Source-receptor relationships for airborne measurements of CO2, CO and O3 above Siberia: a cluster-based approach

Paris J.D., Stohl A., Philippe Ciais, Nédélec P., Belan B.D., Arshinov M.y, Michel Ramonet

**Atmospheric
Chemistry
and Physics**

Source-receptor relationships for airborne measurements of CO$_2$, CO and O$_3$ above Siberia: a cluster-based approach

J.-D. Paris[1], **A. Stohl**[2], **P. Ciais**[1], **P. Nédélec**[3], **B. D. Belan**[4], **M. Yu. Arshinov**[4], and **M. Ramonet**[1]

[1]Laboratoire des Sciences du Climat et de l'Environnement/IPSL, CEA-CNRS-UVSQ, UMR1572, Gif sur Yvette, France
[2]Norwegian Institute for Air Research (NILU), Kjeller, Norway
[3]Laboratoire d'Aérologie, Observatoire Midi Pyrénées, CNRS-UPS, Toulouse, France
[4]Zuev Institute of Atmospherics Optics, SB RAS, Tomsk, Russia

Abstract. We analysed results of three intensive aircraft campaigns above Siberia (April and September 2006, August 2007) with a total of ∼70 h of continuous CO$_2$, CO and O$_3$ measurements. The flight route consists of consecutive ascents and descents between Novosibirsk (55° N, 82° E) and Yakutsk (62° N, 129° E). We performed retroplume calculations with the Lagrangian particle dispersion model FLEXPART for many short segments along the flight tracks. To reduce the extremely rich information on source regions provided by the model calculation into a small number of distinct cases, we performed a statistical clustering – to our knowledge for the first time – into potential source regions of the footprint emission sensitivities obtained from the model calculations. This technique not only worked well to separate source region influences but also resulted in clearly distinct tracer concentrations for the various clusters obtained. High CO and O$_3$ concentrations were found associated with agricultural fire plumes originating from Kazakhstan in September 2006. A statistical analysis indicates that summer uptake of CO$_2$ is largely explained (∼50% of variance) by air mass exposure to uptake by Siberian and sub-arctic ecosystems. This resulted in an average 5 to 10 ppm difference with overlaying air masses. Stratosphere-troposphere exchange is found to strongly influence the observed O$_3$ mixing ratios in spring, but not in summer. European emissions contributed to high O$_3$ concentrations above Siberia in April 2006 and August 2007, while emissions from North Eastern China also contributed to higher O$_3$ mixing ratios in summer, but tend to lower mixing ratios in spring, when the airmass aerosol burden is important. In the lower troposphere, large-scale deposition processes in the boreal and sub-arctic boundary layer is a large O$_3$ sink, resulting in a ∼20 ppb difference with overlaying air masses. Lagrangian footprint clustering is very promising and could also be advantageously applied to the interpretation of ground based measurements including calculation of tracers' sources and sinks.

1 Introduction

Atmospheric transport of pollutants can occur both at low altitude in the boundary layer and in the free troposphere, where it can be much faster. Export of North-East Asian and North American emissions is associated with strong uplift in the warm conveyor belts of mid-latitude cyclones (e.g. Cooper et al., 2001; Liang et al., 2004; Owen et al., 2006). On the opposite, European emissions have a tendency to remain in the lower troposphere and are most frequently channelled to the Arctic or to Siberia (Wild et al., 2004; Stohl et al., 2002, 2007a; Duncan and Bey, 2004; Law and Stohl, 2007; Liu et al., 2002). Faster zonal advection occurs in winter and early spring in association with the Siberian High (Newell and Evans, 2000; Wild et al., 2004; Liu et al., 2003). In April, European pollutants can contribute significantly to tracers' concentrations in the Asian outflow to the Pacific (Liu et al., 2003; Liang et al., 2004). With respect to ozone production, European export increase O$_3$ concentrations over Siberia by 2 to 4 ppb in spring and 2 to 6 ppb in summer (Wild et al., 2004; Duncan and Bey, 2004).

