

Assessment of parameters describing representativeness of air quality in-situ measurement sites

Journal Article**Author(s):**

Henne, S.; Brunner, D.; Folini, Doris S.; Klausen, Jörg; Klausen, Jörg; Buchmann, Brigitte

Publication date:

2009

Permanent link:

<https://doi.org/10.3929/ethz-b-000029069>

Rights / license:

[Creative Commons Attribution 3.0 Unported](#)

Originally published in:

Atmospheric Chemistry and Physics 10(8), <https://doi.org/10.5194/acpd-9-20019-2009>

This discussion paper is/has been under review for the journal *Atmospheric Chemistry and Physics (ACP)*. Please refer to the corresponding final paper in *ACP* if available.

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

Assessment of parameters describing representativeness of air quality in-situ measurement sites

S. Henne¹, D. Brunner¹, D. Folini^{1,3}, S. Solberg², J. Klausen¹, and B. Buchmann¹

¹Empa, Swiss Federal Laboratories for Materials Testing and Research, Dübendorf, Switzerland

²NILU, Norwegian Institute for Air Research, Kjeller, Norway

³ETH, Swiss Federal Institute of Technology, Zürich, Switzerland

Received: 27 August 2009 – Accepted: 10 September 2009 – Published: 24 September 2009

Correspondence to: S. Henne (stephan.henne@empa.ch)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Abstract

The atmospheric layer closest to the ground is strongly influenced by variable surface fluxes (emissions, surface deposition) and can therefore be very heterogeneous. In order to perform air quality measurements that are representative of a larger domain or a certain degree of pollution, observatories are placed away from population centres or within areas of specific population density. Sites are often categorised based on subjective criteria that are not uniformly applied within different administrative domains. A novel approach for the assessment of parameters reflecting site representativeness is presented here, taking emissions, deposition and transport towards 34 sites covering Western and Central Europe into account. These parameters are directly inter-comparable among the sites and can be used to select sites that are, on average, more or less suitable for data assimilation and comparison with satellite and model data. Advection towards these sites was simulated by backward Lagrangian Particle Dispersion Modelling (LPDM) to determine the sites' annual catchment areas for the year 2005 and advection times of 12, 24 and 48 h. Only variations caused by emissions and transport during these periods were considered assuming that these dominate the short-term variability of most but especially short lived trace gases. The parameters of representativeness derived were compared between sites and a novel, uniform and observation-independent categorisation of the sites based on a clustering approach was established. Six groups of European background sites were identified ranging from very remote coastal to polluted rural sites. These six categories explained 50 to 80% of the inter-site variability of median mixing ratios and their standard deviation for NO_2 and O_3 , while differences between group means of the longer lived trace gas CO were insignificant. The derived annual catchment areas strongly depended on the applied LPDM and input wind fields, the catchment settings and the year of analysis. Nevertheless, the parameters of representativeness showed considerably less variability than the catchment geometry, supporting the robustness of the derived station categorisation.

20020

ACPD

9, 20019–20062, 2009

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



1 Introduction

Ground based in-situ measurement sites form the backbone of the atmospheric observing system. They usually provide a much larger number of observational sites than vertical sounding or ground based remote sensing sites and, while subject to ongoing discussion, better precision, accuracy and often long term stability than satellite observations. This is mainly due to the fact that in-situ measurement techniques are in general simpler and less expensive to operate than remote sensing methods and can more easily be traced back to international calibration standards. However, satellite observations are horizontally more homogeneous because they are derived for different regions with the same instrument. Surface measurements are further complicated by the fact that the atmospheric layer close to the ground is strongly influenced by exchange processes at the Earth's surface (momentum, heat, mass fluxes) and can therefore exhibit large horizontal heterogeneities and might deviate strongly from free tropospheric conditions. The positioning of ground based sites is hence critical when addressing a specific scientific objective and the question of what a site is representative of arises.

For air quality (AQ) monitoring one is often interested in the question of how much the population is exposed to concentrations of certain species above national or international limit values. Monitoring networks are therefore often designed to cover different pollution levels, which usually coincides with areas of different emissions, to be representative of different exposure levels. For climate change related problems one is more interested in changes and trends of the atmospheric composition of background air masses. Sites therefore are placed in areas with weak horizontal gradients of the species of interest and thus away from emission sources.

Definitions of site representativeness include the following two concepts. According to Larssen et al. (1999) "the area in which the concentration does not differ from the concentration measured at the station by more than a specified amount can be called the area of representativeness of the station". Typical radii of the area of representative-

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

ness are also given by Larssen et al. (1999) and range from metres, for polluted traffic sites, to hundreds of kilometres for background remote sites. Since these estimates are based on subjective experience, they may not withstand a thorough quantitative evaluation for specific sites.

5 Nappo et al. (1982) define a point measurement to be representative of the average in a larger area (or volume) if the probability that the squared difference between point and area (volume) measurement is smaller than a certain threshold for more than 90% of the time. The maximum tolerable difference has to be assessed for every individual problem; it should not be smaller than the uncertainty of the measurement. In addition,
10 the area (volume) of interest will vary with application. For the inter-comparison of in-situ (point data) and model simulations or remote sensing data (volume data) and for data assimilation purposes it is important that the measurements are representative in the sense of the definition given by Nappo et al. (1982) or that the area of representativeness is at least as large as the satellite or model grid box containing the site.

15 To reliably assess the area of representativeness or the representativeness in the sense of Nappo et al. (1982) knowledge of the 4-D concentration field would be necessary and could be obtained through extensive measurements at many different locations within an area (e.g., Blanchard et al., 1999; Kuhlbusch et al., 2006) or detailed modelling studies (e.g. on the street scale, Scaperdas and Colville, 1999). Factors
20 influencing the concentration of a certain trace species within a certain volume are horizontal and vertical transport and mixing, chemical transformations, surface deposition and emissions. With this and the aforementioned definitions of representativeness in mind it has to be concluded that representativeness will not only vary with time (e.g. season, day-to-day) but also largely depend on the species of interest. In general,
25 species with strong surface sources or sinks and with short atmospheric lifetimes due to photochemistry and deposition show stronger spatial variability and therefore smaller areas of representativeness than species with weak surface fluxes and long lifetimes. The problem of temporal variability of representativeness due to changing advection towards an AQ site and different pollution uptake on the way is often addressed by using

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

sector or cluster analysis of air mass back-trajectories (e.g. Henne et al., 2008). In this study we focus on the question of average representativeness of surface observations of air pollutants with (e-folding) lifetimes of hours to a few days within the atmospheric boundary layer. This includes the most commonly observed levels of O₃ and NO₂.

5 While no absolute quantification of representativeness in terms of the aforementioned definitions can be given, we derive parameters of average representativeness based on emission and deposition proxy fields taking advection towards the sites into account. These parameters are directly inter-comparable among the sites and can be used to select sites that are, on average, more or less suitable for data assimilation and comparison with satellite and model data.

10 Next to a quantification of representativeness an objective site categorisation is very valuable for the purposes just mentioned, data interpretation and also extrapolation of exposure levels to areas not directly covered by an AQ network. In Europe, the European Environment Agency EEA/Airbase database (<http://air-climate.eionet.europa.eu/databases/airbase/>; Mol et al., 2008) as implemented through the Exchange of Information Decision (European Council, 1997) collects data from ~3000 AQ monitoring sites and provides a two-dimensional site categorization (station type: traffic, industrial, residential, background; area type: urban, suburban, rural) based on station meta-data information on population densities and emissions in the surroundings of the sites.

20 However, these classifications are often derived from subjective interpretations of the site's maintainer (due to different levels of available and reliable information). Here we will develop a categorisation method that is objectively based on parameters of representativeness and independent of previously recorded AQ data. For verification, the obtained categorization can then be tested against observed data.

25 The sites selected for this study (Fig. 3 and Table 1) are mainly categorised as “rural” according to EEA/Airbase and thus not directly influenced by local emissions. The site Ispra (IT04) is categorised suburban but was included because it is part of the European Monitoring and Evaluation Programme (EMEP) network, while several of the selected high altitude sites are not included within EEA/Airbase and therefore not

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

categorised. Most of the sites are part of networks or programmes that focus on the observation of the global (WMO Global Atmosphere Watch; GAW) and/or European scale (EMEP) atmospheric background composition. Sites were selected according to data availability of O₃, NO₂, CO, to assure coverage of Western and Central Europe, according to their contributions to international and European programmes and because they are supported within European Commission framework programmes.

The present manuscript is organised as follows. Section 2 focusses on the method to derive parameters of representativeness from Lagrangian transport simulations combined with proxy emission and deposition data and how to use these in a site categorisation. The derived parameters of representativeness together with the site categorisation are presented in Sect. 3 followed by a discussion of the robustness of the parameter estimation in terms of methodological settings and inter-annual variability in Sect. 4. Conclusions and outlook end the manuscript in Sect. 5.

2 Methods

2.1 Parameters of representativeness

For a European-wide analysis of station representativeness high resolution 4-D air quality data are currently not available for any extended periods. However, for most but especially short lived species like NO₂ or O₃ (O₃ is only short lived close to the surface due to strong dry deposition) emissions and deposition largely determine the small scale (~1 km) variability of these gases. Therefore, emission and deposition data are supposed to be appropriate proxies for concentrations and can be used to derive parameters of representativeness. Actually, this is a common procedure within monitoring networks that define thresholds for population and traffic within a certain area around a site for a general site categorization. Unfortunately, no kilometre scale emission data set was available for this study. Therefore, population data was used as a proxy for emissions. A large fraction of NO_x emissions are traffic related, however, traffic outside

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Parameters
describing
representativeness of
air quality sites**S. Henne et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

towns is not reflected in population distributions. Even though the sites considered in this study are not close to any major traffic routes, we might therefore underestimate the influence of traffic in our results. Furthermore, surface dry deposition plays an important role for surface O_3 . Thus, typical deposition velocities were derived from high resolution land-use data. Parameters of representativeness can be obtained by directly investigating total population and deposition influence within certain surroundings of a site (for example circles of 10 and/or 50 km radius). On a local scale this approach would already yield valuable results to uniformly characterize sites. However, for more remote sites advection towards the site and dispersion should be taken into account. This is especially evident for sites with well defined clean and polluted air sectors, as it is often the case for coastal sites or for sites situated on mountain tops that might sample free tropospheric and boundary layer conditions at different times. In the present study Lagrangian Particle Dispersion Models (LPDM) were applied in backward mode, directly yielding the surface area from which an air sample was potentially influenced (Seibert and Frank, 2004).

While focussing on the representativeness of short lived species most relevant to O_3 production here, the presented method is not limited to these substances. As long as the distribution of a substance is mainly driven by emissions and deposition the same approach could be used even if the emissions have a spatial distribution that is different from population. However, the different emission distributions would need to be taken into account which may lead to different parameters of representativeness and hence a different station categorization than obtained in this study. The determined catchment area, nevertheless, is independent of the pollutant and could easily be applied to other source distributions. For species with surface distributions that are not driven by emissions the presented method is not valid and parameters of representativeness could only be assessed from detailed model studies or dense observation networks.

