**Atmospheric
Chemistry
and Physics**

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

S. Henne[1], D. Brunner[1], D. Folini[2], S. Solberg[3], J. Klausen[1], and B. Buchmann[1]

[1]Empa, Swiss Federal Laboratories for Materials Testing and Research, Dübendorf, Switzerland
[2]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[3]NILU, Norwegian Institute for Air Research, Kjeller, Norway

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 by the atmospheric community within different administrative domains yielding an inconsistent global air quality picture. 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' average 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 derived parameters describing representativeness 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 *generally remote* to more polluted *agglomeration* sites. These six categories explained 50 to 80% of the intersite 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 describing representativeness showed considerably less variability than the catchment geometry, supporting the applicability of the derived station categorisation.

1 Introduction

Ground-based in-situ measurement sites form the backbone of the atmospheric observing system dedicated to composition change and air pollution. 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 longterm 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 typically 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 site representativeness 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 in 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 representativeness 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.

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 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, the area (volume) of interest will vary with application. For the inter-comparison of in-situ (point data) and chemistry transport model (CTM) 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.

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 Colvile, 1999). Factors 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. Considering this and the aforementioned definitions of representativeness, 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, 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 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 O_3 and NO_2.

Next to a quantification of representativeness an objective site categorisation would be very valuable for the purposes just mentioned, for data interpretation and also for 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. However, these classifications are often derived subjectively by the site's maintainer (due to different levels of available and reliable information). Here we develop a categorisation method that is objectively based on parameters describing representativeness and independent of previously recorded AQ data. For verification, the obtained categorization can then be tested against observational data.

The sites selected for this study (Table 1 and Fig. 4) 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 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 O_3, NO_2, 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 describing representativeness from Lagrangian transport simulations combined with proxy emission and deposition data and how to use these in a site categorisation. The derived parameters describing 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.

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) *mostly remote*, (3) *agglomeration*, (4) *weakly influenced, constant deposition*, (5) *generally remote*, (6) *weakly influenced, variable deposition*. For sites with Airbase category n.a. no category was available.

Site	ID	GAW ID	Lat. (° N)	Long. (° E)	Altitude (m a.s.l.)	Release alt. (m a.s.l.)	Model	Category (Airbase)	Category (this study)
Bialystok	BIA		53.2	22.75	120	168	**F**		1
Birkenes	NO01	BIR	58.383	8.25	190	190	**F**	rural	2
Cabauw	NL11		51.967	4.933	60	60	**F**, C	rural	3
Campisabalos	ES09		41.283	−3.15	1360	1410	**C**	rural	4
Donon	FR08		48.5	7.133	775	775	**F**, C	rural	1
Finokalia	GR02		35.317	25.667	150	150	**F**	rural	2
Harwell	GB36		51.567	−1.317	137	137	**F**, C	rural	3
Hegyhatsal	HNG	HUN	46.95	16.65	344	344	**F**	n.a.	1
Hohenpeissenberg	HPB	HPB	47.8	11.016	985	985	**F**	n.a.	1
Ispra	IT04	IPR	45.8	8.633	209	960	**F**	suburban	3
Jungfraujoch	CH01	JFJ	46.55	7.983	3580	2650	**C**	n.a.	2
Kollumerwaard	NL09	KMW	53.333	6.283	0	20	**F**	rural	3
Kosetice	CZ03	KOS	49.583	15.083	534	534	**F**, C	rural	1
K-puszta	HU02	KPS	46.967	19.583	125	125	**F**	rural	1
Lampedusa	LMP	LMP	35.517	12.633	60	60	**F**	rural	5
Lough Navar	GB06		54.433	−7.9	126	126	**F**	rural	2
Mace Head	IE31	MHD	53.333	−9.9	25	25	**F**	rural	5
Mahón	ES06	MHN	39.9	4.25	10	20	**F**, C	n.a.	2
Monte Cimone	CIM	CMN	44.167	10.683	2165	1350	**C**	n.a.	4
Monte Velho	PT04	MNH	38.083	−8.8	43	43	**F**	rural	6
Neuglobsow	DE07	NGL	53.15	13.033	62	62	**F**	rural	1
Obs. de H.-Provence	OHP		43.917	5.7	650	620	**C**	n.a.	4
Pic du Midi	PDM		43.067	0.167	2860	810	**C**	n.a.	2
Preila	LT15	PLA	55.35	21.067	5	35	**F**	rural	6
Puy de Dome	PUY		45.75	3	1465	860	**C**	n.a.	4
Roquetas	ES03	ROQ	40.817	−0.5	50	350	**F**	n.a.	6
Schauinsland	DE03	SSL	47.917	7.9	1205	1205	**F**	rural	1
Schmücke	DE08	SMU	50.65	10.767	937	937	**F**	rural	1
Sniezka	PL03	SNZ	50.733	15.733	1604	1040	**C**	rural	1
Sonnblick	AT34	SNB	47.05	12.967	3106	2250	**C**	rural	2
Weybourne	WEY		52.95	1.122	16	16	**F**	rural	3
Zavizan	HR04		44.817	14.983	1594	1150	**C**	n.a.	4
Zingst	DE09	ZGT	54.433	12.733	1	33	**F**	rural	6
Zugspitze	ZUG	ZSF	47.417	10.983	2950	1640	**C**	n.a.	4

2 Methods

2.1 Parameters describing 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 primary species like NO_2, emissions and deposition largely determine the small scale ($\sim$1 km) variability of these gases. The spatial distribution of emissions will largely determine the spatial distribution of the species itself and on average the atmospheric concentrations might scale with emission rates. Therefore, emission and deposition data are considered to be appropriate proxies for concentrations and can be used to derive parameters describing representativeness.

In general we assess representativeness on 2 different axes. First, the total surface flux influence (emissions and deposition) on a site is investigated. On this scale sites with small total burden should on average be representative of larger areas. Second, the variability of surface fluxes within the area influencing a site is assessed. Small variability of surface fluxes again points to larger representativeness of a site. These parameters describing representativeness cannot give an absolute quantification of representativeness in terms of the aforementioned definitions, since they don't directly relate a volume average to a point measurement. However, with a combination of such parameters we aim to characterise different aspects of representativeness and to derive a site's "fingerprint"" of representativeness. Furthermore, the parameters describing representativeness are directly intercomparable 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.

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 towns is not reflected in population distributions. Therefore, we might underestimate the influence of traffic in our results, even though the sites considered in this study are not close to any major traffic route. 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 describing representativeness can be obtained by directly investigating total population and deposition influence within certain areas surrounding 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 surface flux sensitivities and the 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, 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 the population. However, the different emission distributions would need to be taken into account which may lead to different parameters describing representativeness and hence a different station categorization than obtained in this study. The determined surface flux sensitivities, nevertheless, are independent of the pollutant in question and could easily be applied to other source distributions. For species with surface distributions that are not driven by surface fluxes 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 Jungfraujoch (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 (see Table 1). and traced backwards in time for 60 h. Sensitivity tests for the site CH01 showed that a release 80 m above model ground yielded the best performance in terms of simulated CO time series (Folini et al., 2008). Starting 80 m above model ground also ensures that particles (trajectories) are not trapped in the lowest model level. 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. Residence 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 20 m above model ground (see Table 1). 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 intercomparison of the model performance (see Sect. 4.3 and supplementary material, see http://www.atmos-chem-phys.net/10/3561/2010/acp-10-3561-2010-supplement.pdf). For these sites, only FLEXPART results were used for the site categorisation.

Table 2a. Catchment area parameters for 12 h catchment: A_{12} total surface area of catchment, r_{12} equivalent radius, $DD_{\mathrm{max},12}$ main advection direction, T_{12} total residence time, $\sum PT_{12}$ population times total 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. The table entries are sorted by population times total residence time.

