summer, with 12.6 % of the total number of days. Values of the meteorological elements represent a typical summer day in Szeged. During this type the high pressure system from Azores slightly withdraws, while the thermal low of SW Asia develops comparing to their position in the pressure pattern of *Cluster 2*. During this air mass type level of CO increases, while concentration of SO_2 drops.

Cluster 4 This is the most frequent type with 16.5 % of the total number of days and is characteristic for each summer month. Its pressure pattern is very similar to that of *Cluster 3*. The only basic difference is that the extensive low pressure system in Northern Europe withdraws in this cluster. Levels of CO decrease, while a substantial increase of irradiance causes only a slight increase in O_3 concentration. (NO levels practically do not change compared to those in *Cluster 3*.)

Cluster 5 Typical early summer with the lowest value of the total number of days (3.7 %). The high pressure system from Azores develops and extends over Eastern Europe avoiding the Carpathian Basin and, at the same time, a low pressure centre deepens over the North Atlantic. The daily temperature range is the highest, with low irradiance and medium wind speed. There are not substantial differences in concentrations of pollutants comparing to those of *Cluster 4*.

Cluster 6 Typical late summer (10.0 % of the total number of days). The high pressure system from Azores extends over Eastern Europe and, in this air mass type, includes the

Carpathian Basin. There are no weather fronts in Northern Europe. Irradiance is very high, which involves similarly high values of the temperature variables. On the other hand, wind speed is the lowest. Hence, primary pollutants are highly enriched. Though both irradiance and NO concentration (having opposite effect onto the levels of O_3 and O_{3max}) increased comparing to those in *Cluster 5,* higher weight of irradiance is indicated by resulting in slight increase in the levels of the secondary pollutants.

Cluster 7 This is the second most frequent type with 15.6 % of the total number of days. The high pressure centre from Azores withdraws and, at the same time, a low pressure pattern deepens over Northern Europe indicating a more characteristic pressure system than that of *Cluster 6*. However, both weather and pollutants' levels practically do not change considering the former cluster.

Cluster 8 This type occurs with the same frequency in each summer month (10.2 % of the total number of days). The high pressure centre from Azores strengthens and extends over Middle Europe, while the low pressure pattern in Northern Europe is divided into two parts: the Icelandic low and the Baltic low. The Carpathian Basin is under the influence of the Baltic low. Hence, irradiance drops which involves decrease of the temperature parameters and wind speed reaches its maximum of all clusters. This is why both the primary and the secondary pollutants' levels are very low.

Cluster 9 It consists of 11.1 % of the total number of days. Extension of the high pressure centre from Azores does not change, while Northern and Eastern Europe is covered by an extremely large and uniform low pressure pattern. The Carpathian Basin lies in the high pressure edge. As the weather situation between *Clusters 8* and *9* is extremely similar, it involves very little changes in meteorological parameters. Hence, no significant differences occur in levels of the pollutants.

Cluster 10 Typical and late summer with 5.4 % of the total number of days. The high pressure centre from Azores slightly weakens and withdraws. On the other hand, the low pressure pattern over Ukraine and Romania, presented in *Cluster 9*, disappears and a small high pressure centre forms on its place, whilst a largely extended low pressure centre develops over Northern Europe. The Carpathian Basin lies between the two high pressure centres ensuring undisturbed irradiance with very high temperatures and fairly moderate winds. This air mass type results in the highest levels of both primary and secondary pollutants, except SO_2.

Similarly to the winter months, significance of inter – air mass type differences in pollutant levels was determined by using ANOVA. The results are shown in Table 5. Mean concentrations of CO, NO, NO_2, O_3, O_{3max} and TSP present significant inter – air mass type differences at the 0.01 % probability level, while SO_2 at the 0.02 % and NO_2vsNO at the 0.04 % levels, respectively. Performing the pairwise comparisons (Tukey's honestly significant tests), the statistically significant differences are shown in Table 6 at the 0.01 and 0.05 % probability levels, respectively. There are no two air mass types for which inter – air mass type differences in concentrations of all the eight pollutants considered are significant. The highest inter – air mass type difference is indicated by five pollutants for the following comparisons: types 1-10, 6-8, 8-10 and 9-10. Air mass types 1-8, 1-9, 2-3, 2-4, 2-5, 3-4, 3-5, 4-5, 5-6, 5-7, 5-9, 6-7 and 8-9 are the most similar, considering that no significant differences in levels of any pollutants can be detected between them. The pairwise multiple comparisons indicated significant differences in concentrations of at least four pollutants for the following cluster pairs: types 1-6 (CO, NO_2, O_{3max} and TSP); 1-10 (CO, NO_2, O_3, O_{3max}, and TSP); types 2-10 (CO, O_3, O_{3max} and TSP); types 3-6 (NO_2, O_3, SO_2 and TSP); types 3-10 (CO, O_3, O_{3max} and TSP); types 4-6 (CO, NO_2, O_{3max} and TSP); types 4-10 (CO, O_3, O_{3max} and TSP); types 6-8 (CO, NO, NO_2, O_{3max}, and TSP); types 6-9 (CO, NO, NO_2 and TSP); types 8-10 (CO, NO_2, O_3, O_{3max} and TSP) and types 9-10 (CO, NO_2, O_3, O_{3max} and TSP). In general, air mass types 6 and 10 differ mostly from the others, since the pairwise multiple comparisons detected for them significant differences in levels of the most pollutants. Its reason might be the fact that these two types show a considerable difference in wind speed. At the same time, type 5 seems to be an intermediate cluster, since it shows fewer pairwise differences than the others.