2.2 Lagrangian modelling of the catchment area

2.2.1 Model description

An adapted version of the COSMO (Consortium for Small-Scale Modelling) LPDM (Glaab et al., 1998) was applied to sites within complex terrain. Previously, the model was successfully applied in backward mode for the high Alpine site Jungfrauoch (Folini et al., 2008). The model uses input wind data obtained from the operational COSMO weather prediction system operated by MeteoSwiss. The resolution of the meteorological input data is approximately 7 km by 7 km on 45 vertical levels up to 20 hPa. The model grid covers most of Western and Central Europe. While this grid resolution is not sufficient to explicitly represent all vertical exchange processes that are due to thermally induced circulations, it is expected that the major effects (Alpine heat low, plain-to-mountain flow) were correctly simulated (Weissmann et al., 2005). For 15 of the selected sites (see Table 1) the COSMO LPDM was run for the whole year 2005. The model was initialized every 3 h, 25 000 particles were released at the sites 80 m above model ground and traced backwards in time for 60 h. In total 2920 individual simulations were available for each site. The model produced residence time fields between the model surface and 500 m above model ground, indicating where the air had surface contact on its transport path towards the site. The COSMO LPDM is limited in its horizontal extent, since the high resolution grid is not nested into a global domain. This causes problems for receptor sites close to the boundaries of the model domain. For such sites and those in flat terrain a second LPDM was used. The FLEXPART LPDM (Stohl et al., 2005) is a well documented research tool in atmospheric dispersion modeling and can be applied in forward and backward mode (Seibert and Frank, 2004). FLEXPART was operated on 3 hourly global meteorological fields as retrieved from ECMWF analyses and forecasts with a horizontal resolution of 1° by 1° on 60 vertical levels up to 0.2 hPa. The output of residence times was stored on two different domains: first a coarse domain (0.5° by 0.5°) covering Europe, the North Atlantic and eastern North America and second a fine domain (0.1° by 0.1°) covering Europe. Resi-

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



dence times were further sampled for different vertical levels with level tops at 100, 500, 1000, 3000, and 10 000 m above model ground. The model was initialized for 24 of the selected sites (see Table 1) every 3 h for the year 2005 and integrated backwards in time for 120 h. At each site 50 000 particles were released at station altitude above sea level or if this was below model ground at 10 m above model ground. In total 2920 individual simulations are available for each site. In contrast to the COSMO LPDM, more sites could be assessed at the border of the fine grid domain for which residence times are still available on the coarse grid. For five sites in flat terrain both models were run allowing for inter-comparison of the model performance (see Sect. 4.3 and supplementing material, see <http://www.atmos-chem-phys-discuss.net/9/20019/2009/acpd-9-20019-2009-supplement.pdf>). For these sites, only FLEXPART results were used for the site categorisation.

2.2.2 Catchment area definition

For each site a 5-dimensional field of residence times as derived from one of the two LPDMs was stored. To analyse the average region of influence of a site annual total residence times were derived by summing residence times over all start times and over all integration time steps within a selected integration interval for all grid cells

$$T_{i,j,k} = \sum_m \sum_l \tau_{i,j,k,l,m}, \quad (1)$$

where i, j are the horizontal grid indices, k is the vertical level, l is the integration time step in hours ($l=3, 6, \dots, L_{\max}$; $L_{\max}=60$ COSMO LPDM; $L_{\max}=120$ FLEXPART), and $m=1, \dots, M$ ($M=2920$) is the time index of the initialization time. Annual total residence times for integration intervals 12, 24, and 48 h were investigated here.

For a given site surface fluxes within a specific area will significantly alter the chemical composition of an air mass sampled at this site, while surface fluxes elsewhere only cause undetectable variations. To determine this area we adapted the concept of Schmid (1997), originally developed for the analysis of representativeness of flux

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



measurements on the micro scale. We first define the catchment volume of a site as the volume of highest annual residence times $T_{i,j,k} = \sum_m \sum_l \tau_{i,j,k,l,m}$ enclosing 50% of the total residence time $T_{\text{tot}} = \sum_i \sum_j \sum_k T_{i,j,k}$.

In order to represent the influence of surface processes (emissions, deposition etc.) the *catchment area* is then defined as the horizontal projection of the slice of the catchment volume from the surface up to 500 m above model ground. The catchment area thus only contains surface grid points with a significant individual contribution to the total residence time, while the majority of grid points with smaller individual contributions is neglected. The catchment area is the area in which surface fluxes are expected to create a detectable and significant signal at the receptor sites.

The full 3-dimensional domain rather than the surface residence times was used to adequately represent the elevated position of high altitude sites that usually experience large surface sensitivities only close to the site within the elevated area but are characterised by small surface sensitivities over surrounding flat terrain, resulting in rather small total surface residence times. However, if taking 50% of these surface residence times into account a larger area would be selected as catchment area including grid points with small residence times at larger distances. These would only have an insignificant influence on observations at elevated sites.

The threshold value of $f=50\%$ was arbitrarily chosen by Schmid (1997) and could be set to different values. However, the author argues that the influence of a grid cell just outside the 50% area usually is an order of magnitude smaller than the influence of the grid cell with maximum residence time. In our study, $\max(T_{i,j})$ outside the catchment area was 2–3 orders of magnitude smaller than $\max(T_{i,j})$ inside the catchment area. Meaning a source/sink just outside the catchment area would need to be 2–3 orders of magnitude larger to have the same effect as a source/sink close to the site. The sensitivity of the derived parameters of representativeness to the chosen threshold value is further discussed in Sect. 4.1.

To derive the volume of largest residence times on a vertically unequally spaced grid it is necessary to transform residence times to residence time densities:

Parameters describing representativeness of air quality sites

S. Henne et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

$\gamma_{i,j,k} = \tau_{i,j,k} / m_{i,j,k}$ for the individual residence times and $\Gamma_{i,j,k} = T_{i,j,k} / m_{i,j,k}$ for the annual total residence times, with m being the mass of air in each grid cell, assuming international standard atmospheric conditions. All $\Gamma_{i,j,k}$ were then sorted in decreasing order, Γ_n , with $n=1, \dots, IJK$. All $T_{i,j,k}$ were ordered following the same permutation. A threshold $\Gamma_{n_c} = \Gamma_{50}$ was then derived for the smallest index n_c for which $\sum_n^{1, \dots, IJK} T_n \geq 0.5 T_{\text{tot}}$ was fulfilled. Finally, all surface grid cells fulfilling $\Gamma_{i,j}^{500} \geq \Gamma_{50}$ were defined as catchment area. It was necessary to scale the total annual residence times at sites simulated by the COSMO LPDM in order to be comparable to FLEXPART simulated sites by a factor of 0.88, 0.81 and 0.83 for 12, 24 and 48 h total residence times, respectively (see Sect. 4.3 and supplement, see <http://www.atmos-chem-phys-discuss.net/9/20019/2009/acpd-9-20019-2009-supplement.pdf>).

The geometry of the catchment areas can be summarized by a few simple parameters that are given for each site in Table 2a–c. From the total surface area of the catchment, A , an equivalent radius, $r = \sqrt{A/\pi}$ was calculated. Furthermore, the main advection direction DD_{max} of a site was determined from the sector with the farthest extent of the catchment area.

While in micro-meteorological applications of the catchment area concept (see Schmid (2002) for a review) the focus is often on the representativeness of flux measurements and the flux footprint has a more limited horizontal extent as compared to the concentration footprint (Kljun et al., 2002), here the extent of the catchment area is limited by the integration interval of the LPDM which was chosen to be in the range of time scales (<48 h) responsible for most observable short term variability.

2.3 Proxy data

2.3.1 Population data

Fine scale population data, $P_{i,j}$, can be used as a proxy of fine scale emissions. Both the total population and its variability within a certain area around a site can be used

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

to characterize the representativeness of a site. In this study the analysed area is the catchment area of a site but for model comparison the area could be selected equal to the grid box of an air quality model. Low absolute population will indicate that a site can be seen as a remote background site, while low variability within a more populated grid cell allows the conclusion that the site is representative of a certain population density and will not experience large variability due to the direction of advection. To analyse these two factors population data from CIESIN, Columbia University, Center for International Earth Science Information Network (CIESIN) – Columbia University and Centro Internacional de Agricultura Tropical (CIAT) (2005) with a horizontal resolution of 2.5' by 2.5' (arc-minutes, ~3 km by ~4.5 km in central Europe) were used. The reference year for the data set is 2005.

2.3.2 Ozone deposition

The land cover analysis is based on the global land cover data set GLC2000 produced by the Global Environment Monitoring Unit of the Joint Research Centre, Ispra, Italy, European Commission – Joint Research Centre (2003). For Europe the categorisation comprises 23 land cover types as presented in the supplement (Table S1, <http://www.atmos-chem-phys-discuss.net/9/20019/2009/acpd-9-20019-2009-supplement.pdf>). The horizontal resolution of the gridded data is 32'' (arc-seconds, ~0.6 km by ~1 km in central Europe). The reference year for the vegetation categories is 2000.

The land/vegetation cover influences the chemical composition of the air in several ways (emissions of biogenic substances, dry deposition, photolysis rates through albedo). However, here we only focus on the effect of land cover on ozone through surface dry deposition. From the land cover types typical summer day-time ozone deposition velocities, $v_{d,i,j}$, were calculated following the parameterisation of Wesely (1989). Atmospheric conditions were set to 20 °C surface temperature, 800 W m⁻² global radiation and 0.7 m s⁻¹ friction velocity (independent of land cover type). The resulting ozone

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



deposition velocities represent day-time maxima and therefore have to be seen as an upper limit of the deposition influence. Wesely's parameterisation considers 11 different land cover types that differ slightly from the land cover scheme described above. It was therefore necessary to map the two different land cover categorizations. The GLC categories were mapped as fractions of the 11 land cover categories of the deposition parameterisation (see supplement Table S1: <http://www.atmos-chem-phys-discuss.net/9/20019/2009/acpd-9-20019-2009-supplement.pdf>). The resulting typical summer day-time ozone deposition velocities by category are given in the supplement (Table S1). The smallest ozone deposition velocity is experienced over water bodies and ice and snow followed by barren or burned areas. The largest ozone deposition velocities are estimated for managed areas (agriculture) while values are slightly smaller for forested areas and depend on the type and density of the forest. As for population, total deposition influence and its variability in the catchment area were investigated.

2.4 Site categorisation

The parameters chosen for the site categorisation are derived from the population data and ozone deposition velocity combined with total annual residence times in the catchment areas. The total emission burden was represented by the sum of the product of population and total annual residence times, $\sum T_{i,j}P_{i,j}$, in the three investigated catchment areas (12, 24, 48 h). The variability of the emissions within the catchment areas was expressed through the residence time weighted standard deviation (Galassi et al., 2009) of the population density

$$\sigma_{P,T} = \sqrt{\frac{\sum T_{i,j}}{(\sum T_{i,j})^2 - \sum T_{i,j}^2} \sum T_{i,j}(P_{i,j} - \bar{P})^2}, \quad (2)$$

where \bar{P} is the residence time weighted mean population density

$$\bar{P} = \frac{\sum P_{i,j}T_{i,j}}{\sum T_{i,j}}. \quad (3)$$

The total surface deposition influence and its variability were represented in an analogous way. In total, 12 parameters (the 4 mentioned parameters for 3 catchment areas each) were selected to derive a site categorization (compare Fig. 2 and Table 2a–c). COSMO LPDM derived total residence times were scaled by a factor of 0.88, 0.81 and 0.83 for the 12, 24, and 48 h catchment areas, respectively to be comparable to FLEXPART results (as deduced from the model inter-comparison, see Sect. 4.3 and supplementing material).

