ACTA CLIMATOLOGICA ET CHOROLOGICA
Universitatis Szegediensis, Tomus 42-43, 2009, 97-106

LONG-RANGE TRANSPORT OF PM_{10}, PART 1

L. MAKRA[1], T. SÁNTA[1] and I. MATYASOVSZKY[2]

[1]*Department of Climatology and Landscape Ecology, University of Szeged, P.O.Box 653, 6701 Szeged, Hungary*
E-mail: makra@geo.u-szeged.hu
[2]*Department of Meteorology, Eötvös Loránd University, Pázmány Péter st. 1/A, 1117 Budapest, Hungary*

Summary – The purpose of this study is to identify long-range transport patterns that may have an important influence on PM10 levels in four European cities at different latitudes, namely those in Thessaloniki, Szeged, Helsinki and Oulu. Trajectory positions were computed using the HYSPLIT model. 4-day, 6-hourly 3D backward trajectory positions arriving at these locations at 1200 GMT were determined for each day over a 5-year period from 2001 to 2005. Non-hierarchical cluster analysis with the k-means method was applied using a Mahalanobis metric. The efficiency of the 2D and 3D cluster analyses were compared for each city.

Key words: PM_{10}, long-range transport, backward trajectory positions, cluster analysis, Mahalanobis metric

1. INTRODUCTION

Particulates, alternatively referred to as particulate matter (PM) or fine particles, are tiny particles of solid, liquid or mixed phase, suspended in a gas. In contrast, 'aerosol' refers to the particles and the gas together. Sources of particulate matter can be of anthropogenic or natural origin. A particle with an aerodynamic diameter of 10 μm moves in a gas like a sphere of unit density (1 gcm^{-3}) with a diameter of 10 μm. PM diameters range from less than 10 nm to more than 10 μm. These dimensions represent the continuum from a few molecules up to the size where particles can no longer be carried by a gas. The notation PM_{10} is used to describe particles of 10 μm or less.

The short and long-term 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.

The existing PM_{10} regulations for EU countries are found in the document 1999/30/EC and in the Clean Air for Europe Directive 2008/50/EC (EU Web References 1 and 2). The numerical limit values for PM_{10} are 24-h means of 50 μgm^{-3} not to be exceeded for over 35 days (≈10%) of the days of the year and an annual average of 40 μgm^{-3}. Both of these numbers refer to calendar years. The limit values became legally binding on January 1, 2005. In the spring of 2008, the EU decided on the future PM regulations. The details of their decision can be found in EU Web Reference 1. The above limit values have been retained in the new regulations, but some things have changed relating to their application: (1) member states can obtain permission to postpone compliance with the limit values until

2011; (2) the subtraction of "natural" PM has been increased; (3) spatial application criteria have been changed. All three modifications tend to reduce public health protection compared to the level of protection offered by the 1999 regulations, despite the fact that the numbers have not changed. In sharp contrast to the new regulations, the 1999 daughter directive stated that the annual limit value would be reduced to 20 μgm^{-3} by 2010, and that the number of allowed exceedances of the 24-h limit value would be reduced to 7 per year, also by 2010, in accordance with the levels recommended in EU 'Particles Position Paper' of April 8, 1997 as limit values (EU Web Reference 3). These recommendations have disappeared from the new directive thus keeping the exposure levels high and, as a result, raising health concerns (Brunekreef and Maynard 2008). The above-mentioned legal framework underlines the importance of identifying the sources of PM_{10}, and more specifically the exogenous (i.e. long-range transport) contribution of natural sources.

Several authors have published backward trajectory modelling to detect the long-range transport of pollutant air masses that may have an impact on local PM_{10} levels (Sanchez-Ccoyllo et al. 2006, Riccio et al. 2007, Salvador et al. 2008), to better describe the tropospheric circulations (Jorba et al. 2004) or to characterize and identify spatial and temporal trends of pollutants (Johnson et al. 2007, Coury and Dillner 2007). However, it should be stressed that single backward trajectories, generally applied so as to detect source regions of extreme PM episodes at a given site (Hongisto and Sofiev 2004, Cachorro et al. 2006, Bessagnet et al. 2008) are not suitable for establishing the exact path and origin of an air parcel.

