Meteorol Atmos Phys 92, 115–137 (2006)
DOI 10.1007/s00703-005-0143-x

Meteorology and Atmospheric Physics
Printed in Austria

[1] Department of Climatology and Landscape Ecology, University of Szeged, Hungary
[2] Hungarian Meteorological Service, Budapest, Hungary
[3] Department of Physics, Laboratory of Meteorology, University of Ioannina, Ioannina, Greece

An objective classification system of air mass types for Szeged, Hungary, with special interest in air pollution levels

L. Makra[1], **J. Mika**[2], **A. Bartzokas**[3], **R. Béczi**[1], **E. Borsos**[1], and **Z. Sümeghy**[1]

With 3 Figures

Received October 26, 2004; revised January 21, 2005; accepted March 18, 2005
Published online: September 15, 2005

Summary

This paper determines the characteristic air mass types over the Carpathian Basin for the winter (December, January, and February) and summer (June, July and August) months dependant on levels of the main air pollutants. Based on the ECMWF data set, daily sea-level pressure fields analysed at 00 UTC were prepared for each air mass type (cluster) in order to relate sea-level pressure patterns with the level of air pollutants in Szeged. The data comprise daily values of twelve meteorological and eight pollutant parameters for the period 1997–2001. Objective definition of the characteristic air mass types is achieved by using the methods of Factor Analysis and Cluster Analysis. According to the results, during the winter months five air mass types (clusters) were detected based on higher concentrations of primary pollutants that occur with high irradiance and low wind speed. This is the case when an anticyclone is found over the Carpathian Basin and over the region south of Hungary, influencing the weather of the country. Low levels of pollutants occur when zonal currents exert influence over Hungary. During the summer months anticyclones and anticyclone ridge situations are found over the Carpathian Basin. (During the prevalence of anticyclone ridge situations, the Carpathian Basin is found at the edge of a high pressure centre.) As a result of high irradiance and very low NO levels, secondary pollutants are highly enriched.

1. Introduction

Air pollution has become a very important environmental problem, mostly in crowded cities. Most human activities produce some kind of pollutants, which are progressively accumulated. Air pollution has negative effects not only on the surroundings of its source, but can also affect wider regions. Most air pollutants are released in relation to processes involving combustion. The emission sources include: transportation, fuel combustion, industrial processes, solid waste disposal, amongst others. These harmful particles in the air may damage human health, vegetation and the global environment. In general, in many cases they can form brown or hazy clouds at ground level with unpleasant smells. The damages to human health caused by air pollution generally result from repeated exposure to even low concentrations for long periods. Metals tend to corrode faster in polluted environments. Paints do not last as long as in clean conditions; tires and other rubber materials also tend to fail due to ozone cracking, unless they were produced via the utilization of antioxidant additives. As with human health, the degree of damage caused largely depends on the concentration and the duration of exposure. Although most gaseous air pollutants are totally colourless, there are some exceptions such as NO_2, which is brown. Most visible effects of air pollution are the outcome of the interaction of light with suspended particles (De Nevers, 2000).

Air quality and the concentration of air pollutants are influenced not only by physical and chemical processes, but also by meteorological, geographical and social factors. Some weather conditions, for example, moderate winds or calm air with temperature inversions as typically prevailing under anticyclonic conditions, can also significantly influence the rise of extreme concentration rates of pollutants in the air.

In Europe, many air pollution studies, especially for Athens due to its long summers with undisturbed irradiation and calm or weak breezes, have appeared in the international literature. This weather and the mountains, which surround the city to the north, favour extreme accumulation of the air pollutants (Kambezidis et al, 1995, 1996, 1998; Adamopoulos et al, 2002).

According to Péczely (1959), based on his observations made on air pollution rates in Budapest, air pollution tends to peak during extensive anticyclone events characterized by weak easterly breezes prevailing over the city. Conversely, air pollution is relatively low during the prevalence of cyclonic weather systems characterized by strong and turbulent air currents prevailing over the Carpathian Basin (Fig. 1a), especially when Hungary is in the rear part of the cyclone.

The major aim of the present study was to develop an objective, reliable classification of air mass types prevailing over the city of Szeged during the summer and winter months via the application of multivariate statistical methods. For each air mass type, following characterization by

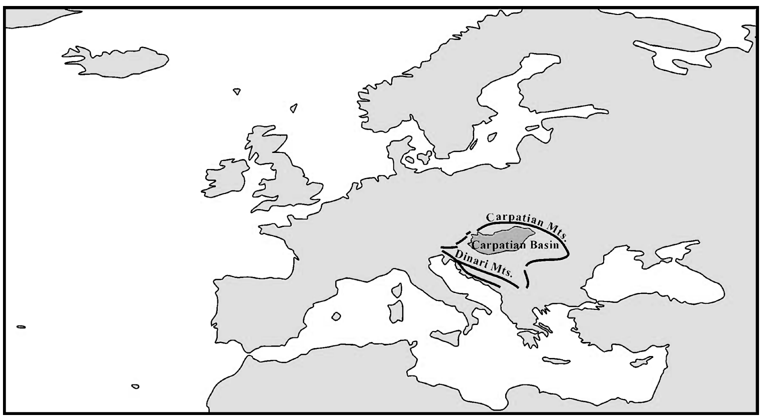


Fig. 1a. Location of the Carpathian Basin

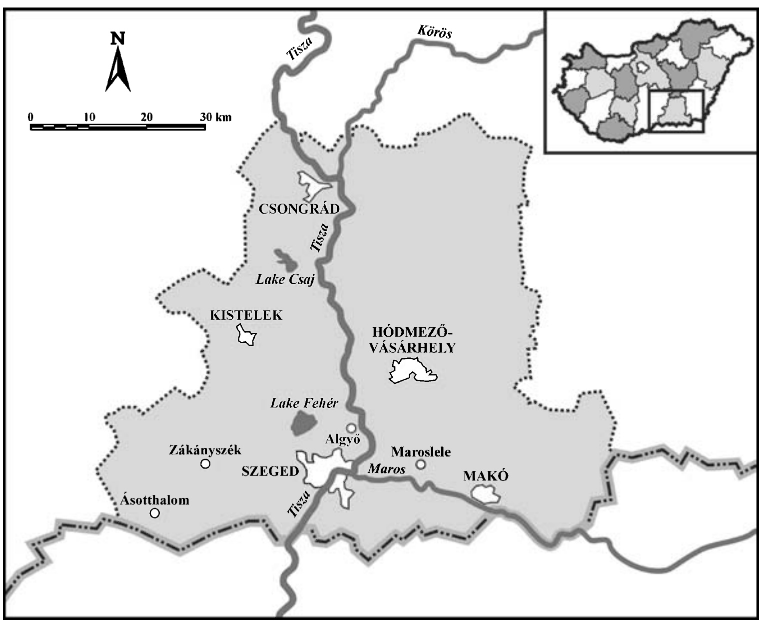


Fig. 1b. Location of Szeged in Csongrád county (centre); Csongrád county in Hungary (top right)

homogenous temperature and humidity conditions, the concentration of the main air pollutants is estimated. Additionally, in order to reveal the possible relationship between the prevailing weather conditions, the spatial distribution of sea-level pressure fields and the concentration of air pollutants in the area of Szeged, mean sea-level pressure fields have been calculated for the different air mass types for the North Atlantic – European region. Unfortunately, very few studies with such scope are known from the literature. The work of Sindosi et al (2003) could be mentioned here as an example of a similar analysis applied to Athens. Alternatively, stability classes are often used, e.g., in air quality modelling, to classify whether the dispersion of air pollutants is high or low due to the prevailing meteorological conditions (determined empirically from wind speed, temperature gradient, cloud cover or solar radiation). Pasquill's method (1962) for diffusion was developed to estimate vertical and lateral plume widths from a continuous source at various distances. These widths depend primarily on the standard deviations of vertical and horizontal wind direction fluctuations, respectively. These quantities are functions of height, roughness length and the Monin-Obukhov length. Where temperature profile measurements are not available, Pasquill suggested a set of six stability classes specified in terms of wind speed and irradiance only. Turner (1964) modified Pasquill's scheme and defined seven stability categories. In this method, net irradiance is estimated from a knowledge of solar altitude and existing conditions of cloud cover and ceiling height. Both the Pasquill and Turner classes are independent of height and surface roughness (Golder, 1972). The method used in this paper is an objective classification compared to the subjectively determined categories of Pasquill and Turner. Furthermore, our method takes into account more meteorological parameters for classifying air mass types and the efficiency of the classes received in grouping pollutant concentrations is statistically evaluated.

The methodology used in the paper is not proposed as a substitute for other chemical transport modelling systems but as a supplement to the existing methodologies, attempting to contribute our efforts to forecast pollution levels and thus adopt appropriate measures when necessary.

This study represents an objective weather classification method, which might also serve as a starting point for the creation of an air pollution monitoring/forecasting system, with the ultimate goal of facilitating debate regarding air pollution in the city of Szeged.

2. Topography, climatology and air quality of Szeged area

2.1 Topography

The city of Szeged is the largest town in SE Hungary. The city (20°06′ E; 46°15′ N) is located at the confluence of the Tisza and Maros Rivers characterized by an extensive flat landscape with an elevation of 79 m a.s.l. (Fig. 1b). The built-up area covers a region of about 46 km^2 with about 155,000 inhabitants.

Szeged and its surroundings are not only characterized by extensive lowlands but it also has the lowest elevation in Hungary and the Carpathian Basin as well. This results in a "double basin" situation. Due to the position of the city in a basin (a smaller one within a larger one), temperature inversions form more easily in the area (e.g., due to cool air flow from the basin slopes) and prevail longer than in flat terrain, leading to an enrichment of air pollutants within the inversion layer.

