



Contents lists available at ScienceDirect

Atmospheric Environment

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

Estimated accuracy of three common trajectory statistical methods

Vitaliy P. Kabashnikov^{a,*}, Anatoli P. Chaikovskiy^a, Tom L. Kucsera^b, Natalia S. Metelskaya^a^a B.I. Stepanov Institute of Physics National Academy of Sciences of Belarus, Nezavisimosti ave., 68, 220072 Minsk, Belarus^b USRA - Goddard Earth Science Technology and Research (GESTAR) Code 613.3, NASA Goddard Space Flight Center Greenbelt, MD 20771, USA

ARTICLE INFO

Article history:

Received 19 April 2011

Received in revised form

28 June 2011

Accepted 4 July 2011

Keywords:

Trajectory statistical method

Back trajectories

Source reconstruction

Atmospheric trace substance

ABSTRACT

Three well-known trajectory statistical methods (TSMs), namely concentration field (CF), concentration weighted trajectory (CWT), and potential source contribution function (PSCF) methods were tested using known sources and artificially generated data sets to determine the ability of TSMs to reproduce spatial distribution of the sources. In the works by other authors, the accuracy of the trajectory statistical methods was estimated for particular species and at specified receptor locations. We have obtained a more general statistical estimation of the accuracy of source reconstruction and have found optimum conditions to reconstruct source distributions of atmospheric trace substances. Only virtual pollutants of the primary type were considered.

In real world experiments, TSMs are intended for application to *a priori* unknown sources. Therefore, the accuracy of TSMs has to be tested with all possible spatial distributions of sources. An ensemble of geographical distributions of virtual sources was generated. Spearman's rank order correlation coefficient between spatial distributions of the known virtual and the reconstructed sources was taken to be a quantitative measure of the accuracy. Statistical estimates of the mean correlation coefficient and a range of the most probable values of correlation coefficients were obtained. All the TSMs that were considered here showed similar close results.

The maximum of the ratio of the mean correlation to the width of the correlation interval containing the most probable correlation values determines the optimum conditions for reconstruction. An optimal geographical domain roughly coincides with the area supplying most of the substance to the receptor. The optimal domain's size is dependent on the substance decay time. Under optimum reconstruction conditions, the mean correlation coefficients can reach 0.70–0.75. The boundaries of the interval with the most probable correlation values are 0.6–0.9 for the decay time of 240 h and 0.5–0.95 for the decay time of 12 h. The best results of source reconstruction can be expected for the trace substances with a decay time on the order of several days.

Although the methods considered in this paper do not guarantee high accuracy they are computationally simple and fast. Using the TSMs in optimum conditions and taking into account the range of uncertainties, one can obtain a first hint on potential source areas.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Trajectory statistical methods (TSM) were first proposed by Ashbaugh (1983) and Ashbaugh et al. (1985) and have been widely used for identifying source areas of atmospheric trace substances (Poirot and Wishinski, 1986; Seibert et al., 1994; Stohl, 1996; Rúa

et al., 1998; Wotawa and Kroeger, 1999; Zeng and Hopke, 1989; Charron et al., 2000; Ferrarese et al., 2002; Apadula et al., 2003; Lee et al., 2004; Zhou et al., 2004; Zhao and Hopke, 2006; Wang et al., 2006; Du and Rodenburg, 2007; Heo et al., 2009). In a number of works, attempts were made to validate TSM (Poirot and Wishinski, 1986; Hopke et al., 1993; Vasconcelos et al., 1996a, 1996b; Wotawa and Kroeger, 1999; Charron et al., 2000; Cheng and Lin, 2001; Ferrarese, 2002; Lupu and Maenhaut, 2002; Zhou et al., 2004; Begum et al., 2005; Lee and Ashbaugh, 2007a,b). Most of the validations were performed through subjective qualitative comparison with known sources. Only a few papers were devoted to the quantitative comparison with the European Monitoring and Evaluation Programme (EMEP) inventory. Papers by Stohl (1996), Seibert (1999) and Scheffinger and Kaiser (2007) dealt

Abbreviations: TSM, trajectory statistical method; CF, concentration field; CWT, concentration weighted trajectory; PSCF, potential source contribution function; EMEP, European Monitoring and Evaluation Programme; PDF, probability density function.

* Corresponding author. Tel.: +375 17 284 22 60; fax: +375 17 284 08 79.