On the other hand, large numbers of trajectories arriving at a given site can be analysed in order to discover the origin of polluted air masses. This can be performed using cluster analysis, a multivariate statistical technique used to determine the structure within a specific data set. Several authors have performed a cluster analysis in order to establish trajectories for a relatively small number of groups. However, it should be mentioned here that cluster analysis does not tell us anything about any cause-effect relationships. In other words, clustering is distinct from classification, because there are no pre-determined characteristics used to define the membership for a cluster, although items in the same cluster are likely to have many characteristics in common. This procedure has been frequently used to interpret the findings of studies concerning the origin and the transport of atmospheric pollution (Kallos et al. 2007, Grivas et al. 2008, Vardoulakis and Kassomenos 2008).

Several examples show that the long-range transport of particulates can considerably worsen air quality and, in this way, may increase the risk of respiratory as well as cardiovascular diseases. For this reason, the future PM regulations introduced by the EU in the spring of 2008 put more emphasis on the trans-boundary transport of particles, compared to those which became legally binding on January 1, 2005.

A change is that the 2008 Directives are complemented by clearly defining natural sources. Moreover, they include sea sprays as additional sources. Another change in article 20 is as follows. "Where the Commission has been informed of an exceedance attributable to natural sources in accordance with paragraph 1, that exceedance shall not be considered as an exceedance for the purposes of this Directive." In other words, days of exceedances traced back to natural origin pollutant sources are to be subtracted from pollution exceedance days, i.e. the regulations have been softened.

Both interpretations in the Directives (article 2.15 1999, article 20 2008) include "atmospheric re-suspension or transport of natural particles from dry regions". In the present study, PM_{10} levels from four cities (Thessaloniki, Southern Europe; Szeged, Central

Europe and Helsinki and Oulu, both Northern Europe) (Fig. 1) have been interpreted using a Mahalanobis metric over a 5-year period. Furthermore, a goodness of cluster analysis is examined for both the two-dimensional (2D) and three-dimensional (3D) trajectory positions for each of the four cities.

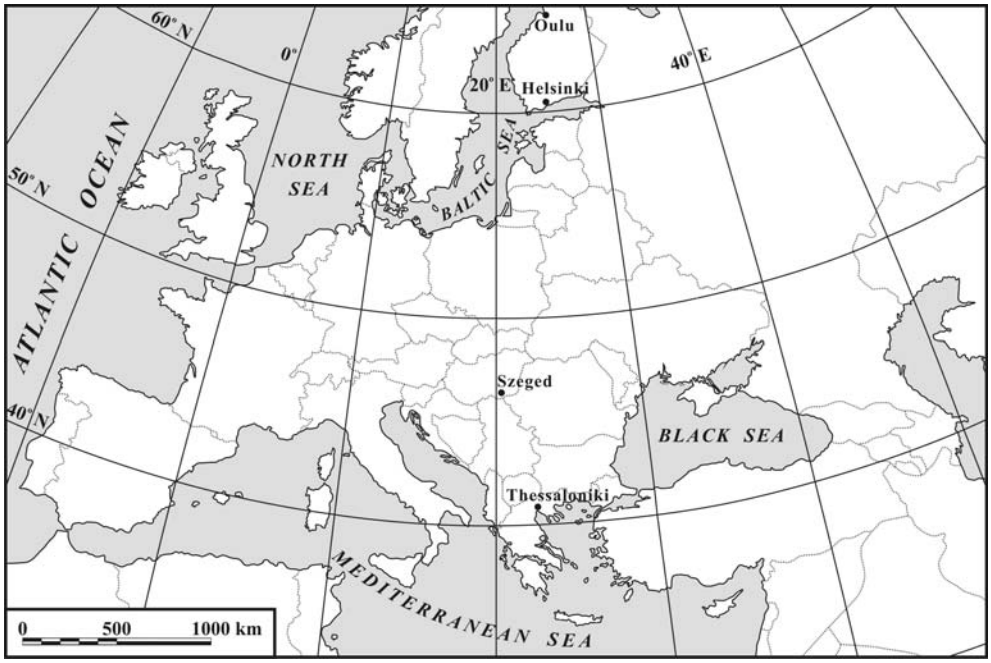

Fig. 1 Geographical position of Thessaloniki, Szeged, Helsinki and Oulu

2. MATERIALS

2.1. Study areas and monitoring data

In this paper five years of daily mean PM_{10} data (2001-2005) taken from four European cities - Thessaloniki (Greece), Szeged (Hungary) and Helsinki and Oulu (Finland) - have been utilised (Fig. 1), and will be presented later on.