Correspondence to: J.-D. Paris
(jean-daniel.paris@lsce.ipsl.fr)

Pollutant export to or across Siberia, however, has been studied almost exclusively through modelling studies and lacks measurement-based assessment. Pochanart et al. (2003) has conducted ground based measurements supported by transport model analysis, showing that European emissions lead to a 1 to 4 ppb increase in O_3 near Baikal Lake. Eneroth et al. (2003), based on backtrajectories analysis of transport to the West-Siberian site of Zotino have shown that stagnant flow conditions were responsible of respiration build-up leading to high CO_2 concentrations (and ruled out European emissions). But measurements of CO_2 and pollutants' distribution are still required over Siberia, along with the possibility to relate them to remote emission sources i.e. source-receptor relationships (SRR).

Contrarily to single back-trajectories, the SRR obtained from a Lagrangian particle dispersion model (LPDM) accounts for the atmospheric turbulence and convection (Stohl et al., 1998). Han et al. (2005) compared the result of regional source apportionment using single backtrajectories and backward dispersion. Their backward dispersion model added individual turbulence-related stochastic component to a large number of HYSPLIT backtrajectories. They demonstrated the better ability of LPDM to identify regional point sources of reactive mercury. Use of LPDM has been proposed (Gerbig et al., 2003; Lin et al., 2003) and eventually demonstrated (Lauvaux et al., 2008) for meso- to regional-scale inversion of CO_2 surface fluxes.

Clustering of backtrajectories has been widely used for the analysis of atmospheric composition measurements at fixed observatories and has been found to be an efficient method for separating air masses with different properties (Moody and Galloway, 1988, Dorling et al., 1992, Sirois and Bottenheim, 1995, Eneroth et al., 2003). Moody and Galloway (1988) were the first to attempt clustering of backtrajectories in order to assess the wet deposition of acid compounds over the Bermuda. The same technique has been used at the continental scale to derive SRRs for O_3 concentrations in central Siberia and supported the conclusion of elevated CO in air masses influenced by European emissions (Pochanart et al., 2003). Traub et al. (2003) applied a simple backtrajectories' partitioning technique to the extensive analysis of measurements obtained from an aircraft. No attempt has been made yet, to our knowledge, to use a clustering technique based on LPDM footprints.

Aircraft measurements only provide a snapshot of the atmosphere at a particular time, but they have the potential to explore transport processes and impact across transported plumes (Takegawa et al., 2004; Fehsenfeld et al., 2006; Stohl et al., 2007b; Methven et al., 2006; Lieleveld et al., 2002) or in the outflow of massively emitting regions (Jacob et al., 2003). Here we investigate data from three recent YAK-AEROSIB (Airborne Extensive Regional Observations in Siberia) intensive campaigns that sampled the Siberian troposphere at different times of the year (Paris et al., 2008). Measurement programmes of different types dedicated to CO_2,

CO, O_3 and other tracers have been implemented in various parts of Siberia using aircraft campaigns equipped with flask sampler (Nakazawa et al., 1997), aircraft vertical profiles (Lloyd et al., 2002; Levin et al., 2002; Ramonet et al., 2002), the Transsiberian railroad (Oberlander et al., 2002) or tall towers (Kozlova et al., 2008).

Three airborne campaigns were made across Siberia in April 2006, September 2006 and August 2007 respectively. Numerous vertical profiles of continuous measurements of CO, CO_2, O_3 and aerosols were collected, extensively sampling the regional troposphere at different times of the year. After describing the data set, we analyse the tracers' concentrations using a clustering technique based on the partitioning of footprints obtained with a LPDM.

More specifically, we focus on the extent to which clusters (groups with similar distributions) of potential emission sensitivity (PES) maps for receptor positions relate to the air masses chemical composition at these receptors. The PES is expressed here as a residence time. The clusters are used to generalize SRR throughout the aircraft trajectory through the Siberian air shed. We try to answer the following questions: (1) can we identify in our data the contribution to CO_2 and CO concentrations of European or other remote anthropogenic sources predicted by models? (2) Is it possible to identify contributions from forest fire or from other types of biomass burning? (3) To what extent can we explain the seasonal variability in CO_2, CO and O_3 concentrations by variations in atmospheric transport patterns through intensive campaigns? (4) Is there a consistent signal about the regional carbon source/sink distribution emerging through model analysis of CO_2 concentrations?

Section 2 describes the campaigns, instruments, model and statistical tools. Section 3 describes the variability of CO_2, CO and O_3 observed above Siberia during the three campaigns. In Sect. 4 we discuss, using cluster analysis, the tracer concentrations observed in selected flights. Section 5 investigates the "seasonal" (inter-campaign) variation of the connection between source regions and trace gas concentrations over Siberia.

