

# Identification of PM10 air pollution origins at a rural background site

Magdalena Reizer<sup>1,\*</sup>, José A. G. Orza<sup>2</sup>

<sup>1</sup>Warsaw University of Technology, Faculty of Building Services, Hydro and Environmental Engineering, 00653 Warsaw, Poland

<sup>2</sup>SCOLab, Física Aplicada, Universidad Miguel Hernández, 03202 Elche, Spain

**Abstract.** Trajectory cluster analysis and concentration weighted trajectory (CWT) approach have been applied to investigate the origins of PM10 air pollution recorded at a rural background site in North-eastern Poland (Diabla Góra). Air mass back-trajectories used in this study have been computed with the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model for a 10-year period of 2006–2015. A cluster analysis grouped back-trajectories into 7 clusters. Most of the trajectories correspond to fast and moderately moving westerly and northerly flows (45% and 25% of the cases, respectively). However, significantly higher PM10 concentrations were observed for slow moving easterly (11%) and southerly (20%) air masses. The CWT analysis shows that high PM10 levels are observed at Diabla Góra site when air masses are originated and passed over the heavily industrialized areas in Central-Eastern Europe located to the south and south-east of the site.

## 1 Introduction

Variability of air pollutant concentrations recorded at a specific site, depending on the species, may vary on time scales from minutes to decades. Over a longer time scale the changes are determined by seasonal or annual cycles in the activity of emission sources, while on a time scale of days or less the concentration variability is closely related to the immediate history of the air before arriving at the sampling point [1]. Back-trajectory analysis comprises a number of methods to identify transport pathways of air masses affecting a study site as well as to determine potential source areas of air pollution. In order to analyse the association between trajectories and concentrations of various species in air arriving at a site, a number of methods to carry out trajectory classifications has been applied. These can generally be split into the following methodological groups: (1) assigning trajectories according to air mass sectors and comparing the air pollution composition in those sectors; (2) clustering the trajectories using a multivariate statistical technique and analysing the concentrations at the receptor site for each trajectory classification; (3) residence time analysis which generates a probability density function identifying the likelihood that an air mass would traverse a given region *en route* to the site

---

\* Corresponding author: [magdalena.reizer@pw.edu.pl](mailto:magdalenareizer@pw.edu.pl)

of interest over a given time period [2]; this often includes conditional probability densities linking residence times and the concentrations measured at the study site.

There has been a number of studies reporting the use of back-trajectory analyses for determining long range transport of air pollution (see e.g. [3, 4]) or characterizing Saharan dust outbreaks in the Mediterranean Basin (see e.g. [5, 6]) and Canary Islands (e.g. [7]). However, there is a lack of studies of air mass back-trajectories for Central-European countries. In scarce previous studies performed for this region, probable sources of particulate matter (PM) air pollution were identified by trajectory classification into wind sectors at urban site for Zagreb, Croatia [8] as well as back-trajectory analysis for selected days with the highest PM concentrations observed in Polish urban areas [9]. A cluster analysis was also carried out for rural background site in Belsk, Poland [10] as well as for urban sites located in Bucharest, Romania and Szeged, Hungary [11]. Different methods of trajectories interpolation were used for Polish urban sites in Gdańsk and Katowice as well as for rural background site in Diabla Góra [12]. Moreover, cluster analysis, potential source contribution function (PSCF) and concentration weighted trajectory (CWT) methods were used to identify the main atmospheric circulation pathways influencing Black Carbon (BC) [13] and aerosol particle number concentration (PNC) [14] observed in Lithuania. Dvorska et al. [15] applied PSCF method in order to identify potential source areas of persistent organic pollutants (POPs) at rural background site in Kosetice (Czech Republic).

The aim of the present paper is to investigate potential source areas of air pollution contributing to PM<sub>10</sub> levels recorded at a Central-European rural background site during the long-term period of 2006–2015. Cluster analysis of back-trajectories has been applied to assess the main transport pathways of air masses, while concentration weighted trajectory approach has been used for identification of potential PM<sub>10</sub> source regions.

## **2 Materials and methods**

### **2.1 Study area**

In this study air mass back-trajectories arriving over the monitoring site of Diabla Góra, Poland were investigated. Diabla Góra is a rural background monitoring site operated within the EMEP air quality network (European Monitoring and Evaluation Programme). This site is located in the north-eastern part of Poland (54° 08' 00" N, 22° 04' 01" E, 157 m a.s.l.), far from dense populated or traffic areas (Fig. 1). The nearest town (>10 000 inhabitants) and road (>50 cars per day) are situated 20 km and 16 km away from this site, respectively.



