Atmospheric Environment 45 (2011) 2630–2641

Contents lists available at ScienceDirect

Atmospheric Environment

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

Monitoring the long-range transport effects on urban PM10 levels using 3D clusters of backward trajectories

László Makra[a,*], István Matyasovszky[b], Zoltán Guba[a], Kostas Karatzas[c], Pia Anttila[d]

[a] Department of Climatology and Landscape Ecology, University of Szeged, P.O. Box 653, H-6701 Szeged, Hungary
[b] Department of Meteorology, Eötvös Loránd University, Pázmány Péter Street 1/A, H-1117 Budapest, Hungary
[c] Department of Mechanical Engineering, Aristotle University, P.O. Box 483, GR-54124 Thessaloniki, Greece
[d] Finnish Meteorological Institute, P.O. Box 503, FI-00101 Helsinki, Finland

ARTICLE INFO

Article history:
Received 22 September 2010
Received in revised form
27 February 2011
Accepted 28 February 2011

Keywords:
Long-range PM10 transport
Backward trajectories
Cluster analysis
Mahalanobis metric

ABSTRACT

The purpose of the study is to identify long-range transport patterns that may have an important influence on PM10 levels in three European cities at different latitudes, namely Thessaloniki, Szeged and Helsinki. A further aim is to separate medium- and long-range PM10 transport for these cities. 4-day, 6-hourly three-dimensional (3D) backward trajectories arriving at these locations at 1200 GMT were computed using the HYSPLIT model over a 5-year period from 2001 to 2005. A k-means clustering algorithm using the Mahalanobis metric was applied in order to develop trajectory types. The 3D delimination of the clusters by the function "convhull" is a novel approach. Two statistical indices were used to evaluate and compare critical daily PM10 exceedances corresponding to the trajectory clusters. For Thessaloniki, the major PM10 transport can be clearly associated with air masses arriving from Central and Southern Europe. Occasional North African dust intrusions over Greece are also found. The transport of particulate matter from North-western Europe to Thessaloniki is of limited importance. For Szeged, Central Europe, Southern Europe and Mid-eastern Europe are the most important sources of PM10. The occasional appearance of North African-origin dust over Hungary is also detected. Local PM10 levels tend to be diluted when air masses arrive at the Carpathian Basin from North-western Europe, the Mid-Atlantic – Western Europe and Northern Europe. For Helsinki, high PM10 concentrations are due to air masses coming from Northern and Eastern Europe including North-western Russia. An occasional Caspian Sea desert influence on particulate levels can also be identified. However, air currents coming from the Northern Atlantics, Northern and North-western Europe tend to dilute PM10 levels. A simple approach is developed in order to separate medium- and long-range PM10 transport for each city.

1. Introduction

PM10 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 PM10 (50 μg m^{-3}) is frequently exceeded in the urban environment. The short- and long-term human exposure to high particulate matter concentrations observed in large cities increases the risk of respiratory (Annesi-Maesano et al., 2007) and also cardiovascular (Analitis et al., 2006) diseases. Therefore, studying potential key regions and long-range transport effects on urban PM10 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 PM10 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 for an overall identification of paths and origins of air parcels that play a role in the contribution to PM10 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). However,

* Corresponding author. Tel.: +36 62 544 856; fax: +36 62 544 624.
E-mail addresses: makra@geo.u-szeged.hu (L. Makra), MATYA@ludens.elte.hu (I. Matyasovszky), h480623@stud.u-szeged.hu (Z. Guba), kkara@eng.auth.gr (K. Karatzas), Pia.Anttila@fmi.fi (P. Anttila).

1352-2310/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2011.02.068

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, although items in the same cluster are likely to have many features in common.

The question may arise whether deposition of coarse particles is meaningful or they play an important role in long-range transport. For Thessaloniki, north-easterly winds can add to the background-urban PM10 level to 10 μg m^{-3} on a monthly basis. Furthermore, south-westerly winds, originating from North Africa, may contribute to a monthly basis up to 20–30 μg m^{-3} to the already existing aerosol load of the city during transition seasons (Katragkou et al., 2009). In Crete, for the low dust period ions dominate the PM10 mass (40–60%), followed by dust (25–35%) and particulate carbonaceous matter (PCM) (20–25%). However, during the identified dust outbreaks the contribution of dust rises to 60–65%, the share of ions declines to 30% and only 10% is PCM (Gerasopoulos et al., 2006).

During a ten-year period (1991–2000) about 50 episodic appearance of Saharan dust in the Hungarian aerosol was assigned and investigated using PIXE analysis (Borbély-Kiss et al., 2004). Local episodes of Saharan origin were selected in the time series of concentration ratios Ti/Ca, Si/Al, Al/Ca, Ti/Fe, Si/Fe, Ca/Fe generally accepted as regional signatures of Saharan dust. These long-range transport episodes of mineral dust from North African deserts caused elevated concentrations of coarse Ca, Mg and K at sites in Central Europe (Borbély-Kiss et al., 2004; Salvador et al., 2010).

For Helsinki, PM10 pollution episodes clearly of local origin were characterized by a low average fraction of PM2.5 (<0.2) in PM10 at the urban traffic monitoring site, low ratio between PM10 concentrations at the regional background site and at the urban traffic site (<0.2), low average ion sums (1.5–2.5 μg m^{-3}) and low accumulation to Aitken mode ratios (0.13–0.26). Furthermore, the episodes of distinct long-range transport characteristics have a high fraction of fine particles in PM10 (0.5–0.6) at the urban traffic site, a high ratio between PM10 concentrations at the regional background site and at the urban traffic site (0.7–0.8), high values for the ion sum (6.6–11.9 μg m^{-3}) and high accumulation to Aitken mode ratios (0.75–0.85) (Aarnio et al., 2008).

The course fraction is considered important not only at these three regions but elsewhere in Europe or even in Hong Kong. In Brussels a dust storm, that occurred 2 days earlier in the Sahara, showed a significant peak in PM10 indicating a clear difference between PM10 and PM2.5 mass concentrations (Vanderstraeten et al., 2008). Most of the mass concentration during this period was due to the presence of the coarse fraction (PM2.5–PM10). Chemical analysis showed a significant similarity to Saharan dust composition (Vanderstraeten et al., 2008). In Hong Kong the coarse size fractions of PM10 concentrations were about 4–8 times higher during dust events (Lee et al., 2010). The concentration ratio of coarse to fine aerosol for the periods before, during and after the dust event here is 1.17, 1.96, and 1.39. About 79% of the mass concentration of PM10 dust is due to the presence of the coarse fraction. Therefore, the size distribution evolution clearly indicates that the coarse mode aerosol is enhanced during the dust period, despite likely scavenging of coarse particles by frontal rain.

As a result, during PM10 pollution episodes the amount of coarse particles increases (Borbély-Kiss et al., 2004; Gerasopoulos et al., 2006; Vanderstraeten et al., 2008; Lee et al., 2010; Salvador et al., 2010), their ratio also increases in Southern Europe (Gerasopoulos et al., 2006), in the temperate regions of Europe (Borbély-Kiss et al., 2004; Vanderstraeten et al., 2008) or in Hong Kong (Lee et al., 2010). However, in Northern Europe the ratio of coarse particles (PM2.5–PM10) increases smaller than that of fine particles (Aarnio et al., 2008). The size distribution of the particles in the target area generally depends on the physical characteristics of the source regions, and on how the dust is picked up. It is probably modified in the initial stages of transport, but after injection into the upper atmosphere, it is not much. In this way, coarse particles can play a substantial role in long-range transport of PM10, hence there is sense to talk about PM10 instead of just PM2.5.