2.2 Climatology

According to the climatic classification system of Köppen, the majority of the Hungarian territory, including Csongrád county and the agglomeration of Szeged, belongs to the *Cf* climate zone characterized by temperate-warm climates with an almost even distribution of precipitation or that of Trewartha's *D.1* climate zone characterized by continental climates with long warm seasons.

The more detailed, higher resolution climatic classification of Hungary is based on the mean temperature values of the growth season (t_{VS}) and the aridity index (H) [where $H = E_S/L \cdot C$ (E_S is the annual mean radiation balance; L is the latent heat of vaporization and C is the total annual mean precipitation)]. Based on the climatological characteristics of the period 1901–1950, the climate of Szeged can be considered as *warm-dry* with $t_{VS} > 17.5\,°C$ and $H > 1.15$ (Péczely, 1979).

The E_S energetic component of the aridity index (H) changes only slightly in Hungary. Its mean annual value, considering the period 1901–1950, is 1760 MJ m^{-2} yr^{-1} (Péczely, 1979). On the other hand, the total annual precipitation amounts may fluctuate quite substantially. The extreme values for Szeged in this period were $P_{min} = 203$ mm; $P_{max} = 867$ mm. On the basis of this information, the aridity index calculated for the most arid year was $H = 3.47$; while that of the most humid year was $H = 0.81$. The typical vegetation assigned to the former value is desert, while that of the latter is woodland. Though persistence of high H values is not observed all-year or all-summer round, the climate of some summers inclines towards semi-arid or arid conditions in the Szeged area. This is reflected in the composition of the natural vegetation. In the southern part of the Great Hungarian Plains native plants include semi-desert species such needle grass (Stipa stenophylla) (Makra et al, 1985).

The climate of Szeged is characterized by hot summers and moderately cold winters. The distribution of rainfall is fairly uniform during the year: with a share of 29 and 19% for the summer (JJA) and the winter (DJF) seasons, respectively. Mean daily summer temperatures are around 22.4 °C, while the mean daily winter temperatures are 2.3 °C. The irradiance values also exhibit large-scale variances with an average of 20.2 and 4.2 MJ m^{-2} in summer and winter days, respectively. The most frequent winds blow along the NNW–SSE axis, with prevailing air currents arriving from NNW (42.3%) and SSW (24%) in the summer and from SSE (32.6%) and NNW (30.8%) during the winter. Thanks to its unique geographical position, Szeged is characterized by relatively low wind speeds with average daily summer and winter values of 2.8 and 3.5 m s^{-1}, respectively. The highest hourly wind speeds have been recorded during the spring with a rate of 5 m s^{-1} (Péczely, 1979).

The dry and warm climate of the region is especially favourable for the Turanian (Aralo-caspian) semi-arid species. One group of plants prefer extremely arid conditions; while another group contains species which indicate aridity, namely, these are xeroindicators (indicate long and dry periods in the climate). The Ellenberg's indicator number (Ellenberg et al, 1991) for the first group is 1, while for the after group is 2. The lower the Ellenberg indicator (min = 1) the more a species prefers a dry and warm climate. Similarly, the higher the number (max = 12) the more it favours a humid environment. Hence, Ellenberg's indicator number is a 12-degree scale, first applied by Borhidi (1995) in Hungary, and was introduced into the Hungarian literature as a WB-value. The WB-value is a relative indicator number for ground water or soil humidity (Water Borhidi). Borhidi also indicated the characteristic species for the Hungarian flora, which Ellenberg dealt with little on the scale.

The analysis is applied to the two extreme seasons of the year because the winter and the summer show the most marked difference in atmospheric circulation over the Carpathian Basin. In the summer, the subtropical (Azores) anticyclone of the Atlantic is most frequent with 20–25% of all weather types. Northerly currents are also characteristic because of the blocking anticyclones. In the winter, southerly currents are most frequent, which are followed by westerly flows. In both seasons, anticyclone centres and anticyclone ridge weather situations are most frequent over the Carpathian Basin, as a result of the basin nature of the region (Péczely, 1983).

2.3 Air quality

Air quality is modified and influenced by the prevailing atmospheric conditions, which are controlled by the prevailing meteorological parameters, mainly the temperature profile close to ground level. The recorded averages for the city of Szeged are the following: annual mean temperature: 11.2 °C; mean January and July temperatures: −1.2 °C and 22.4 °C, respectively; mean annual relative humidity: 71%; mean annual precipitation total: 573 mm; mean annual sunshine duration: 2102 hours; mean annual wind speed: 3.2 m s^{-1}.

The city structure is very simple and is characterized by an intertwined network of boulevards, avenues and streets sectioned by the River Tisza. However, this simplicity also largely contributes to the concentration of traffic as well as air pollution within urban areas.

The industrial area is mainly restricted to the north-western part of the city. Thus, the prevailing westerly and northerly winds tend to carry the pollutants from this area towards the city centre.

Though Szeged is considered to be a centre of light industry in the southern part of the Great Hungarian Plain (wooden- and paprika processing industry, mill industry, hemp processing, salami manufacturing) in general, some elements of heavy industry can also be found in the rural areas (rubber and paint industry, oil and gas mining). After the collapse of the socialist political system in 1989, the industrial potential of the city has undergone a heavy freefall. The textile, cable, clothing and canning factories were all closed down; furthermore, the Light-weight Construction Building and Supplying (KÉSZ) Ltd. moved its headquarters to another city, Kecskemét.

The total urban spread extends well beyond the city limits and includes the largest oil field in Hungary with several oil torches located just north of the town. This oil field is also a significant source of such air pollutants as NO_x and sulphur dioxide. The power station, located in the western part of the city, is also a major source of pollution. Exhaust fumes have also largely contributed to the steady increase of nitrogen oxide and carbon monoxide in Szeged. Besides, as a result of the heavy traffic, deposited dust is often suspended in the atmosphere.

In a detailed analysis, Szeged was ranked 32nd of 88 Hungarian cities, according to the quality of the environment and the level of environmental awareness. [The city ranked 1st was considered to have the best environmental conditions (Makra et al, 2002)].

On the basis of the frequency of pollutant concentrations exceeding air quality limits, measured at the Regional Immission Examining (RIE) network of stations for Hungary in 2001, the air quality of Szeged, according to a three-category classifying system (satisfactory, moderately polluted, polluted), can be listed in the ‘‘polluted’’ category (Mohl et al, 2002).

Nitrogen oxides (NO_x), ozone (O_3) and particulate matter (PM_{10}) concentrations tend to exceed air quality limit values established by EU standards. [Daily (24-hour) concentration of particulate matter is 11–19 times higher, while its annual concentration is twice as much as the EU standard proposed since January 1, 2005!] The information that Szeged was ranked 32nd of 88 Hungarian cities gives the impression that this is a city with rather moderate air quality. Therefore, the information that Szeged is ranked ‘‘polluted’’ according to the RIE database seems surprising. In the above-mentioned analysis, the cities were ranked according to seven different categories (of nineteen environmental indicators), which are as follows: water consumption (1), energy consumption (3), public utilities supply (4), traffic (1), waste management (3), settlement amenities factors (4) and air quality (with average concentration of particulates deposited, sulphur-dioxide and nitrogen-dioxide). As air quality is only one of the seven categories considered and is only represented by three parameters (environmental indicators), it contributes little to the rank of the cities.

The high concentrations of particulate matter are closely connected with the development of respiratory diseases. The annual trend of the air pollutant levels follow a unimodal distribution. Concentrations of NO, NO_2 and PM_{10} are characterized by winter maxima and summer minima. At the same time, ozone reaches its maximum in the summer, in accordance with the annual change of irradiation (Makra and Horváth, 2001; Makra et al, 2001a, b; Mohl et al, 2002; Mayer et al, 2004).

About 50% of the particulate matter comes from the northwest region of Szeged, covered by shifting sands, loess and sandy ridges. At the same time, the industrial district is located in the northwest region of the town. Hence, the dominant northwesterlies transport particulate matter as well as pollutants of industrial origin over Szeged. The remaining PM_{10} originates from traffic. Particulates are produced partly by the engine of vehicles, and partly by air currents generated by vehicles (Mohl et al, 2002).

The traffic system of Szeged is highly overcrowded. Among vehicles participating in the traffic, the ratio of passenger cars is the highest: 84%. In the year 2000, after the modernization of vehicles, CO concentrations in the city were reduced to 36–40% of that measured in 1990. On the other hand, the traffic on the main roads increased to 3–70% during the same period. During a regular day (a 24-hour period) about 70,000–90,000 vehicles, on average, pass through the city (Mohl et al, 2002).

A serious environmental health problem arises during late summer and early autumn caused by the pollen of ragweed with mugwort leaves or short ragweed (Ambrosia artemisiifolia = Ambrosia elatior), considered to be the most dangerous of all pollens. Annual pollen counts of ragweed in Szeged are one or two orders of magnitude higher than those in other European cities (Makra et al, 2004a, b).

The ''double basin'' situation of Szeged favours the longer persistence of anticyclonic weather types both in summer and winter, which contributes to the enrichment of not only Major Air Pollutants (including sulphur dioxide, nitrogen dioxide, carbon monoxide, particles, lead and ozone) but also pollens as biological agents [belonging to Hazardous Air Pollutants (HAPs)]. The role and efficiency of large-scale weather situations established in grouping pollutant concentrations was the main motive behind the present study.

