Science of the Total Environment 458-460 (2013) 36–46

Contents lists available at SciVerse ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

The effect of different transport modes on urban PM_{10} levels in two European cities

László Makra [a,*], Ioana Ionel [b], Zoltán Csépe [a], István Matyasovszky [c], Nicolae Lontis [b], Francisc Popescu [b], Zoltán Sümeghy [a]

[a] Department of Climatology and Landscape Ecology, University of Szeged, HU-6701 Szeged, P.O. Box 653, Hungary
[b] Department of Mechanical Machines, "Politehnica" University, RO-300222 Timişoara, Bv. Mihai Viteazu, No. 1, Romania
[c] Department of Meteorology, Eötvös Loránd University, H-1117 Budapest, Pázmány Péter Street 1/A, Hungary

HIGHLIGHTS

- The 3D delimination of the clusters by the function "convhull" is a novel approach.
- For Bucharest, the most relevant source areas of PM10 transport are Central Europe with the Western Mediterranean.
- For Szeged, Southern and Central Europe are the most important sources of long-range transport of PM10.
- Occasional North-African-origin dust over Romania and Hungary is also detected, respectively.
- A statistical procedure is developed in order to separate medium- and long-range PM10 transport for both cities.

ARTICLE INFO

Article history:
Received 12 March 2013
Accepted 8 April 2013
Available online xxxx

Keywords:
PM_{10} transport
Backward trajectories
Cluster analysis
Mahalanobis metric
Separation of medium- and long-range PM_{10} transport

ABSTRACT

The aim of the study is to identify transport patterns that may have an important influence on PM_{10} levels in two European cities, namely Szeged in East-Central Europe and Bucharest in Eastern Europe. 4-Day, 6-hourly three-dimensional (3D) backward trajectories arriving at these locations at 1200 GMT are computed using the HYSPLIT model over a 5-year period from 2004 to 2008. A k-means clustering algorithm using the Mahalanobis metric is applied in order to develop trajectory types. Two statistical indices are used to evaluate and compare exceedances of critical daily PM_{10} levels corresponding to the trajectory clusters. For Bucharest, the major PM_{10} transport can be clearly associated with air masses arriving from Central and Southern Europe, as well as the Western Mediterranean. Occasional North African dust intrusions over Romania are also found. For Szeged, Southern Europe with North Africa, Central Europe and Eastern Europe with regions over the West Siberian Plain are the most important sources of PM_{10}. The occasional appearance of North-African-origin dust over Hungary is also detected. A statistical procedure is developed in order to separate medium- and long-range PM_{10} transport for both cities. Considering the 500 m arrival height, long-range transport plays a higher role in the measured PM_{10} concentration both for non-rainy and rainy days for Bucharest and Szeged, respectively.

1. Introduction

PM_{10} is a measure of particles in the atmosphere with a diameter of less than or equal to a nominal 10 μm. The 24-h limit value for PM_{10} (50 $\mu g \cdot m^{-3}$) is frequently exceeded in the urban environment. The short- and long-term human exposure to high particulate matter concentrations observed in urban environment increases the risk of respiratory (Schindler et al., 2009) and cardiovascular (Feng and Yang, 2012) diseases. For Bucharest, the predicted average gain in life expectancy for people of 30 years of age for a decrease in average annual PM2.5 level from 38.2 $\mu g \cdot m^{-3}$ (2004–2006) to 10 $\mu g \cdot m^{-3}$ (World Health Organization, 2000) is 22.1 months (Medina et al., 2004). For Szeged, PM_{10} counts a major pollutant influencing respiratory diseases (Matyasovszky et al., 2011); furthermore, a set of explanatory variables including PM_{10} indicates a strong association with allergic asthma emergency room visits (Makra et al., 2012). Therefore, studying potential key regions and long-range transport effects on urban PM_{10} levels is of great importance.

Several authors have published backward trajectory modeling results to help detect the long-range transport of pollutant air masses that may have an impact on local PM_{10} levels (Salvador et al., 2008), to better describe the related tropospheric circulations (Jorba et al., 2004) or to characterize and identify spatial and temporal trends of pollutants (Coury and Dillner, 2007). However, single backward trajectories generally applied to detect key regions of extreme PM episodes for given sites (Hongisto and Sofiev, 2004) are not suitable

* Corresponding author. Tel.: +36 62 544 856; fax: +36 62 544 624.
E-mail addresses: makra@geo.u-szeged.hu (L. Makra), ionel_monica@hotmail.com (I. Ionel), h480623@stud.u-szeged.hu (Z. Csépe), MATYA@ludens.elte.hu (I. Matyasovszky), lontis_nicolae@yahoo.com (N. Lontis), ingfrancisc@gmx.net (F. Popescu), sumeghy@geo.u-szeged.hu (Z. Sümeghy).

0048-9697/$ – see front matter
http://dx.doi.org/10.1016/j.scitotenv.2013.04.021

for an overall identification of paths and origins of air parcels that play a role in the contribution to PM_{10} levels.

Large numbers of trajectories arriving at a given site can be analyzed in order to determine the origin of polluted air masses. Several authors performed cluster analyses in order to place trajectories into a relatively small number of groups (Dorling et al., 1992; Dorling and Davies, 1995). Such procedures have been frequently used to interpret the origin and the transport of atmospheric pollution (Vardoulakis and Kassomenos, 2008; Li et al., 2012). Although cluster analysis alone does not tell us anything about the cause–effect relationships as no pre-determined characteristics are used to define the membership for a cluster, items in the same cluster are likely to have many features in common.

Based on backtrajectory analysis, source areas of long-range PM_{10} transport can be identified. Escudero et al. (2005) and Cabello et al. (2012) found African origin dust episodes over Eastern Spain. Karaca et al. (2009) established that the central part of Northern Africa (Northern Algeria and Libya) is the most significant potential PM_{10} contributors to Istanbul's atmosphere during springtime. Grivas et al. (2008) traced back some severe dust outbreaks in the air of Athens, Greece to the Sahara desert and the Western Mediterranean. Makra et al. (2011) detected an occasional appearance of North-Africanorigin dust even over Hungary, the middle latitudes of the temperate belt. Furthermore, according to their results an occasional Caspian Sea desert influence on particulate levels can also be identified in Northern Europe that is confirmed by findings of Hongisto and Sofiev (2004).

In spite of a vast amount of studies concerning PM_{10} transport only very few papers have been published for Eastern Europe. For instance, Konovalov et al. (2011) used a modified CHIMERE chemistry transport model in order to characterize the surface concentrations of PM_{10} over the Moscow region during the 2010 heat wave. Salvador et al. (2010) and Niemi et al. (2005) found source areas of different pollutants over the European territory of Russia using backtrajectory analysis. Another antecedent work analyzed transport effects on urban PM_{10} levels for three cities including Szeged, Hungary along a north–south axis in Europe (Makra et al., 2011). The present paper, however, involves an analysis for determining source areas of long-range PM_{10} transport in two European cities located at similar latitudes.

Therefore, the aim of this paper is to identify the key geographical regions responsible for PM_{10} levels in two cities (Bucharest, Eastern Europe; and Szeged, East-Central Europe, Fig. 1). Backward trajectories arriving at these sites are clustered using the Mahalanobis metric in order to determine which regions imply high PM_{10} concentrations. The clustering is performed using three-dimensional (3D) backward trajectories. ANOVA is used to determine whether PM_{10} concentrations corresponding to these trajectory clusters differ significantly. Cluster-dependent occurrences, when 24-h mean PM_{10} concentrations exceed the limit value of 50 $\mu g \cdot m^{-3}$ are also analyzed with two statistical indices. Lastly, a statistical procedure is developed in order to separate medium-range PM_{10} transport including local PM_{10} emissions from the long-range transport of PM_{10}.

2. Data and methods

2.1. Study areas and monitoring data

Five years (2004–2008) of daily mean PM_{10} data as well as daily meteorological data (mean temperature, mean global solar flux, mean relative humidity and daily precipitation total) taken from two European cities — Bucharest (Romania) and Szeged (Hungary) (Fig. 1; Table 1) — were analyzed. The reasons for selecting these sites include their fairly big distance (710 km) and their substantial difference in topography and climate. Namely, Bucharest ($\varphi = 44.43°N$; $\lambda = 26.10°E$; $h = 74$ m a.s.l.), the capital of Romania, is located in the southeast of the country. The city lies on the banks of the Dâmbovița River, about 70 km north of the Danube. Szeged ($\varphi = 46.25°N$; $\lambda = 20.10°E$; $h = 79$ m a.s.l.), the largest settlement in SE Hungary, is located at the confluence of the rivers Tisza and Maros.