E-mail addresses: v.kabashnikov@dragon.bas-net.by, v.kabashnikov@rambler.ru (V.P. Kabashnikov).

with SO₂ sources. Wotawa and Kroeger, (1999) compared NO_x sources. Charron et al. (2000) have performed comparison of SO₂, NO_x and NH₃ sources with the EMEP emission data. The most thorough quantitative investigation of TSMs has been undertaken by Scheifinger and Kaiser (2007). They validated TSMs with virtual and real sources under idealized conditions, where the effects of dispersion, deposition and chemical conversion of the trace substance were excluded. Comparison of the EMEP SO₂ emission inventory and trajectory statistical reconstruction was performed. The best spatial correlation of 78% was achieved in an idealized situation. The use of real observed data resulted in correlation coefficients no more than 33%. Scheifinger and Kaiser (2007) have pointed out that the reason for such low accuracy is the simplified transport process described just by trajectory paths. The best results were achieved when TSMs were applied within the frame of the mean residence time of the concerned substance.

In the papers by Stohl (1996), Wotawa and Kroeger, (1999), Seibert (1999), Charron et al. (2000) and Scheifinger and Kaiser (2007), the accuracy of TSMs was estimated for particular species and at specified locations for the receptors. The obtained estimates may not be appropriate for another species or valid at different receptor locations. We have found that the accuracy of source reconstruction can depend on the spatial structure of the source field to be reconstructed. Because of this, the results of the source reconstruction can vary depending on the kind of substance and the locations of the receptors. In this article we focus on this feature of TSMs and consider the problem on the scale that is on the order of one thousand kilometres.

The aim of this work is to estimate typical accuracy values and the range of their variations. In real world experiments, TSMs are intended for application to *a priori* unknown sources. Therefore, we have to test the accuracy of TSMs with all possible patterns of spatial distributions of the sources. For this purpose, we generated an ensemble of fields of known virtual sources. For each particular source field, a time series of the concentration at a receptor point was calculated. Subsequently, we reconstructed a source distribution using the TSMs. The Spearman's rank order correlation coefficient between spatial distributions of the known virtual and the reconstructed sources was taken to be a quantitative measure of the accuracy. The obtained correlation coefficients were then processed to find an ensemble-average correlation and a range of the most probable values of correlations. In fact, the statistical estimation of the source reconstruction accuracy was given.

To clarify the influence of source spatial distribution on the accuracy of the TSMs we examine the simplest single-receptor situation and set aside the problems associated with concentration measurement errors and air trajectory uncertainties. These problems, multi-receptor trajectory source apportionment methods and applications of the TSMs to real world experiments will be discussed later.

2. Method

2.1. Trajectory statistics

Three well-known methods for source localization were studied in this paper, namely the concentration field method (CF) developed by Seibert et al. (1994), the concentration weighted trajectory method (CWT; Hsu et al., 2003) and the method of potential source contribution function (PSCF; Ashbaugh et al., 1985; Zeng and Hopke, 1989). In methods of trajectory statistics, the geographical region surrounding a receptor is divided in cells by a grid. Let the indices i and j enumerate geographical coordinates (longitudes and latitudes) of the cells centers. The conditional mean of some function $F(c)$ of the tracer concentration c on a receptor site under the condition of passing the trajectory over the cell ij is taken as an

indicator of a source in the cell ij for all the above-mentioned methods.

$$P_{ij} = \sum_{l=1}^L F(c_l) \tau_{ij}(l) / \sum_{l=1}^L \tau_{ij}(l), \quad (1)$$

where l is the trajectory index, L the total number of trajectories, c_l the tracer concentration at the receptor point measured on arrival of trajectory l and $\tau_{ij}(l)$ the time spent by trajectory l in grid cell ij . Values P_{ij} can be interpreted as the intensity of sources in the cell ij .

If $F(c)$ is the logarithm of the concentration we have the CF method (Seibert et al., 1994). If $F(c) = c$, formula (1) represents the CWT method (Hsu et al., 2003). If $F(c)$ is the step function that equals 0 when the concentration is below a certain threshold, and is equal to 1 when the concentration is above the threshold, formula (1) gives the PSCF method (Ashbaugh et al., 1985; Zeng and Hopke, 1989).