ID	Altitude (m)	A_{12} (km^2)	r_{12} (km)	$DD_{max,12}$ (°)	T_{12} (s)	$\sum PT_{12}$[a] (s)	$\sigma_{P,T}$[a] ()	$\sum v_d T_{12}$[a] (cm)	σ_{v_d}[a] (cms^{-1})	Land Cover Type	(%)
NL11	60	6.84×10^4	148	SW	4.63×10^7	2.28×10^{10}	519	3.06×10^7	0.288	16	41.2
GB36	137	8.11×10^4	161	W	4.54×10^7	1.47×10^{10}	539	3.69×10^7	0.208	16	57
IT04	209	5.76×10^3	42.8	N	4.31×10^7	1.47×10^{10}	253	2.74×10^7	0.122	2	32.6
HU02	125	3.43×10^4	105	NW	5.16×10^7	7.58×10^9	333	4.93×10^7	0.0677	16	90.9
PL03	1604	4.32×10^4	117	NW	4.22×10^7	7.06×10^9	142	3.38×10^7	0.115	16	46.6
NL09	0	8.79×10^4	167	SW	4.56×10^7	6.34×10^9	233	2.32×10^7	0.355	20	42.3
HPB	985	1.88×10^4	77.5	W	4.26×10^7	6.04×10^9	150	3.06×10^7	0.0815	13	34.8
WEY	16	9.06×10^4	170	W	4.87×10^7	5.86×10^9	310	2.31×10^7	0.409	20	51.7
DE07	62	6.64×10^4	145	W	4.93×10^7	5.8×10^9	339	3.76×10^7	0.159	16	43
HNG	344	3.7×10^4	109	N	4.94×10^7	5.23×10^9	288	4.52×10^7	0.0747	16	62.8
FR08	775	3.96×10^4	112	SW	4.35×10^7	5.08×10^9	207	3.35×10^7	0.12	4	28.9
DE03	1205	2.47×10^4	88.7	SW	2.03×10^7	4.77×10^9	231	1.6×10^7	0.104	2	36.1
CZ03	534	5.05×10^4	127	W	4.47×10^7	4.34×10^9	215	3.89×10^7	0.081	16	68.2
ZUG	2950	1.15×10^4	60.5	W	4.8×10^7	4.17×10^9	64.9	3.06×10^7	0.0802	4	42.6
PT04	43	5.45×10^4	132	N	4.94×10^7	3.82×10^9	230	2.63×10^7	0.316	20	39
BIA	120	5.75×10^4	135	SE	5.06×10^7	3.27×10^9	131	4.16×10^7	0.0868	16	36.8
CMN	2165	5.23×10^3	40.8	N	3.41×10^7	3.18×10^9	130	2.82×10^7	0.0325	2	55.5
PUY	1465	2.83×10^4	94.9	N	4.79×10^7	2.8×10^9	148	3.91×10^7	0.0511	13	46.6
DE09	1	8.01×10^4	160	W	4.79×10^7	2.35×10^9	92	2.23×10^7	0.365	20	50
PDM	2860	5.98×10^3	43.6	W	2.93×10^7	2.25×10^9	114	2.42×10^7	0.132	16	42.5
ES03	50	1.85×10^4	76.7	NW	4.07×10^7	1.86×10^9	63.5	2.94×10^7	0.278	16	47.9
LT15	5	6.55×10^4	144	W	5×10^7	1.46×10^9	110	1.94×10^7	0.38	20	57.9
AT34	3106	8.91×10^3	53.3	NW	3.63×10^7	1.45×10^9	18.5	2.29×10^7	0.0955	4	32.8
OHP	650	1.04×10^4	57.5	N	4.57×10^7	1.38×10^9	92.4	3.72×10^7	0.0754	16	44.9
ES09	1360	3.1×10^4	99.3	W	4.45×10^7	1.34×10^9	193	3.77×10^7	0.103	16	54
CH01	3580	2.64×10^3	29	N	2.16×10^7	1.25×10^9	63.6	1.38×10^7	0.114	13	35.7
GB06	126	1.26×10^5	200	SW	4.03×10^7	1.08×10^9	44.5	2.18×10^7	0.265	13	40.9
GR02	150	5.19×10^4	128	N	4.49×10^7	9.46×10^8	31.5	1.04×10^7	0.201	20	74.1
NO01	190	7.27×10^4	152	S	4.43×10^7	7.38×10^8	35.7	1.89×10^7	0.247	4	43.7
HR04	1594	1.48×10^4	68.5	NE	3.26×10^7	7.14×10^8	47.8	2.53×10^7	0.203	2	41.7
IE31	25	1.2×10^5	195	SW	4.24×10^7	2.9×10^8	13.8	1.06×10^7	0.247	20	73.8
LMP	60	5.73×10^4	135	NW	5.04×10^7	2.28×10^7	2.29	2.67×10^6	0.0089	20	99.8

[a] Used for 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

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

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

12, 24, and 48 h were investigated here. The residence times at the surface are also often called "footprints" and we use these terms interchangeably.

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 measurements at 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 $T_{\mathrm{tot}}=\sum_i\sum_j\sum_k T_{i,j,k}$. To derive the volume of largest residence times it is necessary to transform residence times to mass specific residence times: $\gamma_{i,j,k}=\tau_{i,j,k}/m_{i,j,k}$ for the

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

ID	Altitude (m)	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] (cms^{-1})	Land Cover Type	(%)
NL11	60	3.35×10^5	327	SW	8.4×10^7	3.09×10^{10}	504	5.01×10^7	0.349	16	38.9
IT04	209	2.52×10^4	89.6	N	7.6×10^7	2.42×10^{10}	458	4.93×10^7	0.152	2	27.9
GB36	137	3.09×10^5	313	SW	7.6×10^7	1.95×10^{10}	492	4.9×10^7	0.349	16	40.9
PL03	1604	2.49×10^5	281	W	8.09×10^7	1.28×10^{10}	192	6.63×10^7	0.116	16	52.6
NL09	0	4.04×10^5	358	SW	8.36×10^7	1.27×10^{10}	301	3.87×10^7	0.378	20	48.5
WEY	16	3.54×10^5	336	SW	8.56×10^7	1.25×10^{10}	385	3.73×10^7	0.393	20	53.2
HU02	125	1.57×10^5	224	NW	9.25×10^7	1.23×10^{10}	303	8.8×10^7	0.0691	16	87.3
DE07	62	3.32×10^5	325	W	8.89×10^7	1.13×10^{10}	297	6.62×10^7	0.226	16	45.9
HPB	985	8.29×10^4	162	W	7.15×10^7	1.11×10^{10}	203	5.11×10^7	0.114	13	28.7
FR08	775	1.82×10^5	241	SW	7.7×10^7	1.06×10^{10}	226	6.23×10^7	0.119	16	33.9
DE03	1205	1.33×10^5	206	SW	4.56×10^7	9.58×10^9	253	3.68×10^7	0.104	2	32.4
HNG	344	1.73×10^5	235	N	8.52×10^7	9.26×10^9	278	7.61×10^7	0.105	16	58.3
CZ03	534	2.23×10^5	267	W	7.73×10^7	8.94×10^9	215	6.59×10^7	0.101	16	61.7
CMN	2165	3.67×10^4	108	NE	5.59×10^7	7.93×10^9	196	4.59×10^7	0.14	2	43.2
ZUG	2950	3.17×10^4	100	W	6.9×10^7	7.54×10^9	135	4.41×10^7	0.102	4	36.3
BIA	120	2.73×10^5	295	SE	9.15×10^7	6.84×10^9	193	7.56×10^7	0.108	16	43.1
DE09	1	3.78×10^5	347	W	8.67×10^7	6.35×10^9	197	4.39×10^7	0.37	20	44
OHP	650	4.45×10^4	119	N	8.04×10^7	5.47×10^9	207	6.37×10^7	0.104	16	39.6
PT04	43	2.24×10^5	267	N	8.58×10^7	5.41×10^9	186	4.33×10^7	0.348	20	43.2
ES03	50	5.35×10^4	131	NW	1.11×10^8	4.57×10^9	69.2	7.64×10^7	0.314	16	45.1
PUY	1465	1.27×10^5	201	N	8.12×10^7	4.25×10^9	126	6.7×10^7	0.0648	13	42.4
ES09	1360	9.12×10^4	170	N	6.29×10^7	4.14×10^9	343	5.32×10^7	0.132	16	55.8
HR04	1594	1.02×10^5	180	NE	6.14×10^7	3.26×10^9	229	4.76×10^7	0.245	2	36.9
LT15	5	3.01×10^5	309	W	9.33×10^7	3.25×10^9	113	4.09×10^7	0.386	20	51.1
AT34	3106	3.13×10^4	99.9	NW	5.48×10^7	2.96×10^9	52.9	3.46×10^7	0.0934	4	34.7
CH01	3580	1.63×10^4	72	W	3.3×10^7	2.73×10^9	125	2.11×10^7	0.137	13	33.8
PDM	2860	1.81×10^4	75.9	NW	3.96×10^7	2.61×10^9	105	3.29×10^7	0.139	16	44.5
GB06	126	5.83×10^5	431	SW	7.39×10^7	1.5×10^9	49	2.77×10^7	0.308	20	54.2
NO01	190	3.62×10^5	339	SW	7.91×10^7	1.46×10^9	48.5	2.91×10^7	0.282	20	45.9
GR02	150	2×10^5	252	N	7.64×10^7	1.24×10^9	28.9	1.45×10^7	0.183	20	79.6
LMP	60	2.15×10^5	262	NW	9.22×10^7	9.22×10^8	55	6.6×10^6	0.108	20	96.5
IE31	25	5.87×10^5	432	SW	7.97×10^7	5.57×10^8	25.2	1.6×10^7	0.238	20	79.9

[a] Used for site categorisation.