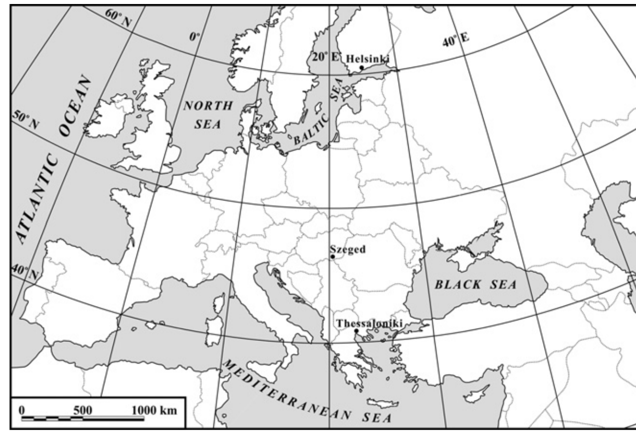

Fig. 1. The geographical positions of Thessaloniki, Szeged and Helsinki.

The main objective of this paper is to identify the key geographical regions responsible for PM10 levels in three cities (Thessaloniki, southern Europe; Szeged, Central Europe; and Helsinki, northern Europe, Fig. 1). Backward trajectories arriving at these sites are clustered using the Mahalanobis metric in order to determine which regions imply high PM10 concentrations. The clustering is performed using three-dimensional (3D) backward trajectories. ANOVA is used to determine whether PM10 concentrations corresponding to these trajectory clusters differ significantly. Cluster-dependent occurrences, when 24-h mean PM10 concentrations exceed the limit value of 50 μg m^{-3} are also analyzed with two statistical indices. Lastly, a simple approach is developed in order to separate medium-range PM10 transport including local PM10 emissions from the long-range transport of PM10.

2. Data and methods

2.1. Study areas and monitoring data

Five years (2001–2005) of daily mean PM10 data as well as daily meteorological data (mean temperature, mean global solar flux, mean relative humidity and daily precipitation total) taken from three European cities – Thessaloniki (Greece), Szeged (Hungary) and Helsinki (Finland) (Fig. 1; Table 1) – were analyzed. The reasons

Table 1
Parameters of basic meteorological data and mean 24-h PM10 concentrations.

City	Parameter	Jan	Jul	Year
Thessaloniki	Mean temperature (°C)	4.6	26.4	15.8
	Precipitation total (mm)	44	20	449
	Mean 24-h PM10 concentration (μg m^{-3})	65.3	51.5	58.9
Szeged	Mean temperature (°C)	−1.2	22.4	11.2
	Precipitation total (mm)	32	51	573
	Mean 24-h PM10 concentration (μg m^{-3})	53.2	38.1	47.0
Helsinki	Mean temperature (°C)	−6.1	17.2	4.9
	Precipitation total (mm)	57	68	640
	Mean 24-h PM10 concentration (μg m^{-3})	14.0	15.5	15.7

for selecting these sites include their fairly big distance and their substantial difference in topography and climate.

Thessaloniki (40.64 N; 22.94 E) is the second largest city of Greece where emissions originate mainly from the local traffic. Szeged (46.25 N; 20.10 E), the largest settlement in SE Hungary, is located at the confluence of the Rivers Tisza and Maros. Helsinki (60.25 N; 25.05 E), the capital of Finland, forms the core of the Helsinki Metropolitan Area.

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 Thessaloniki, Szeged and Helsinki corresponding to the Hybrid Single-Particle Lagrangian Integral Trajectory (HYSPLIT, version 4.8; http://www.arl.noaa.gov/ready/hysplit4.html) model (Draxler and Hess, 1998) were obtained from National Centres for Environmental Prediction/National Centre 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 three locations at 1200 Greenwich Mean Time (GMT) at heights $h = 500$, 1500 and 3000 m above mean sea level (AMSL) for each day over a 5-year period from 2001 to 2005 were taken in order to describe the horizontal and vertical movements of an air parcel for the above-mentioned three cities. Hence, trajectories arriving at these three sites were extended backwards in time for 4 days with 6-h temporal resolutions. 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 PM10 loadings of this region than another air mass traveling at much higher levels over this same area.

2.3. Cluster analysis

Cluster analysis groups 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 AMSL) of the 4-day 3D backtrajectories for each city 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 centre 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 CMRSD (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) arrival height because backtrajectories at this arrival height are expected to have the largest influence on the PM10 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; www.qhull.org) gathers the extreme trajectory positions (positions farthermost from the centre) 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.

2.4. Analysis of variance (ANOVA)

ANOVA is used to test whether the means of PM10 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 PM10 concentrations under different trajectory types may tell us about the origin and transport of air masses on local PM10 levels. There are several versions available for comparing means calculated from sub samples 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 PM10 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 PM10 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 PM10 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 PM10 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 PM10 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 8 or 9 variables (3 or 4 climatic and 5 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 PM10 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 special transformation of the retained factors was performed to discover to what degree the above-mentioned explanatory variables (3 or 4 climatic and 5 trajectory variables) affect the resultant variable (daily mean PM10 concentration), and to give a rank of their influence (Jahn and Vahle, 1968).

2.6. Statistical characterization of PM10 exceedance episodes

The role of long-range transport is studied by analyzing the cluster occurrences on days when 24-h mean PM10 concentrations exceed the limit value of 50 μg m^{-3}. Two statistical indices of daily PM10 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

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

where D_i is the number of occurrences of cluster i, and $D_{(>50)i}$ is the number of 24-h PM10 exceedances. INDEX1 gives the likelihood of an exceedance for a given cluster. INDEX2 is defined as

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

where E is the total number of 24-h PM10 exceedance days recorded at a given site. INDEX2 can be interpreted as the likelihood of certain trajectory being present on a PM10 exceedance day.