To assure that each parameter had a similar influence on the clustering solution the following normalisation was used

$$x_n = \frac{(x - \bar{x})}{\sigma_x}, \quad (4)$$

where \bar{x} represents the parameter mean and σ_x its standard deviation. Furthermore, the parameters used in the clustering should be normally distributed. For the population products this was clearly not the case. Therefore, these were log-transformed prior to normalisation. Recognizing that surface deposition will be of lesser importance for most species monitored at the selected sites than emissions/population, we attributed additional weights 2 and 1 to the parameters describing emissions/population and deposition, respectively.

The applied weighting factor can be justified considering the chemical budget of O_3 . The ratio of chemical production/destruction, which is largely driven by anthropogenic precursor emissions, to surface dry deposition can be obtained from model studies. While for the global tropospheric domain the deposition term dominates the budget (ratio: ~ 3.5 , Wild, 2007), it becomes less important within the continental troposphere (ratio: ~ 0.8 , von Kuhlmann et al., 2003) and the ratio decreases to 0.4–0.6 in the summer-time European boundary layer (Memmesheimer et al., 1997; Derwent and Davies, 1994). For other species, for example NO_x , the importance of surface dry deposition in comparison to chemical processing was estimated to be even smaller in the European boundary layer (ratio: ~ 0.1 , Memmesheimer et al., 1997). By choosing

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

a factor of 0.5 between deposition and emission influence in our clustering approach we consider the lower limit of this factor for the O₃ budget, but are above the upper limit for NO₂ and therefore use a compromise that should represent an average importance of these processes for different species. The influence of the weighting factor is further discussed in the results section (Sect. 3.4).

We applied Ward's hierarchical clustering method (Ward, 1963) to the normalised parameters, which allows for the estimation of the number of significant clusters by evaluating the change in inter-cluster difference when clusters are subsequently merged. Here we selected a threshold of the inter-cluster difference change of 5%. This procedure is similar to the one applied by Henne et al. (2008) for air mass back-trajectories.

2.5 Observations

To test the station categorisation and the performance of the dispersion models (see supplementing material), in-situ observations of O₃, NO₂ and CO at the selected sites were used. The data were obtained from the EMEP database (<http://www.emep.int/>) and the GAW world data centre for greenhouse gases (WDCGG, <http://gaw.kishou.go.jp/wdcgg/>). Furthermore, station PIs were asked to provide additional data where these were missing in the databases. In this manner data were gathered for the French sites from the Pollution Atmosphérique à Echelle Synoptique (PAES) network (<http://paes.aero.obs-mip.fr/>) and for Cabauw (NL11), Weybourne (WEY) and for Monte Velho (PT04).

3 Results

The results are presented in the following sequence. First, some examples for derived catchment areas are presented. Second, the parameters of representativeness are discussed leading to the novel site categorisation and the comparison with observations.

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

3.1 Catchment area examples

The total annual footprints and corresponding catchment areas (12 and 48 h) for the sites Cabauw (NL11) and Ispra (IT04) are compared in Fig. 1. These sites represent the upper and lower extremes of derived catchment area size (compare Table 2a–c) and demonstrate the dominating influence of different advection regimes on the representativeness of surface sites even on short time scales (12 h). Cabauw, situated within a coastal area that often experiences high wind speeds, shows catchment areas with equivalent radii of $r_{12}=148$ km and $r_{48}=575$ km, while Ispra, situated in the foothills of the Alps at the northern edge of the Po Valley, is often dominated by stagnant conditions, indicated by catchment area radii as small as $r_{12}=43$ km and $r_{48}=179$ km. Total annual footprints of all other sites and 12, 24, and 48 h backward integration can be accessed in form of interactive station report cards through the GEOMon project website (<http://www.geomon.eu/science/act2/SciAct2.CHE.html>).

3.2 Parameters describing population/emission influence

The parameters describing total emission burden, $\sum PT$, and variability, $\sigma_{P,T}$, are depicted in Fig. 2a, c, e as scatter plots for all sites and the three analysed catchment areas. The total and variability of population were strongly correlated, especially for the 12 h catchment, however, there were also exceptions to this correlation. The sites with the largest population burden and variability are Harwell (GB36), Cabauw (NL11) and Ispra (IT04) for all three catchment areas. At the lower end of the distribution were the sites Lampedusa (LMP), Mace Head (IE31) and Finokalia (GR02). It is interesting to note that these rankings varied slightly from one to the other catchment area displaying different ratios of local to regional scale emission influence on the sites. For example the site Lampedusa (LMP) was the most remote when considering the 12 h catchment, however, when looking at the 48 h catchment Mace Head (IE31) stood out as being most remote, displaying the growing influence of distant sources in the Mediterranean in contrast to the absence of sources over the North Atlantic. Some

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

sites were characterised by relatively small variability (for example Sonnblick (AT34, central Alps) and Roquetas (ES03, sparsely populated coastal area)) as compared to their total population burden, while others (for example Campisabalos (ES09, vicinity of Madrid, in otherwise relatively sparsely populated area)) experienced strong variability.

Furthermore, for most of the sites the influence due to population was accumulated mainly within the last 24 h before arrival, as indicated by the smaller increase of the population – residence time product in the second 24 h as compared to the first 24 h (Fig. 2c, e). Although total and variability of population were strongly correlated, especially the 24 and 48 h variability contains some independent information that should not be neglected in the site clustering. We also tested the use of relative variability $\sigma_{P,T}/\bar{P}$. However, its distribution was not normal or log-normal, but characterized by individual extremes caused by close to zero total population. During clustering this parameter created one member clusters and was therefore not suited for the approach.

3.3 Parameters describing deposition influence and land use

The parameters describing total deposition, $\sum v_d T$, and its variability, $\sigma_{v_d T}$ are displayed in Fig. 2b, d, f. In contrast to the population parameters the deposition parameters showed no significant correlation between totals and variability for any of the catchment areas. Total deposition influence was largest for sites with large total residence time that are situated in agricultural areas (for example Hegyhatsal (HNG), K-pusztá (HU02) and also Roquetas (ES03) for 24 and 48 h catchment areas). Main land cover types within the catchment areas are given in Table 2a–c The largest deposition variability was estimated for sites in coastal areas that are also characterized by extended agricultural activity (for example Weybourne (WEY), Preila (LT15), Zingst (DE09) and Kollumerwaard (NL09)), while for coastal sites in relatively barren or dry environments (Mace Head (IE31), Finokalia (GR01)) the variability remained at average levels. For the continental sites with large total deposition influence the variability remained small.

For the 12 h catchment (Table 2a) the most frequent dominating land cover categories were 16 (Cultivated and managed areas) and 20 (water bodies), followed by the

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



forest types 2 (tree cover, broadleaved, deciduous, closed) and 4 (tree cover, needle-leaved, evergreen). Two sites showed particularly small heterogeneity (percentage of main class >90%) of the land cover in the catchment area: Lampedusa (LMP) and K-pusztá (HU02). For one site the dominating land cover type made up less than 30% of the total land cover (Donon, FR08) indicating heterogeneous conditions. For the 24 and 48 h catchments (Table 2b–c) more sites are dominated by either land cover type 16 (cultivated and managed areas) or 20 (water bodies), while only 7 sites are dominated by other land cover types.

3.4 Station categorisation

Six groups of sites resulted from the clustering procedure as estimated by the inter-cluster distance method (see Sect. 2.4). The selection of 6 different groups could further be justified by rather long vertical branches of the cluster tree (dendrogram, not shown) between 6 and 7 groups and the rather short branches (reduction of cluster distance) for more than 6 groups. It is further visible that the subgroups 3 and 4 were split at almost the same vertical level of the cluster tree, indicating that either the selection of 4 or 6 groups is meaningful.

Figure 3 identifies the groups on a map of Europe and, together with Fig. 2, allows for a detailed description of the groups' characteristics.

- The first group contains 10 sites and can be described by moderate to large total population and population variability and by large total deposition influence but small deposition variability. This characterisation holds for all catchment areas. The group comprises sites of continental character that can be identified as classical *rural* background sites.
- The second group (7 members) showed small population sums and variability. The total deposition influence was also small while the deposition variability was moderate. These sites can clearly be identified as *remote coastal or high altitude* sites.

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

- Total population influence was large for the third group (5 members), however showing large spread. The population variability was large as well for most sites within this group but increased strongly from the 12 h to the 24 and 48 h catchments. Total deposition influence was moderate but deposition variability was large for all catchment areas. The group contains sites with a large pollution burden with a bias towards sites in coastal areas and might be called regionally *polluted rural* sites.
- The fourth group (6 members) showed rather small total population influence and population variability for the 12 h catchment area. However, the influence was systematically larger for the 24 and 48 h catchment areas than for the remote coastal or high altitude cluster. The total deposition influence was moderate, yet with a large spread in the deposition variability and, again, systematically larger than for the remote sites for the 24 and 48 h catchment area. The group selected *semi-remote coastal and high elevation* sites.
- The fifth group was characterized by extremely low population influence (sums and standard deviations) and low deposition sums but large deposition variability in the case of Mace Head (IE31). The group only contains the two *very remote coastal* sites Mace Head (IE31) and Lampedusa (LMP). When the number of groups of the clustering was limited to 4, this group was merged with the second group (remote coastal or high altitude sites).
- For the sixth group population sums and variability were moderate. The total deposition influence was moderate, while the deposition variability was large. This group comprises 4 *rural coastal* sites and would be merged with the fourth group (semi-remote coastal and high elevation sites) when reducing the number of groups to 4.

While for most of the characterised sites the clustering result supports an intuitive site categorisation, it is interesting to note that the high altitude sites Jungfraujoch

(CH01) and Sonnblick (AT34) were characterised as remote coastal or high altitude sites while the third high Alpine observatory at Zugspitze (ZUG) was within group 4 (semi-remote coastal and high elevation sites). However, this can be explained by the more central Alpine location and higher elevation of Jungfrauoch (3580 m a.s.l.) and Sonnblick (3106 m a.s.l.) compared to position and elevation of Zugspitze (summit) (2950 m a.s.l.) at the northern flank of the Alps.