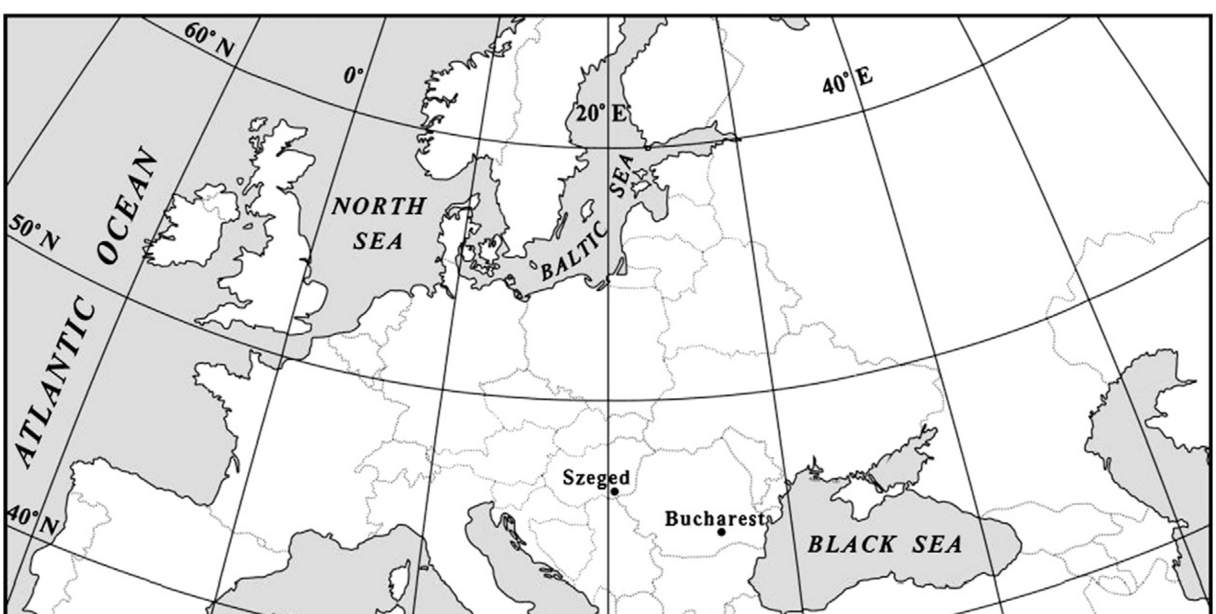

Fig. 1. The geographical positions of Bucharest and Szeged.

Table 1
Parameters of basic meteorological data and mean 24-hr PM_{10} concentrations.

City	Parameter	Jan	Jul	Year
Bucharest	Mean temperature (°C)	−1.0	22.6	11.0
	Precipitation total (mm)	46	72	655
	Mean 24-hr PM_{10} concentration ($\mu g \cdot m^{-3}$)	61.5	55.5	61.0
Szeged	Mean temperature (°C)	1.9	24.0	12.7
	Precipitation total (mm)	19	62	577
	Mean 24-hr PM_{10} concentration ($\mu g \cdot m^{-3}$)	54.9	34.4	43.0

2.2. Backward trajectories

In the frame of an ETEX (European Tracer Experiment) research, efficacy of three large-scale Lagrangian dispersion models (CALPUFF 5.8, FLEXPART 6.2 and HYSPLIT 4.8) was compared. As the HYSPLIT model has the best performance according to four statistical scores (Anderson, 2008) we decided to use the HYSPLIT model (Draxler and Hess, 1998).

Backward trajectories for Szeged and Bucharest corresponding to the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT, version 4.8; http://www.arl.noaa.gov/ready/hysplit4.html) model (Draxler and Hess, 1998) were obtained from the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR; http://dss.ucar.edu/datasets/ds090.0/).

Since a single backward trajectory has a large uncertainty and is of limited significance (Stohl, 1998), a three-dimensional (3D) representation of the synoptic air currents in the given regions was made via the reconstruction and analysis of a large number of atmospheric trajectories. 4-Day, 6-h 3D backward trajectories arriving at the two locations at 1200 Greenwich Mean Time (GMT) at heights $h = 500$, 1500 and 3000 m above ground level (a.g.l.) for each day over a 5-year period from 2004 to 2008 were taken in order to describe the horizontal and vertical movements of an air parcel for the above-mentioned two cities. These three arrival heights were selected in order to understand the behavior of the air masses circulating in the boundary layer (BL) and the free troposphere (FT): 500 m (typical for the near surface), 1500 m (representative for the BL top) and 3000 m (characteristic for the FT heights) (Córdoba-Jabonero et al., 2011). The actual heights of trajectories may act as an indicator of the atmosphere–surface interactions. For instance, an air mass moving over a source area at low vertical levels might be more affected by PM_{10} loadings of this region than another air mass traveling at much higher levels over this same area.

2.3. Cluster analysis

Cluster analysis classifies trajectories with similar paths. The aim of any clustering technique is to maximize the homogeneity of elements (in our case, backward trajectories) within the clusters and also to maximize the heterogeneity among the clusters. Here a non-hierarchical cluster analysis with the k-means method (Anderberg, 1973) was applied using the Mahalanobis metric (Mahalanobis, 1936) available in MATLAB 7.5.0. Input data as clustering variables include the 6-hourly co-ordinate values (φ — latitude, λ — longitude and h — height above ground level (a.g.l.)) of the 4-day 3D backtrajectories for both cities and the three given heights.

The homogeneity within clusters was measured by RMSD defined as the sum of the root mean square deviations of cluster elements from the corresponding cluster center over clusters. As the RMSD will usually decrease with an increasing number of clusters this quantity is not very useful for deciding about the optimal number of clusters. However, the change of RMSD (CRMSD) versus the change of cluster numbers, or rather the change of CRMSD (CCRMSD) is much more informative. Here, working with cluster numbers from 15 to 1, an optimal cluster number was selected so as to maximize the change in CRMSD. The rationale behind this approach is that the number of clusters producing the largest improvement in cluster performance compared to that for a smaller number of clusters is considered optimal.

The results of our cluster analysis are discussed and presented only for the lowest ($h = 500$ m a.g.l.) arrival height because backtrajectories at this arrival height are expected to have the largest influence on the PM_{10} concentration of the target site. The separation of the backward trajectory clusters and preparation of figures for clusters of backward trajectories were performed using a novel approach that employs a function called “convhull”. The algorithm (qhull procedure; http://www.qhull.org) gathers the extreme trajectory positions (positions farthermost from the center) belonging to a cluster, which are then enclosed. Specifically, the procedure creates the smallest convex hull with minimum volume covering the backtrajectories of the clusters (Preparata and Hong, 1977).

Borge et al. (2007) used a two-stage clustering procedure. They observed that the original one-stage cluster analysis including all trajectories was strongly influenced by the trajectory length. Long trajectories representing fast-moving air masses were highly disaggregated, even though they often came from the same geographical region. Many short trajectories representing slow-moving air masses, however, were grouped together, although they came from very heterogeneous regions. Therefore, only the short trajectories were reanalyzed by identifying new clusters (second stage). However, a second-stage analysis is not necessary if the metric in the clustering procedure is non-Euclidean. The problem of justifying the two steps vanishes when a Mahalanobis metric is used. The issue of a two-stage cluster analysis (Borge et al., 2007) arises from different standard deviations of the co-ordinates of the trajectory points being far and near in time. In order to demonstrate the role of different standard deviations, let us take a difference of 200 km in the position of a given trajectory point. Such a difference some 1500 km from us seems relatively insignificant, while the same difference is considered very large when close to the arriving point of the trajectory.

Trajectory clusters are projected on a stereographic polar plane supported by HYSPLIT (Taylor, 1997).

2.4. Analysis of variance (ANOVA)

ANOVA is used to test whether the means of PM_{10} values under different trajectory types (clusters) differ significantly for a given city. If ANOVA, based on the F-test, detects significant difference among these means another test is then applied to determine which means differ significantly from the others. Significant differences among mean PM_{10} concentrations under different trajectory types may tell us about the origin and transport of air masses on local PM_{10} levels. There are several versions available for comparing means calculated from subsamples of a sample. A relatively simple but effective way is to use the Tukey test. It performs well in terms of both the accumulation of first order errors of the test and the test power (Tukey, 1985).