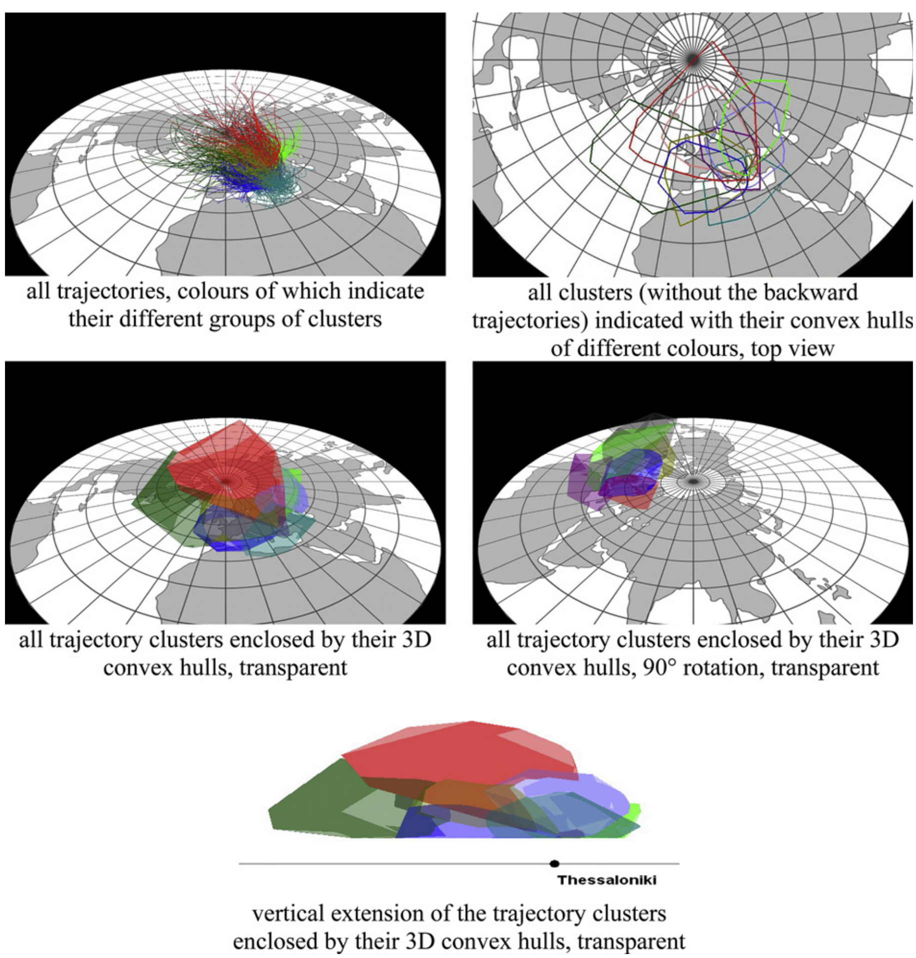

Fig. 2. 3D clusters of the backward trajectories retained, Thessaloniki, $h = 500$ m.

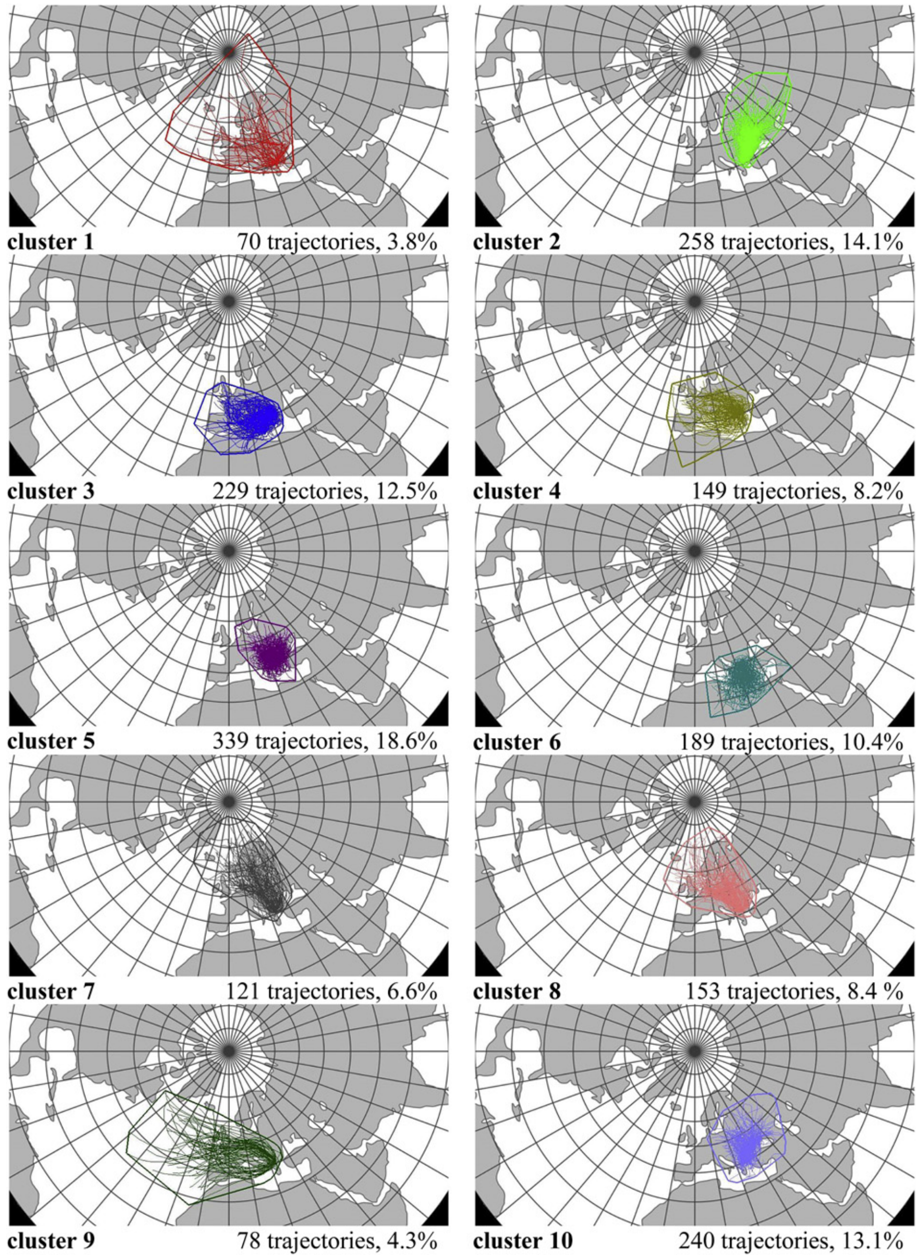

Fig. 3. The individual clusters of the backward trajectories retained, enclosed by their convex hulls, Thessaloniki, top view, $h = 500$ m.

3. Results

3.1. Thessaloniki

The 3D clustering produced ten clusters based on CCRMSD. All of the trajectories with colour-coded clusters, all of the clusters without trajectories but with their 3D convex hulls for the top view and all trajectory clusters enclosed by their transparent 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.

The individual clusters (Fig. 3), with the name of the source regions and their standardized average PM10 concentrations are tabulated (Table 2). Pairwise comparisons of the cluster averages found 29 significant differences among the possible 45 cluster pairs (64.4%). Hereafter only clusters of the above-mentioned significantly different cluster-averaged PM10 levels were then considered and analyzed (Fig. 3; Table 2).

Clusters 1, 4 and 10 have the highest INDEX1 values, namely 37.1%, 38.9% and 34.6%, respectively (Fig. 8). The high INDEX1 value of these three clusters is in agreement with their high mean PM10 concentrations. Furthermore, the high standard deviation of cluster 4 implies a higher chance of extreme PM10 episodes for this cluster. The first and third highest INDEX2 values (18.2% and 14.7%) belonging to cluster 10 and cluster 3 are characterized by high mean PM10 levels. Note that INDEX1 and INDEX2 are not independent parameters. When a cluster is frequent, a high INDEX1 value means a high INDEX2 value (see cluster 10). Although cluster 6 has a high mean PM10 level, it has medium-valued indices. The high frequency of clusters 3, 6 and 10 and their low-moving backward trajectories emphasize the importance of these clusters in PM10 transport (Fig. 3; Table 2).

Table 2
The individual clusters with the name of the source regions and their standardized average PM10 concentrations for each city, $h = 500$ m (bold: maximum; italic: minimum) (The standardized PM10 values are dimensionless).