The robustness of the site categorisation was tested by modifying different parameters used in the clustering procedure. First, the clustering was repeated with equal weights for all selected cluster variables. However, the results did not yield a reasonable categorisation of the continental sites. The obtained categories explained less of the observed inter-site variability of NO₂ and O₃ than the reference clustering (see Sect. 3.5). The categorisation was the same as in the reference case for weights 1.9 to 2.4. Giving more importance to the emission related parameters (weights larger 2.4) did also not yield a reasonable clustering and again less inter-site variability could be explained. These results indicate that the selected scaling factor of 2 between emission and deposition influence is well suited for this application. Second, the clustering was repeated without the COSMO sites because total residence times as obtained with the COSMO LPDM had been scaled (see Sect. 4.3). The remaining FLEXPART sites were clustered in the same way as in the reference clustering. Third, when the COSMO LPDM residence times were not scaled the clustering yielded only 5 groups. The sites within the aforementioned group 4 were split up and merged with the rural background categorisation (Puy de Dome (PUY), Observatoire de Haute-Provence (OHP), Monte Cimone (CMN), Campisabalos (ES09), and Zugspitze (ZUG)) and the remote coastal or high altitude sites (Zavizan (HR04)). Since such a categorisation does not seem to give sufficient credit to the special situation of elevated sites, we conclude that the correction of COSMO LPDM residence times is necessary to inter-compare results between the sites and models. A fourth sensitivity test of the clustering was done using only the parameters derived from the 12 and 48 h catchment areas. The resulting groups changed only slightly from the reference categorisation, probably due to the suf-

**Parameters
describing
representativeness of
air quality sites**S. Henne et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

5 efficient correlation between the results for different catchments. Including a correlated variable in the clustering process would be identical to increasing the weight of the original variable. However, when only the parameters derived from the 12 h catchment areas were used in the clustering, the categorisation changed considerably. The 12 h only categories did not show such a clear distinction between high altitude sites and sites in flat terrain. Furthermore, the resulting categorisation did not show significant differences between observed group mean concentrations as it was the case for the original clustering (see Sect. 3.5). This indicates the importance of including advection within the last 48 h even if looking at species with lifetimes in a similar range. Finally, weighted mean population and deposition (\bar{P} and \bar{v}_d) instead of totals were used in the clustering. Only four groups were selected by the algorithm in this case. Again, high altitude stations were not well separated from rural sites. This selection does not take into account the generally weaker surface influence on high altitude sites as compared to sites in flat terrain, as reflected by smaller total residence times in the catchment area.

3.5 Observations versus categorisation

To test the obtained site categorisation, observational data from the sites were considered. Median mixing ratios and standard deviations of NO_2 , O_3 and CO are plotted against station category in Fig. 4. Medians and standard deviations were derived from yearly available data in the period 1995–2006 if the availability for any individual year was larger than 75%. The observational data was not constrained to the year 2005, for which footprints were calculated, in order to obtain values for a sufficiently large number of sites. For NO_2 the remote and semi-remote high altitude and coastal sites (category 2 and 4) showed the smallest mixing ratios, followed by the rural (category 1) and rural coastal sites (category 6), while the largest mixing ratios were observed at the polluted rural sites (category 3). A fraction of 75% of the variance within station NO_2 medians was explained by the categorisation (significantly different group means, probability of error $\alpha < 0.01$). Similar rankings were obtained for NO_2 standard devia-

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Parameters
describing
representativeness of
air quality sites**S. Henne et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

tions with an even larger fraction of explained inter-site variance (85%). For O₃ the ranking between the sites is contrary to NO₂. Highest O₃ mixing ratios were observed at high altitude sites within category 2 and 4, while values were in general smaller for the coastal sites in these categories. Average mixing ratios were obtained at rural and very remote coastal sites, while lowest O₃ mixing ratios were reported for rural coastal and for rural polluted sites (due to NO titration). A fraction of 55% of the inter-station O₃ variability was explained by the categorisation ($\alpha < 0.05$). In contrast to median levels, ozone variability was largest for rural sites (category 1), and similar for rural polluted, rural coastal and semi-remote sites. Smallest variability was observed at the remote and very-remote sites. For CO, unfortunately, only 10 observational data sets were available. Relatively low CO values were obtained at the remote and semi-remote high altitude and coastal sites (category 2 and 4). Nevertheless, there was large spread in category 1 and 2 (rural and polluted rural). The categorisation explained 54% of the variance between station medians, however the differences between the group means were not significant ($\alpha > 0.1$). CO variability closely followed the rankings for median mixing ratios.

From this observational proof we conclude that our categorisation yielded meaningful results for species with (boundary layer) lifetimes in the order of 0.5–2 d, while the results for CO with a much longer lifetime were inconclusive.

3.6 Station categorisation based on pre-defined circular surrounding area

The categorisation presented above is based on intensive advection calculations and the method is therefore only feasible for a limited number of sites given limited computing resources. Alternatively, parameters of representativeness can be derived in defined areas around a site instead of the catchment area, neglecting surface emission sensitivities (footprints). Obviously, such a method would largely ignore the influence of transport and dilution which was shown to be significantly different for different sites (Fig. 1). Nevertheless, we derived total population and deposition burdens and their variability in circular areas around the sites with radii of 10 and 50 km, respectively. To

have some handle on the relative vertical position of a site we included an additional parameter describing the altitude difference between the site and the median surface altitude in the selected area. Topographic data were taken from the approx. 1 km by 1 km GLOBE data set (<http://www.ngdc.noaa.gov/mgg/topo/globe.html>). In total, these 10 variables were then treated in a similar way as described in Sect. 2.4 and processed by the same clustering algorithm. Altitude difference and population parameters were given weight 2, while deposition parameters were assigned weight 1.

Only 5 different groups of sites were identified by the clustering algorithm (see Figs. S4 and S5 in the supplement). These groups were identified as: high altitude, rural, rural coastal, rural polluted, and semi-remote coastal/high altitude. 17 of the 34 sites ended up in similar groups as obtained with the catchment area approach. Differences are especially apparent for polluted rural sites when advection is ignored. On the one hand, several elevated sites that are close to population centres (Puy de Dome (PUY), Donon (FR08), Schauinsland (DE03)) fell into this group as well, since the population burden dominated the altitude difference parameters, while in reality these sites often sample outside the polluted boundary layer. On the other hand, the four sites that were identified as most polluted by the catchment area approach fell into three different groups in the simpler approach. In contrast to the catchment area approach, the categorisation derived with the surrounding area approach explained less of the inter-site variability of medians and standard deviations of NO₂ and O₃ (see Fig. S6 in the supplement). For CO slightly higher amounts of variability were explained than by the reference categorisations.

A clustering method based solely on parameters of representativeness derived from the surrounding area of a site is more amenable to the categorisation of a larger number of sites but it suffers from ignoring detailed advective transport. While in flat terrain total annual footprints might be similar for sites close to each other and it might therefore be valid to apply the total footprint derived at one site to other sites in the vicinity, this is certainly not possible for sites in more complex terrain. The same needs to be said about bulk footprints that could be applied to any site. A bulk footprint could be

**Parameters
describing
representativeness of
air quality sites**S. Henne et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

parameterised for example as decreasing residence times with the inverse square distance from the site, possibly combined with information on average wind speed and wind direction distribution at the site. These would consider the distance to emissions for all sites in the similar manner, again neglecting the significantly different transport regimes experienced by different sites.

4 Discussion

4.1 Influence of total residence time threshold on catchment area

The catchment area was defined with an arbitrary total residence time threshold, f , of 0.5 which describes the fraction of residence time contained within the catchment volume (see Sect. 2.2.2). To test the robustness of the derived parameters of representativeness we evaluated these for a range of f between 0.1 and 0.9 for all sites. By definition total residence times within the catchment area increase monotonically with increasing f . This is also reflected in total population and deposition burdens (Figs. 5a, c). However, it is worth noticing that for most sites the differences of $\sum P T$ and $\sum v_d T$ for $f=0.4$ and $f=0.6$ remained within the range of $\pm 25\%$ of their reference values for all considered catchment areas. For the variability parameters (Figs. 5b, d) the dependence on the threshold f was in general smaller and for most sites remained within $\pm 25\%$ of its reference for f between 0.3 and 0.7. Rank correlations between the parameters of representativeness obtained for the reference value of $f=0.5$ and for the sensitivity values were larger 0.9 for f between 0.3 and 0.7, showing that a station ranking or clustering based on these parameters is relatively insensitive to the selected threshold.

4.2 Inter-annual variability of catchment areas and representativeness

Catchment areas were derived for the individual reference year 2005. In order to quantify the inter-annual variability of the catchment area and the parameters of represen-

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Parameters
describing
representativeness of
air quality sites**S. Henne et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

tativeness we performed additional simulations using FLEXPART for the years 2003 and 2004 for the site Hohenpeissenberg (HPB). The catchment area was derived for each year individually. The same population and deposition maps as for the base year 2005 were used. Figure 6 compares the derived catchment geometric parameters for the investigated years and the 3 catchment areas. While the total surface area in the catchment, A , did not vary strongly (<20%) for the 12 h catchment, the area covered was 25% and 40% smaller in 2003 and 2004 as compared to 2005 for the 24 and 48 catchment area, respectively. The shape of the catchment areas was similar for different years as also indicated by the catchment's circularity¹ (Fig. 6b). In contrast to the surface area, total residence times within the catchment area were larger by 60% and 120% for the years 2003, 2004 and the 24 and 48 h catchment areas, respectively. This observation points to faster transport and stronger diffusion in 2005 as compared to the years 2003 and 2004. Meteorological conditions in the summer 2003 were rather exceptional (e.g., Schär et al., 2004) with extended high pressure periods and heat wave development both favourable for weak diffusion conditions.

Despite the large differences of the catchment area and its total contained residence time, the inter-annual variability in the derived parameters of representativeness remained in general below 10% (Fig. 7). This can be understood because residence times decrease almost quadratically from the receptor site giving the strongest weight to population and deposition close to the receptor site. Therefore, these parameters were relatively unaffected by inter-annual variability in advection conditions.

4.3 Model inter-comparison

For the catchment area approach, products of total residence times and population/deposition were used to derive total population and deposition influence. In order

¹ Circularity describes the deviation of a shape from a circle by the ratio between the shape's surface area, A , and the surface area of a circle with the same perimeter as the length of the contour line, L , enclosing the shape $c=4\pi A/L^2$.

to assure similar scales of these parameters for the two different models used in this study, residence times for five sites in rather flat terrain were derived by both models (more details can be found in the supplementing material). This inter-comparison indicated the need to scale the COSMO LPDM residence times with respect to the FLEXPART results by a factor of 0.88, 0.81 and 0.83 for 12, 24 and 48 h total residence times, respectively.

The parameters of representativeness used for the station categorisation as derived by the two different models are displayed in Fig. 8. While there is generally close agreement between both simulation results, which is also indicated by Spearman rank correlation coefficients (given in the figure legend) close to or equal to 1, there remained a positive bias for the parameters representing total burdens as derived by the COSMO LPDM. However, after the aforementioned correction had been applied the root mean square difference between both simulations was largely reduced and the positive bias vanished (compare open symbols in Fig. 8a, c). For $\sum PT$ the reductions in root mean square difference were 52, 75 and 68% and for $\sum v_d T$ 73, 83, and 79% for the 12, 24 and 48 h catchment area, respectively.

From this inter-comparison we conclude that although the residence time maps themselves showed differences between the two models (see supplement) the derived parameters of representativeness were similar and, after a scale conversion, can be used in a combined station categorization through clustering.