3. Data collection

3.1 The air pollution data

The air pollution monitoring station, mentioned above, is located in downtown Szeged in a crossroad with heavy traffic (Kossuth Avenue and Damjanich Street), about 10 m distance from the Kossuth Avenue. This is one of the busiest crossroads of Szeged. The monitoring station was put into operation in September 1, 1996. At a distance of 10 m from the station there is a two-storey building which affects wind and irradiance parameters. Sensors, measuring concentrations of the air pollutants, are placed 3 m above the surface.

Concentrations of CO, NO, NO_2, SO_2, O_3 and TSP at the monitoring station are measured by 5 different analyzers. The concentration of CO is measured by non-dispersive infrared absorption (type of the instrument: CO11M-LCD). The measurement of NO and NO_2 concentrations is based on the principle of chemiluminescence; the concentrations of NO_x are obtained so that the instrument adds up latest NO and NO_2 values automatically (type of the instrument: TE 42C). The measurement method of SO_2 is UV fluorescence emission (type of the instrument: FHAF21M-LCD). O_3 concentration measurements are based on the UV absorption of wavelength of 254 nm (type of the instrument: TE 49C). TSP concentrations are recorded by absorption of β-radiation (type of the instrument: FH 62 I-N).

Gas analyzers are calibrated in two ways. One of these is the 0-point, adjustment which occurs automatically every 24 hours. The other calibration point is adjusted by a verified sample every two weeks. Measurements of TSP are verified once in every quarter year. The instruments are controlled and the gathered data are stored on a personal computer. From the 10-second measurements one-minute averages are derived; then, from the latter data, 30-minute averages are calculated.

The type of instruments measuring meteorological parameters at the station are as follows: temperature: THS-611, humidity: THS-611, irradiation: RS 81-I and wind speed: WS-12 H+. Temperature and humidity values are measured 3 m above the surface, while wind direction, wind speed and irradiation are recorded at a height of 6 m above the ground level.

3.2 The meteorological data

The data consist 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 the average mass concentrations of the main air pollutants [CO, NO, NO_2, SO_2, O_3 and TSP ($\mu g\, m^{-3}$)]; and the daily values of the main climatic parameters (temperature, humidity, atmospheric pressure, global irradiance and 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, hPa), water vapour pressure (VP, hPa), potential evaporation (PE, mm), dew point temperature (T_d, °C) and atmospheric pressure (P, hPa).

The 8 pollution parameters considered are the average diurnal mass concentrations of the following pollutants: CO ($mg\, m^{-3}$); 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 ($\mu g\, m^{-3}$).

Daily sea-level pressure fields measured at 00 UTC (Universal Time Centre) were acquired from the ECMWF (European Centre for Medium-Range Weather Forecasts) Re-Analysis ERA 40 project, in the frame of which daily data have been re-analyzed since September 1st, 1957. The procedure has been performed with a uniform method from the data being available in the investigated period. Data for the ECMWF Re-Analysis ERA 40 project are verified, dynamically correct, the pressure field is true even over the Atlantic Ocean and there are no missing data. When using this method, the measured false input data are omitted. On the other hand, if original station data are used, false data can frequently be encountered.

The investigated area is the North-Atlantic – European region between 30° N–70.5° N latitudes and 30° W–45° E longitudes. The grid network is selected with a density of $1.5° \times 1.5°$, which indicates $28 \times 51 = 1428$ grid points across the region.

4. Methods

4.1 Cartographical background

For each objective type, average daily isobar maps on the basis of daily sea-level pressure data calculated at each grid point of the investigated region were constructed by applying the Surfer 7.00 GIS software. Isobars for an average day, i.e., for an average objective type, were drawn using $28 \times 51 = 1428$ grid data on the basis of the standard Kriging method without increasing element number of data and with maximum smoothing. As a result of the procedure, the curved surface on the Earth as a spherical trapezoid with 40.5° difference of latitude and 75° difference of longitude was transformed into a plane rectangle with equal spacing both horizontally and vertically. Isobar maps produced in this way can only be fitted to those informative background maps which are prepared with the same projection. For this reason, the background map of the investigated region was produced in an equidistant cylindrical projection. The major advantage of this map is that it is free from longitudinal distortion along each meridian and, therefore, the determination of the points of the compass is simple at any location of the map. Namely, the north–south and east–west directions are parallel with the vertical and horizontal sides of the rectangle; hence, geographical co-ordinates of various locations and air pressure formations can easily be determined with the help of the spaces indicated on the rectangle both horizontally and vertically (and, if required, by applying linear interpolation). The only drawback of its use is that the background map becomes longer at higher latitudes in the east–west direction. The grid denotes a horizontal distance of about 107.3 km at the latitude of 50° N.

4.2 χ^2-test, independence analysis

In order to decide whether or not the sea level pressure fields examined differ significantly from each other, the χ^2-test independence analysis was applied. This method determines whether two random variables (ξ and η) are independent. According to the 0-hypothesis, ξ and η are not independent.

4.3 Factor analysis

In order to reduce the dimensionality of the above collected meteorological data sets and thus to explain the relations among the 12 meteorological variables, the multivariate statistical method of factor analysis is used. The main object of factor analysis is to describe the initial variables $X_1, X_2, \ldots, X_p$ in terms of m linearly independent indices ($m < p$), the so called factors, measuring different "*dimensions*" of the initial data set. Each variable X can be expressed as a linear function of the m factors, which are the main contributors to the climate of Szeged:

$$X_i = \sum_{j=1}^{m} \alpha_{ij} F_j, \tag{1}$$

where α_{ij} are constants termed factor loadings. The square of α_{ij} represents the part of the variance of X_i that is accounted for by the factor F_j.

One important stage of this method is the decision for the number (m) of the retained factors. On this matter, many criteria have been proposed. In some studies, the *Guttmann criterion* or *Rule 1* is used, which determines to keep the factors with eigenvalues >1 and neglect those that do not account for at least the variance of one standardised variable X_i. Perhaps the most common method is to specify a least percentage (80% in this paper) of the total variance in the original variables that has to be achieved (Jolliffe, 1993; Sindosi et al, 2003). Extraction was performed by *Principal Component Analysis* (*k*th eigenvalue is the variance of the *k*th principal component). There is an infinite number of equations alternative to Eq. (1). In order to select the best or the most desirable factors, the so-called "*factor rotation*" is applied, a process, which either maximises or minimizes factor loadings for a better interpretation of the results. In this study, the "*varimax*" or "*orthogonal factor rotation*" is applied, which ensures that the factors remain uncorrelated (Jolliffe, 1990, 1993; Bartzokas and Metaxas, 1993, 1995; Sindosi et al, 2003).

Factor analysis was applied to the data set consisting of 12 columns (12 meteorological variables) and 450 rows (450 days) both for the winter and summer months, in order to reduce the 12 interrelated meteorological parameters and reveal the main independent meteorological factors which are responsible for the formation of the weather in Szeged.

4.4 Cluster analysis

Cluster analysis is applied to the factor scores time series in order to group objectively days with similar weather conditions. The aim of this method is to maximize the homogeneity of objects within the clusters and also to maximize the heterogeneity between the clusters. Each observation (day) corresponds to a point in the *m*-dimensional space and each cluster consists of those observations, which are "*close*" to each other in this space. The characterization of a distance between two observations k and l as "*close*" or "*far*" is determined by the square of their Euclidean distance:

$$D_{kl}^2 = \sum_{i=1}^{m} (x_{ki} - x_{li})^2, \qquad (2)$$

where x_{ki} is the value of the *i*-th factor for the *k*-th day and x_{li} is the value of the *i*-th factor for the *l*-th day.

There are two main clustering techniques: hierarchical and non-hierarchical. Their basic difference is that in the non-hierarchical technique the number of clusters must be known *a priori*. In contrast, in the hierarchical method the ultimate number of clusters is determined by a variety of statistical criteria. In this paper, the hierarchical technique is applied because of the lack of an objective classification of weather types for the Szeged region. The hierarchical technique can be applied by using various methods. Here, the "*average linkage*" method is used, since it does not depend on extreme values, besides it produces more realistic groupings and properly combines extreme weather days into distinct meteorological units (Anderberg, 1973; Kalkstein et al, 1987; Hair et al, 1998; Sindosi et al, 2003).

Then, for each of the derived clusters of days, the mean value for every meteorological and pollution parameter is computed. In this way, the relationships between weather conditions and the corresponding concentration levels of air pollutants are revealed. Finally, for each weather type, the composite maps of the mean sea level pressure distribution over the North-Atlantic – European region (00 UTC) are constructed. The aim of these maps is to associate atmospheric circulation patterns and pollution levels in the Szeged region. The classification of synoptic patterns into distinct groups enables us to describe the most important synoptic types considering the position of the Szeged region.

4.5 ANOVA and Tukey's honestly significant difference test

When determining the synoptic types, only meteorological parameters are taken into account, excluding pollution data. Hence, the differences of the mean pollution levels calculated for each synoptic type need further statistical evaluation. This is performed by the method of one-way Analysis of Variance (ANOVA) for each pollutant. By using this method, significant differences in pollutant concentrations of different synoptic types (clusters) can be determined. Finally, the Tukey's honestly significant difference test is applied in order to quantitatively compare the mean air 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.

5. Results

5.1 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). Table 2 displays the factor loadings after orthogonal rotation.