2 Data and method

2.1 Campaigns overview and instruments

The airborne platform is an Antonov-30 dubbed "Optik-E" chartered by Tomsk's Institute of Atmospheric Optics. It was equipped in collaboration between French and Russian laboratories for the measurement of CO_2, CO, O_3, aerosols and meteorological parameters. Intensive campaigns took place over Siberia between 11 and 14 April 2006, between 7 and 10 September 2006 and between 17 and 20 August 2007. The flight route consists in a large, continental-scale loop from Novosibirsk in central Siberia to Yakutsk in eastern Siberia (Fig. 1) that is done in four days. Each campaign
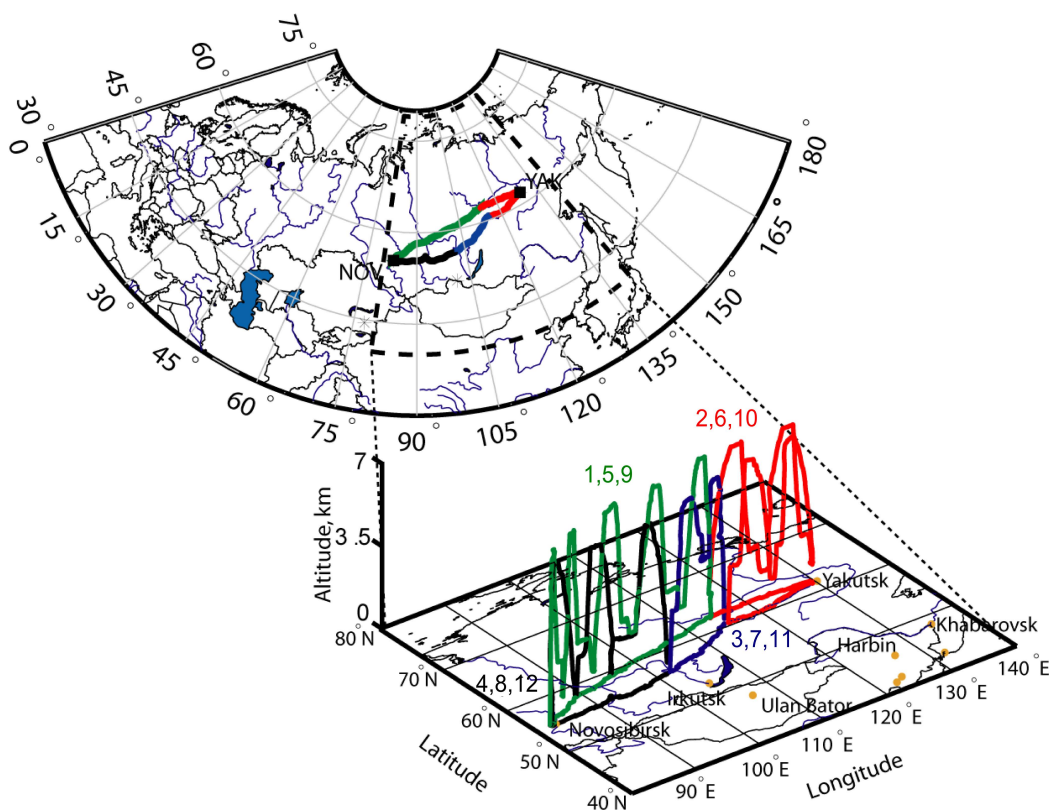
Fig. 1. Horizontal (top panel) and 3-D (bottom panel) flight tracks. Each campaign followed a similar path. Flights are numbered in the bottom panel. The different flights of each campaign (four each, typically one per day) are shown in different colors.

consists of 4 flights (Fig. 1) lasting up to 8 h (limited by refuelling needs and airports availability). These flights are numbered 1 to 4, 5 to 8 and 9 to 12 for the April 2006, September 2006 and August 2007 campaigns respectively. The flight route is repeated with minor modifications for each campaign. Flights were conducted in all weather conditions, usually between 09:00 and 17:00 LT (Paris et al., 2008). Profiles are collected as often as possible as the flight tracks consist mainly of ascents up to 7 km altitude and descents. Thick cloud decks were flown over instead of conducting normal ascent and descent through them.