4.4 Comparison with other studies

Several studies for the categorisation of AQ stations based on reported measurements were conducted in recent years. Snel (2004) used cluster analysis of weekly NO/NO₂ ratios to verify site categories for Dutch AQ sites. In addition, threshold values for NO/NO₂ ratios were used to categorise all EEA/Airbase sites with available NO and NO₂ data. Only 6 sites were common between their and our study and both studies indicated the rural character of these sites, confirming the original EEA/Airbase categorisation (see Table 1). Flemming et al. (2005) derived species specific site cat-

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



5 egorisation of 650 air quality monitoring sites in Germany based on O₃, NO₂, SO₂ and PM₁₀ concentrations applying Ward's clustering on median concentrations and daily variance. Using a similar approach, Tarasova et al. (2007) categorized EMEP and GAW O₃ monitoring sites by their seasonal variation of the diurnal cycle, applying a clustering approach to the resulting matrix of 24×12 aggregates for each site. They identified 6 categories of ozone monitoring sites: clean background, rural, semi-polluted non-elevated, semi-polluted semi-elevated, elevated, and polar-remote. Their categories were available for 18 of the 34 sites discussed here. While for the more remote sites our categorisation resembles theirs, for rural sites the two methods yield 10 substantial variability within the rural subcategories. All three studies yielded meaningful categories for existing stations. In contrast, the method presented here can be used for sites where no data are available (yet) and therefore presents a tool for network design and evaluation independent of available observations.

15 Likewise, Spangl et al. (2007) apply a method for station categorisation and applied it to Austrian AQ stations based on the amount of and the distance to emissions (considered explicitly by species and category) in a 1 and 10 km environment. In contrast to the present study, their approach is more focused on the local scale and implies constant dilution of the emissions independent of station climatologies. Instead of a clustering approach category thresholds were defined based on the distribution of derived parameters of representativeness. They report good consistency of their categorisation 20 based on local road emissions and average NO₂ concentrations.

5 Conclusions

25 An analysis of parameters characterising the representativeness of 34 European AQ sites based on population (emission proxy) and deposition influences within the sites' catchment area was presented. A sites' catchment area is that area in which surface fluxes have a large influence on trace gas concentrations at the site. These areas were derived by explicit backward dispersion simulations using Lagrangian Particle Disper-

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



sion models for a one year period. Emissions and deposition (total and variability) were evaluated within 12, 24 and 48 h catchment areas to focus on the representativeness of species with similar lifetimes in the atmospheric boundary layer. In addition to the catchment area that yields valuable information about the dispersion and advection characteristics of each site, the analysis resulted in a set of 12 parameters of representativeness that can be compared between the sites. These parameters can be used, for example, for the selection of sites suitable for satellite inter-comparison or data assimilation in air quality models. Taking a very short lived species with lifetimes on the order of 12 h that is mainly influenced by emissions into account, it would be reasonable to sort the available sites by $\sum PT_{12}$ and select only those sites below a certain threshold for inter-comparison. When looking at a species with longer lifetimes $\sum PT_{48}$ might be more suitable for a site selection.

Furthermore, the parameters of representativeness were used in a clustering approach to categorise the sites. Six categories were distinguished by the clustering, extending the current EEA/Airbase categorisation (mainly rural). A significant part of the inter-site variability of median O_3 and NO_2 was explained by the new categorisation. The large spread of the parameters of representativeness strongly emphasizes the need for such an additional categorisation, otherwise such remote sites as Mace Head (IE31) would be treated in the same manner as a site as polluted as Kollumerward (NL11) by the incautious data user. While developed for sites focussing on surface O_3 , the presented categorisation is not limited to O_3 and NO_2 . Basically the categorisation is valid for any substance with a horizontal distribution that is driven by emissions proportional to population density. For species with very different emission distributions it would, however, be necessary to derive another set of parameters of representativeness (by calculating totals and variability within the catchment areas) and also a different site categorization.

The robustness of the categorisation was tested by varying the residence time threshold used to derive the catchment area. While the extent and shape of the catchment area was strongly influenced by this choice, the parameters of representativeness

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Parameters
describing
representativeness of
air quality sites**S. Henne et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

remained relatively stable. Year-to-year variations of the catchment area were investigated at one site (Hohenpeissenberg) and resulted in the same conclusions as for the sensitivity test. However, with changing emission and land use pattern this kind of representativeness analysis needs to be redone on a regular basis to account for changes of surface fluxes in the catchment areas. Changes in the local environment (up to 1 km) will have an even stronger impact on the selected rural and remote sites and should thus be avoided whenever possible.

When comparing the categorisation as derived from parameters of representativeness calculated from the catchment areas with a categorisation that was determined from parameters that were derived with a simpler method, not taking advection into account, the value of the advection calculation is emphasised and justifies the computation effort. The categorisation based on parameters of the surrounding only was less capable of handling sites in more complex terrain and in general explained less of the observed inter-site concentration differences.

As discussed by Spangl et al. (2007), the inclusion of many parameters in site categorisation might lead to an over-categorisation of sites with too many subgroups for straight forward data interpretation. The clustering approach used here, however, has the strength of finding groups of stations in a multi-dimensional space of parameters of representativeness and thereby reducing the number of categories to a reasonable number. In addition, no threshold values have to be defined. Nevertheless, redoing the clustering with additional sites might considerably change the detected groups. Alternatively, additional sites can be compared to the current cluster medians and added to the cluster with the smallest distance. However, similar studies with a larger set of sites should be performed, so that the groups will become more robust. Future studies would benefit from the inclusion of actual high resolution emission data. The categories derived here and in future studies should help select sites that match the representativeness requirements of satellites and models.

Acknowledgements. We acknowledge financial support by the EU FP6 project GEOmon, project number 036677. Computations using COSMO LPDM were performed at the Swiss

National Supercomputing Center (CSCS) and input wind fields were provided by MeteoSwiss. Additional observations not included in the EMEP, GAW, EEA AirBase databases were kindly provided by F. Gheusi, Y. Meyerfeld, A. Vermeulen, B. Bandy, K. Uhse and P. Beliche.

References

- 5 Blanchard, C. L., Carr, E. L., Collins, J. F., Smith, T. B., Lehrman, D. E., and Michaels, H. M.: Spatial representativeness and scales of transport during the 1995 Integrated Monitoring Study in California's San Joaquin Valley, *Atmos. Environ.*, 33, 4775–4786, 1999. 20022
- Center for International Earth Science Information Network (CIESIN) – Columbia University and Centro Internacional de Agricultura Tropical (CIAT): Gridded Population of the World Version 3 (GPWv3), online available at: <http://sedac.ciesin.columbia.edu/gpw/>, (last access: August 2007) 2005. 20030
- 10 Derwent, R. G., and Davies, T. J.: Modeling the impact of NO_x or hydrocarbon control on photochemical ozone in Europe, *Atmos. Environ.*, 28, 2039–2052, 1994. 20032
- European Commission – Joint Research Centre: Global Land Cover 2000 database, online available at: <http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php>, last access: September 2009. 20030
- 15 European Council: Exchange on Information decision (Eol) on Air, Tech. Rep., European Council, 1997. 20023
- Flemming, J., Stern, R., and Yamartino, R. J.: A new air quality regime classification scheme for O₃, NO₂, SO₂ and PM₁₀ observations sites, *Atmos. Environ.*, 39, 6121–6129, 2005. 20044
- 20 Folini, D., Ubl, S., and Kaufmann, P.: Lagrangian particle dispersion modeling for the high Alpine site Jungfraujoch, *J. Geophys. Res.*, 113, D18111, doi:10.1029/2007JD009558, 2008. 20026
- Galassi, M., Davies, J., Theiler, J., Gough, B., Jungman, G., Alken, P., Booth, M., and Rossi, F.: GNU scientific library reference manual, Third edition, for Vers. 1.12, Network Theory, Bristol, Third Edn., 2009. 20031
- 25 Glaab, H., Fay, B. and Jacobsen, I.: Evaluation of the emergency dispersion model at the deutscher wetterdienst using ETEX data, *Atmos. Environ.*, 32, 4359–4366, 1998. 20026
- Henne, S., Klausen, J., Junkermann, W., Kariuki, J. M., Aseyo, J. O., and Buchmann, B.: Representativeness and climatology of carbon monoxide and ozone at the global GAW station

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

- Mt. Kenya in equatorial Africa, *Atmos. Chem. Phys.*, 8, 3119–3139, 2008,
<http://www.atmos-chem-phys.net/8/3119/2008/>. 20023, 20033
- Kljun, N., Rotach, M. W., and Schmid, H. P.: A three-dimensional backward lagrangian footprint model for a wide range of boundary-layer stratifications, *Bound.-Lay. Meteorol.*, 103, 205–226, 2002. 20029
- 5 Kuhlbusch, T., John, A., Hugo, A., Peters, A., von Klot, S., Cyrus, J., Wichmann, H.-E., Quass, U., and Bruckmann, P.: Analysis and design of local air quality measurements, *Tech. Rep.*, European Commission, 2006. 20022
- Larsen, S., Sluiter, R., and Helmis, C.: Criteria for EUROAIRNET – The EEA Air Quality Monitoring and Information Network, *Tech. Rep.*, European Environment Agency, 1999. 20021, 20022
- 10 Memmesheimer, M., Jakobs, H. J., Hass, H., Feldmann, H., Ebel, A., Bock, H. J., and Beck, J.: Budget calculations for a summer-smog episode as simulated with the EURAD model, *Proceedings of Eurotrac Symposium '96 – Transport and Transformation of Pollutants in the Troposphere*, Vol. 1, 583–587, 1057, 1997. 20032
- 15 Mol, W., Hooydonk, P., and de Leeuw, F.: European exchange of monitoring information and state of the air quality in 2006, *Tech. Rep.*, ETC/ACC, 2008. 20023
- Nappo, C. J., Caneill, J. Y., Furman, R. W., Gifford, F. A., Kaimal, J. C., Kramer, M. L., Lockhart, T. J., Pendergast, M. M., Pielke, R. A., Randerson, D., Shreffler, J. H., and Wynaard, J. C.: The Workshop on the Representativeness of Meteorological-Observations, June 1981, Boulder, CO, *B. Am. Meteorol. Soc.*, 63, 761–764, 1982. 20022
- 20 Scaperdas, A. and Colvile, R. N.: Assessing the representativeness of monitoring data from an urban intersection site in central London, UK, *Atmos. Environ.*, 33, 661–674, 1999. 20022
- Schmid, H. P.: Experimental design for flux measurements: matching scales of observations and fluxes, *Agr. Forest Meteorol.*, 87, 179–200, 1997. 20027, 20028
- 25 Schmid, H. P.: Footprint modeling for vegetation atmosphere exchange studies: a review and perspective, *Agr. Forest Meteorol.*, 113, 159–183, 2002. 20029
- Schär, C., Vidale, P. L., Luthi, D., Frei, C., Haberli, C., Liniger, M. A., and Appenzeller, C.: The role of increasing temperature variability in European summer heatwaves, *Nature*, 427, 332–336, 2004. 20043
- 30 Seibert, P. and Frank, A.: Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode, *Atmos. Chem. Phys.*, 4, 51–63, 2004,
<http://www.atmos-chem-phys.net/4/51/2004/>. 20025, 20026

**Parameters
describing
representativeness of
air quality sites**S. Henne et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