In the general case, concentration of a primary pollutant at the time l can be written as

$$c_l = \sum_{mn} x_{mn} g_{mn}(l), \quad (2)$$

where x_{mn} is stationary emission rate at the grid cell mn , $g_{mn}(l)$ is an influence function which takes into account actual conditions of the trace substance travel from the source mn to the receptor site with the arrival at a moment l . Substitution of the expression (2) into Eq. (1) gives:

$$P_{ij} = \sum_{l=1}^L F \left[\sum_{mn} x_{mn} g_{mn}(l) \right] \tau_{ij}(l) / \sum_{l=1}^L \tau_{ij}(l), \quad (3)$$

It is seen from (3) that the value P_{ij} depends not only on emissions in the cell ij but also on emissions in other cells. Variation of emissions in the cells other than the cell ij may change value P_{ij} even if the emission in the cell ij remains fixed. Distribution of the emissions x_{mn} over the cells depends on the kind of substance under consideration. Variation of the receptor position changes functions $g_{mn}(l)$ and value P_{ij} as well, although emissions x_{mn} may remain the same. Because of this, the result of source reconstruction may be dependent not only on a kind of substance but also on a receptor's position. This dependence is caused by the form of the Eq. (1) and is an inherent property of the considered methods.

2.2. Trajectory calculation

The data from three-dimensional back trajectories (Pickering et al., 2001; <http://croc.gsfc.nasa.gov/aeronet/IMAGES>) were used for the receptor site at the coordinates of 27.5° E and 53.0° N (Minsk, Belarus). The altitude of Minsk is 230 m asl. The trajectories started at the altitude of 950 hPa (approximately 430 m asl) at 0 and 12 h UTC every day throughout the 7-year period from January, 2004 to December, 2010. The trajectory travel time was 168 h. The overall number of trajectories was 5082. Minimal number of trajectories crossing a grid cell was 10.

2.3. Concentration time series modelling

To calculate the l th element of primary pollutant time series at a receptor site we use the approximation by Seibert (1999) based on simplified Lagrangian model

$$g_{ij} = h^{-1} \tau_{ij}(l) \exp(-t_{ij}(l)/T), \quad (4)$$

where h is the height of the mixing layer which is considered to be constant, t_{ij} is the transport time from the grid cell ij to a receptor point and T the decay time. This roughly describes the removal

processes (deposition and transformation). It is apparent from formula (3) that the quantitative characteristics of the TSMs are dependent on the specific form of the influence function. Here, as in the papers by Scheifinger and Kaiser (2007) and Ching-Ho Lin et al. (2004), we use the simplest form (4) of the influence function to calculate the time series. In the future, it is desirable to use a full dispersion model for the receptor concentration instead of (4) for more exact estimation of quantitative characterizations of the TSMs.

2.4. Ensemble of the virtual source fields

All spatial distributions of the emission rate for a trace substance were generated on the geographical domain covering Europe from 43° to 63° N and from 6° to 46° E. The boundaries for the substance supplier domain are shown in Fig. 1. The receptor site Minsk is in the center of the domain.

To deal with distributions that are not too different from the real ones, we used the EMEP emission inventory as a starting point for the generation of the particular distributions. The EMEP inventory for SO₂, NO₂, Pb, Cd and particulate matter was used. The initial distribution of virtual sources coincided with the EMEP inventory for one of the above-mentioned substances. To generate other distributions, we shifted the original EMEP distribution to the north successively in 0.5° steps and transferred the sources that extended past the northern domain boundary to the vacant southern latitudes until we reverted to the initial EMEP spatial distribution. Similarly, we then shifted the original EMEP distribution to the West successively in 1° steps until we reverted back to the initial distribution. After that, we shifted the original EMEP distribution to the northeast with the simultaneous 0.5° steps in the meridional direction and 1° in the zonal direction. Similarly, we then shifted the original EMEP distribution to the northwest. This resulted in 160 spatial distributions of emission rates for each of the substances. The total number of varied source distributions used for statistical calculations was equal to 800. Every source distribution obtained in this way was subsequently averaged on a grid with 1° meridional and 2° zonal resolution. This spatial resolution was used for the time series calculation and source reconstruction.

3. Results and discussion

A virtual pollutant time series at the receptor point was calculated from the formulas (2), (4) for each particular source distribution.