The April campaign was dominated by stagnant flow conditions and low surface temperatures (Paris et al., 2008), with most of the ground covered by snow. In contrast, the September 2006 and August 2007 campaigns were characterized by westerly flow, warmer and highly variable temperatures, and significant frontal activity in the flight area.

Detailed instrument descriptions can be found in Paris et al. (2008); we give here only a short description. CO_2 is measured by a modified Non Dispersive Infrared Analyser based on a commercial Li-Cor 6262, with accuracy and precision of 0.15 ppm, obtained by periodical in-flight calibration against WMO-referenced reference gases bracketing atmospheric concentrations, and regulation of pressure, temperature and flow in the cells. CO is measured by IR

absorption gas correlation with an accuracy of 5 ppb or 5%. The instrument is based on a commercial infrared absorption correlation gas analyser (Model 48C, TEI Thermo Environment Instruments, USA; Nédélec et al., 2003). O_3 is measured by a modified UV commercial fast response ozone analyser (Thermo Instruments Model 49) with a precision of 2 ppb, 2% for an integration time of 4 s.

2.2 Atmospheric backward transport model

Atmospheric transport was investigated using the FLEXPART v6.2 LPDM. FLEXPART calculates the trajectories of tracer particles using the mean winds interpolated from the analysis fields plus random motions representing turbulence (Stohl and Thomson, 1999). Results presented here use ECMWF analysis fields although both ECMWF and GFS (NOAA/NCEP) fields were used for a prior assessment of transport error (see http://zardoz.nilu.no/~andreas/YAK/). For moist convective transport, FLEXPART uses the scheme of Emanuel and Zivkovic-Rothman (1999), as described and tested by Forster et al. (2007). A backward simulation mode is available, described in more detail by Stohl et al. (2005) and Seibert and Frank (2004). Here the backward method is used to analyse transport pathways from potential flux regions to the receptor position. Each simulation consists of
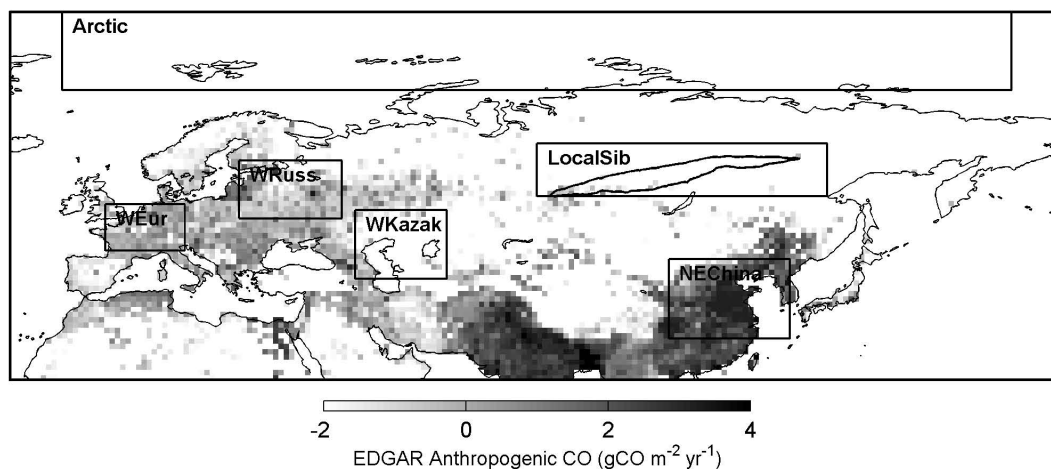
Fig. 2. Regions chosen for data reduction prior to the clustering analysis. The black line shows the fight track during the campaigns. Anthropogenic CO emissions from the EDGAR database are also shown in logarithmic grey scale, in $g(CO)\,m^{-2}\,yr^{-1}$, to illustrate difference between regions.

40 000 particles released whenever the aircraft has moved 0.15° in latitude or longitude, or 10 hPa in altitude. This corresponds to about one minute of sampling or to a layer ∼100 m thick during ascent or descent. Released particles were followed 10 days backward in time. Gridded PESs in three vertical levels (0–300 m, 300–3000 m, and 3000–50 000 m) are stored at 1°×1° resolution every 24 h. The fraction of stratospheric air is recorded as the percentage of all particles originating from the stratosphere (i.e. have resided in the stratosphere any time within the last 10 days).