Thessaloniki (40.64°N; 22.94°E) is the second largest city of Greece, where emissions come mainly from the local traffic, while the formation and transport of pollutants are heavily influenced by the local meteorological and topographical conditions (Fig. 1). The city is the centre of the Thessaloniki Metropolitan Area with 2 million inhabitants and a land area of 677 km^2. The mean temperature is 26.4°C in July and 4.6°C in January, while the annual mean precipitation total is 449 mm. Eleftherio-Kordelio was one of the seven air-pollution monitoring stations selected for our study. The selection was based on the fact that this station characterises an area in the western part of the urban web which greatly suffers from PM_{10} exceedances (Tzima et al. 2009). The data sets used were made available via the public air quality database system of the European Environment

Agency (http://air-climate.eionet.europa.eu/databases/airbase). PM_{10} is monitored with the aid of a β-radiation absorption method (www.pkm.gr) with HORIBA instrumentation.

Szeged (46.25°N; 20.10°E), the largest settlement in SE Hungary, is located at the confluence of the River Tisza and River Maros (Fig. 1). The area is characterised by an extensive flat landscape of the Great Hungarian Plain with an elevation of 79 m AMSL. The city is the centre of the Szeged region with 203,000 inhabitants and a land area of 876 km^2. The mean temperature is 22.4°C in July and -1.2°C in January. The annual mean precipitation total is 573 mm. The air quality monitoring station is located in the inner city area, near one of the busiest crossroads of the city. The station lies some 20 m away from the lead-in section of the M5 motorway with an average weekday traffic volume of 37,100 unit vehicles per day. PM_{10} concentrations are recorded by the absorption of β-radiation (type of instrument: FH 62 I-N; manufacturer: Eberline). The station is under the supervision of the Environmental and Natural Protection and Water Conservancy Inspectorate of the Lower-Tisza Region.

Helsinki (60.25°N; 25.05°E), the capital of Finland, forms the core of the Helsinki Metropolitan Area which has 1 million inhabitants and a land area of 745 km^2. The area is situated on the coast of the Gulf of Finland and is characterized as a lowland plain, which is typical for the whole of southern and western Finland. The mean temperature is 17.2°C in July and -4.9°C in February, while the annual mean precipitation total is 640 mm. The Helsinki Kallio air quality measurement station is an urban background site located in the inner city at the edge of a sports ground (20 m AMSL). The nearest road is some 80 m away with an average weekday traffic volume of about 8000 vehicles per day. The PM_{10} concentrations were recorded by the absorption of ß-radiation (Eberline FH 62 IR). The station is operated by the Helsinki Metropolitan Area Council (Fig. 1).

Oulu (65.02°N; 25.48°E) is the largest city in northern Finland and lies on the shores of the northernmost stretch of the Baltic Sea. The city of Oulu has 130,000 inhabitants and a land area of 370 km^2. The Oulu region is known for its rapidly growing high technology industry sector as well as its traditional pulp, paper and steel industries. The mean temperature is 16.2°C in July and -9.5°C in February, and the annual mean precipitation total is 450 mm. Towards the inland from Oulu, the flat coastal plain gives way to extensive boreal forests, mires and fells that are typical for the sparsely populated part of northern Finland. The Oulu Centre air quality measurement station is an urban traffic site which is located in the inner city of Oulu in a busy (c. 7000 vehicles per day) street canyon only five metres from the edge of the road (5 m AMSL). The PM_{10} concentrations are measured with a tapered element oscillating microbalance (TEOM 1400). The station is run by the Environment Office of the Oulu region (Fig. 1).

2.2. Backward trajectories

Backward trajectories for Thessaloniki, Szeged, Helsinki and Oulu were computed using the Hybrid Single-Particle Lagrangian Integral Trajectory (HYSPLIT, version 4.8; http://www.arl.noaa.gov/ready/hysplit4.html) model (Draxler and Hess 1998). NOAA global reanalysis meteorological data were used in this procedure. These data sets are a subset of sea-level meteorological data received from NCEP/NCAR National Centres for Environmental Prediction / National Centre for Atmospheric Research (http://dss.ucar.edu/datasets/ds090.0/).