- Snel, S.: Improvement of classifications European monitoring stations for AirBase, Tech. Rep., ETC/ACC, 2004. 20044
- Spangl, W., Schneider, J., Moosmann, L., and Nagl, C.: Representativeness and classification of air quality monitoring stations, Tech. Rep., Umweltbundesamt, 2007. 20045, 20047
- 5 Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmos. Chem. Phys., 5, 2461–2474, 2005, <http://www.atmos-chem-phys.net/5/2461/2005/>. 20026
- Tarasova, O. A., Brenninkmeijer, C. A. M., Jöckel, P., Zvyagintsev, A. M., and Kuznetsov, G. I.: A climatology of surface ozone in the extra tropics: cluster analysis of observations and
10 model results, Atmos. Chem. Phys., 7, 6099–6117, 2007, <http://www.atmos-chem-phys.net/7/6099/2007/>. 20045
- von Kuhlmann, R., Lawrence, M. G., Crutzen, P. J., and Rasch, P. J.: A model for studies of tropospheric ozone and nonmethane hydrocarbons: Model description and ozone results, J. Geophys. Res.-Atmos., 108, 4294, doi:10.1029/2002JD002893, 2003. 20032
- 15 Ward, J. H.: Hierarchical grouping to optimize an objective function, J. Am. Stat. Assoc., 58, 236–244, 1963. 20033
- Weissmann, M., Braun, F. J., Gantner, L., Mayr, G. J., Rahm, S., and Reitebuch, O.: The Alpine mountain-plain circulation: Airborne Doppler lidar measurements and numerical simulations, Mon. Weather Rev., 133, 3095–3109, 2005. 20026
- 20 Wesely, M. L.: Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical-models, Atmos. Environ., 23, 1293–1304, 1989. 20030
- Wild, O.: Modelling the global tropospheric ozone budget: exploring the variability in current models, Atmos. Chem. Phys., 7, 2643–2660, 2007, <http://www.atmos-chem-phys.net/7/2643/2007/>. 20032

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 1. Selected sites for detailed assessment of representativeness. In the column Model F stands for FLEXPART and C for COSMO LPDM, a bold letter indicates which model was used for deriving the catchment area of the site. The station categories derived for this study are: (1) rural, (2) remote coastal or high altitude, (3) polluted rural, (4) semi-remote coastal and high elevation, (5) very remote coastal, (6) rural coastal. For sites with Airbase category n.a. no category was available.

Site	ID	GAW ID	Lat. (° N)	Long. (° E)	Alt. a.s.l. (m)	Model	Category (Airbase)	Category (this study)
Bialystok	BIA		53.2	22.75	120	F		1
Birkenes	NO01	BIR	58.383	8.25	190	F	rural	2
Cabauw	NL11		51.967	4.933	60	F, C	rural	3
Campisabalos	ES09		41.283	-3.15	1360	C	rural	4
Donon	FR08		48.5	7.133	775	F, C	rural	1
Finokalia	GR02		35.317	25.667	150	F	rural	2
Harwell	GB36		51.567	-1.317	137	F, C	rural	3
Hegyhatsal	HNG	HUN	46.95	16.65	344	F	n.a.	1
Hohenpeissenberg	HPB	HPB	47.8	11.016	985	F	n.a.	1
Ispra	IT04	IPR	45.8	8.633	209	F	suburban	3
Jungfrauoch	CH01	JFJ	46.55	7.983	3580	C	n.a.	2
Kollumerwaard	NL09	KMW	53.333	6.283	0	F	rural	3
Kosetice	CZ03	KOS	49.583	15.083	534	F, C	rural	1
K-pusztá	HU02	KPS	46.967	19.583	125	F	rural	1
Lampedusa	LMP	LMP	35.517	12.633	60	F	rural	5
Lough Navar	GB06		54.433	-7.9	126	F	rural	2
Mace Head	IE31	MHD	53.333	-9.9	25	F	rural	5
Mahón	ES06	MHN	39.9	4.25	10	F, C	n.a.	2
Monte Cimone	CIM	CMN	44.167	10.683	2165	C	n.a.	4
Monte Velho	PT04	MNH	38.083	-8.8	43	F	rural	6
Neuglobsow	DE07	NGL	53.15	13.033	62	F	rural	1
Obs. de H.-Provence	OHP		43.917	5.7	650	C	n.a.	4
Pic du Midi	PDM		43.067	0.167	2860	C	n.a.	2
Preila	LT15	PLA	55.35	21.067	5	F	rural	6
Puy de Dome	PUY		45.75	3	1465	C	n.a.	4
Roquetas	ES03	ROQ	40.817	-0.5	50	F	n.a.	6
Schauinsland	DE03	SSL	47.917	7.9	1205	F	rural	1
Schmücke	DE08	SMU	50.65	10.767	937	F	rural	1
Sniezka	PL03	SNZ	50.733	15.733	1604	C	rural	1
Sonnblick	AT34	SNB	47.05	12.967	3106	C	rural	2
Weybourne	WEY		52.95	1.122	16	F	rural	3
Zavizan	HR04		44.817	14.983	1594	C	n.a.	4
Zingst	DE09	ZGT	54.433	12.733	1	F	rural	6
Zugspitze	ZUG	ZSF	47.417	10.983	2950	C	n.a.	4

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 2a. Catchment area parameters for 12 h catchment: A_{12} total surface area of catchment, r_{12} equivalent radius, $DD_{max,12}$ main advection direction, T_{12} total residence time, $\sum PT_{12}$ total population times residence time, $\sigma_{P,T}$ standard deviation of population, $\sum v_d T_{12}$ total dry deposition times residence time, σ_{v_d} standard deviation of dry deposition.

ID	A_{12} (km ²)	r_{12} (km)	$DD_{max,12}$ (°)	T_{12} (s)	$\sum PT_{12}$ ^a (s)	$\sigma_{P,T}$ ^a (^o)	$\sum v_d T_{12}$ ^a (cm)	σ_{v_d} ^a (cm s ⁻¹)	Land Cover Type	(%)
BIA	57537	135	SE	5.06e+07	3.27e+09	131	4.16e+07	0.0868	16	36.8
NO01	72717	152	S	4.43e+07	7.38e+08	35.7	1.89e+07	0.247	4	43.7
NL11	68356	148	SW	4.63e+07	2.28e+10	519	3.06e+07	0.288	16	41.2
ES09	30991	99.3	W	4.45e+07	1.34e+09	193	3.77e+07	0.103	16	54
FR08	39595	112	SW	4.35e+07	5.08e+09	207	3.35e+07	0.12	4	28.9
GR02	51852	128	N	4.49e+07	9.46e+08	31.5	1.04e+07	0.201	20	74.1
GB36	81074	161	W	4.54e+07	1.47e+10	539	3.69e+07	0.208	16	57
HNG	37001	109	N	4.94e+07	5.23e+09	288	4.52e+07	0.0747	16	62.8
HPB	18847	77.5	W	4.26e+07	6.04e+09	150	3.06e+07	0.0815	13	34.8
IT04	5757	42.8	N	4.31e+07	1.47e+10	253	2.74e+07	0.122	2	32.6
CH01	2635	29	N	2.16e+07	1.25e+09	63.6	1.38e+07	0.114	13	35.7
NL09	87853	167	SW	4.56e+07	6.34e+09	233	2.32e+07	0.355	20	42.3
CZ03	50531	127	W	4.47e+07	4.34e+09	215	3.89e+07	0.081	16	68.2
HU02	34345	105	NW	5.16e+07	7.58e+09	333	4.93e+07	0.0677	16	90.9
LMP	57254	135	NW	5.04e+07	2.28e+07	2.29	2.67e+06	0.0089	20	99.8
GB06	126044	200	SW	4.03e+07	1.08e+09	44.5	2.18e+07	0.265	13	40.9
IE31	119549	195	SW	4.24e+07	2.90e+08	13.8	1.06e+07	0.247	20	73.8
ES06	40134	113	N	4.71e+07	1.07e+09	45.1	9.00e+06	0.256	20	82.3
CMN	5229	40.8	N	3.41e+07	3.18e+09	130	2.82e+07	0.0325	2	55.5
PT04	54465	132	N	4.94e+07	3.82e+09	230	2.63e+07	0.316	20	39
DE07	66396	145	W	4.93e+07	5.80e+09	339	3.76e+07	0.159	16	43
OHP	10379	57.5	N	4.57e+07	1.38e+09	92.4	3.72e+07	0.0754	16	44.9
PDM	5976	43.6	W	2.93e+07	2.25e+09	114	2.42e+07	0.132	16	42.5
LT15	65470	144	W	5.00e+07	1.46e+09	110	1.94e+07	0.38	20	57.9
PUY	28281	94.9	N	4.79e+07	2.80e+09	148	3.91e+07	0.0511	13	46.6
ES03	18482	76.7	NW	4.07e+07	1.86e+09	63.5	2.94e+07	0.278	16	47.9
DE03	24732	88.7	SW	2.03e+07	4.77e+09	231	1.60e+07	0.104	2	36.1
DE08	41350	115	W	1.64e+07	2.76e+09	148	1.38e+07	0.122	16	61.1
PL03	43187	117	NW	4.22e+07	7.06e+09	142	3.38e+07	0.115	16	46.6
AT34	8911	53.3	NW	3.63e+07	1.45e+09	18.5	2.29e+07	0.0955	4	32.8
WEY	90558	170	W	4.87e+07	5.86e+09	310	2.31e+07	0.409	20	51.7
HR04	14756	68.5	NE	3.26e+07	7.14e+08	47.8	2.53e+07	0.203	2	41.7
DE09	80052	160	W	4.79e+07	2.35e+09	92	2.23e+07	0.365	20	50
ZUG	11487	60.5	W	4.80e+07	4.17e+09	64.9	3.06e+07	0.0802	4	42.6

^a Used for site categorisation

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 2b. Same as Table 2a but for 24 h catchment area.