Cluster no.	Thessaloniki		Szeged		Helsinki	
	Name of the source region	PM10 level	Name of the source region	PM10 level	Name of the source region	PM10 level
1	North-western Europe	0.15	Central and Southern Europe with North Africa	0.28	Scandinavia and the Arctic Sea region	*−0.48*
2	Eastern Europe	−0.14	North-western Europe	−0.09	Scandinavia and the Arctic Sea region	−0.05
3	Central and Southern Europe	0.16	Mid-Atlantic – Western Europe	−0.27	North-western Europe	−0.36
4	Central and Southern Europe	**0.41**	Northern Europe	*−0.56*	Eastern Europe	**1.12**
5	Mid-eastern and Southern Europe	−0.09	Mid-Atlantic – Western Europe	0.35	Northern Scandinavia and North-western Russia	−0.17
6	the Balkan Peninsula and North Africa	0.15	Central Europe	−0.12	Northern Atlantics	−0.01
7	North-western Europe	−0.28	Northern Europe	−0.07	Western Europe	−0.01
8	North-western Europe	*−0.46*	Eastern Europe	0.11	Northern Europe	0.19
9	Mid-Atlantic – Western Europe	0.03	Mid-eastern Europe	0.38	Northern Europe	0.30
10	Mid-eastern Europe	0.14	Mid-Atlantic – Western Europe	**0.49**	Eastern Europe	0.24
11					North-western Europe	−0.47

Based on the above analysis, clusters 3, 4, 6 and 10 play the largest role in PM10 transport to Thessaloniki. A substantial part of clusters 3 and 4 (Central and Southern Europe), as well as cluster 10 (Mid-eastern Europe) and the whole of cluster 6 (the Balkan Peninsula and North Africa) cover arid regions with a negative water balance. These environments are favorable for turbulent air currents to take up and transport particles, and they in this way contribute to the measured PM10 exceedances. The role of the above four clusters in PM10 transport is confirmed by their high mean PM10 levels (Table 2), high frequency (except for clusters 4 and 6) (Fig. 3), high percentage of low-moving backtrajectories (except for cluster 4) (Fig. 2, lower panel) and high INDEX1 and INDEX2 values (Fig. 8). Cluster 6 (the Balkan Peninsula and North Africa) indicates occasional North African dust intrusion over

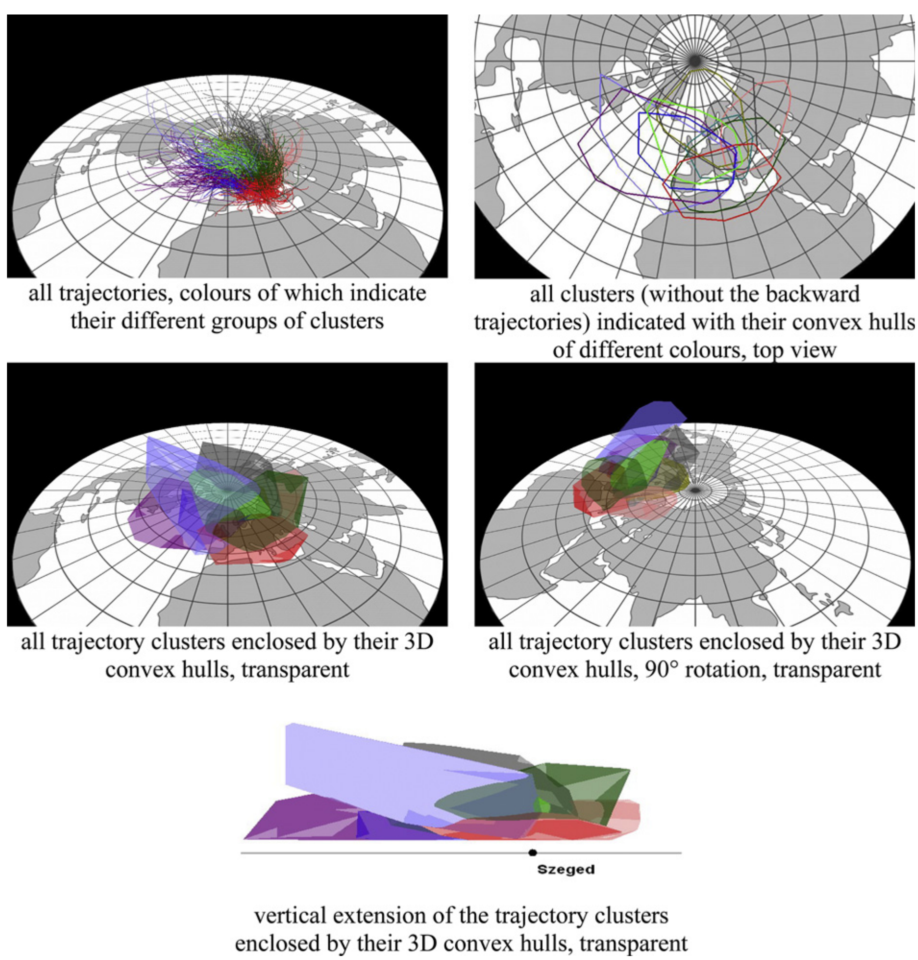

Fig. 4. 3D clusters of the backward trajectories retained, Szeged, $h = 500$ m.

Greece, in accordance with results of earlier studies (Balis et al., 2004; Katragkou et al., 2009). Cluster 10 (Mid-eastern Europe) is also a major source of PM10 transport to Thessaloniki, strengthening the finding of Katragkou et al. (2009). They showed that north-eastern flows (our cluster 10) and southern flows (our cluster 6) are the most frequent, and they can greatly influence the PM10 level of Thessaloniki. Clusters 7 and 8 have the lowest mean PM10 levels; they are both infrequent and include mostly high-moving air masses, which is consistent with the finding of Katragkou et al. (2009) that the transport of particulate matter from North-western Europe to Thessaloniki is of limited importance (Fig. 3; Table 2).

3.2. Szeged

Ten clusters were retained in a 3D analysis based on CCRMSD (Figs. 4 and 5). In Table 2 the individual clusters (Fig. 5) with the name of the source regions and their standardized average PM10 levels are presented. For Szeged, 33 significant differences were detected among the possible 45 cluster pairs (73.3%).

The highest INDEX1 values are associated with cluster 10 (Mid-Atlantic – Western Europe) (48.6%) (Fig. 8), but we can separate this cluster since its frequency is practically negligible (1.9%). The next highest INDEX1 values, in decreasing order, belong to cluster 9 (Mid-eastern Europe) (43.6%), cluster 1 (Central and Southern