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 $n=1,\ldots,IJK$. 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 $\sum_n^{1,\ldots,IJK} T_n \geq fT_{\text{tot}}$ with $f=0.5$ was fulfilled. 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. For this, all surface grid cells fulfilling $\Gamma_{i,j}^{500} \geq \Gamma_{50}$ were defined as catchment area, with $\Gamma_{i,j}^{500}$ being the specific residence time integrated 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 high altitude sites that usually experience large surface sensitivities 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. A large fraction of transport towards a mountain site takes place above the atmospheric boundary layer, therefore the area in which surface fluxes significantly influence a mountain site must be small according to our concept. Folini et al. (2009), using the same LPDM technique as described here, estimated that about 60% and 45% of the observations at Jungfraujoch are unaffected by boundary layer contact in winter and summer, respectively. If, in contrast, taking 50% of surface residence times ($T_{\text{tot},500}=\sum_i\sum_j T_{i,j,500}$) into account

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

ID	Altitude (m)	A_{48} (km^2)	r_{48} (km)	$DD_{max,48}$ (°)	T_{48} (s)	$\sum PT_{48}$[a] (s)	$\sigma_{P,T48}$[a] ()	$\sum v_d T_{48}$[a] (cm)	$\sigma_{v_d 48}$[a] (cms^{-1})	Land Cover Type	(%)
NL11	60	1.04×10^6	575	SW	1.32×10^8	3.83×10^{10}	492	7.61×10^7	0.367	16	36.6
IT04	209	9.92×10^4	178	S	1.17×10^8	3.21×10^{10}	441	7.85×10^7	0.167	2	26.2
GB36	137	7.64×10^5	493	SW	1.13×10^8	2.25×10^{10}	440	6.15×10^7	0.379	20	37.2
PL03	1604	6.3×10^5	448	W	1.21×10^8	2.06×10^{10}	233	1×10^8	0.118	16	54
NL09	0	1.19×10^6	615	SW	1.34×10^8	1.94×10^{10}	334	6.23×10^7	0.385	20	48.1
DE07	62	1.16×10^6	608	W	1.42×10^8	1.92×10^{10}	296	9.91×10^7	0.294	16	44.3
FR08	775	6.37×10^5	450	W	1.2×10^8	1.89×10^{10}	343	9.86×10^7	0.125	16	38.7
HU02	125	5.7×10^5	426	NW	1.46×10^8	1.82×10^{10}	278	1.36×10^8	0.0892	16	77.5
WEY	16	9.34×10^5	545	SW	1.3×10^8	1.74×10^{10}	390	5.31×10^7	0.388	20	55.7
CZ03	534	7.52×10^5	489	W	1.2×10^8	1.69×10^{10}	249	1.02×10^8	0.111	16	57.4
HPB	985	2.83×10^5	300	W	1.06×10^8	1.64×10^{10}	219	7.68×10^7	0.127	2	25.1
HNG	344	6.03×10^5	438	W	1.31×10^8	1.5×10^{10}	264	1.14×10^8	0.145	16	56.5
DE03	1205	5.29×10^5	410	W	8.29×10^7	1.5×10^{10}	268	6.73×10^7	0.12	16	33.8
CMN	2165	2.95×10^5	307	NE	1.08×10^8	1.36×10^{10}	227	7.67×10^7	0.28	2	31.9
DE09	1	1.31×10^6	647	W	1.42×10^8	1.19×10^{10}	213	6.97×10^7	0.378	20	45
ZUG	2950	8.18×10^4	161	W	8.91×10^7	1.14×10^{10}	187	5.87×10^7	0.118	4	32
BIA	120	8.67×10^5	525	W	1.44×10^8	1.11×10^{10}	200	1.14×10^8	0.201	16	43.5
OHP	650	1.14×10^5	190	N	1.12×10^8	9.6×10^9	265	8.48×10^7	0.179	16	35.7
ES03	50	1.56×10^5	223	NW	2.2×10^8	9.13×10^9	109	1.43×10^8	0.337	16	40.9
PUY	1465	3.59×10^5	338	N	1.13×10^8	7.76×10^9	285	9.37×10^7	0.114	13	34.3
PT04	43	8.88×10^5	532	N	1.4×10^8	7.46×10^9	160	6.47×10^7	0.366	20	48.9
HR04	1594	3.89×10^5	352	N	1.06×10^8	6.93×10^9	209	7.78×10^7	0.29	16	36.8
LT15	5	1.04×10^6	576	W	1.53×10^8	6.91×10^9	150	7.59×10^7	0.381	20	43.3
AT34	3106	8.28×10^4	162	W	7.33×10^7	5.35×10^9	133	4.7×10^7	0.104	4	33.5
ES09	1360	1.29×10^5	203	W	6.94×10^7	4.44×10^9	330	5.8×10^7	0.158	16	55.3
CH01	3580	4.94×10^4	125	W	4.27×10^7	4.29×10^9	166	2.77×10^7	0.141	13	31.5
GR02	150	6.22×10^5	445	N	1.13×10^8	3.2×10^9	106	2.37×10^7	0.22	20	77.2
PDM	2860	7.56×10^4	155	NW	5.15×10^7	2.99×10^9	102	4.03×10^7	0.217	16	41.2
GB06	126	2.42×10^6	877	W	1.36×10^8	2.94×10^9	79.7	3.59×10^7	0.292	20	69.8
NO01	190	9.71×10^5	556	SW	1.19×10^8	2.73×10^9	79.4	4.17×10^7	0.296	20	50.3
LMP	60	6.82×10^5	466	NW	1.47×10^8	2.1×10^9	55.5	1.32×10^7	0.153	20	93.1
IE31	25	2.29×10^6	853	W	1.45×10^8	1.17×10^9	39.6	2.25×10^7	0.21	20	85.7

[a] Used for site categorisation.

for mountain sites, 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. However, regional emissions within the catchment area of a mountain site are often small, therefore their influence on concentration measurements is low and signals from outside the catchment area might still be detectable at those sites even though the same signal might not be observable at sites in flat terrain.

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 describing representativeness to the chosen threshold value is further discussed in Sect. 4.1. 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.net/10/3561/2010/acp-10-3561-2010-supplement.pdf).

The geometry of the catchment areas can be summarized by a few simple parameters that are given for each site in Table 2aa–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.

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. The flux footprint has a more limited horizontal extent compared to the concentration footprint (Kljun et al., 2002), which we look at in this study. The extent of the catchment area, as defined in this study, is limited by the integration interval of the LPDM that was chosen to be in the range of time scales (<48h) 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 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 Land cover

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.net/10/3561/2010/acp-10-3561-2010-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 deposion, 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, $800\,Wm^{-2}$ global radiation and $0.7\,ms^{-1}$ friction velocity (independent of land cover type). Summer conditions were chosen because O_3 production is

strongest during summer and also the largest horizontal variability in O_3 can be expected. The resulting ozone 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.net/10/3561/2010/acp-10-3561-2010-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, $\sum T_{i,j} P_{i,j}$ (units number s), 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 (units number)

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

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 categorisation (compare Fig. 2 and Table 2aa–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 supplementary material, http://www.atmos-chem-phys.net/10/3561/2010/acp-10-3561-2010-supplement.pdf).