Table 5. ANOVA statistics for inter – air mass comparison of pollutant concentrations in the summer months (June, July and August)

	CO	NO	NO_2	NO_2vsNO	O_3	O_{3max}	SO_2	TSP
Mean square between groups	332509.51	174.27	1178.53	1873.59	1465.91	4555.88	26.59	2732.57
Mean square within groups	21776.86	37.17	125.86	942.20	253.81	694.29	11.63	134.77
F-Ratio	15.27	4.69	9.36	1.99	5.78	6.56	2.28	20.28
Level of significance	0.01	0.01	0.01	0.04	0.01	0.01	0.02	0.01

5. Conclusions

This paper analyses levels of air pollutants in Szeged during characteristic sea level pressure patterns over the Carpathian Basin. Specific air mass types given by the pressure pattern mentioned both for the winter and summer months were found to play a significant role in the pollutants' concentrations in Szeged downtown. Results received for the winter months revealed that primary pollutants appear with higher concentrations when irradiance is high and wind speed is low (air mass types 1 and 4). This is the case when an anticyclone is found over the Carpathian Basin (*Cluster 1*) and when an anticyclone rules the region south from Hungary influencing weather of the country (*Cluster 4*). Low concentrations of primary pollutants are detected when Hungary lies under the influence of zonal currents (wind speed is the highest) (*Cluster 3*, a transitional type and *Cluster 5*). Pressure patterns for the summer months are not as characteristic as those in the winter, concerning the variability of the pressure fields and the magnitude of the gradients. This is mainly due to the predominance of the anticyclonic and anticyclonic ridge types. Due to high irradiance and very low NO concentrations, rather high levels of secondary pollutants are observed. It is to be noted that O_3 records exhibit about double concentrations than in the winter months.

Prediction of air mass types favours to prevent development of extreme concentrations. Importance of this kind of researches is due to the ever increasing health risk of air pollutants, the increasing morbidity and mortality in exposed regions: cities and industrial areas.

Acknowledgement

This study was supported by the Hungarian National Foundation for Scientific Research (OTKA No. T 034765).

References

Horváth, Sz., Makra, L. and Motika, G., 2002: *An objective assessment of the connection between meteorological elements and the concentrations of the main air pollutants at Szeged, Hungary*. AMS 4th Symposium on the Urban Environment and the 12th Joint Conference on the Applications of Air Pollution Meteorology with the Air and Waste Management Association. Norfolk, Virginia, USA, 20–24 May 2002. Proceedings, J4.3, J58-J59.

McGregor, G.R. and Bamzelis, D., 1995: Synoptic typing and its application to the investigation of weather – air pollution relationships, Birmingham, United Kingdom. *Theor. Appl. Climatol., 51, 223-236.*

Sindosi, O.A., Katsoulis, B.D. and Bartzokas, A., 2003: An objective definition of air mass types affecting Athens, Greece; the corresponding atmospheric pressure patterns and air pollution levels. *Environ. Technol., 24, 947-962.*

Table 6. Air mass type – pollution difference matrix. Each matrix cell represents the comparison between two air mass types. Pollutants appearing in the matrix cells indicate significant inter – air mass difference in concentrations according to Tukey's honestly significant difference test (**bold**: 0.01 level of significance, *italics*: 0.05 level of significance), summer months (June, July and August)

	1	2	3	4	5	6	7	8	9
2	CO **NO_2** TSP								
3	**CO** TSP								
4	CO **TSP**								
5	TSP								
6	**CO** **O_{3max}** **NO_2** **TSP**	**TSP**	NO_2 SO_2 **O_3** **TSP**	CO O_{3max} **NO_2** **TSP**					
7	**CO** **TSP**	**TSP**	**TSP**	**CO** **TSP**					
8		CO **NO** **NO_2**	**CO** **NO_2**	CO **NO_2** TSP	NO_2	**CO** **NO** **O_{3max}** **NO_2** **TSP**	**CO** **NO_2** **TSP**		
9		**CO** NO **NO_2**	**CO**	**CO** NO2vsNO		**CO** NO **NO_2** **TSP**	**CO** NO2vsNO **TSP**		
10	**CO** **O_{3max}** **NO_2**	CO **O_{3max}**	CO **O_{3max}**	**CO** **O_{3max}**	CO		O_{3max}	**CO** **O_{3max}** **NO_2**	**CO** **O_{3max}** **NO_2**