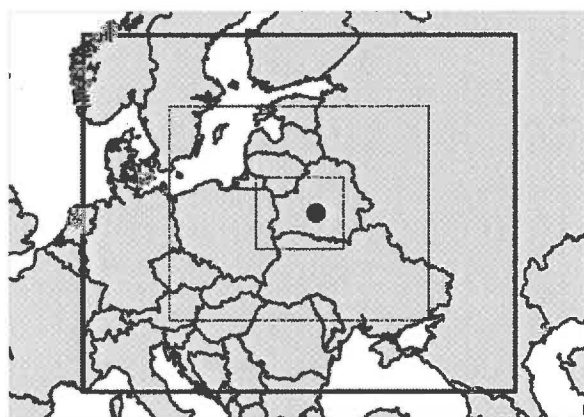


Fig. 1. The outer rectangle depicted by heavy line shows the boundaries of the substance supplier domain. The position of the receptor site is marked by a full circle. The boundaries of two reconstruction domains 1 and 5 (defined in the Table 1) are shown by the thin lines.

Table 1
Boundaries (in degrees) of the geographical domains used for reconstruction of sources. Last column contains mean transport time (in hours) between the receptor site and the domain boundaries.

Number of domain	Latitude minimal	Latitude maximal	Longitude minimal	Longitude maximal	Transport time, hour
1	51	55	22	30	16.7
2	50	56	20	32	24.6
3	49	57	18	34	33.1
4	48	58	16	36	41.0
5	47	59	14	38	48.3
6	46	60	12	40	55.2
7	45	61	10	42	61.8
8	44	62	8	44	67.7
9	43	63	6	46	73.0

Subsequently, every time series was used for the reconstruction of the source distribution with the formula (1). To reconstruct sources, we used the same set of back trajectories that were used for the time series calculation. In the following, values P_i will be conventionally called the reconstructed sources. The sources that were used to calculate concentrations (2) will be called the initial sources. In real world situations, the domain that makes the most contribution to a receptor's concentration is poorly specified and may differ from the domain used for the reconstruction. To study the impact of this factor on the reconstruction quality, the reconstruction was carried out for nine different domains. The boundaries of the domains used for the reconstruction are shown in Table 1. Last column of the Table 1 contains mean transport time between the receptor site and the domain boundaries.

Note that the reconstruction domain number 9 coincides with the substance supplier domain. The Spearman's rank order correlation coefficient between spatial distributions of the initial sources and the reconstructed sources was calculated. For every specific reconstruction domain and decay time, we had 800 different source distributions which resulted in 800 different values of the correlation coefficients distributed over the interval $(-1,1)$. This interval was divided into 50 equal segments and the number of correlation coefficients which fell within each individual segment was calculated. Subsequently, we divided the number of correlation coefficients in the segments by the common number of correlation coefficients (800) and by the segment width (0.04). Probability density function (PDF) of the correlation coefficient was obtained for every specific reconstruction domain and decay time. Two examples of the PDFs are shown in Fig. 2 for two reconstruction domains with decay time of 48 h. Non-monotonicity of the right-hand curve can be explained by the insufficient sample size. As shown in Fig. 2, the PDF maxima are at correlation coefficients 0.75–0.8. The results of other calculations for different reconstruction domains and different

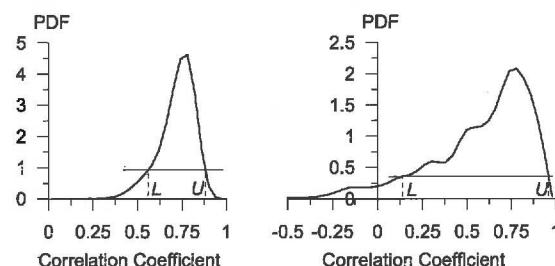


Fig. 2. Probability density function of correlation coefficient obtained for two reconstruction domains designated in the Table 1 by numbers 7 (left) and 1. Decay time is 48 h. The integral of PDF from L to U equals 0.9.

decay times showed that the PDF maximum lies most often in the range of 0.7–0.8.

The accuracy of the TSMs can be considered as a random variable because the accuracy depends on *a priori* unknown source distributions to be reconstructed. To obtain a statistical estimation of the accuracy of the source reconstruction we calculated an ensemble-average correlation and the boundaries of the correlation interval that contains 90% of the most probable correlation values. The boundaries were defined by the condition that the integral of PDF between equally probable points on each side of the PDF maximum must equal 0.9. The upper and lower boundaries of the 90% interval are shown in Fig. 2 by the letters U and L.