Since a single backward trajectory has a large uncertainty and is of limited significance (Stohl 1998), a more reliable representation of the synoptic air currents in the given regions was performed through the reconstruction and analysis of a large number of atmospheric trajectories. In this study 4-day, 6-hourly 3D backward trajectories arriving at the four locations at 1200 GMT were determined in order to describe the horizontal and vertical movements of an air parcel for the above-mentioned four cities. Trajectories arriving at heights h=500, 1500 and 3000 m AMSL for each day over a 5-year period from 2001 to 2005 were computed. Then from the trajectories we selected just the 6-hourly positions of their tracks.

3. METHODS

3.1. Cluster analysis

Cluster analysis is a common statistical technique for grouping elements in an objective way, such as atmospheric trajectories within large data sets. The aim of the method is to maximize the homogeneity of elements (in our case, backward trajectories) within the clusters and also to maximize the heterogeneity between the clusters. It has been shown that clusters of backward trajectories arriving at a given location can be replaced by different synoptic circulation patterns (Dorling et al. 1992). In this paper a non-hierarchical cluster analysis with the k-means method was applied using a Mahalanobis metric (Mahalanobis 1936). Input data, as clustering variables, were the 6-hourly co-ordinate values of the trajectories (φ, λ for 2D and φ, λ, h for 3D) (called trajectory positions) for the three different height levels considered. The trajectories produced for the three arrival heights were collected and their positions were analysed together for each city.

When starting the procedure, k cluster centres are selected automatically by the algorithm. In our case, we chose cluster numbers from 30 to 1, in decreasing order, and the algorithm determines the initial cluster centres for the above cluster numbers randomly. The individual points will join those clusters for which the distance of the point from the cluster centres is a minimum. Next, we add up the distances of the cluster centre and the cluster points cluster by cluster (RMSD = Root Mean Square Deviation) and afterwards we add up the RMSD values in a cluster-by-cluster manner (total RMSD). The total RMSD value usually decreases as the number of clusters increases. We perform the same clustering procedure 9 more times, in order to omit those false cases when empty clusters occur. (Empty clusters may occur such that, when modifying the cluster centres, the points within the clusters approach another cluster centre and, in this way, they will be placed into this other cluster.) Out of the ten clustering procedures analysed, only the one with the minimum total RMSD-value will be retained.

The total RMSD (%) is used in order to select the most compact clustering with the least cluster number from the clustering with 30 to 1 cluster numbers, in decreasing order. The total RMSD (%), associated with a given cluster number, tells us about its change (increase) when selecting one higher cluster number in the clustering procedure. Out of all the pairwise neighbouring comparisons of the clusters, going from the higher cluster number to the smaller ones, we will stop when we reach the highest total RMSD (%) and choose the clustering with one higher cluster number than that indicating the highest total RMSD (%).

Applying a change in CRMSD (namely, a change in the change of RMSD) makes it easier to choose the proper cluster number than that for a change in RMSD. A change in CRMSD (%) for any cluster number is obtained if a change in RMSD (%) associated with a given cluster number is subtracted from that associated with the smaller cluster number. Namely, Change in CRMSD($cluster_{i+1}$) = Change in RMSD($cluster_i$) – Change in RMSD($cluster_{i+1}$). In this case we will choose the cluster number of all the pairwise neighbouring comparisons of the clusters, going from 30 to 1, in decreasing order, for which the change in CRMSD is the highest.

Clustering with the k-means algorithm was performed by MATLAB 7.5.0 software.

3.2. Analysis of variance (ANOVA)

Before applying the ANOVA technique, the annual course of pollution data is removed and standardized data will be used hereafter. Standardized data sets are free of annual cycles and so distinguishing between average pollutant levels corresponding to types of trajectory positions is due to the types themselves and not related to periods of the year.