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

$$x_n = \frac{(x - \overline{x})}{\sigma_x},\tag{4}$$

where $\overline{x}$ represents the parameter mean and σ_x its standard deviation. Furthermore, the parameters used in the clustering should be normally distributed. For the population parameters 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 surface dry deposition to chemical processing, which is largely driven by anthropogenic precursor emissions, can be obtained from model studies. While for the global tropospheric domain the deposition term dominates the budget (ratio: $\sim$3.5, Wild, 2007), it becomes less important within the continental troposphere (ratio: $\sim$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: $\sim$0.1, Memmesheimer et al., 1997). By choosing a factor of 0.5 between deposition and emission influence in our clustering approach we consider the lower limit of this factor for the O_3 budget, but are above the upper limit for NO_2 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 supplementary material, http://www.atmos-chem-phys.net/10/3561/2010/acp-10-3561-2010-supplement.pdf), in-situ observations of O_3, NO_2 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). Whenever possible we included all available station data in our study and only excluded data that was flagged invalid. All flags distinguishing background or non-background data were ignored and all data were included in all derived aggregates.

2.6 Terminology

This section repeats some of the terminology used in the article and gives relations between the different terms.

- Footprint: The term footprint is used here to describe the total annual surface residence times (surface flux sensitivities) of a measurement site as obtained from LPDM backward calculations. The footprint is a quantitative representation, a 2D map, of any ground contact of the air that is sampled at a receptor site.

- Catchment area: That part of the footprint where the ground contact of the air is most substantial, is longest, and hence from where surface fluxes potentially have the most significant impact on the receptor site. This area is not directly connected to the area of representativeness, but is determined by advection towards a site. However, analyses of surface fluxes within the catchment area yields information on representativeness.

- Parameters describing representativeness: These parameters are derived from proxy emission and deposition flux data within the catchment area of a site. Two sets of parameters are evaluated, those that reflect total surface fluxes and those that estimate surface flux variability. For both sets larger values indicate decreasing representativeness. While an individual parameter cannot describe representativeness for various point-to-area geometries and different trace species of interest, a set of parameters is analysed to derive the "fingerprint" of representativeness of a measurement site.

- Representativeness: When using the term representativeness we actually mean the definition given by Nappo et al. (1982) which states that point-to-area (volume) representativeness is the probability that a point measurement lies within a certain threshold of the area (volume) average more than 90% of all times.

- Area of representativeness: This term is used by Larssen et al. (1999) to describe the area in which the concentration of interest does not differ by a certain threshold from the concentration observed at a measurement site. This area is not necessarily continuous, but it represents an area with rather small variability.

If a measurement site is representative of an area in the sense of Nappo et al. (1982), it can nevertheless contain large variabilities that cancel out in the area mean. Such an area could then not be considered the area of representativeness. In contrast, a site will be representative in the sense of Nappo et al. (1982) for any sub-area, containing the site itself, of the area of representativeness, assuming threshold values were chosen similarly.

3 Results

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

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 2aa–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}=148km and r_{48}=575km, 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}=43km and r_{48}=179km. 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 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-puszta (HU02) and also Roquetas (ES03) for 24 and 48 h catchment areas). Main land cover types within the catchment areas are given in Table 2aa–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 2aa) the most frequent dominating land cover categories were 16 (Cultivated and managed areas) and 20 (water bodies), followed by the 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-puszta (HU02). For one site the dominating land cover type made up less than 30% of the total land cover
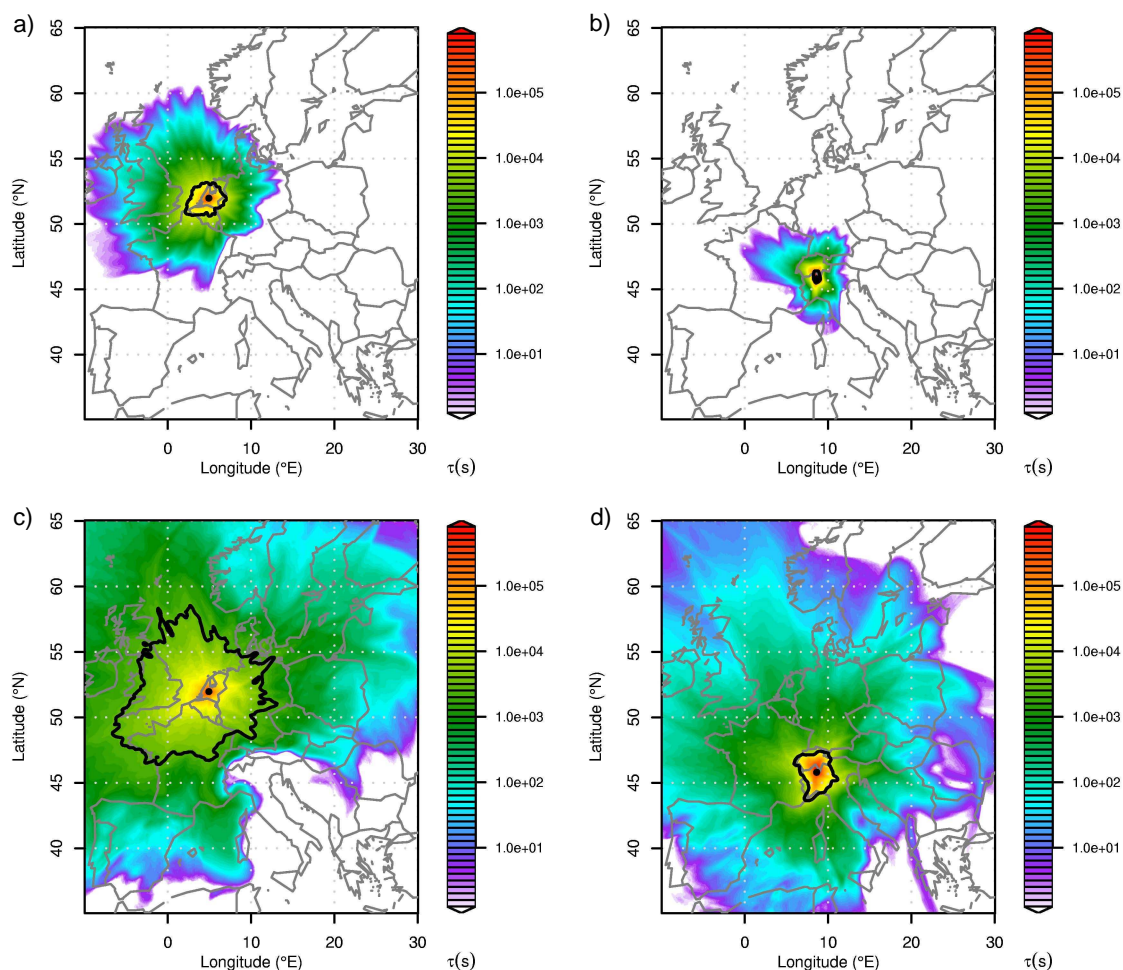
Fig. 1. Total annual surface residence times (footprints) given in units seconds (colour scale) and boundary of catchment area (thick black line) for the sites Cabauw (NL11, **a**, **c**) and Ispra (IT04, **b**, **d**) and two integration intervals, 12 h (a, b) and 48 h (c, d).

(Donon, FR08) indicating heterogeneous conditions. For the 24 and 48 h catchments (Table 2ab–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). From the clustering dendrogram (Fig. 3) it is visible that the subgroups 3 and 4 were split at almost the same height of the cluster tree, indicating that either the selection of 4 or 6 groups is meaningful. With the use of the cluster dendrogram (Fig. 3) we developed category names that are oriented along the observed differences in parameters describing representativeness as observed at each branching in the dendrogram. Starting at the top of the dendrogram the first distinction that is made between sites can clearly be identified as sites *influenced* by surface fluxes and sites with no to weak