2.3 Cluster analysis

Airborne campaigns can deliver detailed information on the atmospheric state at a particular time. To generalize this information and to organize the YAK-AEROSIB data set, a method is sought that groups the data according to common transport properties. Cluster analysis is such an exploratory tool, which sorts multivariate data into groups as dissimilar as possible but whose properties are not known a priori. We seek to investigate to which extent footprints can explain the air mass chemical composition in CO_2, CO and O_3. Footprints are implemented as 10-day summed spatial distribution of PES obtained from FLEXPART, complemented by 10-day averaged relative contributions from the stratosphere, defined as the region with potential vorticity >2 PVU. To facilitate analysis we have reduced the number of variables from the gridded PES by further summing these PES over large regions of interest for our study (Fig. 2). A vector

$$\boldsymbol{x}_j = \left[\sum_{d=1}^{10} r_{1,j}, \sum_{d=1}^{10} r_{2,j}, ..., \sum_{d=1}^{10} r_{M,j} \right]^T$$

of daily PES r_k in box m (m=1, 2, ..., M) is associated with each consecutive receptor position j. This time series $\boldsymbol{x}_j$ will constitute our set of j realizations of M variables to cluster. Defining j over a

single campaign improves the separation capability of the algorithm, whereas defining j over the three campaigns allows inter-campaign comparison.

The M boxes are identified a priori as regions with specific sources or sink for the species relevant to this study. Such identification is based either on case studies from the April and September 2006 campaigns (Paris et al., 2008), or on the questions to be addressed with this dataset, such as the transport of European emissions to Siberia. The regions identified are:

- Europe: European pollutant emissions and outflow;

- Western Russia: European Russian pollutant emissions and outflow;

- NE China and Koreas: extratropical cyclones can lift NE Asian pollutant over central or eastern Siberia;

- Central Asia: Central Asia has a different ecosystem and low anthropogenic sources but it is a potential source of biomass burning (e.g. van der Werf et al., 2006);

- Arctic: Arctic air from beyond the Arctic front can be observed over Siberia during cold air outbreaks;

- Local region: covers the area directly flown over by our aircraft;

- Stratosphere: An extra variable is added which averages the 10-day proportion of stratospheric particles.

The boxes' geographical extent is the result of a trade-off between maximizing differences between transport from the various boxes, and an insufficient number of particles if boxes are chosen to be too small. There are large regions

not covered by any box. Theoretically, this could mean that a footprint is located entirely outside all of the boxes. However, the footprints normally cover quite large regions, so they always overlap with one or more of the boxes. By separating the boxes from each other, differences between the various footprints can be maximized by the clustering algorithm. Choosing boxes that are contiguous make the separation more sensitive to horizontal transport error, and associated PES less contrasted.

More remote regions are less connected (few or no particles reaching the region) by transport to the measurements. As a result, remote regions' average footprint distributions are skewed toward zero. To account for this skewness and different region sizes, normalization is applied to the regions' average footprint time series x according to $x_N = (x - p_{0.05})/\bar{x}$ where $p_{0.05}$ is the 5th percentile of x and was found to optimally separate clusters. The robustness of data reduction into regions and data normalization was tested by running the clustering algorithm for varying region sizes, locations and number of regions (bootstrapping), and various normalization functions.

The K-means algorithm implemented in the MATLAB software's Statistics toolbox is used for clustering. K-means is a classic partitioning (non-hierarchical) algorithm (see e.g. Wilks, 2006) which attempts to separate observations into a fixed number of groups. (1) It defines K centroids with initial random position vector of size M=7 (number of regions); then (2) associates each point x_j to the nearest (in Euclidian metrics) cluster centroid, (3) moves the centroid to the centre of the cluster and (4) repeats steps 2 and 3 until convergence is achieved. In the process, any empty cluster is discarded. In order to maintain the number of clusters, a point having maximum distance to its centroid is singled out and declared centroid of a new cluster.