Factor 1 explains 50.86% of the total variance (Table 1) 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). It can be seen that the temperature variables are not directly related to irradiance, which during the winter depends on the third factor. The reason is that winter air temperature is mainly influenced by the synoptic-scale air mass affecting the area and, to a lesser extent, by the local irradiance. The high loadings of these temperature and humidity parameters show their strong relationship. Namely, the high loading of the water vapor pressure is due to the increase in vapor capacity of the atmosphere as temperature rises. Dew point temperature covaries with the above parameters, since an increase (decrease) in water vapour pressure is explained by higher (lower) temperature at which air becomes saturated (dew point) (Table 2).

Table 1. Initial eigenvalues and cumulative variances, winter months (December, January and February)

Component	Initial eigenvalues		
	Total variance	Relative variance, %	Cumulative variance, %
1	6.10	50.86	50.86
2	2.38	19.85	70.70
3	1.05	8.72	79.42
4	0.85	7.08	86.51
5	0.75	6.29	92.79
6	0.46	3.82	96.61
7	0.34	2.86	99.47
8	0.04	0.32	99.79
9	0.02	0.19	99.97
10	0.00	0.02	100.00
11	0.00	0.00	100.00
12	0.00	0.00	100.00

Factor 2 (19.85% of the total variance; Table 1) 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 (low) potential evaporation is associated with low (high) relative humidity (Table 2).

Factor 3 explains 8.72% of the total variance (Table 1) and comprises the daily temperature range and irradiance. High irradiance values, which indicate low cloud cover, generally cause high T_{max}. Whenever clear skies persist also in the following night, nocturnal long wave radiation leads to larger cooling of the surface and to lower T_{min}; thus, T_{range} is generally larger under these conditions (Table 2). (This is only true, if no advection of, e.g., cooler air masses takes place.)

Table 2. 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 vapor pressure, E	0.93			
Water vapor pressure, VP	0.97			
Potential evaporation, PE		0.74		
Dew point temperature, T_d	0.97			
Atmospheric pressure, P				0.96

Values higher than |0.10| are statistically significant at the 95% level; however, Table 2 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 3. Mean values of the meteorological and pollution parameters for the days belonging to the five dominant clusters of the winter months (December, January and February)

Clusters	1	2	3	4	5
Number of cases (days)	56	134	73	94	90
Freqency (%)	12.4	29.7	16.2	20.8	20.0
T_{mean} (°C)	−4.9	0.24	5.5	2.4	6.0
T_{max} (°C)	−1.0	2.1	9.6	7.6	8.1
T_{min} (°C)	−8.1	−1.6	0.9	−1.4	2.5
$\Delta T = T_{max} - T_{min}$, (°C)	7.0	3.7	8.7	9.0	5.6
WS ($m\,s^{-1}$)	0.3	0.6	1.0	0.4	0.9
RH (%)	81.5	83.5	63.6	80.6	84.3
I ($MJ\,m^{-2}$)	5.4	2.7	6.2	5.4	2.7
E (hPa)	5.9	8.5	12.8	10.1	12.8
VP (hPa)	4.8	7.1	8.1	8.1	10.8
PE (mm)	0.6	0.8	2.0	1.0	1.1
T_d (°C)	−7.4	−2.0	−0.7	−0.4	3.8
P (hPa)	1009.2	1004.0	986.2	1001.2	994.6
CO ($mg\,m^{-3}$)	0.878	0.799	0.565	0.933	0.790
NO ($\mu g\,m^{-3}$)	24.3	20.9	24.1	44.0	28.5
NO_2 ($\mu g\,m^{-3}$)	42.0	34.4	41.0	47.2	39.5
NO_2/NO	3.0	6.1	2.5	1.8	3.3
O_3 ($\mu g\,m^{-3}$)	27.2	23.3	33.4	23.1	20.2
$O_{3\,max}$ ($\mu g\,m^{-3}$)	50.7	39.2	59.5	46.5	39.0
SO_2 ($\mu g\,m^{-3}$)	15.8	10.7	11.6	11.7	9.9
TSP ($\mu g\,m^{-3}$)	61.9	49.1	50.1	61.4	44.7

Factor 4 is slightly weaker than *Factor 3* and explains 7.08% of the total variance (Table 1). It comprises atmospheric pressure only (Table 2).

Next, cluster analysis was applied to the four factor score time series and, as a result of this, six homogenous clusters of days were revealed. However, one cluster included only 4 days (January 5, 10, 18, 19, all in year 2001; a mere 0.89% of the total number of days). These days were related to extreme weather conditions with increased pollutant concentrations, due to an anticyclone. This cluster was therefore omitted from further consideration, since our aim was to analyze days with dominant air mass types. The main characteristics of the other five dominant clusters involving the prevailing air mass types are shown in Table 3, which presents the mean values of their meteorological parameters as well as the mean values of the corresponding pollution parameters.

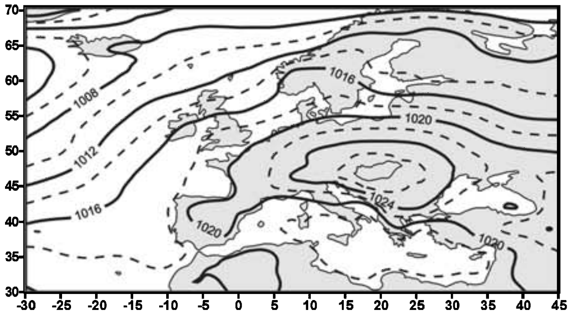
Considering the basic statistical parameters of the pollutants, variation coefficients (standard deviation expressed in the unit of the average) for NO and SO_2 are twice as high as those for other contaminants examined, which denotes their higher variability. The difference of $|median - average|$ remains within the so called interquartile half extent (the interval given by the lower quartile and the upper quartile) for each pollutant. The greatest differences are detected for CO, NO, NO_2/NO and O_3.

The mean sea level pressure distribution belonging to the clusters examined and the number of days in each cluster (air mass type) are shown in Fig. 2.

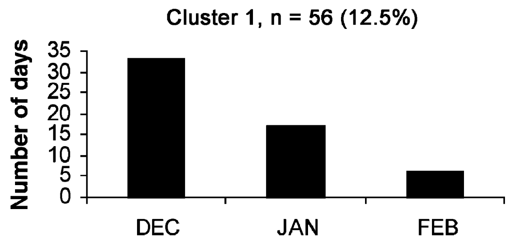
Mean sea level pressure fields for the 5 clusters received in the winter were compared on the basis of the used grid values. In order to decide whether the mean sea level pressure fields of the 5 clusters in the winter differ significantly from each other in the period examined, the χ^2-test was applied. The 0-hypothesis means that there is no significant difference between the mean sea level pressure fields of the clusters compared. On the basis of our computations, probability of the 0-hypothesis in the winter months between the sea level pressure fields of clusters 1–2 and 3–4 is 1. Namely, in these cases the mean cluster

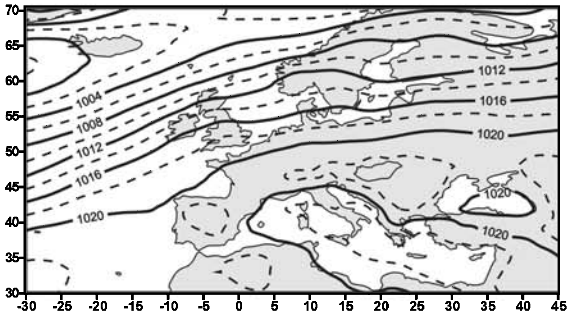
→

Fig. 2. Mean sea-level pressure fields for each air mass type (cluster), and monthly variation of the number of days, North-Atlantic – European region, winter months (December, January and February)