The accuracy of source reconstruction appeared to be only slightly dependent on the method used for the reconstruction. We begin with the CF method which turned out to be a little better than the others. Ensemble-average correlation coefficient between spatial distributions of initial and reconstructed sources as a function of a reconstruction domain number is depicted in Fig. 3 by heavy lines for four values of a trace substance decay time. The geographical position of a domain with a given number is defined by the Table 1. From the data presented in Fig. 3, one can see that the ensemble-average correlation coefficient between spatial distributions of the initial and reconstructed sources reaches about 0.7–0.75. The position of the maximum of the average correlation depends on the trace substance decay time. The maximum shifts to small domains with the decreasing of the decay time. This may be explained by the decreasing influence of the remote sources on receptor concentration.

Thin lines denote the upper and lower boundary of the corridor where the individual correlation finds itself with 90% probability. The distance between the boundaries shows the dispersion of correlations arising from the source distribution variations. The narrow corridor means that accuracy of reconstruction may be well predicted. But as shown in Fig. 3, the corridor may be rather wide. The minimum corridor widths increase when the decay time decreases. The boundaries of the minimal corridors are 0.6–0.9 for the decay time of 240 h and 0.5–0.95 for the decay time of 12 h.

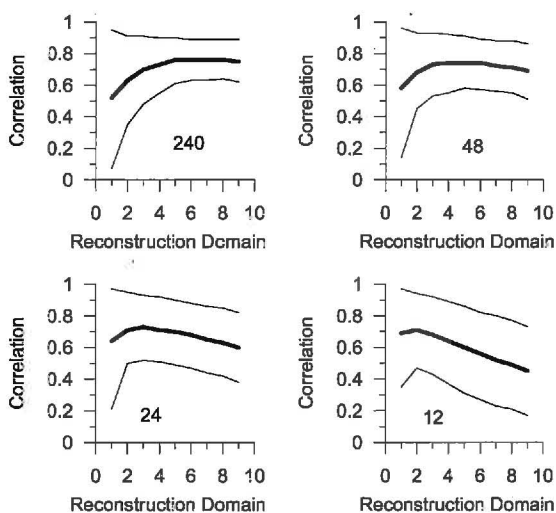


Fig. 3. Ensemble-average correlation coefficient between spatial distributions of the initial and reconstructed sources (heavy lines) and the boundaries of the interval that contains 90% of the most probable correlation values (thin lines) against a reconstruction domain number (defined in the Table 1). Calculations were made by the CF method for four values of decay time $T = 240, 48, 24$ and 12 h.

As Fig. 3 suggests, relatively low values of the average correlations and high dispersion of the correlations take place for the small reconstruction domains when the decay times are 240, 48 and 24 h. Two possible reasons may cause bad reconstruction results for the small domains. First reason occurs when a reconstruction domain is less than the area supplying most of a substance at a receptor site. In this case the monitoring data contain not only information about the sources on the reconstruction domain but the “noise” due to the sources outside of the reconstruction domain. This “noise” will deteriorate the quality of the source reconstruction. Another reason is that the trajectories on small spatial scale are almost straight lines and in the vicinity of the receptor they look like beams outgoing from the receptor. Such trajectories do not cross each other and give only the direction of a source, but can not locate it (Ferrarese, 2002).

The larger the mean correlation and the less the dispersion of correlations, the more effective the reconstruction of the sources is. The ratio of the mean correlation to the width of the 90% corridor may serve as an effectiveness index for source reconstruction. This ratio is depicted on Fig. 4 as a function of a reconstruction domain number. The position of the effectiveness index maximum defines an optimal domain for source reconstruction. Fig. 2 shows PDF for the almost optimal reconstruction domain (left) and for the domain that is far (right) from the optimal one. We compare the optimal domains with the areas supplying most of a substance to a receptor site. For the decay time of 240 h, the whole source domain (number 9) is contributing to a receptor concentration because the transport time from the domain's boundaries to the receptor (see the Table 1) is less than the decay time. In the case when decay time is 48 h and less, the domain creating most of the substance at a receptor site may be determined approximately from the equality of transport time and decay time. From Table 1, these domains numbers are 5, 2, and 1 for the decay times of 48, 24 and 12 h, respectively. In accordance with Fig. 4, a maximum of the effectiveness index is in the vicinity of domain 8 for decay time of 240 h, of domain 6 for decay time of 48 h, of domains 3–4 for decay time of 24 h, of domain 2 for decay time of 12 h. These domains do not coincide with the areas supplying most of a substance at a receptor site, but are close to them. This is in line with results obtained by Scheifinger and Kaiser (2007).

The magnitude of the effectiveness index maximum decreases as the decay time decreases. As can be seen from Fig. 4, one can expect the best results of reconstruction for trace substance with large life time in atmosphere. Note that for a long life time the optimal size of a reconstruction domain also is large.