ANOVA assumes in general that elements of the entire data set represented as random variables are independent, and elements within each group have identical probability distributions. Daily PM_{10} data, however, do not meet these requirements as they have an annual trend in both the expected value and variance. These trends can be removed by standardization. Standardized data are free of annual trends and thus distinguishing between average PM_{10} levels corresponding to trajectory types is due to the types themselves and it is not related to periods of the year. (Note that the standardized PM_{10} values are dimensionless.) The annual trend of the expected value is estimated by fitting sine and cosine waves with periods of one year and half a year to PM_{10} data by the least squares technique. Note that the half a year period is introduced to describe the temporal asymmetry of the annual trend. A subtraction of the estimated trend from data results in centralized data. The annual trend of the variance is estimated by fitting sine and cosine waves with periods of one year and half a year to squared centralized data. Lastly, centralized data are divided by the root of the estimated time-dependent variance in order to get standardized data. Consecutive daily PM_{10} values are correlated and produce higher variances of the means estimated under trajectory types compared to those for uncorrelated data. The autocorrelation structure is modeled via first order autoregressive (AR) processes conditioned on clusters. The classical Tukey-test (Tukey, 1985) is then modified according to the variances of estimated means obtained with the help of the AR models. Note that the omission of this step could systematically overestimate the number of significantly different means.

2.5. Factor analysis and special transformation

Factor analysis (FA) identifies linear relationships among subsets of examined variables, which helps to reduce the dimensionality of the initial database without substantial loss of information. First, a factor analysis was applied to the initial standardized data set consisting of 12 variables (3 climatic and 9 trajectory variables introduced in Section 4) in order to reduce the original set of variables to fewer variables. These new variables called factors can be viewed as the main climate/trajectory features that potentially influence the daily mean PM_{10} concentration. The optimum number of retained factors is determined by the criterion of reaching a prespecified percentage of the total variance (Jolliffe, 1993). This percentage value was set at 80% in our case. Next, a further data manipulation on the retained factors called special transformation (Fischer and Roppert, 1965; Jahn and Vahle, 1968) was performed to discover to what degree the above-mentioned explanatory variables (3 climatic and 9 trajectory variables) affect the resultant variable (daily mean PM_{10} concentration), and to give a rank of their importance.

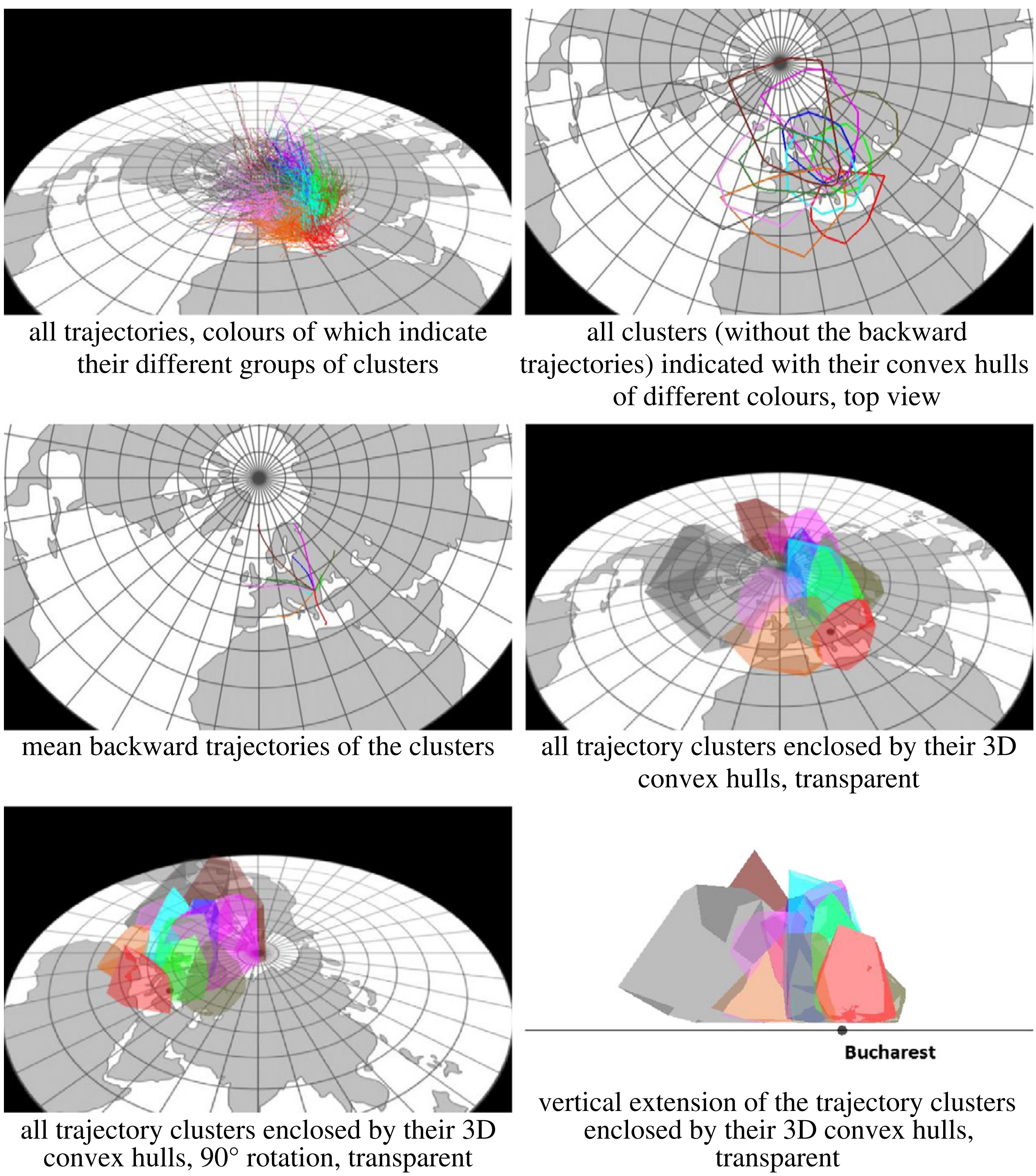

Fig. 2. 3D clusters of the backward trajectories retained, Bucharest, $h = 500$ m.