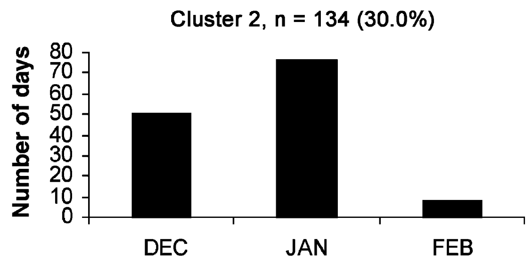


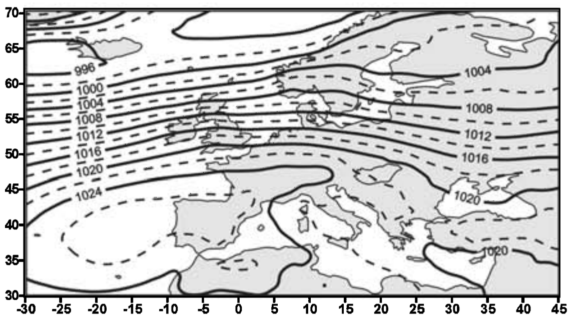
cluster 1



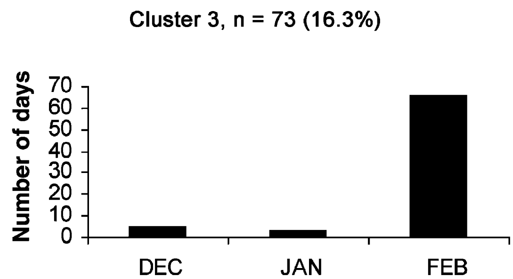


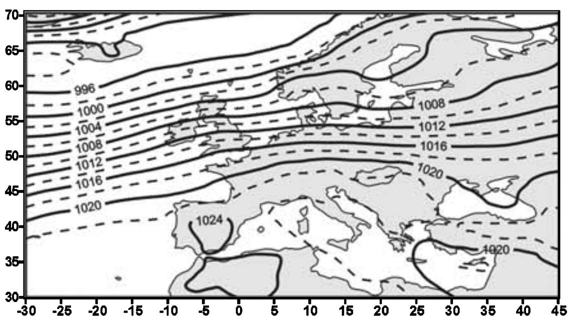
cluster 2



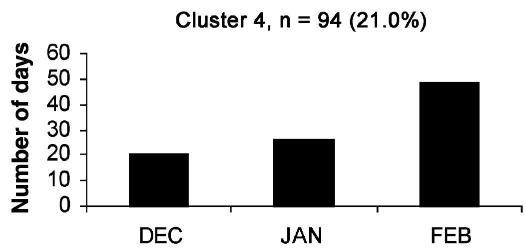


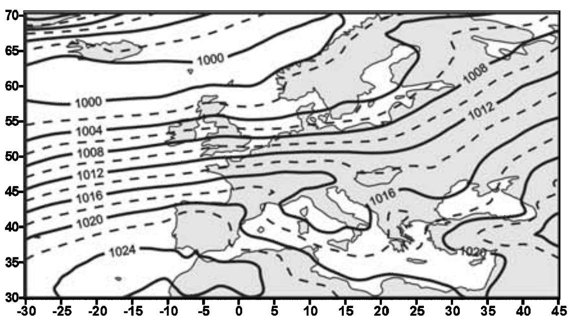
cluster 3





cluster 4





cluster 5

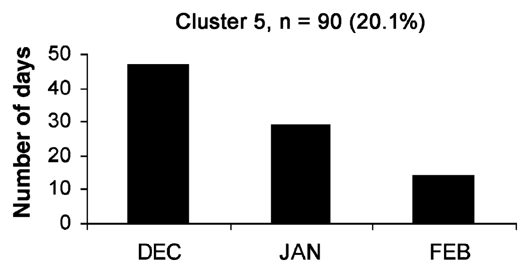


Table 4. χ^2-test, independence analysis of the average sea-level pressure fields for the five clusters derived for the winter months (December, January and February), probability of the 0-hypothesis

Cluster	1	2	3	4	5
1	–	1.0000	0.0000	0.0000	0.0000
2		–	0.0000	0.0000	0.0000
3			–	1.0000	0.0000
4				–	0.0000
5					–

fields mentioned can not be considered independent. On the other hand, the probability of the 0-hypothesis for all the other cluster pairs is 0; namely, the mean cluster fields compared are considered to be independent (Table 4).

The five 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 air mass type accounts for 12.5% of the total number of days and is associated with the following weather characteristics in Szeged: high amounts of irradiance (mean value $= 5.4\,MJ\,m^{-2}$), the lowest values of temperature parameters (mean daily as well as maximum and minimum temperatures), the lowest values of humidity parameters (water vapour pressure, saturation vapour pressure, potential evaporation and dew point temperature) and very low wind speed ($0.3\,m\,s^{-1}$). 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 low cloudiness, concentration of the secondary pollutants (O_3 and $O_{3\,max}$) is relatively high (Fig. 2; Table 3).

Cluster 2: This type is named "anticyclone over the Mediterranean". This cluster contains 30.0% of the total number of days and represents the most frequent situation. This pressure pattern forms an anticyclonic ridge over the Carpathian Basin with calm or weak breezes. Cloudy conditions, on average higher temperatures due to less nocturnal cooling are characteristic during this air mass type. The temperature parameters are significantly higher during this air mass type, compared to those in *Cluster 1*. The ozone levels are lower because of the higher cloudiness. The reason of the lower concentration in primary pollutants might be the higher wind speed (Fig. 2; Table 3).

Cluster 3: Anticyclone stretches out from the region of the Azores. This situation is only characteristic in February. During prevalence of this type an anticyclone can reach Central and even Eastern Europe resulting in calm, sunny weather. This situation causes high temperatures and high winds. Lower concentrations of CO, SO_2 and TSP in this cluster compared to those in *Cluster 1* can be explained with the highest average wind speed in *Cluster 3*. Since average concentrations of NO in *Clusters 1* and *3* are the same, higher enrichment of the ozone in *Cluster 3* relating to that in *Cluster 1* can be explained with the lower cloudiness (This is only true if advection is neglected.) (Fig. 2; Table 3).

Cluster 4: Anticyclone over Southern Europe and North Africa. This cluster does not differ significantly from *Cluster 3*. Here, the high-pressure system from SW Europe extends over the Eastern part of the Mediterranean. Because of the very low wind speeds, concentrations of primary pollutants (except SO_2) are extremely high ($CO = 0.93\,mg\,m^{-3}$; $NO = 44.0\,\mu g\,m^{-3}$; $NO_2 = 47.2\,\mu g\,m^{-3}$; $TSP = 61.4\,\mu g\,m^{-3}$). At the same time, irradiance is also high. However, concentration of ozone 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: $NO + O_3 \longrightarrow NO_2 + O_2$ (Fig. 2; Table 3).

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 strong winds are observed at Szeged. The 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 highest cloudiness ($I = 2.7\,MJ\,m^{-2}$, as in *Cluster 2*) with a medium level of NO results

Table 5a. ANOVA statistics for inter – air mass comparison of pollutant concentrations in the winter months (December, January and February)

	CO	NO	NO_2	NO_2/NO	O_3	$O_{3\,max}$	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, %	99	99	99	78	99	99	99	99

Table 5b. 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 significant difference test (light-faced characters: 95% of significance, **bold** characters: 99% 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**

in the lowest concentration of the ozone parameters ($O_3 = 20.2\,\mu g\,m^{-3}$; $O_{3\,max} = 39.0\,\mu g\,m^{-3}$) (Fig. 2; Table 3).

In order to determine the influence of air mass type on pollutant levels, analyses of variance (ANOVA) were performed on the pollutant parameters. The results are shown in Table 5a. It can be observed that, with the exception of NO_2/NO, all pollutants present significant inter – air mass type differences in mean concentration values at the 99% probability level. Considering that differences are found among the mean pollutant levels, Tukey's tests were applied in order to receive a pairwise multiple assessment of the differences. The statistically significant differences are shown in Table 5b at the 95% and 99% 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 can be considered to be the most different, since levels of most pollutant pairs show substantial difference between them. This can mainly be explained by the fact that these two types show nearly the highest difference in wind speed. On the other hand, *Cluster 2* seems to be an intermediate cluster with respect to 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.

5.2 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 6). Table 7 shows factor loadings of the summer months after orthogonal rotation.

Factor 1, with 47.35% of the total variance (Table 6), includes the same parameters as in the winter case. These are temperature (mean, maximum and minimum temperatures) and humidity variables (saturation vapour pressure, water vapour pressure and dew point temperature). They have the same positive sign, which denotes that higher (lower) temperature values involve higher (lower) values of the humidity parameters.

Table 6. Initial eigenvalues and cumulative variances, summer months (June, July and August)

Component	Initial eigenvalues		
	Total variance	Relative variance, %	Cumulative variance, %
1	5.68	47.35	47.35
2	2.39	19.94	67.29
3	1.06	8.86	76.15
4	0.99	8.22	84.36
5	0.89	7.38	91.74
6	0.59	4.92	96.65
7	0.37	3.07	99.72
8	0.02	0.17	99.89
9	0.01	0.10	99.99
10	0.00	0.01	100.00
11	0.00	0.00	100.00
12	0.00	0.00	100.00

While this factor was classified in the same way as in the winter months, the remaining factors have high loadings with completely different meteorological variables from their winter versions (Table 7).
Factor 2 (19.44% of the total variance; Table 6) 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 7).
Factor 3 (8.86% of the total variance; Table 6) consists of only atmospheric pressure (Table 7).
Factor 4 (8.22% of the total variance; Table 6) is slightly weaker than *Factor 3* and contains only wind speed (Table 7).

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 characterized by a greater number of air mass types compared to the winter period. During the whole summer only two main pressure systems influence the Carpathian Basin: the Icelandic low from NW Europe and the subtropical high from the region of the Azores. Hence, the differences among these dominant types, considering both the mean values of the parameters and the spatial pressure distribution, are rather small. The ten air mass types (clusters) with mean values of both meteorological and air pollution parameters are shown in Table 8.

Basic statistical parameters of the pollutants are calculated for the summer months. Variation

Table 7. 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 vapor pressure, E	0.87			
Water vapor pressure, VP	0.95			
Potential evaporation, PE		0.80		
Dew point temperature, T_d	0.95			
Atmospheric pressure, P			0.92	

Values higher than |0.10| are statistically significant at the 95% level; however, Table 7 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 8. Mean values of the meteorological and pollution parameters for the days belonging to the ten dominant clusters of the summer months (June, July and August)