ID	A_{24} (km ²)	r_{24} (km)	$DD_{max,24}$ (°)	T_{24} (s)	$\sum PT_{24}^a$ (s)	$\sigma_{P,T24}^a$ (°)	$\sum v_d T_{24}^a$ (cm)	$\sigma_{v_d 24}^a$ (cm s ⁻¹)	Land Cover Type	(%)
BIA	272683	295	SE	9.15e+07	6.84e+09	193	7.56e+07	0.108	16	43.1
NO01	361784	339	SW	7.91e+07	1.46e+09	48.5	2.91e+07	0.282	20	45.9
NL11	335406	327	SW	8.40e+07	3.09e+10	504	5.01e+07	0.349	16	38.9
ES09	91186	170	N	6.29e+07	4.14e+09	343	5.32e+07	0.132	16	55.8
FR08	182436	241	SW	7.70e+07	1.06e+10	226	6.23e+07	0.119	16	33.9
GR02	200111	252	N	7.64e+07	1.24e+09	28.9	1.45e+07	0.183	20	79.6
GB36	308637	313	SW	7.60e+07	1.95e+10	492	4.90e+07	0.349	16	40.9
HNG	173256	235	N	8.52e+07	9.26e+09	278	7.61e+07	0.105	16	58.3
HPB	82939	162	W	7.15e+07	1.11e+10	203	5.11e+07	0.114	13	28.7
IT04	25222	89.6	N	7.60e+07	2.42e+10	458	4.93e+07	0.152	2	27.9
CH01	16273	72	W	3.30e+07	2.73e+09	125	2.11e+07	0.137	13	33.8
NL09	403514	358	SW	8.36e+07	1.27e+10	301	3.87e+07	0.378	20	48.5
CZ03	223241	267	W	7.73e+07	8.94e+09	215	6.59e+07	0.101	16	61.7
HU02	157172	224	NW	9.25e+07	1.23e+10	303	8.80e+07	0.0691	16	87.3
LMP	215427	262	NW	9.22e+07	9.22e+08	55	6.60e+06	0.108	20	96.5
GB06	583239	431	SW	7.39e+07	1.50e+09	49	2.77e+07	0.308	20	54.2
IE31	587482	432	SW	7.97e+07	5.57e+08	25.2	1.60e+07	0.238	20	79.9
ES06	165013	229	SE	8.97e+07	1.44e+09	50.5	1.28e+07	0.219	20	88.4
CMN	36737	108	NE	5.59e+07	7.93e+09	196	4.59e+07	0.14	2	43.2
PT04	224252	267	N	8.58e+07	5.41e+09	186	4.33e+07	0.348	20	43.2
DE07	332239	325	W	8.89e+07	1.13e+10	297	6.62e+07	0.226	16	45.9
OHP	44502	119	N	8.04e+07	5.47e+09	207	6.37e+07	0.104	16	39.6
PDM	18105	75.9	NW	3.96e+07	2.61e+09	105	3.29e+07	0.139	16	44.5
LT15	300738	309	W	9.33e+07	3.25e+09	113	4.09e+07	0.386	20	51.1
PUY	127130	201	N	8.12e+07	4.25e+09	126	6.70e+07	0.0648	13	42.4
ES03	53546	131	NW	1.11e+08	4.57e+09	69.2	7.64e+07	0.314	16	45.1
DE03	132816	206	SW	4.56e+07	9.58e+09	253	3.68e+07	0.104	2	32.4
DE08	256583	286	W	4.41e+07	9.83e+09	264	3.71e+07	0.119	16	59.2
PL03	248660	281	W	8.09e+07	1.28e+10	192	6.63e+07	0.116	16	52.6
AT34	31341	99.9	NW	5.48e+07	2.96e+09	52.9	3.46e+07	0.0934	4	34.7
WEY	353916	336	SW	8.56e+07	1.25e+10	385	3.73e+07	0.393	20	53.2
HR04	101769	180	NE	6.14e+07	3.26e+09	229	4.76e+07	0.245	2	36.9
DE09	377531	347	W	8.67e+07	6.35e+09	197	4.39e+07	0.37	20	44
ZUG	31729	100	W	6.90e+07	7.54e+09	135	4.41e+07	0.102	4	36.3

^a Used for site categorisation

Parameters describing representativeness of air quality sites

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 2c. Same as Table 2a but for 48 h catchment area.

ID	A_{48} (km ²)	r_{48} (km)	$DD_{max,48}$ (°)	T_{48} (s)	$\sum PT_{48}^a$ (s)	$\sigma_{P,T48}^a$ (l)	$\sum V_d T_{48}^a$ (cm)	$\sigma_{V_d,48}^a$ (cm s ⁻¹)	Land Cover Type (%)
BIA	8.67e+05	525	W	1.44e+08	1.11e+10	200	1.14e+08	0.201	16 43.5
NO01	9.71e+05	556	SW	1.19e+08	2.73e+09	79.4	4.17e+07	0.296	20 50.3
NL11	1.04e+06	575	SW	1.32e+08	3.83e+10	492	7.61e+07	0.367	16 36.6
ES09	1.29e+05	203	W	6.94e+07	4.44e+09	330	5.80e+07	0.158	16 55.3
FR08	6.37e+05	450	W	1.20e+08	1.89e+10	343	9.86e+07	0.125	16 38.7
GR02	6.22e+05	445	N	1.13e+08	3.20e+09	106	2.37e+07	0.22	20 77.2
GB36	7.64e+05	493	SW	1.13e+08	2.25e+10	440	6.15e+07	0.379	20 37.2
HNG	6.03e+05	438	W	1.31e+08	1.50e+10	264	1.14e+08	0.145	16 56.5
HPB	2.83e+05	300	W	1.06e+08	1.64e+10	219	7.68e+07	0.127	2 25.1
IT04	9.92e+04	178	S	1.17e+08	3.21e+10	441	7.85e+07	0.167	2 26.2
CH01	4.94e+04	125	W	4.27e+07	4.29e+09	166	2.77e+07	0.141	13 31.5
NL09	1.19e+06	615	SW	1.34e+08	1.94e+10	334	6.23e+07	0.385	20 48.1
CZ03	7.52e+05	489	W	1.20e+08	1.69e+10	249	1.02e+08	0.111	16 57.4
HU02	5.70e+05	426	NW	1.46e+08	1.82e+10	278	1.36e+08	0.0892	16 77.5
LMP	6.82e+05	466	NW	1.47e+08	2.10e+09	55.5	1.32e+07	0.153	20 93.1
GB06	2.42e+06	877	W	1.36e+08	2.94e+09	79.7	3.59e+07	0.292	20 69.8
IE31	2.29e+06	853	W	1.45e+08	1.17e+09	39.6	2.25e+07	0.21	20 85.7
ES06	5.25e+05	409	NW	1.55e+08	3.14e+09	139	2.62e+07	0.264	20 85
CMN	2.95e+05	307	NE	1.08e+08	1.36e+10	227	7.67e+07	0.28	2 31.9
PT04	8.88e+05	532	N	1.40e+08	7.46e+09	160	6.47e+07	0.366	20 48.9
DE07	1.16e+06	608	W	1.42e+08	1.92e+10	296	9.91e+07	0.294	16 44.3
OHP	1.14e+05	190	N	1.12e+08	9.60e+09	265	8.48e+07	0.179	16 35.7
PDM	7.56e+04	155	NW	5.15e+07	2.99e+09	102	4.03e+07	0.217	16 41.2
LT15	1.04e+06	576	W	1.53e+08	6.91e+09	150	7.59e+07	0.381	20 43.3
PUY	3.59e+05	338	N	1.13e+08	7.76e+09	285	9.37e+07	0.114	13 34.3
ES03	1.56e+05	223	NW	2.20e+08	9.13e+09	109	1.43e+08	0.337	16 40.9
DE03	5.29e+05	410	W	8.29e+07	1.50e+10	268	6.73e+07	0.12	16 33.8
DE08	8.51e+05	520	W	8.27e+07	1.66e+10	342	6.68e+07	0.198	16 54.2
PL03	6.30e+05	448	W	1.21e+08	2.06e+10	233	1.00e+08	0.118	16 54
AT34	8.28e+04	162	W	7.33e+07	5.35e+09	133	4.70e+07	0.104	4 33.5
WEY	9.34e+05	545	SW	1.30e+08	1.74e+10	390	5.31e+07	0.388	20 55.7
HR04	3.89e+05	352	N	1.06e+08	6.93e+09	209	7.78e+07	0.29	16 36.8
DE09	1.31e+06	647	W	1.42e+08	1.19e+10	213	6.97e+07	0.378	20 45
ZUG	8.18e+04	161	W	8.91e+07	1.14e+10	187	5.87e+07	0.118	4 32

^a Used for site categorisation

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

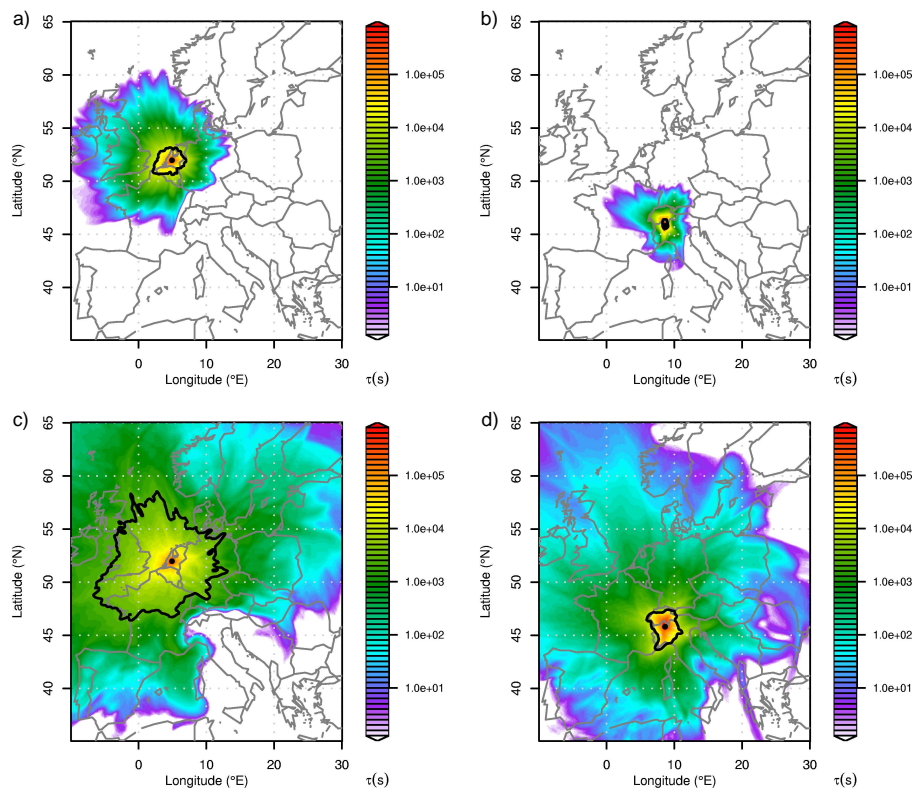


Fig. 1. Total annual footprint (colour scale) and boundary of catchment area (thick black line) for the sites Cabauw (NL11, **a, c**) and Ispra (IT04, **b, d**) and two integration intervals, 12 h (**a, b**) and 48 h (**c, d**).