Figure 3 shows the "silhouette" index ($S \in [-1, 1]$), a measure of the separation capability of the clustering, as a function of the chosen number of clusters. The S index is defined as:

$$S = \frac{1}{N} \sum_{i=1}^{N} \frac{d_{\text{nearest}} - d_{\text{centroid}}}{\max(d_{\text{nearest}}, d_{\text{centroid}})} \quad (1)$$

with N number of points, and where point i has a distance to the nearest cluster d_{nearest} and distance to its associated centroid d_{centroid}. The S index informs about how close a point is to its cluster's centroid, relative to the other nearest cluster centroid (Matlab R2006b documentation, Statistics toolbox, http://www.mathworks.com/access/helpdesk/help/toolbox/stats/silhouette.html). Using the S index, the minimum number k of clusters was found to be optimally set to 6 in the global clustering option, and 4 in the campaign by campaign option (Fig. 3).

As the clustering algorithm is sensitive to a priori (random) position of the centroids, the process is repeated 20 times and the best result only (optimizing the separation between centroids) is used in the analysis. This also ensures a
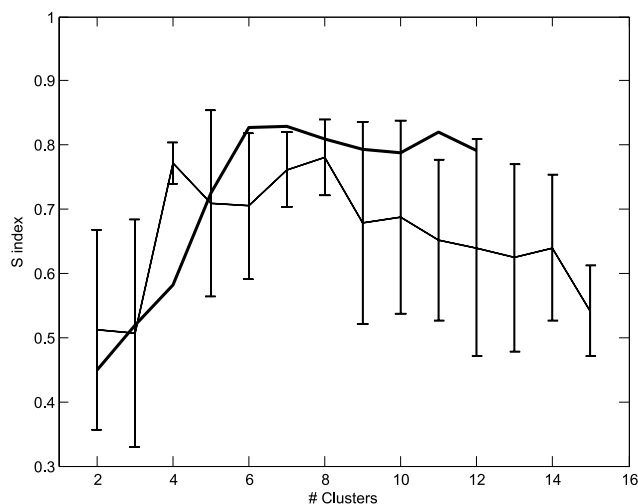

Fig. 3. Silhouette index of the clustering algorithm separation capability as a function of the number of clusters for all data (thick line) and separated campaign by campaign (thin line with error bars showing ±1 std dev).

very high confidence in reproducibility of cluster centroids determination. The normalization and the metrics were also found to have a strong impact on the result. In this respect, cluster analysis is not an objective classification technique as the criteria set for the clustering have to be determined subjectively.

3 Campaign observations

3.1 CO$_2$ measurements

Figure 4a, d and g shows the observed CO$_2$ concentrations averaged for each of the 4 flights of each campaign. Among our surveys in 2006 and 2007, the minimum BL CO$_2$ concentrations have been observed in August 2007 (average 366 ppm below 1 km during Flight 12). While comparing campaigns it should be noted that atmospheric CO$_2$ increases by ~1.4 ppm yr^{-1} due to global anthropogenic emissions. The flight where this CO$_2$ minimum was measured occurred mostly over a forested region, with sparse agricultural landscapes and industrial centres (Kemerovo, Novosibirsk) nearby. This CO$_2$ minimum was coincident with a very high BL top (up to ~3.5 km a.g.l.), as deduced from humidity and CO$_2$ gradients. At the Zotino tall tower, average daytime CO$_2$ mixing ratios measured in the boundary layer were 390 ppm and 382 ppm in April and September 2006, respectively (Kozlova et al., 2008). Kozlova et al.'s September value is significantly higher than the mixing ratio observed in the lower part of our vertical profiles (Fig. 4d; values between 371 and 377 ppm for Flights 5 and 8, the closest to the Zotino site), possibly due to regional variability in sources and sinks of CO$_2$. In the free troposphere, CO$_2$ mixing ratios were
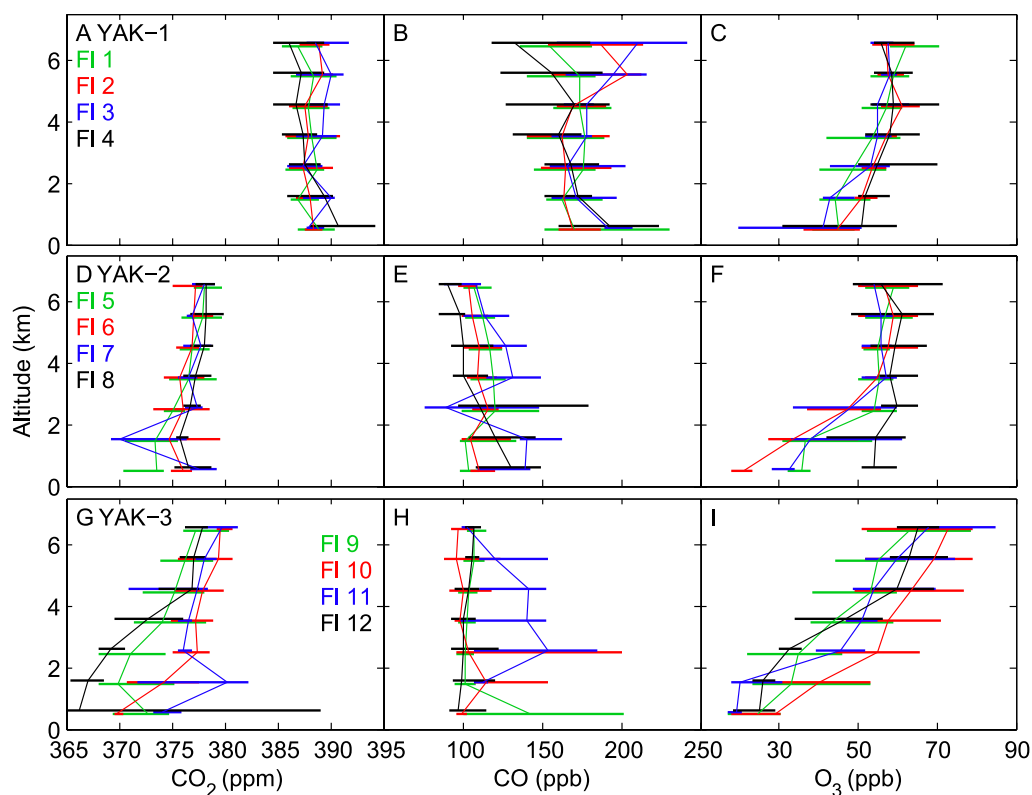
Fig. 4. Profiles of median CO_2, CO and O_3 mixing ratios for each flight over the 3 campaigns (top row: April 2006, middle: September 2006, bottom: August 2007). The horizontal lines join the 10th and 90th percentiles of each altitude level.