Clusters	1	2	3	4	5	6	7	8	9	10
Number of cases (days)	28	40	58	76	17	46	72	47	51	25
Frequency (%)	6.1	8.7	12.6	16.5	3.7	10.0	15.7	10.2	11.1	0.1
T_{mean} (°C)	17.7	20.2	19.7	22.8	21.0	26.1	25.9	20.8	20.0	28.1
T_{max} (°C)	23.9	26.2	25.4	28.3	29.6	32.7	31.8	25.2	25.1	34.7
T_{min} (°C)	12.5	14.9	16.9	17.0	14.0	18.6	20.1	16.5	14.9	21.2
$\Delta T = T_{max} - T_{min}$, (°C)	11.3	11.3	8.5	11.3	15.6	14.1	11.7	8.7	10.2	13.6
WS ($m\,s^{-1}$)	0.8	0.6	0.9	1.0	1.0	0.6	0.8	1.9	1.3	1.5
RH (%)	59.8	68.4	82.2	67.82	66.1	53.5	69.0	74.0	64.9	55.3
I ($MJ\,m^{-2}$)	18.9	18.6	15.6	23.5	18.4	24.5	23.5	18.4	23.4	26.6
E (hPa)	27.7	32.4	31.6	38.3	34.8	46.9	46.4	34.1	32.3	52.7
VP (hPa)	16.5	22.0	26.0	25.9	23.1	25.1	31.9	24.9	20.7	28.9
PE (mm)	4.4	4.4	2.8	5.2	4.9	8.1	6.1	4.0	4.7	8.9
T_d (°C)	10.0	14.4	16.9	16.9	14.7	16.4	20.2	16.2	13.5	18.7
P (hPa)	1018.0	1016.1	1013.4	1015.0	1015.7	1017.4	1015.6	1013.7	1013.8	1015.7
CO ($mg\,m^{-3}$)	0.239	0.361	0.396	0.348	0.323	0.440	0.443	0.249	0.241	0.484
NO ($\mu g\,m^{-3}$)	6.2	10.0	7.1	6.4	8.7	9.9	6.7	3.5	5.6	7.9
NO_2 ($\mu g\,m^{-3}$)	20.1	31.4	26.6	25.4	28.5	33.9	27.7	17.5	22.5	33.2
NO_2/NO	6.1	7.4	9.8	9.0	8.0	10.5	9.6	14.4	27.5	6.4
O_3 ($\mu g\,m^{-3}$)	54.7	55.4	53.8	56.1	58.9	64.2	58.9	57.4	60.4	77.5
$O_{3\,max}$ ($\mu g\,m^{-3}$)	89.7	103.4	99.8	98.5	113.7	114.7	108.2	92.7	98.8	129.8
SO_2 ($\mu g\,m^{-3}$)	5.0	4.8	2.9	4.3	5.6	5.3	4.6	3.6	4.3	5.0
TSP ($\mu g\,m^{-3}$)	25.1	35.5	34.2	38.2	38.0	47.3	46.7	31.3	32.6	53.4

coefficients for O_3 and $O_{3\,max}$ decreased substantially compared to their values in the winter months. The difference of $|median - average|$ for NO_2/NO is found beyond the interquartile half extent in clusters 2, 3, 4, 5, 7 and 9. Furthermore, the averages for TSP and SO_2 are also detected beyond this interval in *Cluster 9*. This indicates that distribution functions of the pollutant concentrations in the clusters mentioned above are distorted; namely, the means of the samples are not representative of the data sets.

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 in Fig. 3a, b.

In order to decide whether the mean sea-level pressure fields of the 10 clusters in the summer differ significantly from each other, the χ^2-test was applied. According to our results, only the mean sea-level pressure fields of clusters 5–9 can be considered independent. For all the other cluster pairs independence is not realized (Table 9).

The pressure patterns and the corresponding pollution levels in Szeged are discussed below.

Cluster 1: Comprises 6.1% of the total number of days. It is characterized by a high pressure system extending over Europe except Scandinavia and includes the Carpathian Basin. 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 is that most of these days belong to June. Therefore, the primary (CO, NO, NO_2, NO_2 vs NO and TSP except SO_2) and the secondary pollutants (O_3 and $O_{3\,max}$) have the lowest concentrations (Fig. 3a; Table 8).

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 weaken. Wind speed is the lowest of the all clusters. The concentration of pollutants increases, apart from SO_2, while the levels of NO reach highest values (Fig. 3a; Table 8).

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 the pressure pattern in *Cluster 2*. During this air mass type, the level of CO increases, while the concentration of SO_2 falls (Fig. 3a; Table 8).

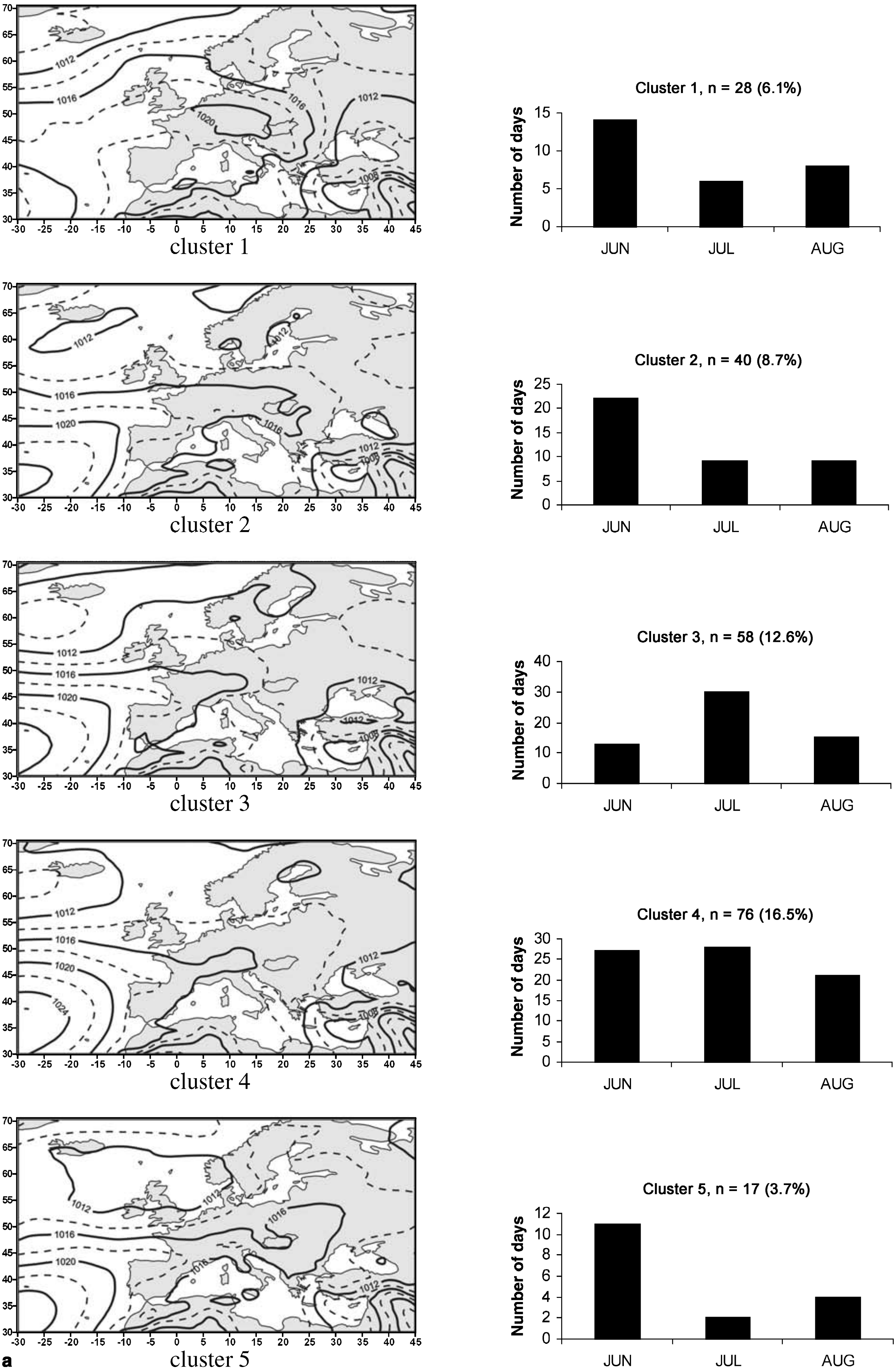
Cluster 4: This is the most frequent type with 16.5% of the total number of days and is characteristic of each summer month. Its pressure pattern is very similar to that of *Cluster 3*. The only basic difference is that the extensive low pressure system in Northern Europe is absent in this cluster. The levels of CO decrease, while a substantial decrease in cloudiness causes a slight increase in O_3 concentration. (NO levels practically do not change compared to those in *Cluster 3*.) (Fig. 3a; Table 8). The fact that a substantial decrease in cloudiness causes only a slight increase in O_3 concentration, might be explained by the change of transport processes. Namely, a decrease in cloudiness might be the result of change in circulation, which transports less ozone over Szeged. On the other hand, the smaller ozone concentration is only partly compensated by the photochemical processes accelerated by increased irradiance. Long-range transport can also influence the local ozone concentration and, in this way, the rate of local ozone formation depending little on local irradiance.

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 high, with cloudy conditions and moderate winds. There are not substantial differences in the concentration of pollutants compared to those of *Cluster 4* (Fig. 3a; Table 8).