Accuracy of source reconstruction by the CWT and the PSCF methods is a little worse than by the CF method although both of these methods give results close to the CF method. A comparison of

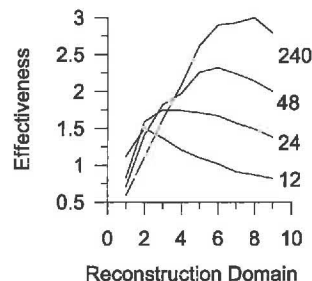


Fig. 4. Effectiveness index of source reconstruction depending on reconstruction domain number (defined in the Table 1) for four values of the trace substance decay time: 240, 48, 24 and 12 h.

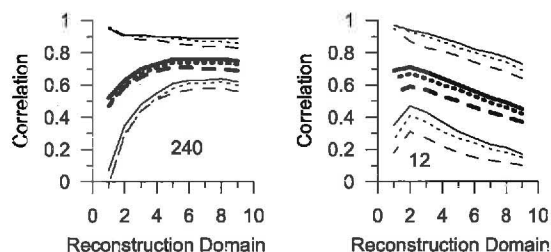


Fig. 5. Ensemble-average correlation coefficient between spatial distributions of the initial and reconstructed sources (heavy lines and boldface dots) and boundaries of the interval that contains 90% of the most probable correlation values (thin lines and small dots) against a reconstruction domain number (defined in the Table 1). Calculations were made by the CF (continuous lines), the CWT (dash lines) and the PSCF with 75th percentile criteria (dots) methods for two values of decay time $T = 240$ and 12 h.

the results obtained by the three methods is shown in Fig. 5 for two limited cases of large and small decay time. Maximum discrepancy takes place for the decay time of 12 h, minimal for the decay time of 240 h. For intermediate decay times, the discrepancies lie between these two limiting cases.

Spearman rank correlations 0.77–0.84 for the CF and the PSCF methods were obtained by Scheffinger and Kaiser (2007) using sources of the EMEP SO₂ emission inventory and formulas (2), (4) for a concentration time series. Wotawa and Kroeger (1999) did a similar test with EMEP NO_x emissions using the Lagrangian box model to calculate concentration at receptor locations. They obtained rank correlations between 0.57 and 0.70 for CF method. In the papers by Stohl (1996), Seibert (1999) and Charron et al. (2000), the reconstruction was performed on the base of real world monitoring data. Correlation values obtained by the CF method ranged from 0.22 to 0.81 (Stohl, 1996) and from 0.27 to 0.63 (Charron et al., 2000). Seibert (1999) obtained the correlation coefficient 0.75 using the CWT method. All these results are within the range of the data shown in Figs. 3 and 5.

As was shown in Section 2.1, the result of source reconstruction by TSMs may be dependent not only on the distribution of initial sources but also on where a receptor is located. In the paper by Lupu and Maenhaut (2002), the sources of several pollutants were reconstructed by the CF and PSCF methods with the use of concentration time series measured in Finland, Norway, and Israel. They found that the reconstructed source distributions for the same region in Europe were different for different monitoring sites. So the inference about the dependence of source reconstruction results on a receptor position is in line with data obtained by Lupu and Maenhaut (2002).

4. Conclusions

Three well-known TSMs for reconstruction of sources of atmospheric trace substances were considered in this work under conditions of negligible measurement and transport errors. It was found that source reconstruction accuracy is dependent upon the spatial structure of the source distribution to be reconstructed. This is a special feature of the TSMs that has never been discussed in the literature before.

Rank order correlation coefficient between initial and reconstructed sources was taken as a quantitative measure of the accuracy. The correlation coefficient may be considered as a random variable because the source distribution to be reconstructed is *a priori* unknown. Therefore, statistical estimation of the accuracy of source reconstruction was given. For this purpose, an ensemble of fields of virtual sources was generated. For each particular source

field, a time series of the concentration at a receptor point was calculated. Subsequently, source distribution was reconstructed using the TSMs and the correlation coefficient between the initial and reconstructed field was calculated.