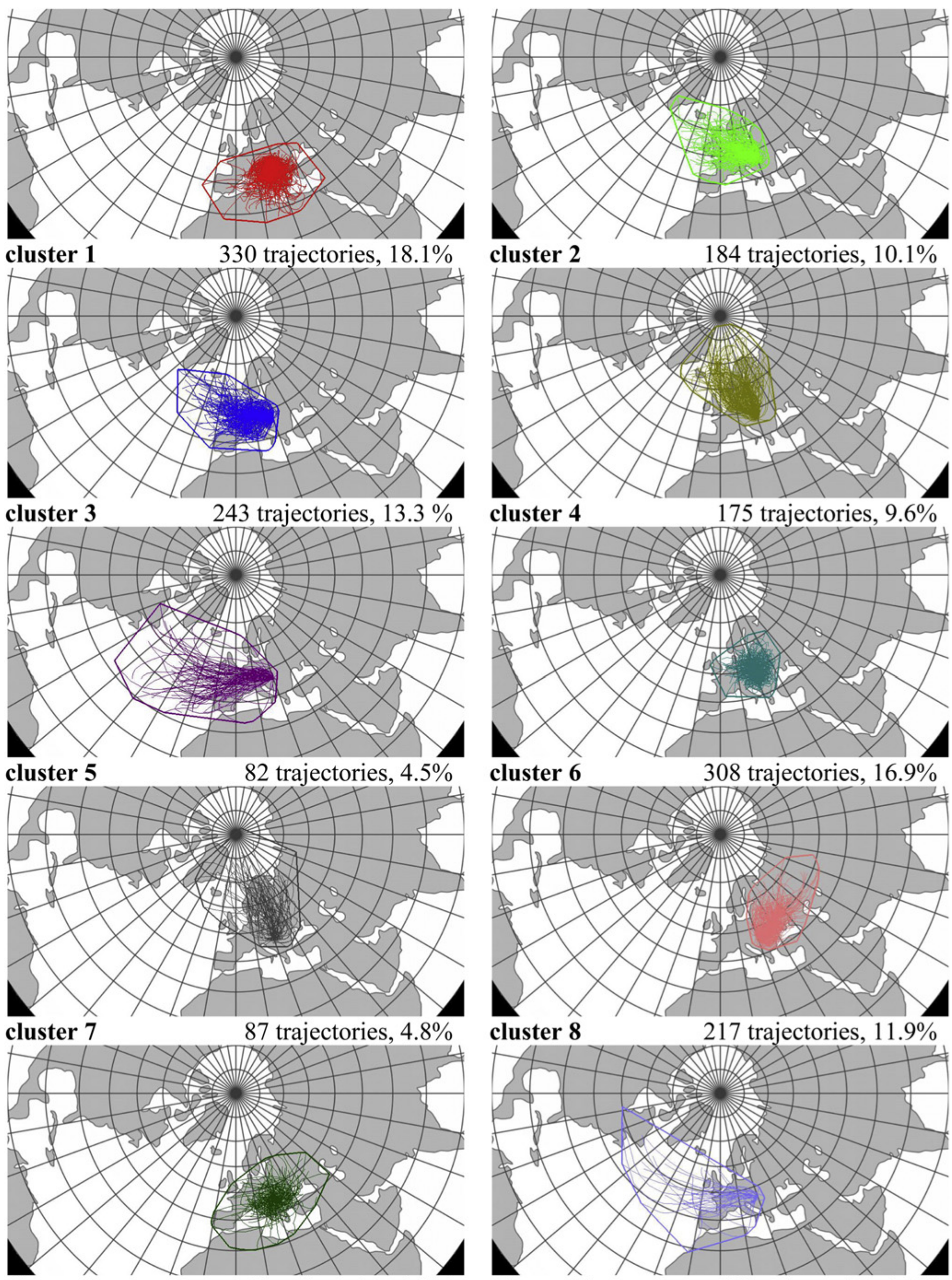

Fig. 5. The individual clusters of the backward trajectories retained, enclosed by their convex hulls, Szeged, top view, $h = 500$ m.

Europe with North Africa) (34.5%) and cluster 5 (Mid-Atlantic – Western Europe) (31.7%). This is in agreement with the fact that these clusters have very high mean PM10 levels (Fig. 5; Table 2).

Accordingly, clusters 1, 5 and 9 are the most relevant in terms of PM10 transport to Szeged, which is confirmed by their highest mean PM10 levels (Table 2), their high frequency (except for cluster 5) (Fig. 5), high portions of low-moving backtrajectories (Fig. 4, lower panel) and high INDEX1 and INDEX2 values (except for cluster 5) (Fig. 8). Cluster 1 (Central and Southern Europe with North Africa), being the most frequent, includes short backtrajectories and, hence, slow-moving air masses. The region enclosed 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 PM10. Furthermore, an important part of this cluster is over North Africa, and this shows that high PM10 exceedance episodes can occasionally be related to low-moving air masses coming from North Africa (Borbély-Kiss et al., 1999; Koltay et al., 2006). Cluster 5 (together with cluster 10) has the longest (fastest) backtrajectories. Even though they originate over a pollutant free region (the Mid-Atlantics), they take up and transport a considerable amount of dust over Hungary after crossing Southern and Central Europe. This cluster is only of minor importance due to its infrequent occurrence. Together with cluster 1 (Central and Southern Europe with North Africa), cluster 9 (Mid-eastern Europe) is the most important for transporting PM10 to Szeged. Consequently, Central-, Southern- and Mid-eastern Europe are the most important sources of PM10 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; Table 2).

3.3. Helsinki

Eleven clusters were identified for Helsinki based on CCRMSD (Figs. 6 and 7). The individual clusters (Fig. 7), with the name of the source regions and their standardized average PM10 concentrations are indicated (Table 2).

The highest INDEX1 values are associated with clusters having the highest PM10 levels; namely, in decreasing order, cluster 4 (69.1%), cluster 10 (37.9%), cluster 8 (30.9%) and cluster 9 (29.6%) (Fig. 8). But the highest INDEX2 values, in decreasing order, are indicated by cluster 4 (18.4%), cluster 8 (16.7%) and cluster 10 (14.7%). The INDEX2 value of cluster 9, having a high mean PM10 level, is low due to its very low frequency. Clusters 1, 3 and 11 with the lowest particulate concentrations have the lowest INDEX1 and INDEX2 values (Fig. 7; Table 2).

In summary, clusters 4 (Eastern Europe), 8 (Northern Europe) and 10 (Eastern Europe) are considered the most important. These clusters have the highest PM10 levels (Table 2), have a high frequency (except for cluster 4) (Fig. 7) and have the highest INDEX1 and INDEX2 values (Fig. 8). Furthermore, these clusters have significant portions of low-moving backtrajectories (Fig. 6, lower panel). Note that cluster 9 (Northern Europe) also has a high

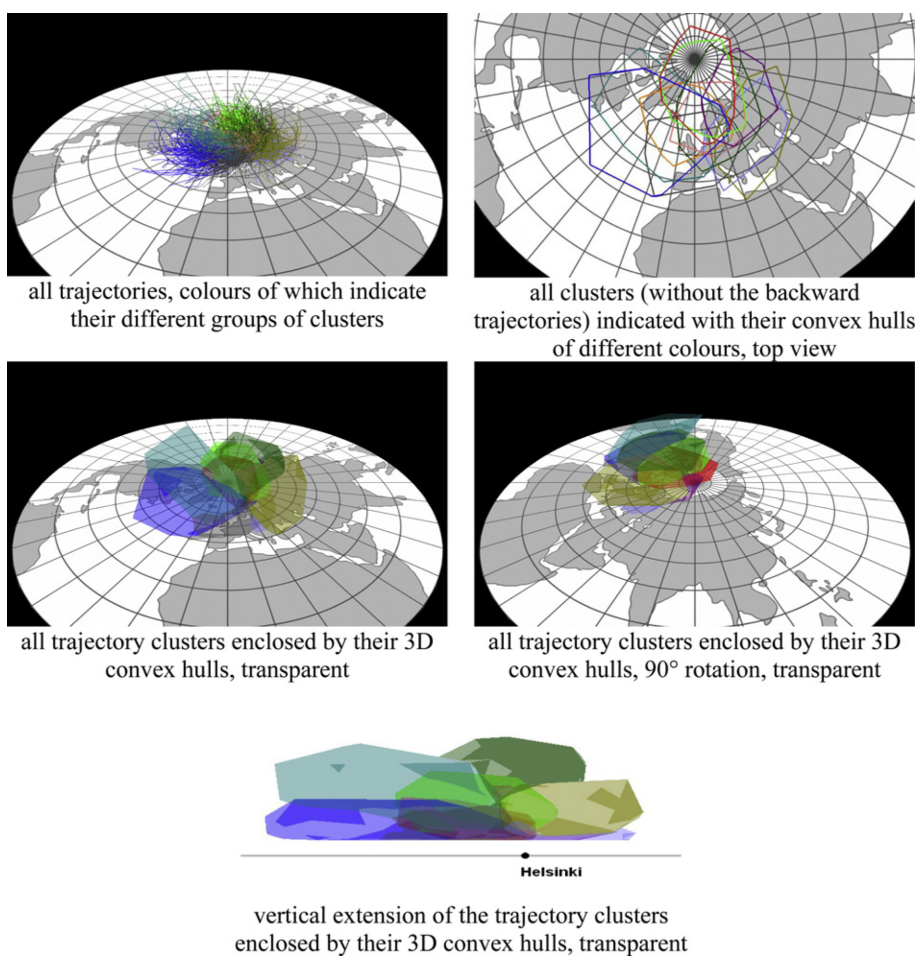

Fig. 6. 3D clusters of backward trajectories retained, Helsinki, $h = 500$ m.