surface fluxes, which are commonly called *remote*. The next separation is along the same dimension of surface flux influence and splits the influenced sites into two sub-categories, which can be called *weakly influenced* and *strongly influenced*. The strongly influenced sites are again split according to smaller and larger surface fluxes and we identify these two groups as *rural* and *agglomeration*. Moving from 4 to 5 groups the remote category decomposes into a group with *generally* very low influence of surface fluxes and a group showing intermittent influence of surface fluxes, which thus was called *mostly* remote. Sites in this category are for example the well-established high altitude sites Jungfraujoch, Sonnblick and Pic Du Midi that are known to be characterised by mainly free tropospheric air masses interrupted by transport events from the European atmospheric boundary layer. The last subdivision that yields a total of 6 groups separates sites within the weakly influenced category according to the amount of deposition variability. Therefore, these sub-categories were called *constant deposition* and *variable*
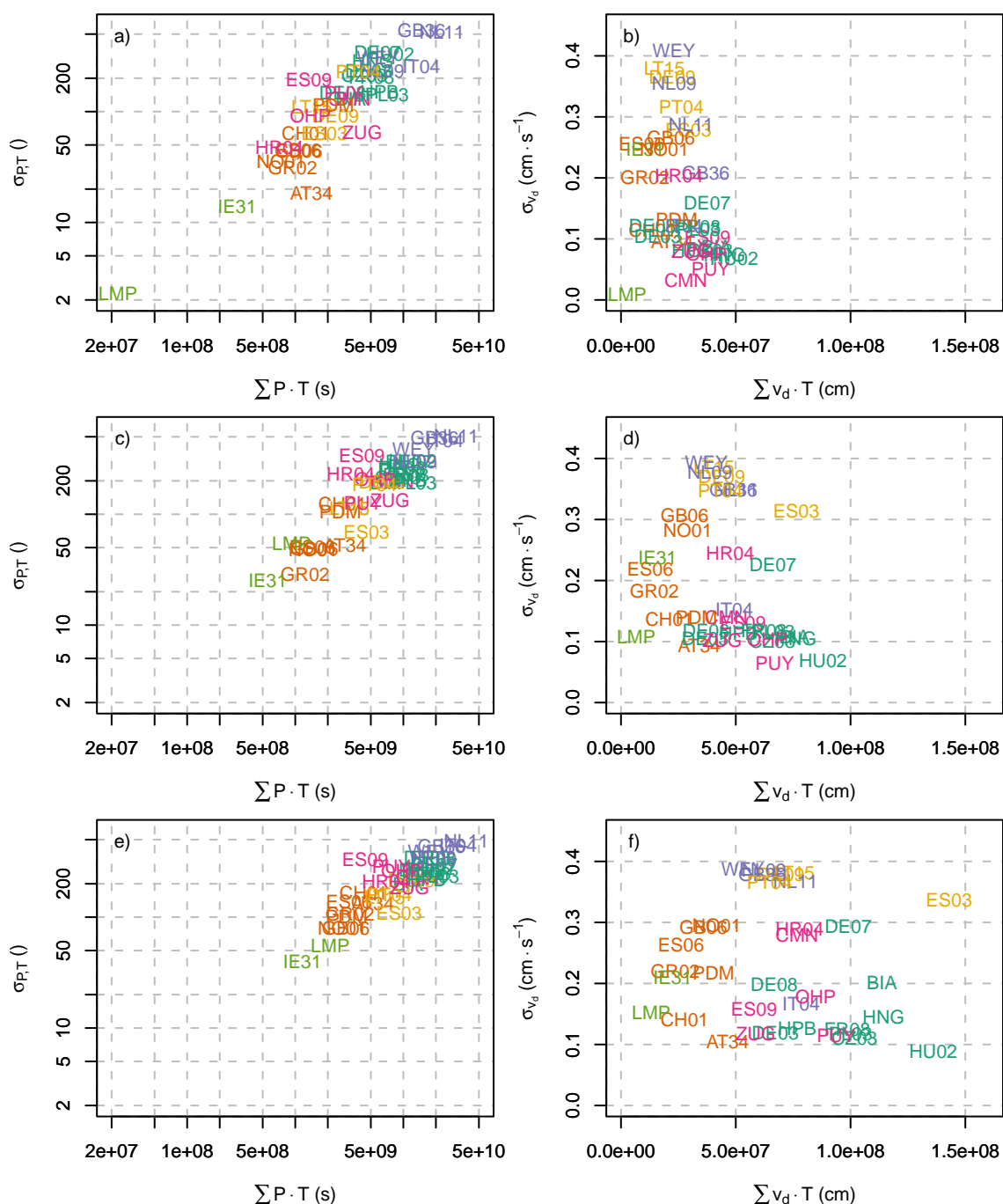
Fig. 2. Scatter plots of population variability $\sigma_{P,T}$ versus population sum $\sum P\ T$ for **(a)** 12 h, **(c)** 24 h, **(e)** 48 h catchment area and deposition variability $\sigma_{v_d,T}$ versus deposition sum $\sum v_d\ T$ for **(b)** 12 h, **(d)** 24 h) **(f)** 48 h catchment area. The colours refer to the categories identified by the site categorisation, compare Fig. 4.

deposition. The presented cluster dendrogram offers the possibility to reduce the 6 categories discussed here to whatever seems most applicable to any user of this categorisation.

Figure 4 identifies the groups on a map of Europe and, together with Fig. 2, allows for a further description of the groups' characteristics.

– The *rural* group contains 10 sites and is characterised 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 in general should be valuable for
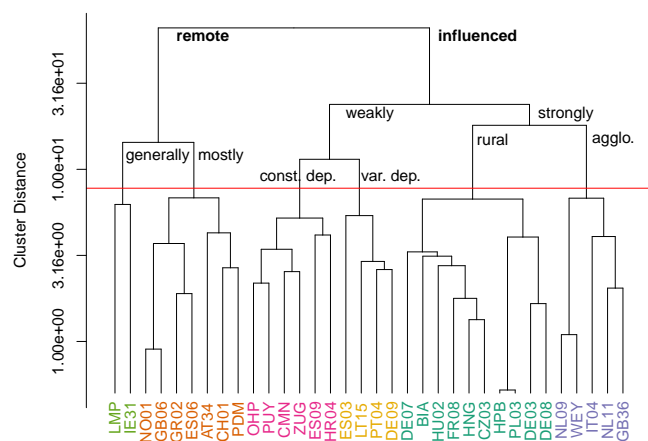
Fig. 3. Dendrogram of cluster analysis of parameters describing representativeness. Note that the y-axis (cluster distance or simply height) is logarithmic.
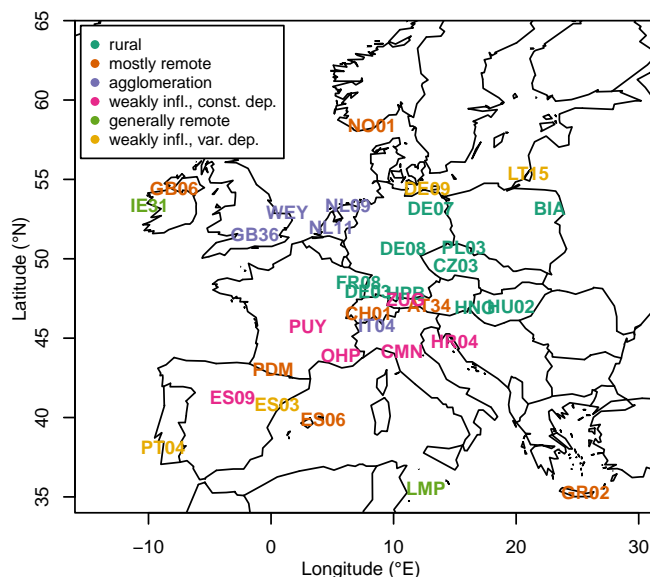
Fig. 4. Map of sites showing categorisation as obtained from clustering of parameters describing representativeness in the catchment areas.

the validation of European scale CTMs and higher resolution satellite observations.

– The *mostly remote* category (7 members) showed small population sums and variability. The total deposition influence was also small while the deposition variability was moderate. The category comprises high altitude and coastal/island sites. While these sites should in general be suitable for comparison with larger scale CTMs and satellite data, care must be taken considering the vertical position of the high altitude sites in comparison to the model topography.

– Total population influence was large for the 5 sites in the *agglomeration* category, however showing large spread. The population variability was large as well and 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 the coastal areas of the Netherlands and south-eastern England. These sites are considered less representative for larger areas and therefore are only suited for comparison with higher resolution CTMs or satellite data.

– The 6 sites in the *weakly influenced, constant deposition* category 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 mostly remote 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. Like the *rural* sites these sites should be suited for validation of European scale CTMs and satellite data. However, additional care needs to be taken for the more elevated sites.

– The two sites Mace Head (IE31) and Lampedusa (LMP) were put into the *generally remote* category that 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). These sites are without further restrictions well suited for validation of larger scale CTMs.

– For the 4 sites in the *weakly influenced, variable deposition* category population sums and variability were moderate. The total deposition influence was moderate, while the deposition variability was large. In general, sites in this category are adequate for European scale CTM validation or satellite comparison, however, due to the large variability in space of the deposition flux the representativeness of these sites might also vary strongly with time depending on the direction of advection.