With a data set x_t, $t = 1, \ldots, n$ the expected value function m(t) of x_t is approximated by

$$m(t) = a_0 + \sum_{k=1}^{2} a_k \cdot \cos(w_k \cdot t) + b_k \cdot \sin(w_k \cdot t), \quad (1)$$

where

$$w_1 = 2 \cdot \pi / 365.25, \quad w_2 = 2 \cdot w_1. \quad (2)$$

According to the least squares concept, the linear system of equations

$$\sum_{k=1}^{5} c_k \cdot \left(\sum_t u(t) \cdot f_j(t) \cdot f_k(t)\right) = \sum_t u(t) \cdot x_t \cdot f_j(t), \quad j = 1, \ldots, 5 \quad (3)$$

should be solved for the vector $c = (c_1, c_2, c_3, c_4, c_5) = (a_0, a_1, a_2, b_1, b_2)$, where

$$f_1(t) = 1, \quad f_2(t) = \cos(w_1 \cdot t), \quad f_3(t) = \cos(w_2 \cdot t), \quad f_4(t) = \sin(w_1 \cdot t), \quad f_5(t) = \sin(w_2 \cdot t), \quad (4)$$

Furthermore, $u(t) = 0$, if x_t is absent and $u(t) = 1$ otherwise. Then the annual course free and standardized data set is

$$y_t = (x_t - m(t)) / d(t), \quad t = 1, \ldots, n, \quad (5)$$

where $d^2(t)$ is estimated as $m(t)$ by equations (1) to (4) but with x_t replaced by $x_t^* = (x_t - m(t))^2$, $t = 1, \ldots, n$.

The variability of the daily PM_{10} concentrations for each city was analysed in order to see whether the clusters retained had any influence on the PM_{10} levels. More precisely, based on the F-test, we examined whether the differences between the cluster-averaged PM_{10} concentrations were significant. If, after performing ANOVA, we get a significant difference among cluster averages of PM_{10} concentrations, then the Tukey-test should be

performed in order to detect those cluster pairs that are associated with significantly different PM_{10} averages. In the reverse case if no significant difference is found, the Tukey test is omitted (Tukey 1985).

The results of ANOVA should help clarify the possible role of long-range transport on local pollutant levels. Significant differences between PM_{10} concentrations of different cluster pairs may indicate that there is an important influence of the origin of air masses on local PM_{10} levels.

There are several versions available for comparing means calculated from the 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. Indeed, when the null-hypothesis on the equality of two expected values corresponding to two subsamples is rejected, the Tukey test provides a better certainty compared, for instance, to the Fisher test. This is why we chose to use the Tukey test. This or any other similar test assumes a statistical independence of the data. Consecutive air pollution data, however, may be correlated and give higher variances of the estimated means than those for uncorrelated data. The autocorrelation structure is modelled via first order autoregressive processes conditioned on the types of trajectory positions. For this, estimates of type-dependent one lag autocorrelations a are required.

In order to compare the averages corresponding to the ith and jth clusters (i.e. M_i and M_j), the test statistic

$$t = \frac{M_i - M_j}{\sqrt{MSE / n_h}}, \tag{6}$$

is used, where MSE is the empirical variance calculated from concentration data associated with both clusters. (Note that the Fisher test uses the entire data set when calculating the MSE.) If n_i and n_j denote the values of data corresponding to individual clusters, then

$$\frac{1}{n_h} = \frac{1+a_i}{1-a_i} \cdot \frac{1}{n_i} + \frac{1+a_j}{1-a_j} \cdot \frac{1}{n_j}. \tag{7}$$

Here Eq. (7) becomes $1/n_h = 1/n_i + 1/n_j$ for uncorrelated data (a=0), resulting in the classical Tukey test statistic.

Next we applied a simplification for estimating type dependent one lag autocorrelations. As the average time step between two consecutive concentration data sets associated with the same type is n/n_i we will write $a_i = R(n/n_i)$, where $R(k)$ is the autocorrelation for lag k estimated from the entire data set via

$$R(k) = B(k) / B(0), \tag{8}$$

where

$$B(k) = \frac{1}{N_k} \cdot \sum_{t=1}^{n-k} u(t) \cdot u(t+k) \cdot (y_t - \bar{y}_u) \cdot (y_{t+k} - \bar{y}_u), \quad k = 0,\ 1,\ 2,\ \ldots. \tag{9}$$