**Fig. 1.** Location of the study area.

Daily PM10 concentrations were determined by the reference gravimetric method (Standard EN 12341). PM10 data for the period from 1 January 2006 to 31 December 2015 was collected from the EBAS database (<http://ebas.nilu.no>).

## 2.2 Back-trajectories calculations

Three dimensional 96-hour back-trajectories of air parcels arriving at Diabla Góra site at the height of 200 m above sea level were computed at 00, 06, 12 and 18 UTC each day during the period 2006–2015. The trajectories were calculated by the PC version of the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPPLIT) model [16]. The meteorological data used for trajectory calculations was provided by the ECMWF (European Centre for Medium-Range Weather Forecasts).

## 2.3 Cluster analysis

The k-means procedure, which is the most commonly used form of non-hierarchical clustering for trajectory studies, is an iterative algorithm that uses a specified number of clusters,  $k$ , to partition the data by comparing each object to the arithmetic mean of all the members of each of the clusters (cluster centroids). The selection of the optimal number of clusters that best describes the different air flow patterns is performed by computing the percentage change in within-cluster variance, as a function of the number of clusters [17]. The assignment of trajectories to a given cluster is carried out by minimising the internal variability within the group of trajectories and maximising the external variability between different groups based on the trajectory co-ordinates based on the Root Mean Square Deviation (RMSD) of each trajectory from its cluster centroids [2].

In the present study the k-means cluster analysis of back-trajectories following the procedure described in detail by Orza et al. [18] was carried out. Hourly latitude and longitude were used as input variables in the clustering procedures, while the similarity metrics was based on great-circle distances. The selection of the appropriate number of clusters was done by obtaining the best solution for each number of clusters (from a large number of initializations) and then by the analysis of the percentage change in the total RMSD when decreasing one by one the number of clusters.

Daily PM10 data was assigned to a cluster for a given day, if three out of four trajectories for that day fell into the same cluster. Non-parametric pairwise Mann-Whitney

tests with the Dunn-Sidak adjustment for multiple comparisons were used to test the significance of inter-cluster variation in PM10 concentration.

## 2.4 Concentration weighted trajectory

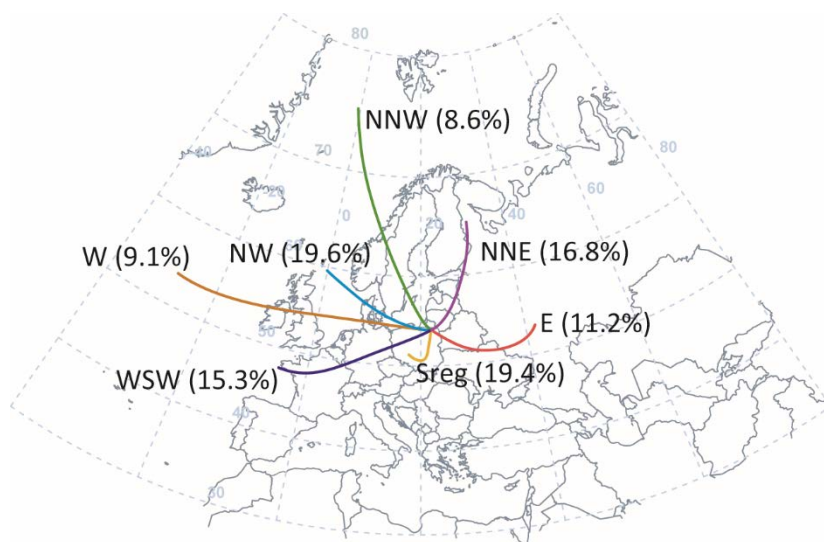
The concentration weighted trajectory (CWT) is a method of weighting trajectory residence times with associated air pollutant concentrations [19]. This procedure assigns to each grid cell a weighted concentration obtained by averaging sample concentrations associated to trajectories that crossed that grid cell, i.e. each concentration is used as a weighting factor for the residence times of all the trajectories in each grid cell which correspond to that concentration and then it is divided by the cumulative residence time from all trajectories. CWT method employs an arbitrary weight function to minimize the inaccuracy caused by the small number of polluted trajectories.

In summary, weighted concentration fields show concentration gradients across potential sources and helps to identify the relative significance of potential sources [19].

## 3 Results and discussion

### 3.1 Cluster analysis

Trajectories arriving at Diabla Góra at 200 m are found to be clustered in 7 groups (Fig. 2). The groups correspond to the major advection patterns at the study site are named according to the main direction they come from.