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 *mostly remote* sites while the third high Alpine observatory at Zugspitze (ZUG) was within group 4 (*weakly influenced, constant deposition*). However, this can be explained by the more central Alpine location and higher elevation of Jungfraujoch (3580 m a.s.l.) and Sonnblick (3106 m a.s.l.) compared to the position and elevation of Zugspitze (summit station) (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 both groups of 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 NO_2 and O_3 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 than 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* category (Puy de Dome (PUY), Observatoire de Haute-Provence (OHP), Monte Cimone (CMN), Campisabalos (ES09), and Zugspitze (ZUG)) and the *mostly remote* 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 sufficient 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 daily mean NO_2, O_3 and CO mixing ratios are plotted against station category in Fig. 5. 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 *mostly remote* and *weakly influenced, constant deposition* (category 2 and 4) showed the smallest mixing ratios, followed by the *rural* (category 1) and *weakly influenced, variable deposition* sites (category 6), while the largest mixing ratios were observed at the *agglomeration* sites (category 3). A one-way analysis of variance (e.g., Dalgaard, 2002) was performed to determine if category means were significantly different from each other. The fraction of explained variance was estimated as the variation within groups divided by total variance. 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 deviations with an even larger fraction of explained inter-site variance (85%). For O_3 the ranking between the sites is contrary to NO_2. Highest O_3 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 *generally remote* (category 5) sites, while lowest O_3 mixing ratios were reported for *weakly influenced, variable deposition* (category 6) and for *agglomeration* (category 3) sites (due to NO titration). A fraction of 55% of the inter-station O_3 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 *agglomeration* (category 3), *weakly influenced, variable deposition* (category 6) and *weakly influenced, constant deposition* (category 4) sites. Smallest variability was observed at the *generally remote* (category 5) and *mostly remote* (category 2) sites. For CO, unfortunately, only 10 observational data sets were available. Relatively low CO values were obtained at the *mostly remote* and *weakly influenced, constant deposition* sites (category 2 and 4). Nevertheless, there was large spread in category 1 and 2 (*rural* and *agglomeration*). 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.
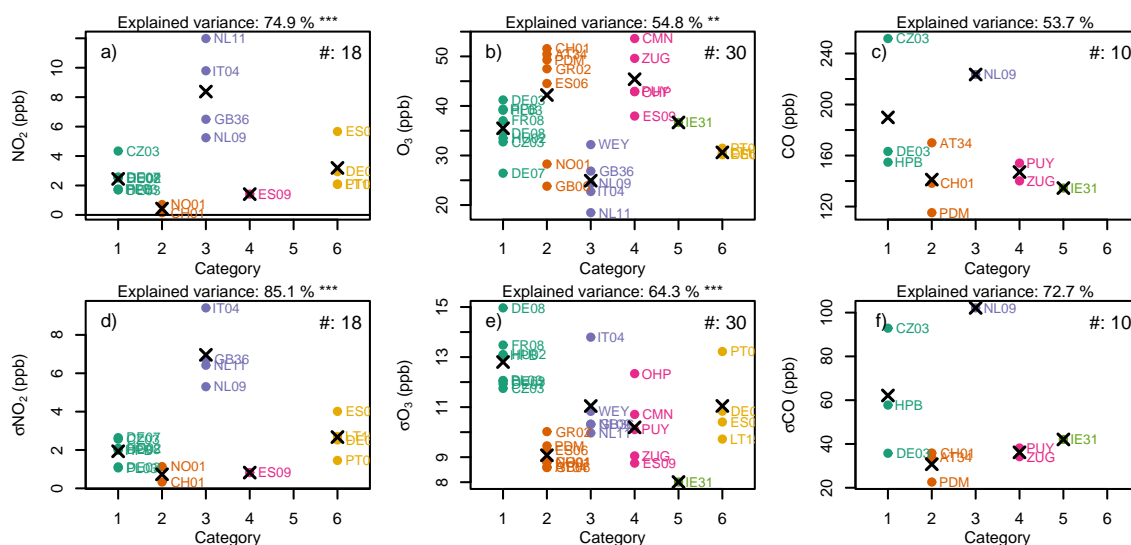
Fig. 5. Sites' median (upper row) and standard deviation (lower row) of observed daily mean mixing ratios of **(a, d)** NO_2, **(b, e)** O_3, **(c, f)** CO by site plotted versus site category. Black crosses represent the category mean. The star notation in each panel represents the confidence level of differences between category means as derived from ANOVA f statistics (*: $\alpha < 0.1$, **: $\alpha < 0.05$, ***: $\alpha < 0.01$).

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 describing 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 consider 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, http://www.atmos-chem-phys.net/10/3561/2010/acp-10-3561-2010-supplement.pdf). These groups were identified as: high altitude, rural, weakly influenced/variable deposition, agglomeration, and remote. Seventeen of the 34 sites ended up in similar groups as ob-

tained by the catchment area approach. Differences are especially apparent for agglomeration 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 NO_2 and O_3 (see Fig. S6 in the supplement, http://www.atmos-chem-phys.net/10/3561/2010/acp-10-3561-2010-supplement.pdf). For CO slightly higher amounts of variability were explained than by the reference categorisations.

A clustering method based solely on parameters describing 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 and at larger distances (see Fig. 1). The same needs to be said about bulk footprints that could be applied to any site. A bulk footprint could be 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 Sensitivity tests

The catchment area was defined with an arbitrary total residence time threshold, f, of 0.5 which describes the fraction of total residence time contained within the catchment volume (see Sect. 2.2.2). To test the robustness of the derived parameters describing 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. 6a, c). However, it is worth noting 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. 6b, d) the dependence on the threshold f was in general smaller and for most sites remained within $\pm 25\%$ of its reference for $f = 0.3 - 0.7$. 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 $f = 0.3 - 0.7$, showing that a station ranking or clustering based on these parameters is relatively insensitive to the selected threshold.

To assess the influence of different atmospheric stability regimes dominating the day- and night-time footprints we estimated catchment areas separately for day- and night-time (09:00, 12:00, 15:00, 18:00 and 21:00, 00:00, 03:00, 06:00 UTC, respectively) simulations. Considerable differences in size and total residence time within the catchment were only observed for the 12 h catchments. Night-time catchment areas were somewhat smaller and total residence times larger for sites in flat terrain as could be expected from generally smaller wind speeds in shallow night-time surface inversions accompanied by little vertical mixing. For the elevated sites the picture was not as conclusive. While some spread was observed between day- and night-time parameters describing representativeness, no clear tendency to smaller or larger values could be estimated for the population parameters and the deposition variability. Total deposition influence within the 12 hour catchment area was increased at night for sites with generally large deposition influence. However, this estimate might be misleading, since we took typical day-time deposition velocities for the calculations, while night-time values are usually much smaller. For 24 and 48 hour catchments the differences in catchment area size total residence time and parameters describing representativeness, were minor.

Our method was not intended to analyse representativeness on the local ($< \sim 1$ km) scale since a) detailed advection is not resolved by the meteorological input for the LPDM calculations and b) the proxy data used have limited resolution as well (1 and 4 km, respectively). Nevertheless, we performed additional FLEXPART calculations for two urban background sites that are close to two of the already selected sites: Munich Lohstrasse (total population 1 400 000, 55 km from Hohenpeissenberg) and Freiburg Mitte (total population 200 000, 10 km from Schauinsland). The same set of parameters describing representativeness was derived for these additional sites and both sites were added to the clustering procedure. While the catchment areas were very similar for both pairs of urban vs. non-urban sites, the parameters describing representativeness differed largely for Munich compared to Hohenpeissenberg but were similar for Freiburg and Schauinsland, though showing slightly larger total burdens and variability for the urban site. When the two additional urban sites were included in the clustering all previous categories remained unaltered. Only the site Munich was put into an additional category, while the site of Freiburg was categorised as "rural", the same as Schauinsland. This finding corroborates the general performance of our categorisation method but also shows its limitations to distinguish between rural and urban sites for medium sized cities like Freiburg on spatial scales smaller than 10 km. Hence, we again emphasize that the method with its current resolution of the underlying LPDMs and emission proxies is not suited for urban sites.