Here

$$N_k = \sum_{t=1}^{n-k} u(t) \cdot u(t+k) \tag{10}$$

and

$$\bar{y}_u = \frac{1}{n_k} \cdot \sum_{t=1}^{n} u(t) \cdot y_t \tag{11}$$

with

$$n_k = \sum_{t=1}^{n} u(t) \quad . \tag{12}$$

3.3. Goodness of cluster analysis and comparison of 2D – 3D clustering

When comparing a goodness of cluster analysis in 2D – 3D clustering, the procedure should give good results if concentration values associated with the 2D or 3D clusters are well separated. For this purpose, two methods can be applied. 1) We can count the number of cluster pairs having significantly different average concentrations both for 2D and 3D clustering procedures. Then these values for both clustering procedures should be divided by the total number of cluster pairs. 2) We can compare the internal and external distances of the two cluster systems. In this case we should work with the same metric (Mahalanobis metric) used in the cluster analysis. Here distances should be divided by 96 (4 days x 4 moments/day x 2 dimensions x 3 levels) and 144 (4 days x 4 moments/day x 3 dimensions x 3 levels), respectively to allow one to make a better comparison. Let this ultimate distance be $d(x,y)$ between vectors x and y. Take the internal distance for the ith cluster as the average of distances between the cluster elements and the M_i cluster centre using distance d. Next, consider the average of the internal distances of the individual clusters. This will be the internal distance ID. The smaller the ID the better the cluster analysis is, because the clusters will be compact. Take distances $d_{ij}=d(M_i,M_j)$, $i,j=1,\ldots,N$ forming a symmetric NxN matrix (with zero diagonal elements), where N is the number of clusters. The higher the d_{ij} values, the better the cluster analysis since each cluster will be well separated. The average

$$\frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j,j \neq i} d_{ij} \tag{13}$$

of these distances is called external distance ED.

Although an ideal system of clusters has a small ID with a large ED, increasing the cluster number usually results in a smaller ID with a smaller ED for a given data set to be clustered. In other words, when we use more and more clusters the price to be paid for improving compactness is a worsening average separation. However, with two different data sets (2D and 3D data sets) it may happen that a) clustering with a smaller cluster number produces a smaller ID with a larger ED; b) clustering with a bigger cluster number gives a smaller ID with a smaller ED; c) any other possible case. In the first case, the clustering with fewer clusters is made more efficient, as it produces more compact and

more separated clusters. In the second and third cases (especially in the second case) subjective judgements are needed when comparing the internal and external distances for the 2D and 3D situations.

3.4. Statistical characterization of PM_{10} episodes

The role of long-range transport was studied by analysing cluster occurrence on days when 24-h mean PM_{10} concentrations exceeded the limit value of 50 μgm^{-3}. For this reason, two statistical indices related to the probability (INDEX1) and frequency (INDEX2) of daily PM_{10} episodes associated with different clusters of trajectory positions were calculated in the same manner as Borge et al. (2007). For a given site and cluster, INDEX1 is defined as

$$INDEX1_i(\%) = \frac{D_{(\rangle 50)i} \cdot 100}{D_i}, \tag{14}$$

where D_i is the number of days for which a backward trajectory associated with cluster i is present (namely, number of occurrences of cluster i), and $D_{(>50)i}$ is the number of 24-h PM_{10} exceedances (namely, days with 24-h mean PM_{10} > 50 μgm^{-3} with occurrence of cluster i. INDEX1 tells us the likelihood of an exceedance occurring for a given cluster. In a similar way, INDEX2 is defined as

$$INDEX2_i(\%) = \frac{D_{(\rangle 50)i} \cdot 100}{E}, \tag{15}$$

where E is the total number of 24-h PM_{10} exceedances recorded at a given site. INDEX2 can be interpreted as the likelihood of certain trajectory position being present on a PM_{10} exceedance day. These two cluster-related indices whose values range from 0 to 100% provide complementary information on the influence of different atmospheric circulation patterns on PM_{10} levels (Borge et al. 2007).

Acknowledgements – The authors would like to thank Roland Draxler for his useful advice and consultations on the HYSPLIT model, version 4.8; to Gábor Motika (Environmental Protection Inspectorate of Lower-Tisza Region, Szeged), to Helsinki Metropolitan Area Council and to Environment Office of the Oulu region for providing PM_{10} data and to Zoltán Sümeghy for the digital mapping in Fig. 1. This study was supported by the EU-6 Quantify project [No. 003893 (GOCE)].