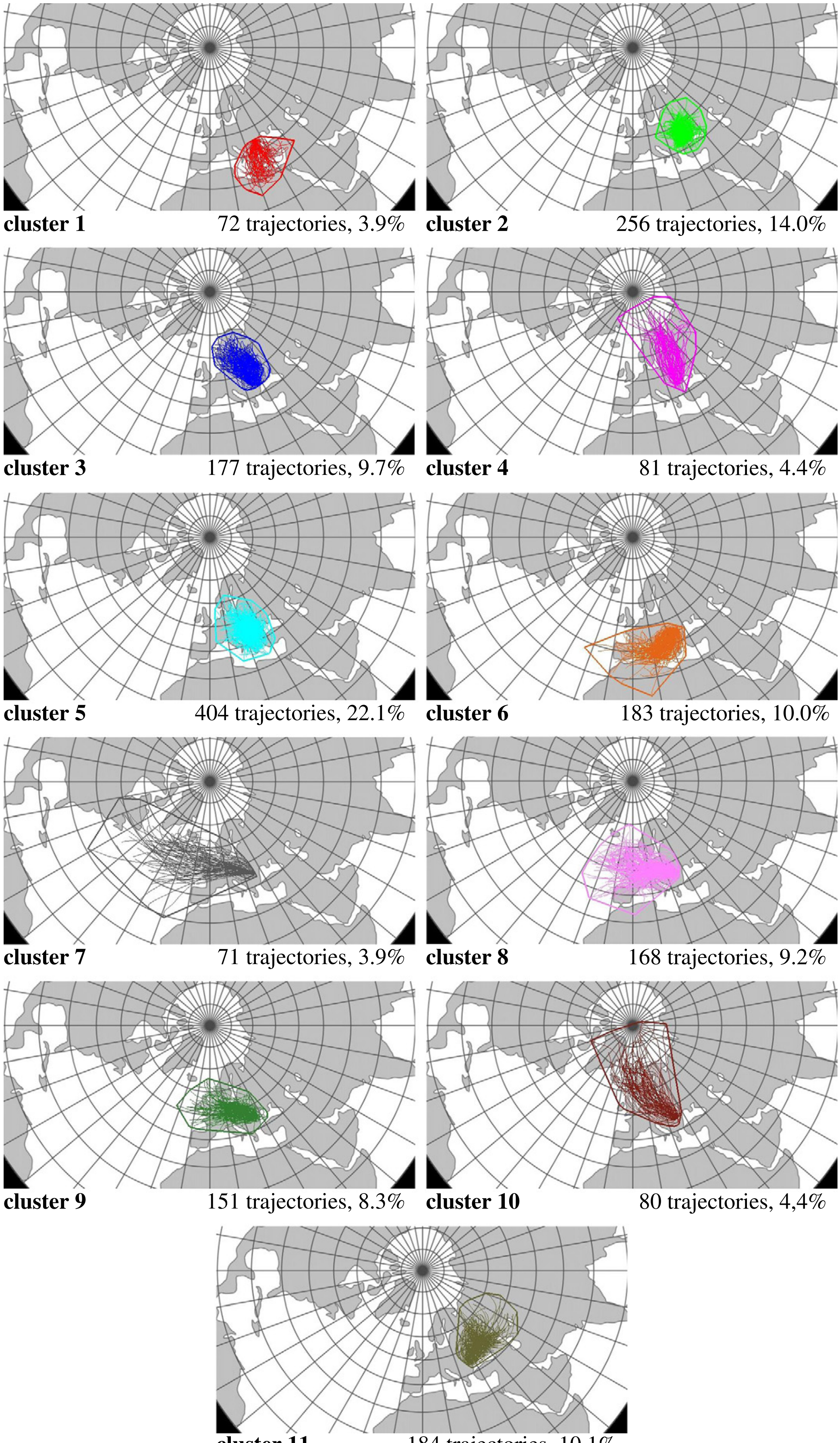

Fig. 3. The individual clusters of the backward trajectories retained, enclosed by their convex hulls, Bucharest, top view, $h = 500$ m.

Table 2
The individual clusters with the name of the source regions and their standardized average PM_{10} concentrations for both cities, $h = 500$ m (bold: maximum; italic: minimum). (The standardized PM_{10} values are dimensionless.)

Cluster no.	Bucharest		Szeged	
	Name of the source region	PM_{10} level	Name of the source region	PM_{10} level
1	South-eastern Europe with North Africa	−0.01	Southern Europe with North Africa	0.53
2	Eastern European plain	0.01	Northern Europe	−0.24
3	North-Western Europe	−0.10	North-western Europe	0.02
4	Northern Europe	−0.17	Western and Southern Europe	0.09
5	Central and Southern Europe	0.21	Mid-Atlantic – South-western Europe	−0.05
6	Southern Europe and the Western Mediterranean	0.20	North-Atlantic – Western Europe	−0.15
7	Western Europe with the North Atlantic	−0.23	Northern Mid-Atlantic – North-western Europe	−0.40
8	Western Europe with the Northern Mid-Atlantic	−0.14	Arctic – North-western Europe	−0.42
9	Western Europe	−0.06	Central Europe	0.25
10	North-western Europe with the Arctic	−0.33	Eastern Europe with regions over the West Siberian Plain	0.31
11	Eastern Europe with regions beyond the Ural mountains	−0.07		

2.6. Statistical characterization of PM_{10} exceedance episodes

The role of long-range transport is studied by analyzing the cluster occurrences on days when 24-h mean PM_{10} concentrations exceed the limit value of 50 $\mu g \cdot m^{-3}$. Two statistical indices of daily PM_{10} exceedance episodes associated with trajectory clusters are calculated in the same manner as in Borge et al. (2007). For a given site and cluster i, INDEX1 is defined as

$$\mathrm{INDEX1}_i(\%) = \frac{D_{(>50)i} \cdot 100}{D_i}, \quad (1)$$

where D_i is the number of occurrences of cluster i, and $D_{(>50)i}$ is the number of 24-h PM_{10} exceedances. INDEX1 gives the likelihood of an exceedance for a given cluster. INDEX2 is defined as

$$\mathrm{INDEX2}_i(\%) \frac{D_{(>50)i} \cdot 100}{E}, \quad (2)$$

where E is the total number of 24-h PM_{10} exceedance days recorded at a given site. INDEX2 can be interpreted as the likelihood of certain trajectory being present on a PM_{10} exceedance day.

3. Results

3.1. Bucharest

The 3D clustering produced eleven clusters based on CCRMSD. All of the trajectories with color-coded clusters, all of the clusters without trajectories but with their 3D convex hulls for the top view, in addition with the mean backward trajectories of the clusters for the top view, and all trajectory clusters enclosed by their transparent 3D convex hull as well as their 90° rotated version are presented in Fig. 2. A vertical view of the trajectory clusters enclosed by their transparent 3D convex hulls has also been added (Fig. 2). The individual clusters (Fig. 3) with the name of the source regions and their standardized average PM_{10} concentrations are tabulated (Tables 2 and 3). Pairwise comparisons of the cluster averages found 5 significant differences among the possible 55 cluster pairs (9.1%) (Table 4). Hereafter only clusters of the above-mentioned significantly different cluster-averaged PM_{10} levels were then considered and analyzed (Fig. 3; Table 4).

Clusters 5 and 6 have the highest INDEX1 values, namely 70.1% and 74.2%, respectively (Fig. 6). The high INDEX1 values of these clusters (Fig. 6) are in agreement with their high mean PM_{10} concentrations (Table 3). The high standard deviation corresponding to cluster 5 (Table 3) implies a higher chance of extreme PM_{10} episodes for this cluster. Note that INDEX1 and INDEX2 are not independent parameters. When a cluster is frequent, a high INDEX1 value involves a high INDEX2 value (see cluster 5) (Fig. 6). The highest frequency of cluster 5 and a relatively high occurrence of cluster 6 (having the 4th highest frequency of the clusters) emphasize the importance of these clusters in PM_{10} transport. The low-moving backward trajectories of cluster 6 further raise the significance of this cluster in long-range transport. Cluster 5 comprises high-moving backtrajectories that weakens the role in transporting particulates (Fig. 2, upper left panel, middle right panel, as well as the lower left and right panels; Fig. 3; Table 3).

The above analysis shows that clusters 5 and 6 play the largest role in PM_{10} transport to Bucharest. A substantial part of cluster 5 (Central and Southern Europe) and the whole cluster 6 (Southern Europe and the Western Mediterranean) cover arid regions with a negative water balance. These environments are favorable for turbulent air currents to take up and transport particles contributing to the observed PM_{10} exceedances. The role of the above two clusters in PM_{10} transport is confirmed by their high mean PM_{10} levels (Table 3), high frequency (Fig. 3; Table 3), high percentage of low-moving backtrajectories (except for cluster 5) (Fig. 2, upper left panel, middle right panel, as well as the lower left and right panels; Fig. 3; Table 3) and high INDEX1 and INDEX2 values (Fig. 6). Primarily cluster 1, and partly cluster 6 indicate occasional North African dust intrusions over Romania, which is confirmed by Saharan dust episodes in Hungarian aerosol detected over higher latitudes than Bucharest (Borbély-Kiss et al., 2004; Koltay et al., 2006). Clusters 7 (Western Europe with the North Atlantic) and 10 (North-western Europe with the Arctic) have the lowest mean PM_{10} levels; they are both infrequent and include mostly high-moving air masses (Fig. 2, upper left panel, middle right panel, as well as the lower left and right panels; Fig. 3; Table 3), which is consistent with the finding of Makra et al. (2011) that the transport of particulate matter from Northern and North-western Europe to East-Central Europe is of limited importance (Fig. 3; Tables 2 and 3).

3.2. Szeged

Ten clusters were retained in a 3D analysis based on CCRMSD (Figs. 4 and 5). The individual clusters (Fig. 5) with the name of the

Table 3
Parameters of standardized PM_{10} concentrations for the individual clusters, Bucharest, $h = 500$ m (bold: maximum; italic: minimum).

Cluster	1	2	3	4	5	6	7	8	9	10	11
Mean ($\mu g \cdot m^{-3}$)	−0.01	0.01	−0.10	−0.17	**0.21**	0.20	−0.23	−0.14	−0.06	*−0.33*	−0.07
Standard deviation ($\mu g \cdot m^{-3}$)	0.57	0.49	0.53	0.54	**1.86**	0.63	0.47	0.48	0.51	0.47	*0.45*
Number of trajectories	72	256	177	81	**404**	183	*71*	168	151	80	184
%	3.9	14.0	9.7	4.4	**22.1**	10.0	*3.9*	9.2	8.3	4.4	10.1

Table 4
Significant differences between the standardized cluster averages of PM_{10} concentrations, based on the Tukey test for Bucharest, h = 500 m (in *X*: significant at p < 0.05, in **X**: significant at p < 0.01).