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 describing representativeness 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 the base year 2005 were used. Figure 7 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, respectively, compared to 2005, for the 24 and 48 catchment area. The shape of the catchment areas was similar for different years as also indicated by the catchment's circularity[1] (Fig. 7b). 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

---

[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 $c = 4\pi A/L^2$.
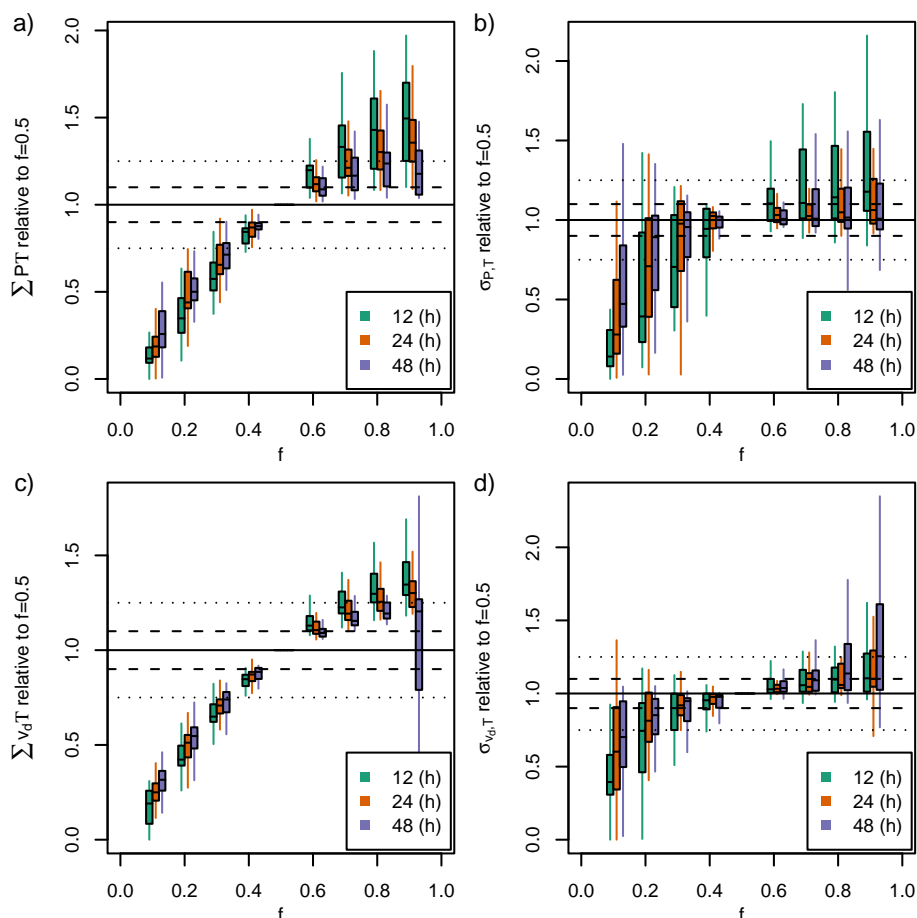
Fig. 6. Boxplots of catchment area parameters for 34 sites as derived for different total residence time thresholds f and 12, 24 and 48 h catchment areas; **(a)** population sum $\sum P\,T$, **(b)** population variability $\sigma_{P,T}$, **(c)** deposition sum $\sum v_d\,T$ and **(d)** deposition variability $\sigma_{v_d,T}$.
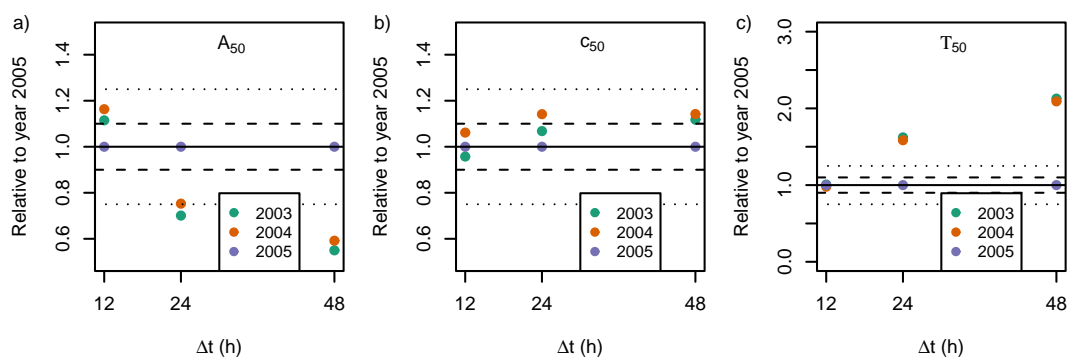

Fig. 7. Annual variability for catchment geometry parameters, **(a)** area A, **(b)** circularity c and **(c)** total annual residence time T as derived for the site Hohenpeissenberg (HPB) and the period 2003–2005.

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 favouring weak diffusion conditions.

Despite the large differences in the catchment area and its total contained residence time, the inter-annual variability in the derived parameters describing representativeness remained in general below 10% (Fig. 8). This can be understood because residence times decrease almost quadratically from the receptor site leading to strongest population and de-
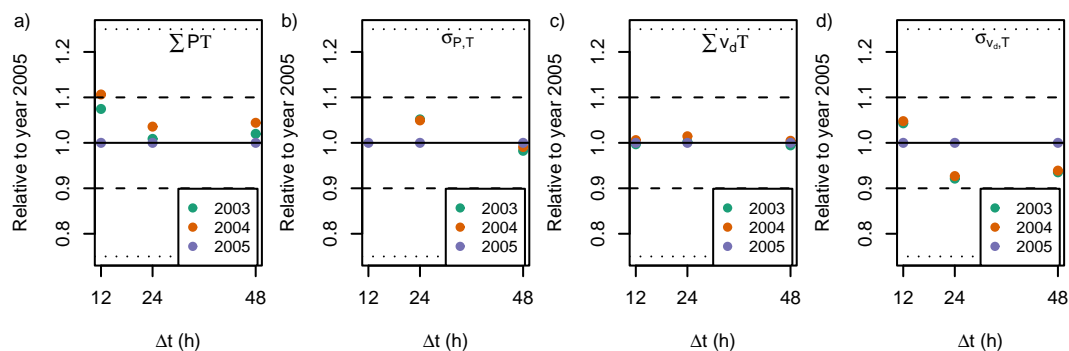
Fig. 8. Annual variability for catchment area parameters **(a)** population sum $\sum P\,T$, **(b)** population variability $\sigma_{P,T}$, **(c)** deposition sum $\sum v_d\,T$ and **(d)** deposition variability $\sigma_{v_d,T}$ as derived for the site Hohenpeissenberg (HPB) and the period 2003–2005.

position 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 to assure similar scales for the parameters of 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 supplementary 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 describing representativeness used for the station categorisation as derived by the two different models are displayed in Fig. 9. While there is generally close agreement between results from both simulations, which is also indicated by Spearman rank correlation coefficients close to or equal to 1 (see figure legend), 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. 9a, c). For $\sum P\,T$ 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 areas, 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 describing 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/NO_2 ratios to verify site categories for Dutch AQ sites. In addition, threshold values for NO/NO_2 ratios were used to categorise all EEA/Airbase sites with available NO and NO_2 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 categorisations of 650 air quality monitoring sites in Germany based on O_3, NO_2, SO_2 and PM_{10} concentrations applying Ward's clustering on median concentrations and daily variance. Using a similar approach, Tarasova et al. (2007) categorized EMEP and GAW O_3 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 substantial variability within the rural subcategories. All three previous 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.