Cluster 6: Typical late summer (10.0% of the total number of days). The high pressure system from Azores extends over Eastern Europe and includes the Carpathian Basin. There are no weather fronts in Northern Europe. Irradiance is very high, which involves high values of the temperature variables. On the other hand, wind speed is

→

Fig. 3a, b. Mean sea-level pressure fields for each air mass type (cluster), and monthly variation of the number of days, North-Atlantic – European region, summer months (June, July and August)


cluster 1
Cluster 1, n = 28 (6.1%)
Number of days
JUN
JUL
AUG
cluster 2
Cluster 2, n = 40 (8.7%)
cluster 3
Cluster 3, n = 58 (12.6%)
cluster 4
Cluster 4, n = 76 (16.5%)
cluster 5
Cluster 5, n = 17 (3.7%)
a

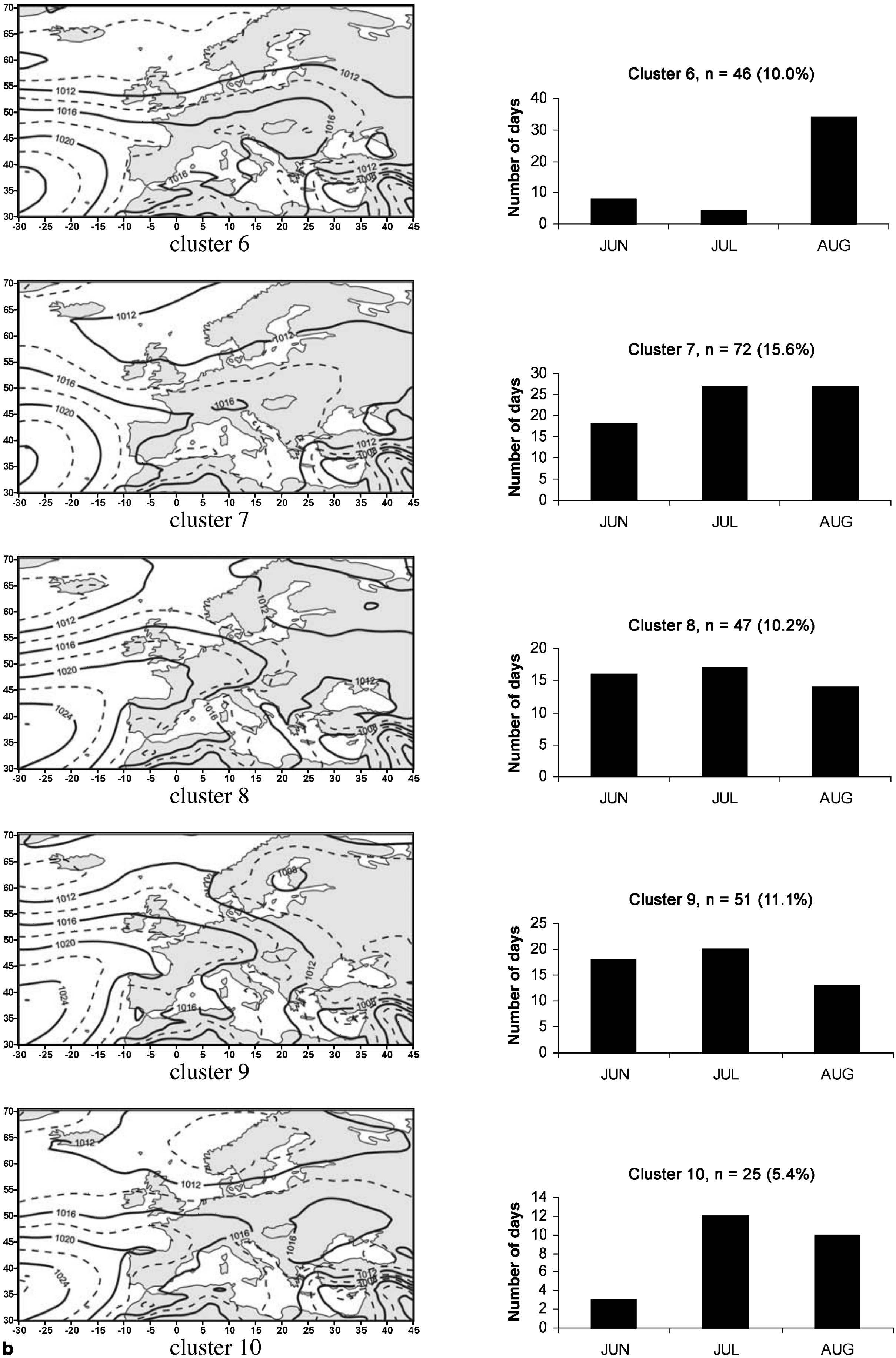

cluster 6
Cluster 6, n = 46 (10.0%)
Number of days
JUN
JUL
AUG
cluster 7
Cluster 7, n = 72 (15.6%)
Number of days
JUN
JUL
AUG
cluster 8
Cluster 8, n = 47 (10.2%)
Number of days
JUN
JUL
AUG
cluster 9
Cluster 9, n = 51 (11.1%)
Number of days
JUN
JUL
AUG
b
cluster 10
Cluster 10, n = 25 (5.4%)
Number of days
JUN
JUL
AUG

Table 9. χ^2-test of the mean sea-level pressure fields for the ten clusters of the summer months (June, July and August), probability of the 0-hypothesis

Cluster	1	2	3	4	5	6	7	8	9	10
1	–	1.0000	1.0000	1.0000	1.0000	0.9998	1.0000	1.0000	1.0000	0.9806
2		–	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3			–	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4				–	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5					–	1.0000	1.0000	1.0000	0.0113	1.0000
6						–	1.0000	1.0000	1.0000	1.0000
7							–	1.0000	1.0000	1.0000
8								–	1.0000	1.0000
9									–	1.0000
10										–

low. Hence, primary pollutants are highly enriched. Though both irradiance and NO concentration (having opposite effect on the levels of O_3 and $O_{3\,max}$) are higher than in *Cluster 5*, higher weight of irradiance is indicated by resulting in a slight increase of the secondary pollutants (Fig. 3b; Table 8).

Cluster 7: This is the second most frequent type with 15.6% of the total number of days. The high pressure centre from Azores is not present 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 pollutant levels practically do not change when compared to the former cluster (Fig. 3b; Table 8).

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 central 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, cloudiness increases which involves a decrease in the temperature parameters and wind speed reaches its maximum of all clusters. This is why both the primary and the secondary pollutant levels are so low (Fig. 3b; Table 8).

Cluster 9: Consists of 11.1% of the total number of days. The location 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 at the edge of the high pressure system. As the weather situation between *Clusters 8* and *9* is extremely similar, there are only minor differences in the meteorological parameters. Hence, no significant differences occur in levels of the pollutants (Fig. 3b; Table 8).

Cluster 10: Typical and late summer with 5.4% of the total number of days. The high pressure centre from the Azores weakens and gradually disappears. On the other hand, the low pressure pattern over Ukraine and Romania, presented in *Cluster 9*, disappears and a small high pressure centre forms in its place, whilst an 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 (Fig. 3b; Table 8).

Table 10a. ANOVA statistics for inter – air mass comparison of pollutant concentrations in the summer months (June, July and August)

	CO	NO	NO_2	NO_2/NO	O_3	$O_{3\,max}$	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, %	99	99	99	96	99	99	98	99

Similarly to the winter months, the significance of inter – air mass type differences in pollutant levels was determined by ANOVA. The results are shown in Table 10a. Mean concentrations of CO, NO, NO_2, O_3, $O_{3\,max}$ and TSP present significant inter – air mass type differences at the 99% probability level, while SO_2 at the 98% and NO_2/NO at the 96% levels, respectively. Performing the pairwise comparisons (Tukey's significant difference tests), the statistically significant differences are shown in Table 10b at the 95% and 99% 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, $O_{3\,max}$ and TSP); types 1–10 (CO, NO_2, O_3, $O_{3\,max}$, and TSP); types 2–10 (CO, O_3, $O_{3\,max}$ and TSP); types 3–6 (NO_2, O_3, SO_2 and TSP); types 3–10 (CO, O_3, $O_{3\,max}$ and TSP); types 4–6 (CO, NO_2, $O_{3\,max}$ and TSP); types 4–10 (CO, O_3, $O_{3\,max}$ and TSP); types 6–8 (CO, NO, NO_2, $O_{3\,max}$, and TSP); types 6–9 (CO, NO, NO_2 and TSP); types 8–10 (CO, NO_2, O_3, $O_{3\,max}$ and TSP) and types 9–10 (CO, NO_2, O_3, $O_{3\,max}$ and TSP). In general, air mass type 6 and type 10 differ mostly from the others, since the pairwise multiple comparisons detected significant differences

Table 10b. 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 (light-faced characters: 95% of significance, **bold** characters: 99% 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** NO_2vsNO		**CO** NO **NO_2** **TSP**	**CO** NO_2vsNO **TSP**		
10	**CO** **O_{3max}** **NO_2** **O_3** **TSP**	CO **O_{3max}** **O_3** **TSP**	CO **O_{3max}** **O_3** **TSP**	**CO** **O_{3max}** **O_3** **TSP**	CO **O_3** **TSP**	O_3	O_{3max} **O_3**	**CO** **O_{3max}** **NO_2** **O_3** **TSP**	**CO** **O_{3max}** **NO_2** **O_3** **TSP**

for them in levels of the most pollutants. The reason for this 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 pair-wise differences than the others.

6. Discussion

In order to assess the effect of different air mass types on air pollutant levels in Szeged, objective multivariate statistical methods were applied to meteorological and air pollution data. By defining objective pressure patterns over the North-Atlantic – European region, air mass types dominating the Carpathian Basin were determined.