	1									
2		2								
3			3							
4				4						
5			*X*		5					
6						6				
7					*X*		7			
8					**X**			8		
9									9	
10					**X**	**X**				10
11										

source areas and their standardized average PM_{10} levels are presented (Tables 2 and 5). For Szeged, 20 significant differences were detected among the possible 45 cluster pairs (44.4%) (Table 6).

The highest INDEX1 value (57.3%) is associated with cluster 1 (Southern Europe with North Africa) with relatively low frequency (7.2%) (Figs. 5 and 6; Table 5). The next highest INDEX1 values, in decreasing order, belong to cluster 10 (Eastern Europe with regions over the West Siberian Plain) (38.4%) and cluster 9 (Central Europe) (37.8%). This is in agreement with the fact that these clusters have high mean PM_{10} levels (Figs. 5 and 6; Table 5).

Accordingly, clusters 1, 9 and 10 are the most relevant in terms of PM_{10} transport to Szeged, which is confirmed by their highest mean PM_{10} levels (Table 5), their high frequency (except for clusters

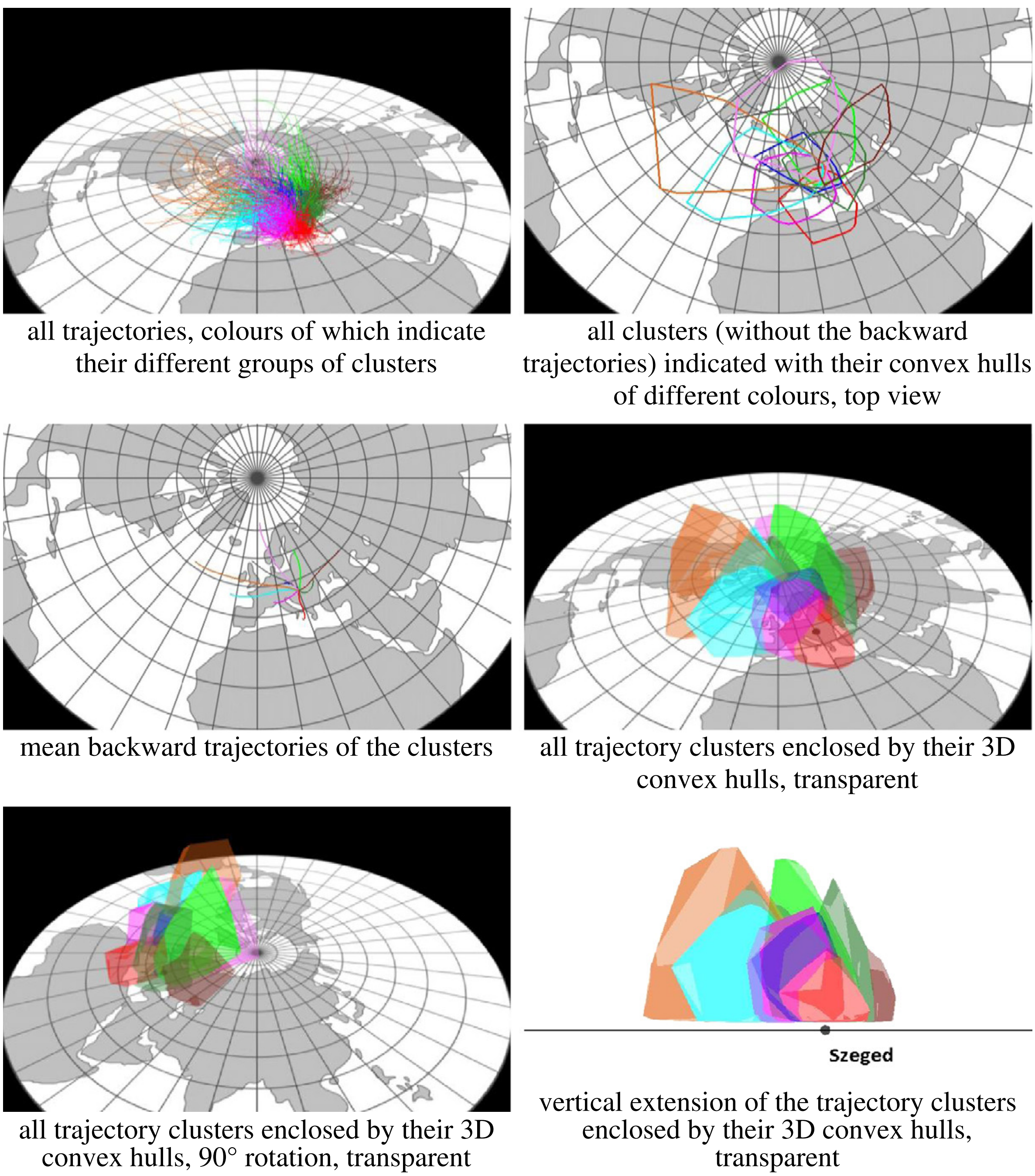

Fig. 4. 3D clusters of the backward trajectories retained, Szeged, h = 500 m.

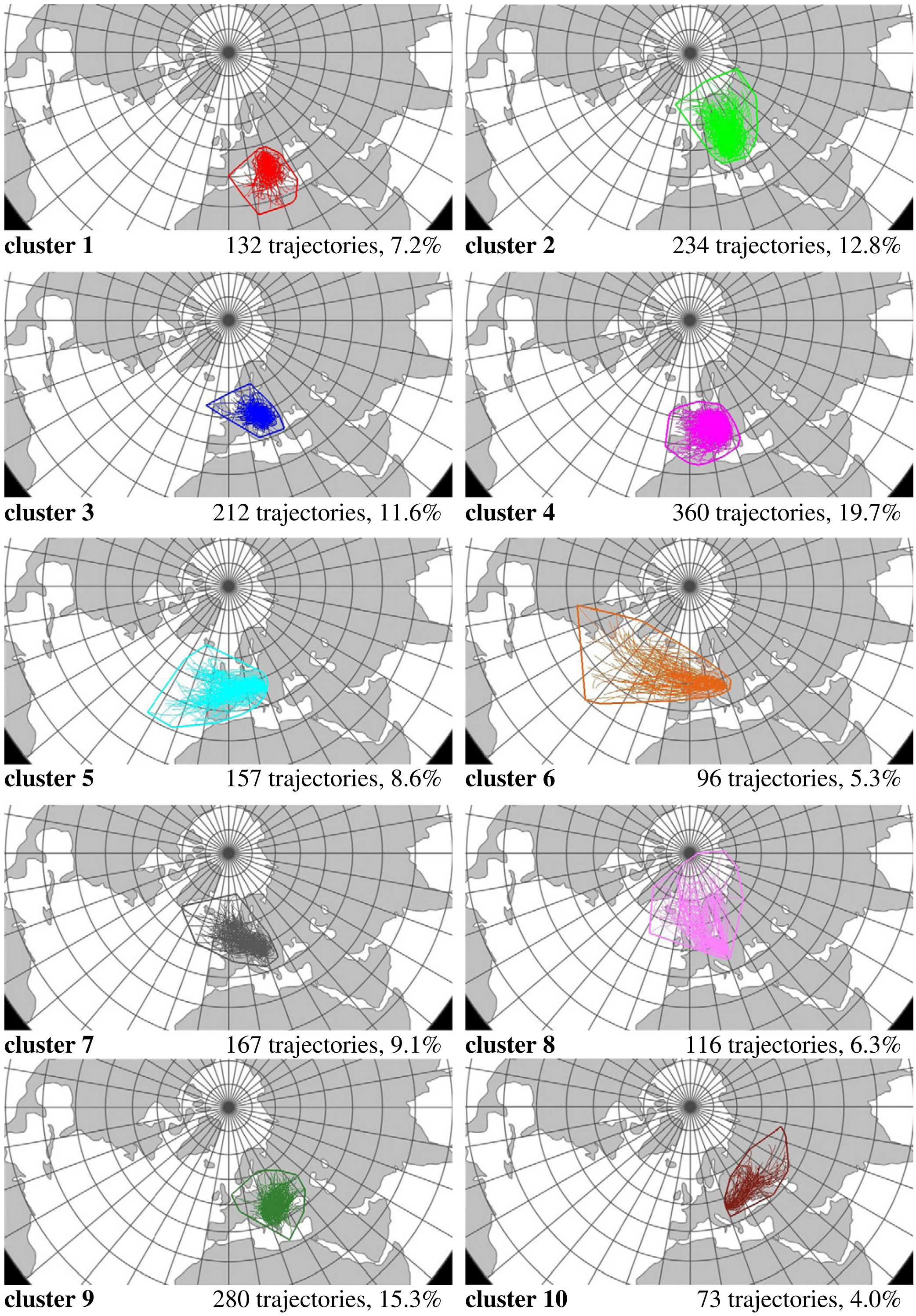

Fig. 5. The individual clusters of the backward trajectories retained, enclosed by their convex hulls, Szeged, top view, $h = 500$ m.