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

**Parameters
describing
representativeness of
air quality sites**

S. Henne et al.

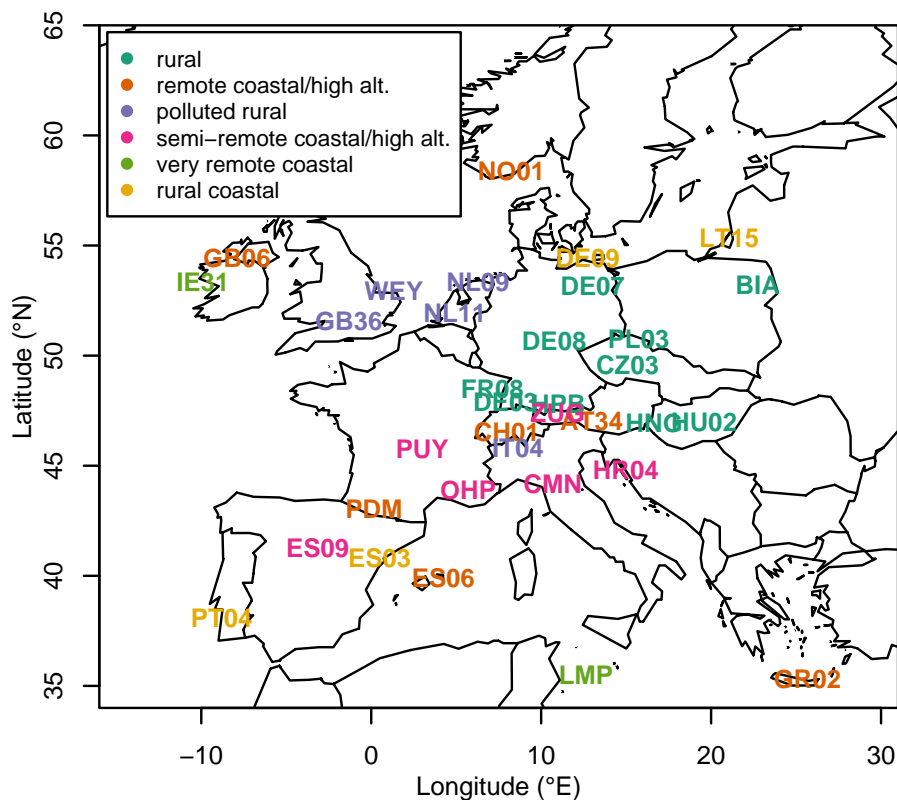


Fig. 3. Map of sites showing categorisation as obtained from clustering of parameters of representativeness in the catchment areas.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Parameters describing representativeness of air quality sites

S. Henne et al.

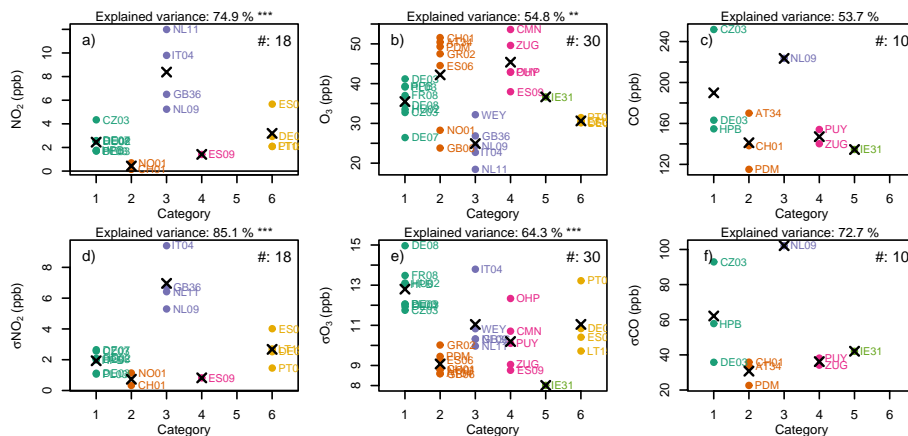


Fig. 4. Sites' median (upper row) and standard deviation of observed (lower row) mixing ratios of (a, d) NO_2 , (b, e) O_3 , (c, f) CO by site plotted versus site category. Black crosses represent the category mean. The star notation in each panel represents the confidence level of differences between category means as derived from ANOVA f statistics (*: $\alpha < 0.1$, **: $\alpha < 0.05$, ***: $\alpha < 0.01$).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Parameters describing representativeness of air quality sites

S. Henne et al.

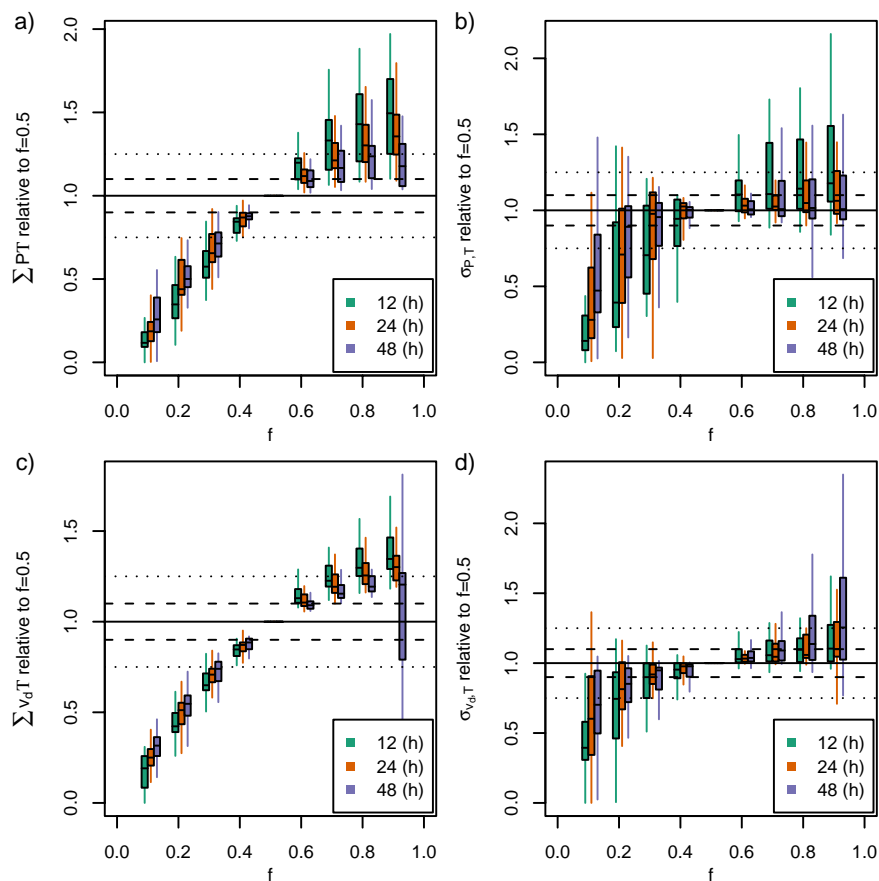


Fig. 5. Boxplots of catchment area parameters for 34 sites as derived for different total residence time thresholds f and 12, 24 and 48 h catchment areas; **(a)** population sum $\sum P T$, **(b)** population variability $\sigma_{P,T}$, **(c)** deposition sum $\sum v_d T$ and **(d)** deposition variability $\sigma_{v_d,T}$.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Parameters
describing
representativeness of
air quality sites

S. Henne et al.

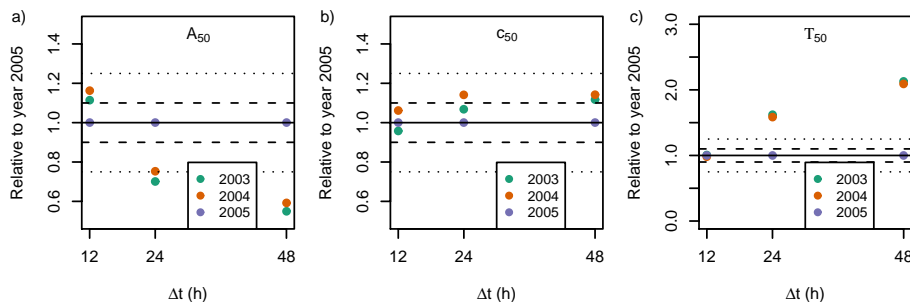


Fig. 6. Annual variability for catchment geometry parameters, **(a)** area A , **(b)** circularity c and **(c)** total annual residence time T as derived for the site Hohenpeissenberg (HPB) and the period 2003–2005.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Parameters describing representativeness of air quality sites

S. Henne et al.

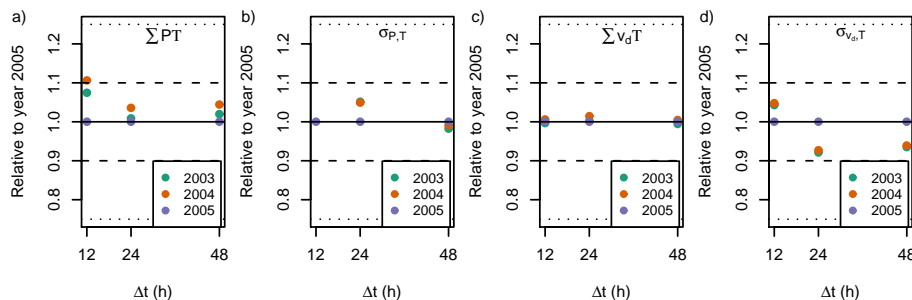


Fig. 7. Annual variability for catchment area parameters **(a)** population sum $\sum P T$, **(b)** population variability $\sigma_{P,T}$, **(c)** deposition sum $\sum v_d T$ and **(d)** deposition variability $\sigma_{v_d,T}$ as derived for the site Hohenpeissenberg (HPB) and the period 2003–2005.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

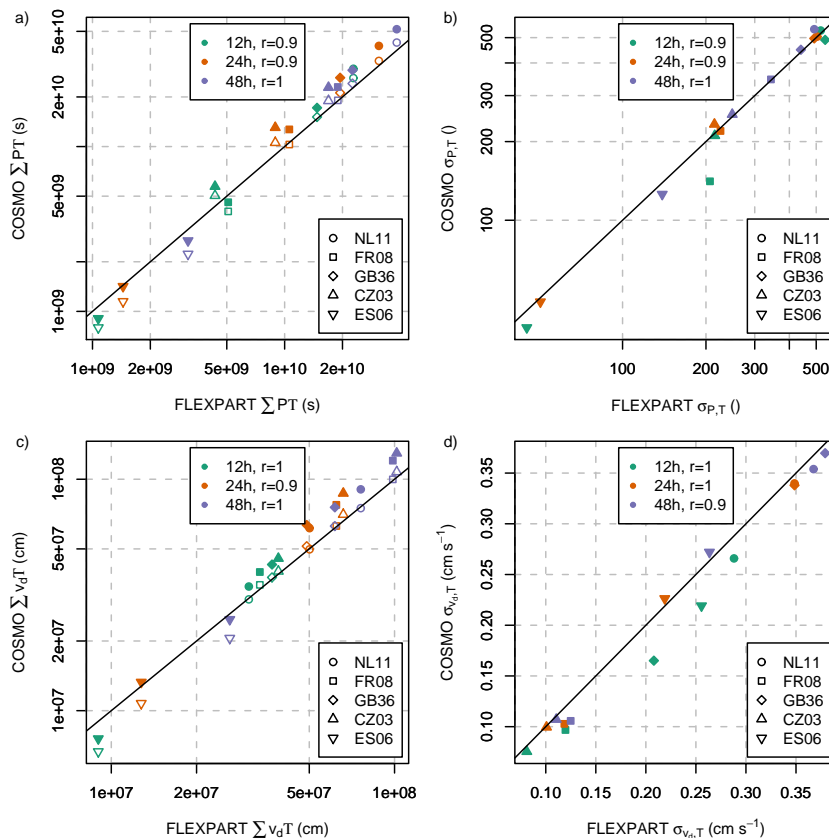


Fig. 8. Catchment area parameters **(a)** population sum $\sum PT$, **(b)** population variability $\sigma_{P,T}$, **(c)** deposition sum $\sum v_d T$ and **(d)** deposition variability $\sigma_{v_d,T}$ as derived by COSMO LPDM versus those derived by FLEXPART. Solid symbols represent original COSMO LPDM results, open symbols represent parameters derived with scaled COSMO LPDM residence times. r gives the Spearman rank correlation coefficient.