The obtained set of correlation coefficients was then used to find ensemble-average correlation and the interval of the most probable values of correlations for each decay time and reconstruction domain. The larger the mean correlation and the less the dispersion of correlations, the more effective the reconstruction of sources is. The ratio of mean correlation to the interval with the most probable correlation values may serve as an effectiveness index for source reconstruction. The position of the effectiveness index maximum indicates the optimal reconstruction domain. This domain is dependent on a substance decay time and is close to the area that supplies most of the substance to the receptor. This is in line with the results of Scheffinger and Kaiser (2007). For the optimal reconstruction domains, average correlation coefficients between reconstructed sources and the initial ones can reach 0.70–0.75. The boundaries of the interval containing 90% of the most probable correlation values are 0.6–0.9 for decay time of 240 h and 0.5–0.95 for decay time of 12 h. The best results in the source reconstruction can be expected for the trace substances whose decay time amounts approximately to several days.

The quantitative characterizations of the TSMs can be defined more exactly, using a full dispersion model for the receptor concentration instead of (4) and taking into consideration concentration measurement errors and air trajectory uncertainties. But in the frame of the TSMs, the results of source reconstruction are uncertain even in optimal conditions because of the simplified consideration of the trace substance transport process. An exact source distribution can be found by solving inverse problem (2) if the influence functions $g_{mn}(l)$ are known.

Although the methods considered in this paper do not guarantee high accuracy they are computationally simple and fast. Using the TSMs under optimum conditions and taking into account the bounds of uncertainties, one can obtain a first hint on potential source areas.

Acknowledgements

We would like to thank the Belarusian Republican Foundation for Fundamental Research for financial support.

References

- Apadula, F., Gotti, A., Pignini, A., Longhetto, A., Rocchetti, F., Cassardo, C., Ferrarese, S., Forza, R., 2003. Localization of source and sink regions of carbon dioxide through the method of the synoptic air trajectory statistics. *Atmospheric Environment* 37, 3757–3770. N18.
- Ashbaugh, L.L., 1983. A statistical trajectory technique for determining air pollution source regions. *Journal of Air Pollution Control Assessment* 33, 1096–1098.
- Ashbaugh, L.L., Malm, W.C., Sadeh, W.Z., 1985. A residence time probability analysis of sulfur concentrations at Grand Canyon National Park. *Atmospheric Environment* 19, 1263–1270.
- Begum, B.A., Kim, E., Jeong, C.H., Lee, D.W., Hopke, P.K., 2005. Evaluation of the potential source contribution function using the 2002 Quebec forest fire episode. *Atmospheric Environment* 39, 3719–3724.
- Charron, A., Plaisance, H., Sauvage, S., Coddeville, P., Galoo, J.C., Guillermo, R., 2000. A study of the source–receptor relationships influencing the acidity of precipitation collected at a rural site in France. *Atmospheric Environment* 34, 3665–3674.
- Cheng, M., Lin, C., 2001. Receptor modelling for smoke of 1998 biomass burning in Central America. *Journal of Geophysical Research* 106 (D19), 22821–22886.
- Du, S., Rodenburg, L.A., 2007. Source identification of atmospheric PCBs in Philadelphia/Camden using positive matrix factorization followed by the potential source contribution function. *Atmospheric Environment* 41, 8596–8608.
- Ferrarese, S., 2002. Sensitivity test of a source–receptor model. *Il Nuovo Cimento* 25C (4), 501–511.
- Ferrarese, S., Longhetto, A., Cassardo, C., Apadula, F., Berton, D., Giraud, C., Gotti, A., 2002. A study of seasonal and yearly modulation of carbon dioxide sources and