Table 5
Parameters of standardized PM_{10} concentrations for the individual clusters, Szeged, $h = 500$ m (bold: maximum; italic: minimum).

Cluster	1	2	3	4	5	6	7	8	9	10
Mean ($\mu g \cdot m^{-3}$)	**0.53**	−0.24	0.02	0.09	−0.05	−0.15	−0.40	*−0.42*	0.23	0.31
Standard deviation ($\mu g \cdot m^{-3}$)	0.92	0.83	0.90	1.08	1.16	0.82	*0.78*	0.79	0.98	**1.29**
Number of trajectories	132	234	212	**360**	157	96	167	116	280	*73*
%	7.2	12.8	11.6	**19.7**	8.6	5.3	9.1	6.3	15.3	*4.0*

Table 6
Significant differences between the standardized cluster averages of PM_{10} concentrations, based on the Tukey test for Szeged, h = 500 m (in *X*: significant at p < 0.05, in **X**: significant at p < 0.01).

	1								
2	**X**	2							
3	**X**		3						
4	**X**	**X**		4					
5	**X**				5				
6	**X**					6			
7	**X**		**X**	**X**	*X*		7		
8	**X**		**X**	**X**				8	
9		**X**				*X*	**X**	**X**	9
10		**X**					**X**	**X**	

1 and 10) (Fig. 5), high portions of low-moving backtrajectories (Fig. 4, upper left panel, middle right panel, as well as the lower left and right panels) and high INDEX1 and INDEX2 values (except for the low INDEX2 value of cluster 10) (Fig. 6). Cluster 1 (Southern Europe with North Africa), showing relatively low frequency, includes short backtrajectories and hence slow-moving air masses (Fig. 5; Table 2). The region covered by this cluster is generally dry, especially in spring, and wind erosion is frequently significant (e.g. the Hungarian Great Plain within the Carpathian Basin; Mezosi and Szatmari, 1998) creating a source of PM_{10}. An important part of this cluster is over North Africa showing that high PM_{10} exceedance episodes can occasionally be related to low-moving air masses coming from North Africa (Fig. 4, upper left panel, middle right panel, as well as the lower left and right panels; Fig. 5; Table 5) (Borbély-Kiss et al., 1999; Koltay et al., 2006). Among these three clusters, cluster 10 has the longest (fastest) backtrajectories. This cluster is only of minor importance due to its infrequent occurrence (Fig. 5; Table 5). Together with cluster 1 (Southern Europe with North Africa), cluster 9 (Central Europe) is the most important for transporting PM_{10} to Szeged. Consequently, Southern Europe with North Africa and Central Europe are the most important sources of PM_{10} for Hungary. Most of these regions, especially the Mediterranean, are warm and arid for a substantial part of the year making it easier to uplift and transport particulates to the target area (Fig. 5; Tables 2 and 5). Clusters 2 (Northern Europe), 7 (Northern Mid-Atlantic — North-western Europe) and 8 (Arctic — North-western Europe) are accompanied with the lowest PM_{10} levels, they occur rarely (except for cluster 2) and all three clusters comprise mostly high-moving air masses (Fig. 4, upper left panel, middle right panel, as well as the lower left and right panels; Fig. 5; Tables 2 and 5).

4. Discussion and conclusions

A cluster analysis was applied to 4-day, 6-hourly backward trajectories arriving at Bucharest and Szeged over a 5-year period in order to identify the main atmospheric circulation pathways influencing PM_{10} levels at these sites. When performing ANOVA, the decision on the significance of two cluster averages is based on a modified t-test because the test is performed using standardized data instead of the original data. The Mahalanobis metric was used in order to avoid the need for a two-stage cluster analysis introduced in Borge et al. (2007). The 3D delimination of the clusters by the function "convhull" is a novel approach. Furthermore, the presentation of vertical extension of the trajectory clusters was enclosed by their 3D convex hulls and, in this way, delimiting low-moving backtrajectories is a novel procedure. Furthermore, no papers have been published so far studying PM_{10} transport for Eastern European target stations using backward trajectories.

When determining important clusters that mainly influence PM_{10} levels, the following aspects were considered: 1) the average PM_{10} level of a given cluster should differ significantly from that of another cluster, 2) the average of the given cluster should be high, 3) the INDEX1 value and/or INDEX2 value of the given cluster should be high. Two other factors could be important, namely whether the given cluster has a high frequency and whether the given cluster has low-level backward trajectories.

For Bucharest, the major PM_{10} transport can be clearly associated with air masses coming from Central and Southern Europe (cluster 5), as well as Southern Europe and the Western Mediterranean (cluster 6). The importance of these clusters is justified by large regions that have a negative water balance in a substantial part of the year. Clusters 1 and 6 indicate occasional North African dust intrusions over Romania confirmed by Saharan dust episodes in Hungarian aerosol detected over higher latitudes than Bucharest (Borbély-Kiss et al., 2004; Koltay et al., 2006). Clusters 7 (Western Europe with the North Atlantic) and 10 (North-western Europe with the Arctic) have the lowest mean PM_{10} levels; both have low frequency and comprise mostly high-moving air masses, which is confirmed with the finding of Makra et al. (2011) that PM_{10} transport from Northern and North-western Europe to East-Central Europe is of limited importance.

For Szeged, clusters 1, 9 and 10 are the most relevant in PM_{10} transport. Cluster 1, corresponding to Southern Europe with North Africa, includes the occasional appearance of North African-origin dust over Hungary and corroborates earlier studies (Borbély-Kiss et al., 1999; Koltay et al., 2006). Though cluster 10 (Eastern Europe with regions over the West Siberian Plain) has high PM_{10} concentrations; it is generally of little importance due to its infrequent occurrence. Together with cluster 1, cluster 9 (Central Europe) is the most important for transporting PM_{10} to Szeged. Accordingly, Southern Europe with North Africa (cluster 1) as well as Central Europe (cluster 9) are the most important sources of PM_{10} over Hungary. Most of these regions, especially the Mediterranean, are warm and arid for a substantial part of the year making it easier to

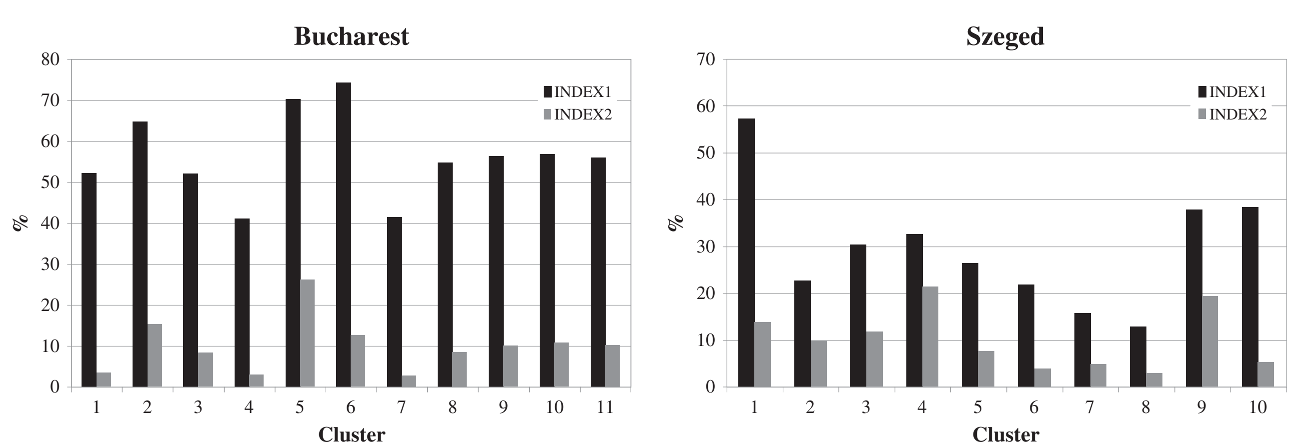

Fig. 6. Indices 1 and 2 for 3D clusters of the backward trajectories, h = 500 m.

uplift and transport particulates over the target area. Clusters 2 (Northern Europe), 7 (Northern Mid-Atlantic — North-western Europe) and 8 (Arctic — North-western Europe) are marked by the lowest PM_{10} levels, in accordance with the results of Makra et al. (2011) that these regions are of typical low PM_{10} level areas.