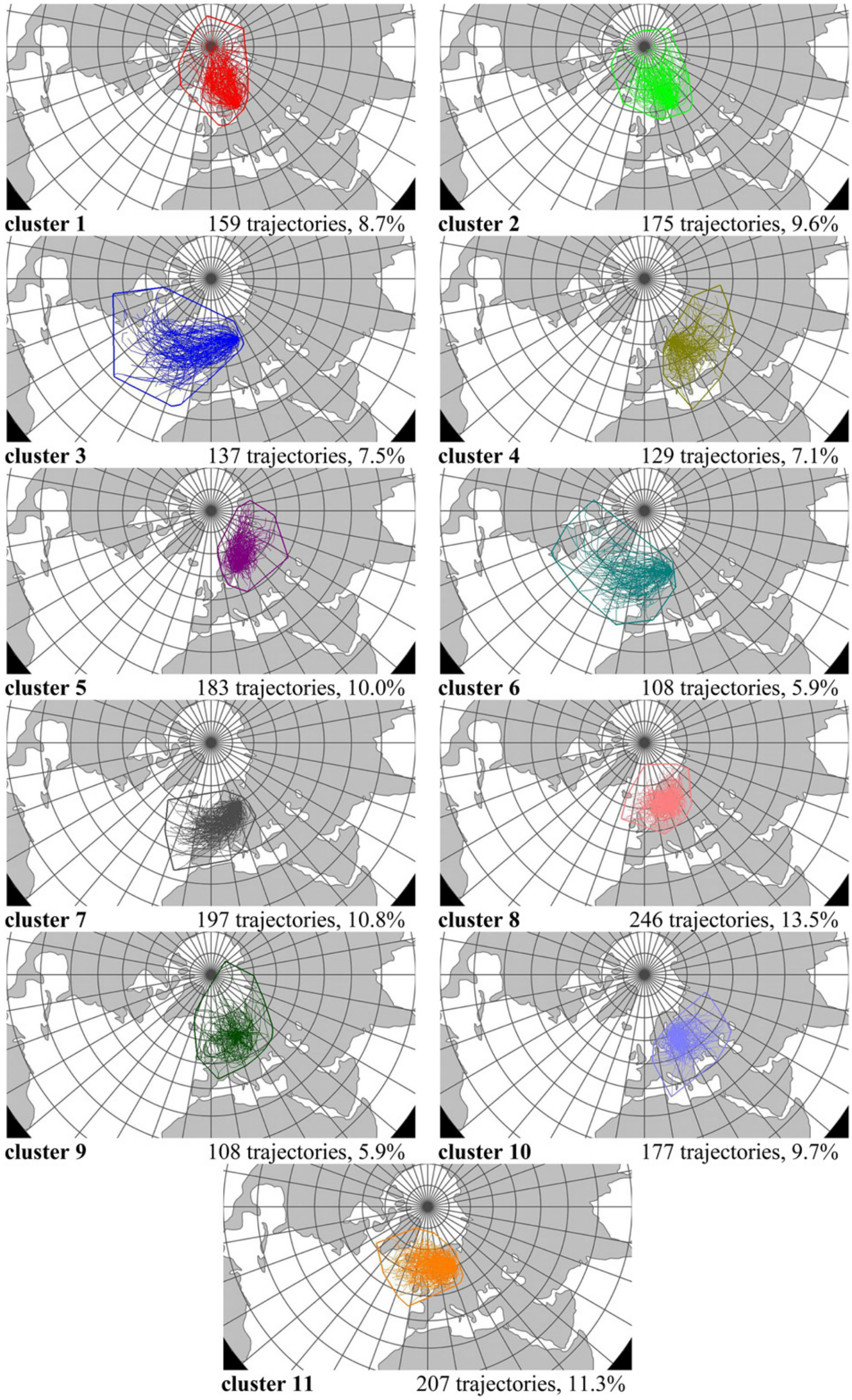

Fig. 7. The individual clusters of the backward trajectories retained, enclosed by their convex hulls, Helsinki, top view, $h = 500$ m.

PM10 level, but its INDEX2 value is small due to its low frequency and it has mostly high-moving backtrajectories. These characteristics substantially weaken its role in increasing local PM10 concentrations. The substantial role of clusters 4 and 10 (Eastern Europe) in contributing to local PM10 load in Helsinki was found by Hongisto and Sofiev (2004) and Niemi et al. (2005). These two clusters cover arid/semiarid regions with insufficient annual precipitation amounts in a substantial part of the year, contributing to frequent PM10 exceedance episodes. Evidence for the recurrent long-range transport of elevated PM10 concentrations from the Baltic States and North-western Russia (our cluster 8) over southern Finland was reported by Hongisto and Sofiev (2004) and Niemi et al. (2005). Furthermore, the occasional Caspian Sea desert influence on aerosol concentrations over the target area indicated

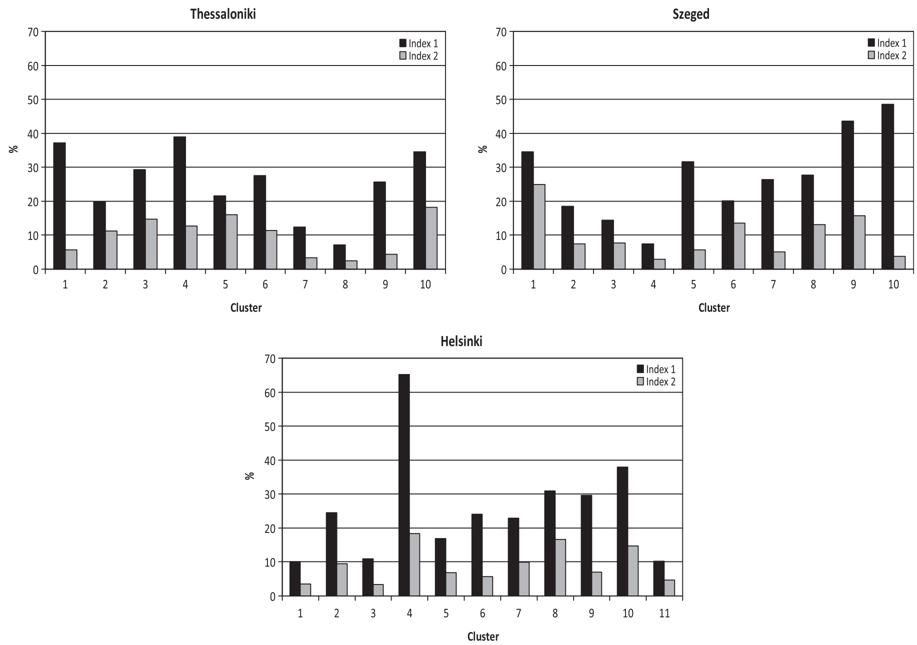

Fig. 8. Indices 1 and 2 for 3D clusters of the backward trajectories, $h = 500$ m.

by some backtrajectories of clusters 4 and 10 was previously mentioned by Hongisto and Sofiev (2004).