REFERENCES

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

Annesi-Maesano I, Forastiere F, Kunzli N, Brunekref B (2007) Particulate matter, science and EU policy. Eur Respir J 29:428-431

Borge R, Lumbreras J, Vardoulakis S, Kassomenos P, Rodriguez E (2007) Analysis of long-range transport influences on urban PM_{10} using two-stage atmospheric trajectory clusters. Atmos Environ 41:4434-4450

Brunekreef B, Maynard RL (2008) A note on the 2008 EU standards for particulate matter. Atmos Environ 42:6425-6430
Coury C, Dillner AM (2007) Trends and sources of particulate matter in the Superstition Wilderness using air trajectory and aerosol cluster analysis. Atmos Environ 41:9309-9323
Dorling SR, Davies TD, Pierce CE (1992) Cluster analysis: a technique for estimating the synoptic meteorological controls on air and precipitation chemistry – method and applications. Atmos Environ 26:2575-2581
Draxler RR Hess GD (1998) An overwiew of the HYSPLIT_4 modelling system for trajectories, dispersion and deposition. Aust Meteorol Mag 47:295-308
EU Web Reference 1 <http://wwweuroparleuropaeu/sides/getDocdo?TypeĽTA&ReferenceĽP6-TA-2007-0596&languageĽEN> (accessed 290208)
EU Web Reference 2 <http://eur-lexeuropaeu/LexUriServ/LexUriServdo?uriĽOJ:L:1999:163:0041:0060:EN:PDF> (accessed 290208)
EU Web Reference 3 <http://eceuropaeu/environment/air/pdf/pp_pmpdf> (accessed 290208)
Grivas G, Chaloulakou A, Kassomenos P (2008) 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 389:165-177
Hongisto M, Sofiev M (2004) Long-range transport of dust to the Baltic Sea region. Int J Environ Pollut 22:72-86
HYSPLIT_4 Model version 4.8 (1997) (HYbrid Single-Particle Lagrangian Integrated Trajectory) http://wwwarlnoaagov/ready/hysplit4html, NOAA, Air Resources Laboratory, Silver Spring, MD
Johnson D, Mignacca D, Herod D, Jutzi D Miller H (2007) Characterization identification of trends in average ambient ozone and fine particulate matter levels through trajectory cluster analysis in eastern Canada. J Air Waste Manage 57:907-918
Jorba O, Pere, C, Rocadenbosch F, Baldasano JM (2004) Cluster analysis of 4-day backward trajectories arriving in the Barcelona area, Spain, from 1997 to 2002. J Appl Meteorol 43:887-901
Kallos G, Astitha M, Katsafados P, Spyrou C (2007) Long-range transport of anthropogenically and naturally produced particulate matter in the Mediterranean and North Atlantic: Current state of knowledge. J Appl Meteorol Climatol 46:1230-1251
Mahalanobis PC (1936) On the generalized distance in statistics Proc Nat Inst Sci India 2(1), 49-55
Riccio A, Giunta G, Chianese E (2007) The application of a trajectory classification procedure to interpret air pollution measurements in the urban area of Naples (Southern Italy). Sci Total Environ 376:198-214
Salvador P, Artinano 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. Sci Total Environ 390:495-506
Sanchez-Ccoyllo OR, Silva Dias PL, Andrade MD, Freitas SR (2006) Determination of O-3-, CO- and PM_{10}-transport in the metropolitan area of Sao Paulo, Brazil through synoptic-scale analysis of backward trajectories. Meteorol Atmos Phys 92:83-93
Stohl A (1998) Computation, accuracy and application of trajectories – A review and bibliography. Atmos Environ 32:947-966
Tukey JW (1985) The problem of multiple comparisons (1953) (unpublished manuscript) In: The Collected Works of John W Tukey, Volume II: Time Series, 1965-1984 Wadsworth Advanced Books & Software, Monterey, CA
Tzima F, Niska H, Kolehmeinenen M, Karatzas K Mitkas P (2009) An experimental evaluation of ZCS-DM for the prediction of urban air quality 4th International Symposium on Information Technologies in Environmental Engineering, Thessaloniki, 28-29 May (accepted)
Vardoulakis S, Kassomenos P (2008) Sources and factors affecting PM_{10} levels in two European cities: Implications for local air quality management. Atmos Environ 42:3949-3963