After classifying objective groups of backtrajectories and, in this way, detecting the main circulation pathways for the cities in question, it is important to separate local and transported components of the actual PM_{10} levels. In other words, it is necessary to determine the relative weight of these two components in the measured PM_{10} concentration. There are several case studies available that allow one to distinguish the long-range PM_{10} transport episodes from local PM_{10} pollution episodes (Escudero et al., 2006; Aarnio et al., 2008). Masiol et al. (2012) applied a chemometric analysis and a source apportionment model for discriminating local processes and long-range transport on particulate matter levels. Wong et al. (2013) discerned short- and long distance sources on the types of aerosols. Juda-Rezler et al. (2011) developed a combination of different methods to distinguish long-range transport and regional transport from local pollution sources. Analyses of local meteorological conditions and air-mass backtrajectories for a given city play an important role in developing methods for the above purpose (Aarnio et al., 2008). An attempt is made here to discriminate these two pollution modes (i.e. local PM_{10} emission and long-range PM_{10} transport) in the entire 5-year data set using local meteorological parameters and components of the backtrajectories. Local PM_{10} pollution is characterized next via the daily mean temperature, daily relative humidity and daily global solar flux. Long-range PM_{10} transport is described by (1) the real 3D length of the backtrajectories, (2) the length of the 3D backtrajectories as the crow flies, (3) their ratio, (4) the average daily highest and (5) lowest positions of the backward trajectories based on their 4-day, 6-hourly positions. In order to take into account further characteristics of long-range PM_{10} transport, stereographic plane projection of each backtrajectory is considered. The target station is located into the origin of an imaginary frame of reference. Further parameters of the long-range transport are as follows: x coordinates belonging to the (6) easternmost and the (7) westernmost points of the given backtrajectory, as well as the y coordinates belonging to the (8) northernmost and the (9) southernmost points of the same given backtrajectory. The average daily highest and lowest positions of the backward trajectories refer to the vertical transport of PM_{10} in the atmosphere, which comes from either turbulent transport dominating the vertical exchange of PM_{10} in the boundary layer or intense convective upwelling, which results in large amounts of particulates being transported from near the surface to high elevations (Ansmann et al., 2003). The latter four (6–9) characteristics represent the extreme points of a backward trajectory both to east–west and north–south directions on a horizontal plane, representing the east–west and north–south extension of the long-range transport.

As the PM_{10} level on a given day is substantially influenced by weather conditions such as precipitation, the backward trajectories are divided into two groups, i.e. non-rainy and rainy days of the arriving sites. This kind of classification of days reveals the role of precipitation in the quantity of transported PM_{10} (Querol et al., 2009). Factor analysis with special transformation was carried out for both cities with the two groups (rainy or non-rainy days) and the 500 m, 1500 m and 3000 m arrival heights of the backward trajectories, separately. Thus, altogether $2 \times 2 \times 3 = 12$ procedures gave information about the weights of the local source and long-range transport reflected by the 12 explanatory variables. The main conclusions are as follows.

Considering the 500 m arrival height, long-range PM_{10} transport plays a higher role compared to local PM_{10} emission both for non-rainy and rainy days for Bucharest and also for Szeged. The predominance of long-range transport compared to local emission is higher in Bucharest than in Szeged on non-rainy days, while it is equally higher for both cities on rainy days. As regards the components of the two different transport modes on non-rainy days, the local variables are equally important for both cities and all three heights. The components associated to the length of the backward trajectories have equally high weights for both cities, furthermore their east–west components have also substantial role for both cities and all three heights. In addition, the role of the north–south components is more important for Szeged. For rainy days, components of neither the local nor the long-range transport are important for Bucharest, while temperature and global solar flux as well as the east–west components of the long range transport are the most relevant for Szeged. Adding up the weights of the local pollution and long-range transport, the average value of the two weights is larger for both Bucharest and Szeged on non-rainy days and for Bucharest on rainy days at 500 m height compared to the higher levels. Hence, the twelve variables contain more information on PM_{10} when using backtrajectories arriving at 500 m height in these three cases. For the remaining case (Szeged on rainy days) the results (the average weight of the local pollution and long-range transport is the lowest at 500 height compared to the remaining two arrival heights) disagree with our preliminary expectations because near surface air currents might be affected by several factors that substantially modify the ratio of the local and transported particulates. Moreover, the variables contain more information at higher levels on the transported PM_{10} for Szeged on rainy days.

In a subsequent examination, another factor analysis with special transformation was performed for rainy and non-rainy days including collectively the 500 m, 1500 m and 3000 m arrival heights of the backward trajectories. In this way, altogether $2 \times 2 = 4$ procedures were implemented. Each procedure comprised 30 explanatory variables in all (3 parameters gave information about the weights of the local source, while the long-range transport was reflected by $3 \times 9 = 27$ explanatory variables, due to the three arrival heights of the backtrajectories). The major results are as follows.

For non-rainy days, components of the long-range transport for the 500 m arrival height have the most important role in determining PM_{10} concentration both for Bucharest and Szeged. Furthermore, global solar flux and relative humidity, as components of the medium-range PM_{10} transport including local PM_{10} emission, as well as the east–west components of the long-range transport for all three heights are within the first ten most important explanatory variables for both cities. For rainy days, only real 3D length of the backtrajectories at 3000 m height is in an important association with the PM_{10} concentration for Bucharest. For Szeged, both temperature and global solar flux have again an important role. In addition, east–west components of the backtrajectories at 500 m height, as well as average daily highest and lowest positions of the backward trajectories at both 1500 m and 3000 m heights are the most relevant explanatory variables.

For both kinds of factor analysis with special transformation, temperature and global solar flux are in significant negative, while relative humidity is in significant positive association with PM_{10} concentration. These associations assume an anticyclone ridge weather situation, when descending air currents prevent vertical mixing of the polluted urban and, hence, air pollution can accumulate. These situations with cloudy weather involve a decrease of temperature and global solar flux and an increase of relative humidity.

For both cities and all three heights, components of the backtrajectories are directly proportional to the resultant variable. Namely, bigger length as well as more extreme horizontal and vertical components of the backtrajectories involves higher PM_{10} concentrations.

Note that these findings are valid only for variations of the daily PM_{10} concentrations accounted for by the explanatory variables and nothing is known about the variance portion not explained by these variables.

Acknowledgments

The authors would like to thank Roland Draxler for his useful advice and consultations on the HYSPLIT model, version 4.8; to Gabriel Ciuiu (Environmental Protection Agency, Bucharest) and to Gábor Motika (Environmental Protection Inspectorate of Lower-Tisza Region, Szeged) for providing daily PM_{10} and meteorological data. This study was supported by the TRANSAIRCULTUR project (No. HURO/1001/139/1.3.4) Hungary-Romania Cross-Border Co-operation Programme 2007–2013, under the auspices of the European Union and the European Regional Development Fund.

References

Aarnio P, Martikainen J, Hussein T, Valkama I, Vehkamäki H, Sogacheva L, et al. Analysis and evaluation of selected PM_{10} pollution episodes in the Helsinki Metropolitan Area in 2002. Atmos Environ 2008;42:3992–4005.

Anderberg MR. Cluster analysis for applications. New York: Academic Press; 1973.

Anderson B. Evaluation methods for long range transport. 9th Modeling Conference Presentations. US Environmental Protection Agency, Technology Transfer Network Support Center for Regulatory Atmospheric Modeling; 2008.

Ansmann A, Bosenberg J, Chaikovsky A, Comeron A, Eckhardt S, Eixmann R, et al. Long-range transport of Saharan dust to northern Europe: the 11–16 October 2001 outbreak observed with EARLINET. J Geophys Res-Atmos 2003;108(D24). http://dx.doi.org/10.1029/2003JD003757. [Article No. 4783].

Borbély-Kiss I, Koltay E, Szabó GY, Bozó L, Tar K. Composition and sources of urban and rural atmospheric aerosol in eastern Hungary. J Aerosol Sci 1999;30:369–91.

Borbély-Kiss I, Kiss ÁZ, Koltay E, Gy Szabó, Bozó L. Saharan dust episodes in Hungarian aerosol: elemental signatures and transport trajectories. J Aerosol Sci 2004;35(10): 1205–24.

Borge R, Lumbreras J, Vardoulakis S, Kassomenos P, Rodríguez E. Analysis of long-range transport influences on urban PM_{10} using two-stage atmospheric trajectory clusters. Atmos Environ 2007;41:4434–50.