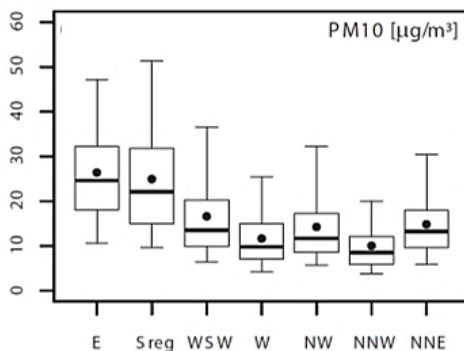
**Fig. 2.** Cluster centroids of the trajectories reaching Diabla Góra, Poland at 200 m in the period 2006–2015. The percentage of trajectories occurring in each cluster is shown in parentheses.

Most of the trajectories correspond to westerly flows (44% of the cases), bringing air masses from the northern part of the North Sea, through Denmark and North Poland (NW, 19.6%), North France, Central Germany and western part of Poland (WSW, 15.5%), as well as from the Atlantic Ocean, through Great Britain, southern part of Denmark and northern

part of Poland (W, 9.1%). The second type of trajectories groups the northerly flows (over 25%), bringing air masses from Scandinavian Peninsula, passing over northern part of Russia, Latvia and Lithuania (NNE, 16.8%), as well as from the southern part of Greenland, through Scandinavia and the Southern Baltic Sea (NNW, 8.6%). The third type of flow includes short (regional) trajectories, bringing air masses from highly polluted Southern Poland (Upper Silesia and Lesser Poland) through central part of the country (Sreg, 19.4%). The last type of trajectories includes the air masses coming from western part of Russia, through the Belarusian-Ukrainian border and the eastern part of Poland (E, 11.2%).

The highest mean PM10 concentrations ( $26 \mu\text{g}/\text{m}^3$ ) observed at Diabla Góra site are associated with slow moving easterly flows (Fig. 3). As these flows are mainly associated to the westward extension of the Siberian high located to the east of the study site, cluster E is also associated with the highest mean air pressure – 1 001 hPa (with maximum of 1 044 hPa and one of the lowest mean air temperatures of  $5.0 \text{ }^\circ\text{C}$  (with minimum of  $-21.5^\circ\text{C}$ ). Quite high PM10 concentrations ( $25 \mu\text{g}/\text{m}^3$ ) are also related with much warmer ( $9.2^\circ\text{C}$ , with minimum of  $-11.5^\circ\text{C}$ ) slow moving air masses from the South (Sreg). In opposite, the lowest mean PM10 concentrations ( $10 \mu\text{g}/\text{m}^3$ ) are recorded when air masses arrive with fast moving NNW flows (Fig. 3).

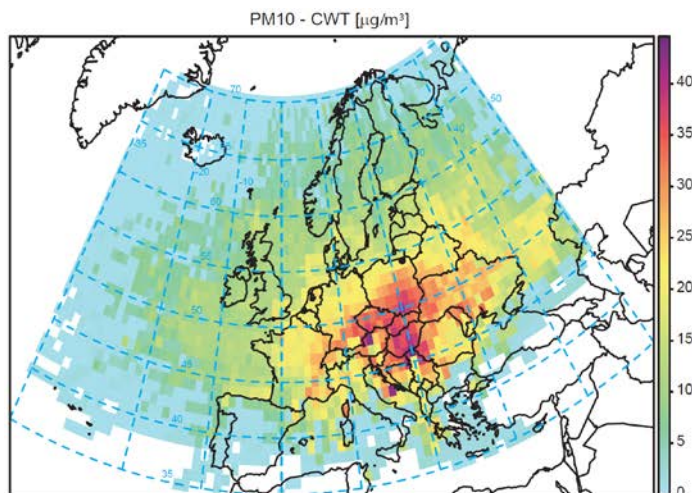
Overall, the Mann-Whitney test results have shown that mean PM10 concentrations associated with slow moving easterly and southerly air masses (clusters E and Sreg) were significantly higher compared to those observed for other clusters.



**Fig. 3.** Boxplots of PM10 concentrations ( $\mu\text{g}/\text{m}^3$ ) for each trajectory cluster arriving at Diabla Góra, Poland. The horizontal lines and dots indicate the median and mean values, respectively, the boxes cover the 25th-75th percentiles. The length of the whiskers is 1.5 times the interquartile range.

### 3.2 Concentration weighted trajectory

The concentration weighted trajectory map for PM10 in Diabla Góra during study period is shown in Fig. 4.