The procedure itself has been applied in the literature (Sindosi et al, 2003); however, this is a new approach for classifying air mass types in the region examined, since for Hungary only the subjective categorization of the prevailing pressure patterns made by Péczely (1957, 1983) is known. The basis of his classification is the same as with the objective categorization: daily sea-level pressure fields measured at 00 UTC. Péczely determined 13 large-scale weather situations for the Carpathian Basin. As regards the winter months, four groups of the Péczely's macrotypes are characteristic over the Carpathian Basin: (1) types connected with southerly currents, (2) an anticyclone extending from the west, (3) an anticyclone in the north from Hungary and (4) an anticyclone over the Carpathian Basin. These weather types account for more than 70% of the total number of days in this season. On the other hand, the five objective clusters detected in this paper for winter are basically characterized by zonal currents (87.5% of the total number of days). In more detail they are as follows: an anticyclone south of Hungary (*Clusters 2* and *4*), an anticyclone extending from the west (*Cluster 3*) and a zonal cyclonic type (*Cluster 5*). These types are completed with an anticyclonic cluster over the Carpathian Basin (*Cluster 1*) (12.5% of the total number of days). Concerning the summer months, four Péczely types are emphasized: (1) Hungary lies in the rear part of an East-European cyclone, (2) an anticyclone extending from the west, (3) an anticyclone north of Hungary and (4) an anticyclonic type over the Carpathian Basin. These air mass types comprise a total of more than 60% of the total number of days. At the same time, the ten objective clusters for summer are determined mainly by the following groups: an anticyclone extending from the west (*Clusters 2–5*, *7–9*) (zonal currents), an anticyclone over the Carpathian Basin (*Clusters 1* and *6*) and an anticyclone east of Hungary (*Cluster 10*) (meridional currents). Predominance of the anticyclonic and anticyclonic ridge situations in the summer is very clear both in the Péczely types and in the objective clusters, too.

The air mass types determined for the winter and summer months were also related to levels of air pollutants in Szeged downtown. Relation of the objectively determined air mass types and air quality of Szeged detected that pollution levels can be connected to different pressure patterns ruling the region examined. Hence, with knowledge based on weather forecasts, expected pollution levels can be indicated, which contributes to the abatement of severe air pollution episodes. However, it has to be underlined that atmospheric circulation is not the only factor controlling the air pollution in Szeged. The revealed pressure patterns can only influence the concentration of the pollutants, which in their vast majority are of human origin. Thus, for a precise air pollution forecast, apart from a good weather forecast, a good knowledge of human activities is necessary. For example, days with traffic peaks, days of vacation or holidays must also be considered before certain restrictions on the emissions are imposed. Finally, another factor, which cannot be ignored either, is weather persistence. It must be kept in mind that the presence of pressure patterns favouring high concentrations of pollutants for a long period of time may cause even worse air quality conditions.

7. Conclusions

The paper analyzed the levels of air pollutants in Szeged recorded under characteristic sea level pressure patterns over the Carpathian Basin. Specific air mass types given by the pressure patterns for both the winter and summer months were found to play a significant role in the concentration of pollutants in downtown Szeged. Results for the winter months revealed that primary pollutants appear in higher concentration when both cloudiness and wind speed are low (air mass types 1 and 4; Fig. 2, Table 3). This is the case when an anticyclone is found over the Carpathian Basin (*Cluster 1*) and when an anticyclone rules the region south of Hungary influencing the 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 in the summer months are not as easily grouped into clusters 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 conditions and anticyclonic ridge types. Due to low cloudiness 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 the development of extreme concentrations.

Acknowledgements

The authors thank the Department of Analysis and Methodology, Hungarian Meteorological Service for providing the sea-level pressure data for the investigated period, Gábor Motika (Environmental Protection Inspectorate of Lower-Tisza Region, Szeged, Hungary) for handling meteorological and air pollution data, László Haszpra and László Horváth (Hungarian Meteorological Service) for valuable advice on ozone related processes and Áron Deák József for advice on plant ecology. This study was supported by the Hungarian National Foundation for Scientific Research (OTKA No. T 034765).

References

Adamopoulos AD, Kambezidis HD, Zevgolis D (2002) Case studies of total NO_2 column in the atmosphere of Athens, Greece: comparison between summer and winter. Fresen Environ Bull 11: 484–487. Sp. Iss. SI

Anderberg MR (1973) Cluster analysis for applications. New York: Academic Press, 353 pp

Bartzokas A, Metaxas DA (1993) Covariability and climatic changes of the lower troposphere temperatures over the Northern Hemisphere. Nouvo Cimento C 16C: 359–373

Bartzokas A, Metaxas DA (1995) Factor analysis of some climatological elements in Athens, 1931: Covariability and climatic change. Theor Appl Climatol 52: 195–205

Borhidi A (1995) Social behaviour types, the naturalness and relative ecological indicator values of the higher plants in the Hungarian Flora. Acta Bot Hun 39: 97–181

De Nevers N (2000) Air pollution control engineering. Singapore: McGraw-Hill, 2nd ed., 586 pp

Ellenberg H, Weber HE, Düll R, Wirth V, Werner W, Paulissen D (1991) Zeigerwerte von Pflanzen in Mitteleuropa. Scripta Geobot 18: Göttingen: Goltze-Verlag (in German)

Golder D (1972) Relations among stability parameters in the surface layer. Bound Layer Meteor 3: 47–58

Hair JF, Anderson RE, Tatham RL, Black WC (1998) Multivariate data analysis, 5th ed. New Jersey: Prentice Hall, 730 pp

Horváth F, Dobolyi ZK, Morschhauser T, Lökös L, Karas L, Szerdahelyi T (1995) Flora database 1.2, Vácrátót: Ecological and Botanical Institute of the Hungarian Academy of Sciences (in Hungarian)

Horváth Sz, Makra L, 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 Symp. on the Urban Environment and the 12th Joint Conf. on the Applications of Air Pollution Meteorology with the Air and Waste Management Association, Norfolk, VA, USA, 20–24 May 2002. Proc. J4.3: J58–J59

Jolliffe IT (1990) Principal component analysis: a beginner's guide – I. Introduction and application. Weather 45: 375–382

Jolliffe IT (1993) Principal component analysis: a beginner's guide – II. Pitfalls, myths and extensions. Weather 48: 246–253

Kalkstein LS, Tan G, Skindlov JA (1987) An evaluation of three clustering procedures for use in synoptic classification. J Climate Appl Meteorol 26: 717–730

Kambezidis HD (1996) Daylight levels in Athens during spring and autumn. Fresen Environ Bull 5: 61–66

Kambezidis HD, Tulleken R, Amanatidis GT, Paliatsos AG, Asimakopoulos DN (1995) Statistical evaluation of selected air-pollutants in Athens, Greece. Environmetrics 6: 349–361

Kambezidis HD, Weidauer D, Melas D, Ulbricht M (1998) Air quality in the Athens basin during sea breeze and non-sea breeze days using laser-remote-sensing technique. Atmos Environ 32: 2173–2182

Makra L, Horváth Sz (2001) Assessment of air pollution in Szeged. Légkör 46: 14–18 (in Hungarian)

Makra L, Kiss Á, Palotás J (1985) The spatial and temporal variability of drought in the southern part of the great hungarian plain. Acta Clim Univ Szegediensis 18–20: 65–85

Makra L, Horváth Sz, Zempléni A, Csiszár V, Rózsa K, Motika G (2001a) Air quality trends in southern hungary. EURASAP Newsletter 42: 2–13

Makra L, Horváth Sz, Zempléni A, Csiszár V, Fodré Zs, Bucsiné Kapocsi I, Motika G, Sümeghy Z (2001b) Analysis of air quality parameters in Csongrád county. Acta Clim Chorol Univ Szegediensis 34–35: 23–44

Makra L, Horváth Sz, Sümeghy Z (2002) An objective analysis and ranking of cities on environmental and social factors. In: IGU 2002. Geographical Renaissance at the Dawn of the Millennium. Durban, South-Africa, 2002. (Nkemdirim LC, ed), Climates in Transition Minuteman Press, pp 161–172

Makra L, Juhász M, Béczi R, Borsos E (2004a) The history and impacts of airborne Ambrosia (Asteraceae) pollen in Hungary. Grana 43: 1–8

Makra L, Juhász M, Borsos E, Béczi R (2004b) Meteorological variables connected with airborne ragweed pollen in Southern Hungary. Int J Biometeorol 49: 37–47

Mayer H, Makra L, Kalberlach F, Ahrens D, Reuter U (2004) Air stress and air quality indices. Meteorol Z 13: 395–403

McGregor GR, 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

Mohl M, Gaskó B, Horváth Sz, Makra L, Szabó F (2002) 2nd Environmental Programme of Szeged, 2003–2007. Manuscript (in Hungarian) (Mayor's Office, 6720 Szeged, Széchenyi tér 10, Hungary)

Pasquill F (1962) Atmospheric diffusion. London: Van Nostrand, 209 pp

Péczely G (1957) Grosswetterlagen in Ungarn. Kleinere Veröffentlichungen der Zentralanstalt für Meteorologie Budapest 30, 86 pp

Péczely G (1959) Air pollution of Budapest in different large scale weather situations. Idöjárás 63: 19–27 (in Hungarian)

Péczely G (1979) Climatology. Budapest: Tankönyvkiadó, 336 pp (in Hungarian)

Péczely G (1983) Catalogue of the macrosynoptic types for Hungary (1881–1983). Budapest: Hungarian Meteorological Service 53: 116 pp (in Hungarian)

Sindosi OA, Katsoulis BD, 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

Turner DB (1964) A diffusion model for an urban area. J Appl Meteor 3: 83–91

Authors' addresses: L. Makra (E-mail: makra@geo.u-szeged.hu), R. Béczi (E-mail: beczir@geo.u-szeged.hu), E. Borsos (E-mail: borsosemoke@yahoo.com), Z. Sümeghy (E-mail: sumeghy@geo.u-szeged.hu), Department of Climatology and Landscape Ecology, University of Szeged, 6701 Szeged, P.O. Box 653, Hungary; J. Mika (E-mail: mika.j@met.hu), Hungarian Meteorological Service, 1525 Budapest, P.O. Box 38, Hungary; A. Bartzokas (E-mail: abartzok@cc.uoi.gr), Department of Physics, Laboratory of Meteorology, University of Ioannina, 45110 Ioannina, Greece