- sinks, with a particular attention to the Boreal Atlantic Ocean. *Atmospheric Environment* 36, 5517–5526.
- Heo, J.-B., Hopke, P.K., Yi, S.-M., 2009. Source apportionment of PM_{2.5} in Seoul, Korea. *Atmospheric Chemistry and Physics* 9, 4957–4971.
- Hopke, P., Gao, N., Cheng, M., 1993. Combining chemical and meteorological data to infer source volumes of airborne pollutants. *Chemometrics and Intelligent Laboratory Systems* 19 (2), 187–199.
- Hsu, Y.-K., Holsen, T., Hopke, P., 2003. Comparison of hybrid receptor models to locate PCB sources in Chicago. *Atmospheric Environment* 37, 545–562.
- Lee, S., Ashbaugh, L., 2007a. Comparison of multi-receptor and single-receptor trajectory source apportionment (TSA) methods using artificial sources. *Atmospheric Environment* 41 (6), 1119–1127.
- Lee, S., Ashbaugh, L., 2007b. Comparison of the MURA and an improved single-receptor (SIRA) trajectory source apportionment (TSA) method using artificial sources. *Atmospheric Environment* 41 (21), 4466–4481.
- Lee, L.Y.L., Kwok, R.C.W., Cheung, Y.P., Yu, K.N., 2004. Analyses of airborne ⁷Be concentrations in Hong Kong using back-trajectories. *Atmospheric Environment* 38, 7033–7040.
- Lin, Ching-Ho, Wu, Yee-Lin, Chang, Ken-Hui, Lai, Chin-Hsing, 2004. A method for locating influential pollution sources and estimating their contributions. *Environmental Modeling and Assessment* 9, 129–136.
- Lupu, A., Maenhaut, W., 2002. Application and comparison of two statistical trajectory techniques for identification of source regions of atmospheric aerosol species. *Atmospheric Environment* 36, 5607–5618.
- Pickering, K.E., Thompson, A.M., Kim, H., DeCaria, A.J., Pfister, L., Kucsera, T.L., Witte, J.C., Avery, M.A., Blake, D.R., Crawford, J.H., Heikes, B.G., Sachse, G.W., Sandholm, S.T., Talbot, R.W., 2001. Trace gas transport and scavenging in PEM-Tropics B South Pacific Convergence Zone convection. *Journal of Geophysical Research* 106, 32591–32602.
- Poirot, R.L., Wishinski, P.R., 1986. Visibility, sulfate and air mass history associated with the summertime aerosol in Northern Vermont. *Atmospheric Environment* 20, 1457–1469.
- Rúa, A., Hernández, E., de las Parras, J., Martín, I., Gimeno, L., 1998. Sources of SO₂, SO₄²⁻, NO_x and NO₃ in the air of four Spanish remote stations. *Journal of Air and Waste Management Association* 48, 838–845.
- Scheffinger, H., Kaiser, A., 2007. Validation of trajectory statistical methods. *Atmospheric Environment* 41, 8846–8856.
- Seibert, P., 1999. Inverse modelling of sulfur emissions in Europe based on trajectories. In: Kasibhatla, P., Heimann, M., Rayner, P., Mahowald, N., Prinn, R.G., Hartley, D.E. (Eds.), *Inverse Methods in Global Biogeochemical Cycles*, vol. 114. AGU Geophysical Monograph, Washington, ISBN 0-87590-097-6, pp. 147–154.
- Seibert, P., Kromp-Kolb, H., Baltensperger, U., Jost, D.T., Schwikowski, M., Kasper, A., Puxbaum, H., 1994. Trajectory analysis of aerosol measurements at high alpine sites. In: Borrel, P.M., Borrel, P., Cvitas, T., Seiler, W. (Eds.), *Transport and Transformation of Pollutants in the Troposphere*. Academic Publishing, Den Haag, pp. 689–693.
- Stohl, A., 1996. Trajectory statistics – a new method to establish source-receptor relationships of air pollutants and its application to the transport of particulate sulfate in Europe. *Atmospheric Environment* 30, 579–587.
- Vasconcelos, L.A., Kahl, J.D.W., Liu, D., Macias, E.S., White, W.H., 1996a. A tracer calibration of back trajectory analysis at the Grand Canyon. *Journal of Geophysical Research* 101 (D14), 19329–19335.
- Vasconcelos, L.A., Kahl, J.D.W., Liu, D., Macias, E.S., White, W.H., 1996b. Spatial resolution of a transport inversion technique. *Journal of Geophysical Research* 101 (D14), 19337–19342.
- Wang, Y.Q., Zhang, X.Y., Arimoto, R., 2006. The contribution from distant dust sources to the atmospheric particulate matter loadings at Xi'an, China during spring. *Science of the Total Environment* 368, 875–883.
- Wotawa, G., Kroeger, H., 1999. Testing the ability of trajectory statistics to reproduce emission inventories of air pollutants in cases of negligible measurement and transport errors. *Atmospheric Environment* 33, 3037–3043.
- Zeng, Y., Hopke, P., 1989. Study of the sources of acid precipitation in Ontario, Canada. *Atmospheric Environment* 23, 1499–1509.
- Zhao, W., Hopke, P.K., 2006. Source investigation for ambient PM_{2.5} in Indianapolis. In: *Aerosol Science and Technology* 40, 898–909.
- Zhou, L., Hopke, P.K., Liu, W., 2004. Comparison of two trajectory based models for locating particle sources for two rural New York sites. *Atmospheric Environment* 38 (13), 1955–1963.