**Fig. 4.** CWT map for PM10 levels at Diabla Góra, Poland.

Cells with high (above  $30 \mu\text{g}/\text{m}^3$ ) and moderate (above  $20 \mu\text{g}/\text{m}^3$ ) CWT values were found south and south-east of Diabla Góra site, respectively, in good accordance with the previous analysis of PM10 levels by advection pattern. The high CWT values for short lengths of the S and E trajectories indicates stable situations with reduced removal of pollutants and slow air movement over Central-Eastern Europe due to the influence of high-pressure centres. As a result, polluted air masses with a high PM10 content, originating from southern and eastern part of Europe may accumulate during several days under anticyclonic situations and they may remain there for several days before reaching the study site.

## 4 Summary and conclusions

A database of 96-h back-trajectories arriving to North-eastern part of Poland (Diabla Góra) at 200 m computed for the 10-year period from January 2006 to December 2015 was applied to describe the main flow patterns over the study site, as well as to identify PM10 air pollution origins at the rural background site. A cluster analysis grouped back-trajectories into 7 clusters. Most of the trajectories correspond to fast and moderately moving westerly (45% of the cases) and northerly (25%) flows, followed by the short (regional) trajectories, bringing air masses from the South (20%) and from the East (11%). Mean PM10 concentrations associated with slow moving easterly and southerly air masses are significantly higher compared to those observed for other clusters. The CWT approach has identified the heavily industrialized areas in Central-Eastern Europe as probable PM10 source areas.

This work was partially supported by the Dean of the Faculty of Building Services, Hydro and Environmental Engineering, Warsaw University of Technology within the Grant no. 504/02637/1110/42.000100. The authors also gratefully acknowledge the NOAA Air Resources Laboratory (ARL) for the provision of the HYSPLIT model and the European Centre for Medium-Range Weather Forecasts (ECMWF) for making available the ERA-Interim database.

## References

1. J.N. Cape, J. Methven, L.E. Hudson, *Atmos. Environ.* **34**, 3651-3663 (2000)
2. Z.L. Fleming, P.S. Monks, A.J. Manning, *Atmos. Res.* **104-105**, 1-39 (2012)
3. K. Dimitriou, P. Kassomenos, *Atmos. Environ.* **96**, 154-162 (2014)
4. M. Reizer, K. Juda-Rezler, *Air Qual. Atmos. Health* **9**, 517-531 (2016)
5. X. Querol, J. Pey, M. Pandolfi, A. Alastuey, M. Cusack, N. Pérez, T. Moreno, M. Viana, N. Mihalopoulos, G. Kallos, S. Kleanthous, *Atmos. Environ.* **43**, 4266-4277 (2009)
6. M. Cabello, J.A.G. Orza, V. Galiano, G. Ruiz, *Adv. Sci. Res.* **2**, 65-70 (2016)
7. S. Alonso-Pérez, E. Cuevas, X. Querol, *Meteorol. Atmos. Phys.* **113**, 109-124 (2011)
8. I. Bešlić I., K. Šega, M. Čacković, Z. Bencetić, A. Bajić, *Bull. Environ. Contam. Toxicol.* **80**, 270-273 (2008)
9. K. Juda-Rezler, M. Reizer, J.-P. Oudinet, *Atmos. Environ.* **45**, 6557-6566 (2011)
10. A. Pietruczuk, J. Jarosławski, *Acta Geophys.* **61**, 445-462 (2013)
11. L. Makra, I. Ionel, Z. Csépe, I. Matyasovszky, *Sci. Total Environ.* **458-460**, 36-46 (2013)
12. J. Godłowska, M.J. Hajto, M. Tomaszewska, *Arch. Environ. Prot.* **41**, 28-35 (2015)
13. S. Byčenkienė, V. Dudoitis, V. Ulevicius, *Adv. Meteorol.*, 137694 (2014)
14. S. Byčenkienė, K. Plauškaitė, V. Dudoitis, V. Ulevicius, *Atmos. Res.* **143**, 279-292 (2014)
15. A. Dvorska, G. Lammel, J. Klanova, I. Holoubek, *Environ. Pollut.* **156**, 403-408 (2008)
16. A.F. Stein, R.R. Draxler, G.D. Rolph, B.J.B. Stunder, M.D. Cohen, F. Ngan, *Bull. Am. Meteorol. Soc.* **96**, 2059-2077 (2015)
17. S.R. Dorling, T.D. Davies, C.E. Pierce, *Atmos. Environ.* **26**, 2575-2581 (1992)
18. J.A.G. Orza, M. Cabello, V. Galiano, A.T. Vermeulen, A.F. Stein, *Lagrangian Modeling of the Atmosphere*, 127-141 (2012)
19. A. Stohl, *Atmos. Environ.* **30**, 579-587 (1996)