Likewise, Spangl et al. (2007) developed 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 focussed on the local scale, implying constant dilution of the emissions independent of station climatologies. Instead of a clustering approach, category thresholds were
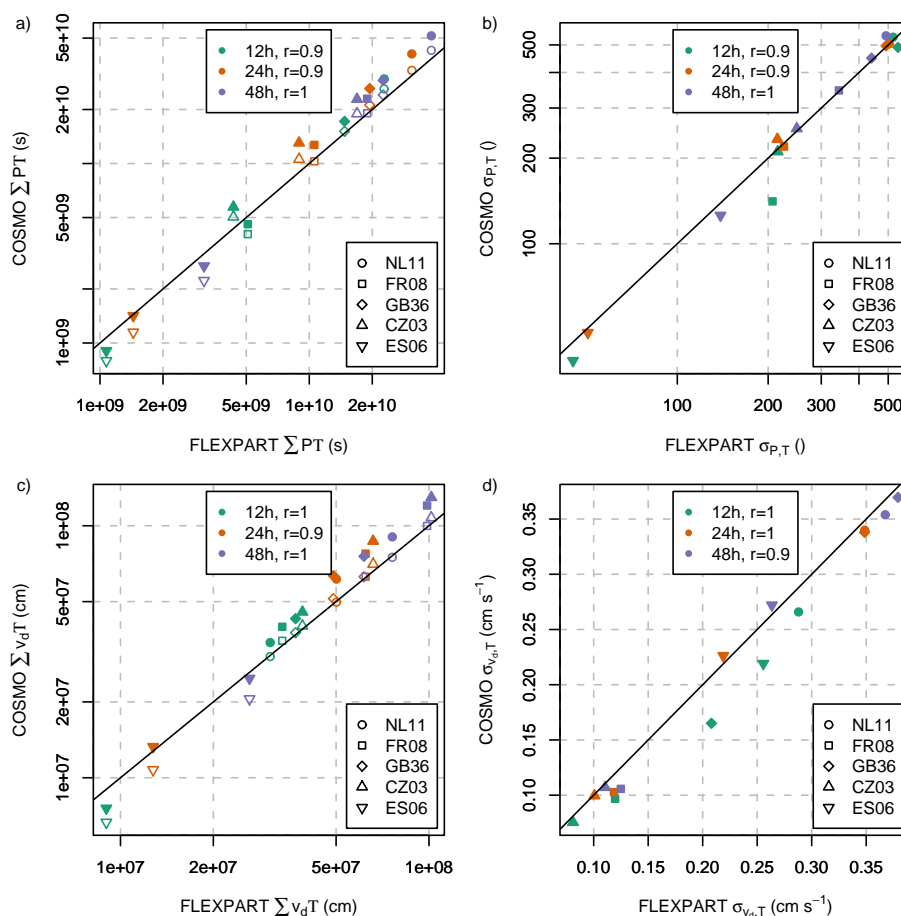
Fig. 9. Catchment area parameters **(a)** population sum $\sum P\,T$, **(b)** population variability $\sigma_{P,T}$, **(c)** deposition sum $\sum v_d\,T$ and **(d)** deposition variability $\sigma_{v_d,T}$ as derived by COSMO LPDM versus those derived by FLEXPART. Solid symbols represent original COSMO LPDM results, open symbols represent parameters derived with scaled COSMO LPDM residence times. r gives the Spearman rank correlation coefficient.

defined based on the distribution of derived parameters describing representativeness. They report good consistency of their categorisation based on local road emissions and average NO_2 concentrations.

5 Conclusions

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 site's catchment area is the 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 Dispersion 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 describing 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 lifetime on the order of 12 h that is mainly influenced by emissions into account, it would be reasonable to sort the available sites by $\sigma_{P,T_{12}}$ and $\sum PT_{12}$ and select only those sites below a certain threshold for inter-comparison. When looking at a species with longer lifetimes $\sigma_{P,T_{48}}$ and $\sum PT_{48}$ might be more suitable for site selection.

Furthermore, the parameters describing 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 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 describing representativeness (e.g. by calculating totals and variability within the catchment areas) and also a different site categorization. Using 6 (or fewer, if merging is preferred) site categories can be of help in any comparison study: categories that are less influenced by surface fluxes would be expected to agree best with model or satellite data. Should this not be the case, it is probably an indication of a specific problem, such as inaccurate surface deposition treatment indicated by disagreement at sites experiencing large deposition fluxes. Therefore, this type of grouped comparison provides an efficient way of double-checking.

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 describing representativeness remained relatively stable. Year-to-year variations in 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 patterns this kind of representativeness analysis needs to be redone on a regular basis to account for changes in 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 computational effort. In contrast, the categorisation based on parameters of the surroundings was less capable of handling sites in more complex terrain and in general explained less of the observed inter-site concentration differences. However, for typical air pollution observatories such as those of the European Airbase network, which does not include remote mountain top and remote coastal sites, such a simplified approach would probably yield reasonable results without taking detailed dispersion simulations into account.

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 straightforward data interpretation. The clustering approach used here, however, has the strength of finding groups of stations in a multi-dimensional space of parameters describing 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 characteristics and number of the detected groups. Alternatively, additional sites can be compared to the current cluster medians and added to the cluster for hat they show smallest distance. Similar studies with a larger set of sites should be performed, so that the groups will become more robust. The parameters describing representativeness presented here can only give a general and temporal average estimate. There is potential to further validate these parameters by independent surface measurements, high resolution model studies or from high-resolution remote sensing data. The categories derived here and in future studies should help select sites that match the representativeness requirements of satellites and models.

References

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.

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), available online at: http://sedac.ciesin.columbia.edu/gpw/, 2005.

Dalgaard, P.: Introductory statistics with R, Springer, New York, USA, 267 pp., 2002.

Derwent, R. G. and Davies, T. J.: Modeling the Impact of Nox or Hydrocarbon Control on Photochemical Ozone in Europe, Atmos. Environ., 28, 2039–2052, 1994.

European Commission – Joint Research Centre: Global Land Cover 2000 database, available online at: http://www-tem.jrc.it/glc2000/, 2003.

European Council: Exchange on Information decision (EoI) on Air, Tech. rep., European Council, 1997.

Flemming, J., Stern, R., and Yamartino, R. J.: A new air quality regime classification scheme for O_3, NO_2, SO_2 and PM_{10} observations sites, Atmos. Environ., 39, 6121–6129, 2005.

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.

Folini, D., Kaufmann, P., Ubl, S., and Henne, S.: Region of infuence of 13 remote European measurement sites based on modeled CO mixing ratios, J. Geophys. Res., D08307, doi: 10.1029/2008JD011125, 2009.

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, UK, third edition edn., 2009.

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.

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 Mt. Kenya in Equatorial Africa, Atmos. Chem. Phys., 8, 3119–3139, 2008, http://www.atmos-chem-phys.net/8/3119/2008/.

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.

Kuhlbusch, T., John, A., Hugo, A., Peters, A., von Klot, S., Cyrys, J., Wichmann, H.-E., Quass, U., and Bruckmann, P.: Analysis and design of local air quality measurements, Tech. rep., European Comission, 2006.

Larssen, S., Sluyter, R., and Helmis, C.: Criteria for EUROAIRNET – The EEA Air Quality Monitoring and Information Network, Tech. rep., European Environment Agency, 1999.

Memmesheimer, M., Ebel, A., and Roemer, M: Budget calculations for ozone and its precursors: Seasonal and episodic features based on model simulations, J. Atmos. Chem., 28(1–3), 283–317, 1997.

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.

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 Wyngaard, J. C.: The Workshop on the Representativeness of Meteorological-Observations, June 1981, Boulder, Colorado, USA, B. Am. Meteorol. Soc., 63, 761–764, 1982.

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.

Schmid, H. P.: Experimental design for flux measurements: matching scales of observations and fluxes, Agr. For. Meteorol., 87, 179–200, 1997.

Schmid, H. P.: Footprint modeling for vegetation atmosphere exchange studies: a review and perspective, Agr. For. Meteorol., 113, 159–183, 2002.

Schr, 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.

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/.

Snel, S.: Improvement of classifications European monitoring stations for AirBase, Tech. rep., ETC/ACC, 2004.

Spangl, W., Schneider, J., Moosmann, L., and Nagl, C.: Representativeness and classification of air quality monitoring stations, Tech. rep., Umweltbundesamt, 2007.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEX-PART version 6.2, Atmos. Chem. Phys., 5, 2461–2474, 2005, http://www.atmos-chem-phys.net/5/2461/2005/.

Tarasova, O. A., Brenninkmeijer, C. A. M., Jockel, P., Zvyagintsev, A. M., and Kuznetsov, G. I.: A climatology of surface ozone in the extra tropics: cluster analysis of observations and model results, Atmos. Chem. Phys., 7, 6099–6117, 2007, http://www.atmos-chem-phys.net/7/6099/2007/.

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.

Ward, J. H.: Hierarchical Grouping to Optimize an Objective Function, J. Am. Stat. Assoc., 58, 236–244, 1963.

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.

Wesely, M. L.: Parameterization of Surface Resistances to Gaseous Dry Deposition in Regional-Scale Numerical-Models, Atmos. Environ., 23, 1293–1304, 1989.

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/.