4. Discussion and conclusions

A cluster analysis was applied to 4-day, 6-hourly backward trajectories arriving at Thessaloniki, Szeged and Helsinki over a 5-year period in order to identify the main atmospheric circulation pathways influencing PM10 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, presentation of vertical extension of the trajectory clusters enclosed by their 3D convex hulls and, in this way, delimiting low-moving backtrajectories is a novel procedure.

When determining important clusters that mainly influence PM10 levels, the following aspects were considered: 1) the average PM10 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 Thessaloniki, the major PM10 transport can be clearly associated with air masses coming from Central and Southern Europe, the Balkan Peninsula and North Africa and Mid-eastern Europe. The importance of these clusters is justified by large regions that have a negative water balance in a substantial part of the year. The Balkan Peninsula and North Africa suggest occasional North African dust intrusions over Greece, which are in accordance with results obtained from related studies (Balis et al., 2004; Katragkou et al., 2009). Mid-eastern Europe is also a major source of PM10 transport to Thessaloniki. This conclusion is in accordance with the results of Katragkou et al. (2009) who found that north-eastern flows (our cluster 10) and southern flows (our cluster 6) are the most frequent and they can influence the PM10 level of Thessaloniki. Although clusters 7 and 8 have the lowest PM10 levels, they are both infrequent and include mostly high-moving air masses, suggesting that transport of particulate matter from North-western Europe to Thessaloniki is of little importance (Katragkou et al., 2009).

For Szeged, clusters 1, 5 and 9 are the most relevant in PM10 transport. Cluster 1, corresponding to Central and 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). Cluster 5 (Mid-Atlantic – Western Europe) consisting of trajectories that pass over Southern and Central Europe, transports a considerable amount of dust over Hungary. However, this cluster is generally of little importance due to its infrequent occurrence. Together with cluster 1, cluster 9 (Mid-eastern Europe) is the most important for transporting PM10 to Szeged. Consequently, Central-, Southern- and Mid-eastern Europe are the most important sources of PM10 over 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 over the target area.

As regards Helsinki, clusters 4 (Eastern Europe), 8 (Northern Europe) and 10 (Eastern Europe) are considered the most important. The substantial role of clusters 4 and 10 in contributing to local PM10 load in Helsinki is in agreement with the findings of Hongisto and Sofiev (2004) and Niemi et al. (2005). These two clusters cover arid/semiarid regions with insufficient annual precipitation amounts for a substantial part of the year and contribute to frequent PM10 exceedance episodes. Evidence for the recurrent long-range transport of elevated PM10 concentrations from the Baltic States and North-western Russia over southern Finland was reported by Hongisto and Sofiev (2004) and Niemi et al. (2005). An occasional Caspian Sea desert influence on aerosol concentrations over the target area indicated by some backtrajectories of clusters 4 and 10 strengthens earlier conclusions (Hongisto and Sofiev, 2004).

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 PM10 levels. In other words, it is necessary to determine the relative weight of these two components in the measured PM10 concentration. There are several case studies available that allow one to distinguish the long-range PM10 transport episodes from local PM10 pollution episodes (Escudero et al., 2006; Aarnio et al., 2008). Analyses of local meteorological conditions and air mass back-trajectories for a given city play an important role in developing methods for the above purpose (Aarnio et al., 2008). An attempt is made here using local meteorological parameters and components of the trajectories to discriminate these two pollution modes in the entire 5-year data set. Local PM10 pollution is characterized next via the daily mean temperature, daily relative humidity and daily global solar flux. Snow depth is also considered for Helsinki, because the spring peak of PM10 typically covers the whole snowmelt period for this city (Anttila and Salmi, 2006). Long-range PM10 transport is described by the real 3D length of the trajectories, the length of the 3D trajectories as the crow flies, their ratio, as well as the average daily highest and lowest positions of the backward trajectories based on their 4-day, 6-hourly positions. The latter two characteristics refer to the vertical transport of PM10 in the atmosphere, which comes from either turbulent transport, dominating the vertical exchange of PM10 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).

As the PM10 level on a given day is substantially influenced by whether it corresponds to a non-rainy or rainy day, the backward trajectories are divided into two groups, i.e. non-rainy and rainy days. This kind of classification of days reveals the role of precipitation in the quantity of transported PM10 (Querol et al., 2009). After performing a factor analysis, a special transformation was carried out for each city with the two groups (rainy or non-rainy days) and the 500 m, 1500 m and 3000 m arrival heights of the backward trajectories. Thus, altogether $3 \times 2 \times 3 = 18$ procedures gave information about the weights of the local source and long-range transport reflected by the 8 (9 for Helsinki) explanatory variables. The main conclusion is as follows. Local PM10 pollution is greater, both for non-rainy and rainy days, for Thessaloniki and Helsinki, while long-range transport plays a higher role both for non-rainy and rainy days for Szeged. Adding up the weights of the local pollution and long-range transport, the average value of the two weights is larger for Szeged on non-rainy days and for Szeged and Helsinki on rainy days at the 500 m height compared to the higher levels. Hence, the eight (nine) variables contain more information on PM10 when using backtrajectories arriving at the 500 m height in these three cases. For the other three cases (Thessaloniki on non-rainy and rainy days, as well as Helsinki on non-rainy days) the results 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 PM10 for Thessaloniki on rainy days and for Helsinki on non-rainy days. The role of both kinds of variables is essentially the same when considering rainy or non-rainy days. Note that these findings are valid only for variations of daily PM10 concentrations accounted for by the eight (nine for Helsinki) 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 and to Zoltán Sümeghy for the digital mapping in Fig. 1. This study was supported by the EU-6 Quantify project [No. 003893 (GOCE)], as well as the European Union and the European Social Fund [No. TAMOP 4.2.1/B-09/1/KMR-2010-0003 and TAMOP-4.2.1/B-09/1/KONV-2010-0005].

References

Aarnio, P., Martikainen, J., Hussein, T., Valkama, I., Vehkamäki, H., Sogacheva, L., Härkönen, J., Karppinen, A., Koskentalo, T., Kukkonen, J., Kulmala, M., 2008. Analysis and evaluation of selected PM10 pollution episodes in the Helsinki Metropolitan Area in 2002. Atmospheric Environment 42 (17), 3992–4005.

Analitis, A., Katsouyanni, K., Dimakopoulou, K., Samoli, E., Nikoloulopoulos, A.K., Petasakis, Y., Touloumi, G., Schwartz, J., Anderson, H.R., Cambra, K., Forastiere, F., Zmirou, D., Vonk, J.M., Clancy, L., Kriz, B., Bobvos, J., Pekkanen, J., 2006. Short-term effects of ambient particles on cardiovascular and respiratory mortality. Epidemiology 17, 230–233.

Anderberg, M.R., 1973. Cluster Analysis for Applications. Academic Press, New York.

Anderson, B., 2008. Evaluation Methods for Long Range Transport. 9th Modeling Conference Presentations. US Environmental Protection Agency, Technology Transfer Network Support Center for Regulatory Atmospheric Modeling.

Annesi-Maesano, I., Forastiere, F., Kunzli, N., Brunekref, B., 2007. Particulate matter, science and EU policy. European Respiratory Journal 29, 428–431.