Cabello M, Orza JAG, Barrero MA, Gordo E, Berasaluce A, Cantón L, et al. Spatial and temporal variation of the impact of an extreme Saharan dust event. J Geophys Res-Atmos 2012;117. http://dx.doi.org/10.1029/2012JD017513. [Article No. D11204].

Córdoba-Jabonero C, Sorribas M, Guerrero-Rascado JL, Adame JA, Hernández Y, Lyamani H, et al. Synergetic monitoring of Saharan dust plumes and potential impact on surface: a case study of dust transport from Canary Islands to Iberian Peninsula. Atmos Chem Phys 2011;11:3067–91.

Coury C, Dillner AM. Trends and sources of particulate matter in the Superstition Wilderness using air trajectory and aerosol cluster analysis. Atmos Environ 2007;41: 9309–23.

Dorling SR, Davies TD. Extending cluster analysis — synoptic meteorology links to characterise chemical climates at six northwest European monitoring stations. Atmos Environ 1995;29:145–67.

Dorling SR, Davies TD, Pierce CE. Cluster analysis: a technique for estimating the synoptic meteorological controls on air and precipitation chemistry — method and applications. Atmos Environ 1992;26:2575–81.

Draxler RR, Hess GD. An overview of the HYSPLIT_4 modelling system for trajectories, dispersion and deposition. Aust Meteorol Mag 1998;47:295–308.

Escudero M, Castillo S, Querol X, Avila A, Alarcón M, Viana MM, et al. Wet and dry African dust episodes over Eastern Spain. J Geophys Res-Atmos 2005;110(D18). http://dx.doi.org/10.1029/2004JD00473. [Article No. D18S08].

Escudero M, Stein A, Draxler RR, Querol X, Alastuey A, Castillo S, et al. Determination of the contribution of northern Africa dust source areas to PM_{10} concentrations over the central Iberian Peninsula using the Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPLIT) model. J Geophys Res-Atmos 2006;111(D6). http://dx.doi.org/10.1029/2005JD006395. [Article No. D06210].

Feng J, Yang W. Effects of particulate air pollution on cardiovascular health: a population health risk assessment. PLoS One 2012;7. http://dx.doi.org/10.1371/journal.pone.0033385. [Article No. e33385].

Fischer G, Roppert J. Ein Verfahren der Transformationsanalyse faktorenanalytischer Ergebnisse. Lineare Strukturen in Mathematik und Statistik unter besonderer Berücksichtigung der Faktoren- und Transformationsanalyse. Arbeiten aus dem Institut für höhere Studien und wissenschaftliche ForschungWien-Würzburg: Wien Verlag Physica; 1965. [in German].

Grivas G, Chaloulakou A, Kassomenos P. An overview of the PM_{10} pollution problem, in the metropolitan area of Athens, Greece. Assessment of controlling factors and potential impact of long range transport. Sci Total Environ 2008;389:165–77.

Hongisto M, Sofiev M. Long-range transport of dust to the Baltic Sea region. Int J Environ Pollut 2004;22:72–86.

Jahn W, Vahle H. Die Faktoranalyse und ihre Anwendung. Berlin: Verlag die Wirtschaft; 1968 [in German].

Jolliffe IT. Principal component analysis: a beginner's guide—II. Pitfalls, myths and extensions. Weather 1993;48:246–53.

Jorba O, Pérez C, Rocadenbosch F, Baldasano JM. Cluster analysis of 4-day backward trajectories arriving in the Barcelona area, Spain, from 1997 to 2002. J Appl Meteorol 2004;43:887–901.

Juda-Rezler K, Reizer M, Oudinet JP. Determination and analysis of PM_{10} source apportionment during episodes of air pollution in Central Eastern European urban areas: the case of wintertime 2006. Atmos Environ 2011;45:6557–66.

Karaca F, Anil I, Alagha O. Long-range potential source contributions of episodic aerosol events to PM_{10} profile of a megacity. Atmos Environ 2009;43:5713–22.

Koltay E, Borbély-Kiss I, Kertész Z, Kiss ÁZ, Szabó Gy. Assignment of Saharan dust sources to episodes in Hungarian atmosphere by PIXE and TOMS observations. J Radioanal Nucl Chem 2006;267:449–59.

Konovalov IB, Beekmann M, Kuznetsova IN, Yurova A, Zvyagintsev AM. Atmospheric impacts of the 2010 Russian wildfires: integrating modelling and measurements of an extreme air pollution episode in the Moscow region. Atmos Chem Phys 2011;11: 10031–56.

Li MM, Huang X, Zhu L, Li JF, Song Y, Cai XH, et al. Analysis of the transport pathways and potential sources of PM_{10} in Shanghai based on three methods. Sci Total Environ 2012;414:525–34.

Mahalanobis PC. On the generalized distance in statistics. Proc Natl Inst Sci India 1936;12:49–55.

Makra L, Matyasovszky I, Guba Z, Karatzas K, Anttila P. Monitoring the long-range transport effects on urban PM_{10} levels using 3D clusters of backward trajectories. Atmos Environ 2011;45:2630–41.

Makra L, Matyasovszky I, Bálint B. Association of allergic asthma emergency room visits with the main biological and chemical air pollutants. Sci Total Environ 2012;432: 288–96.

Masiol M, Squizzato S, Ceccato D, Rampazzo G, Pavoni B. Determining the influence of different atmospheric circulation patterns on PM_{10} chemical composition in a source apportionment study. Atmos Environ 2012;63:117–24.

Matyasovszky I, Makra L, Bálint B, Guba Z, Sümeghy Z. Multivariate analysis of respiratory problems and their connection with meteorological parameters and the main biological and chemical air pollutants. Atmos Environ 2011;45:4152–9.

Medina S, Plasencia A, Ballester F, Mucke HG, Schwartz J. Apheis: public health impact of PM_{10} in 19 European cities. J Epidemiol Community Health 2004;58:831–6.

Mezosi G, Szatmari J. Assessment of wind erosion risk on the agricultural area of the southern part of Hungary. J Hazard Mater 1998;61:139–53.

NCEP/NCAR National Centers for Environmental Prediction/National Center for Atmospheric Research. http://dss.ucar.edu/datasets/ds090.0/.

Niemi JV, Tervahattu H, Vehkamaki H, Martikainen J, Laakso L, Kulmala M, et al. Characterization of aerosol particle episodes in Finland caused by wildfires in Eastern Europe. Atmos Chem Phys 2005;5:2299–310.

Preparata FP, Hong SJ. Convex hulls of finite sets of points in two and three dimensions. Commun ACM 1977;20:87–93.

Querol X, Pey J, Pandolfi M, Alastuey A, Cusack M, Pérez N, et al. African dust contributions to mean ambient PM_{10} mass-levels across the Mediterranean Basin. Atmos Environ 2009;43:4266–77.

Salvador P, Artiñano B, Querol X, Alastuey A. A combined analysis of backward trajectories and aerosol chemistry to characterise long-range transport episodes of particulate matter: the Madrid air basin, a case study. Sci Total Environ 2008;390: 495–506.

Salvador P, Artiñano B, Pio C, Afonso J, Legrand M, Puxbaum H, et al. Evaluation of aerosol sources at European high altitude background sites with trajectory statistical methods. Atmos Environ 2010;44:2316–29.

Schindler C, Keide D, Gerbase MW, Zemp E, Bettschart R, Brandli O, et al. Improvements in PM_{10} exposure and reduced rates of respiratory symptoms in a cohort of Swiss adults (SAPALDIA). Am J Respir Crit Care Med 2009;179:579–87.

Stohl A. Computation, accuracy and application of trajectories — a review and bibliography. Atmos Environ 1998;32:947–66.

Taylor AD. Conformal map transformations for meteorological modelers. Comput Geosci 1997;23:63–75.

Tukey JW. The problem of multiple comparisons (1953). Time series, 1965–1984, vol. II. Monterey, CA: Wadsworth Advanced Books & Software; 1985.

Vardoulakis S, Kassomenos P. Sources and factors affecting PM_{10} levels in two European cities: implications for local air quality management. Atmos Environ 2008;42: 3949–63.

Wong MS, Nichol JE, Lee KH. Estimation of aerosol sources and aerosol transport pathways using AERONET clustering and backward trajectories: a case study of Hong Kong. Int J Remote Sens 2013;34:938–55.

World Health Organization. Air quality guidelines for Europe. WHO Regional Publications, European Series, No. 91World Health Organization Regional Office for Europe Copenhagen 0378-225592 890 1358 3; 2000 [NLM Classification: WA 754].

Web references

http://www.arl.noaa.gov/ready/hysplit4.html.

http://dss.ucar.edu/datasets/ds090.0/.

http://www.qhull.org.