Ansmann, A., Bosenberg, J., Chaikovsky, A., Comeron, A., Eckhardt, S., Eixmann, R., Freudenthaler, V., Ginoux, P., Komguem, L., Linne, H., Marquez, M.A.L., Matthias, V., Mattis, I., Mitev, V., Muller, D., Music, S., Nickovic, S., Pelon, J., Sauvage, L., Sobolewsky, P., Srivastava, M.K., Stohl, A., Torres, O., Vaughan, G., Wandinger, U., Wiegner, M., 2003. Long-range transport of Saharan dust to northern Europe: the 11–16 October 2001 outbreak observed with EARLINET. Journal of Geophysical Research – Atmosphere 108 (D24), 4783.

Anttila, P., Salmi, T., 2006. Characterizing temporal and spatial patterns of urban PM10 using six years of Finnish monitoring data. Boreal Environment Research 11, 463–479.

Balis, D.S., Amiridis, V., Nickovic, S., Papayannis, A., Zerefos, C., 2004. Optical properties of Saharan dust layers as detected by a Raman lidar at Thessaloniki, Greece. Geophysical Research Letters 31 (13), L13104.

Borbély-Kiss, I., Koltay, E., Szabó, G.Y., Bozó, L., Tar, K., 1999. Composition and sources of urban and rural atmospheric aerosol in eastern Hungary. Journal of Aerosol Science 30, 369–391.

Borbély-Kiss, I., Kiss, ÁZ., Koltay, E., Szabó, Gy, Bozó, L., 2004. Saharan dust episodes in Hungarian aerosol: elemental signatures and transport trajectories. Journal of Aerosol Science 35 (10), 1205–1224.

Borge, R., Lumbreras, J., Vardoulakis, S., Kassomenos, P., Rodríguez, E., 2007. Analysis of long-range transport influences on urban PM_{10} using two-stage atmospheric trajectory clusters. Atmospheric Environment 41, 4434–4450.

Coury, C., Dillner, A.M., 2007. Trends and sources of particulate matter in the Superstition wilderness using air trajectory and aerosol cluster analysis. Atmospheric Environment 41, 9309–9323.

Dorling, S.R., Davies, T.D., 1995. Extending cluster analysis – synoptic meteorology links to characterise chemical climates at six northwest European monitoring stations. Atmospheric Environment 29, 145–167.

Dorling, S.R., Davies, T.D., Pierce, C.E., 1992. Cluster analysis: a technique for estimating the synoptic meteorological controls on air and precipitation chemistry – method and applications. Atmospheric Environment 26, 2575–2581.

Draxler, R.R., Hess, G.D., 1998. An overwiew of the HYSPLIT_4 modelling system for trajectories, dispersion and deposition. Australian Meteorological Magazine 47, 295–308.

Escudero, M., Stein, A., Draxler, R.R., Querol, X., Alastuey, A., Castillo, S., Avila, A., 2006. Determination of the contribution of northern Africa dust source areas to PM10 concentrations over the central Iberian Peninsula using the Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPLIT) model. Journal of Geophysical Research-Atmosphere 111, D06210.

Gerasopoulos, E., Kouvarakis, G., Babasakalis, P., Vrekoussis, M., Putaud, J.P., Mihalopoulos, N., 2006. Origin and variability of particulate matter (PM10) mass concentrations over the Eastern Mediterranean. Atmospheric Environment 40 (25), 4679–4690.

Hongisto, M., Sofiev, M., 2004. Long-range transport of dust to the Baltic Sea region. International Journal of Environment and Pollution 22 (1–2), 72–86.

Jahn, W., Vahle, H., 1968. Die Faktoranalyse und ihre Anwendung. Verlag die Wirtschaft, Berlin (in German).

Jolliffe, I.T., 1993. Principal component analysis: a beginner's guide – II. Pitfalls, myths and extensions. Weather 48, 246–253.

Jorba, O., Pérez, C., Rocadenbosch, F., Baldasano, J.M., 2004. Cluster analysis of 4-day backward trajectories arriving in the Barcelona area, Spain, from 1997 to 2002. Journal of Applied Meteorology 43, 887–901.

Katragkou, E., Kazadzis, S., Amiridis, V., Papaioannou, V., Karathanasis, S., Melas, D., 2009. PM10 regional transport pathways in Thessaloniki, Greece. Atmospheric Environment 43 (5), 1079–1085 (special issue).

Koltay, E., Borbély-Kiss, I., Kertész, Z., Kiss, ÁZ., Szabó, Gy, 2006. Assignment of Saharan dust sources to episodes in Hungarian atmosphere by PIXE and TOMS observations. Journal of Radioanalytical and Nuclear Chemistry 267 (2), 449–459.

Lee, Y.C., Yang, X., Wenig, M., 2010. Transport of dusts from East Asian and non-East Asian sources to Hong Kong during dust storm related events 1996–2007. Atmospheric Environment 44 (30), 3728–3738.

Mahalanobis, P.C., 1936. On the generalized distance in statistics. Proceedings of the National Institute of Sciences India 12, 49–55.

Mezosi, G., Szatmari, J., 1998. Assessment of wind erosion risk on the agricultural area of the southern part of Hungary. Journal of Hazardous Material 61, 139–153.

Niemi, J.V., Tervahattu, H., Vehkamaki, H., Martikainen, J., Laakso, L., Kulmala, M., Aarnio, P., Koskentalo, T., Sillanpaa, M., Makkonen, U., 2005. Characterization of aerosol particle episodes in Finland caused by wildfires in Eastern Europe. Atmospheric Chemistry and Physics 5, 2299–2310.
Preparata, F.P., Hong, S.J., 1977. Convex hulls of finite sets of points in two and three dimensions. Communications of the ACM 20, 87–93.
Querol, X., Pey, J., Pandolfi, M., Alastuey, A., Cusack, M., Pérez, N., Moreno, T., Viana, M., Mihalopoulos, N., Kallos, G., Kleanthous, S., 2009. African dust contributions to mean ambient PM10 mass-levels across the Mediterranean Basin. Atmospheric Environment 43 (28), 4266–4277.
Salvador, P., Artiñano, B., Querol, X., Alastuey, A., 2008. 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. Science of the Total Environment 390, 495–506.
Salvador, P., Artiñano, B., Pio, C., Afonso, J., Legrand, M., Puxbaum, H., Hammer, S., 2010. Evaluation of aerosol sources at European high altitude background sites with trajectory statistical methods. Atmospheric Environment 44 (19), 2316–2329.
Stohl, A., 1998. Computation, accuracy and application of trajectories – a review and bibliography. Atmospheric Environment 32, 947–966.
Tukey, J.W., 1985. The problem of multiple comparisons (1953). In: The Collected Works of John W. Tukey. Time Series, 1965–1984, vol. II. Wadsworth Advanced Books & Software, Monterey, CA.
Vanderstraeten, P., Lénelle, Y., Meurrens, A., Carati, D., Brenig, L., Delcloo, A., Offer, Z.Y., Zaady, E., 2008. Dust storm originate from Sahara covering Western Europe: a case study. Atmospheric Environment 42 (21), 5489–5493.
Vardoulakis, S., Kassomenos, P., 2008. Sources and factors affecting PM10 levels in two European cities: implications for local air quality management. Atmospheric Environment 42, 3949–3963.
http://www.arl.noaa.gov/ready/hysplit4.html.
http://dss.ucar.edu/datasets/ds090.0/.
http://www.qhull.org.