between 377 and 379 ppm in August 2007, comparable (up to the interannual trend) to the 377–378 ppm range measured in September 2006 but much lower than in April 2006 (387–388 ppm), reflecting the hemispheric CO_2 seasonal cycle. CO_2 mixing ratios in April 2006 were as high as 392 ppm in polluted filaments encountered between 5 and 6 km altitude (Paris et al., 2008).

3.2 CO and O_3 measurements

Figure 4b shows the average CO profiles for each flight during the April 2006 campaign. Throughout this campaign the average CO mixing ratio was 175 ppb. A number of plumes with higher CO mixing ratios were repeatedly observed. A plume with CO up to 220 ppb was found to be associated to uplift of a polluted air mass from NE China by a warm conveyor belt (Paris et al., 2008). During this campaign, the O_3 profile shows a relatively flat vertical profile (from ∼60 ppb at 6 km to ∼45 ppb at 1 km, Fig. 4c) compared to summer campaigns.

During the September 2006 campaign the average CO concentration was about 100 ppb, with a large variability (Fig. 4e), due to the presence of air masses with very distinct origins. For example during Flight 5 the aircraft crossed a warm front across which strong composition gradients were found, with CO mixing ratios up to 150 ppb immediately

above the front. Except for local emissions in the lower troposphere, the August 2007 campaign has shown low CO mixing ratios (background 100±5 ppb) with little variation, although satellite fire detection has shown numerous fires in temperate and boreal Eurasia. High ozone values (up to 90 ppb) were found near the flight ceiling during the August 2007 campaign (Fig. 4i), and the O_3 vertical gradient was the steepest of the three campaigns.

4 Cluster-based SRR relationships: case studies

In this section we apply the clustering algorithm and discuss the results to examine the measurements for a selected flight from each of the three campaigns (Flight 1 on 11 April 2006, Flight 5 on 7 September 2006 and Flight 9 on 17 August 2007). All three flights are going from Novosibirsk to Myrni (green track in Fig. 1). The three flights took off at 12:00, 09:00 and 09:30 LT, and landed at 21:00, 19:00 and 18:15 LT respectively. The clustering is applied on a single campaign-basis. The clustering being independent from the observed concentrations, statistical separation of the concentrations across different clusters is interpreted as a validation of the SRR.
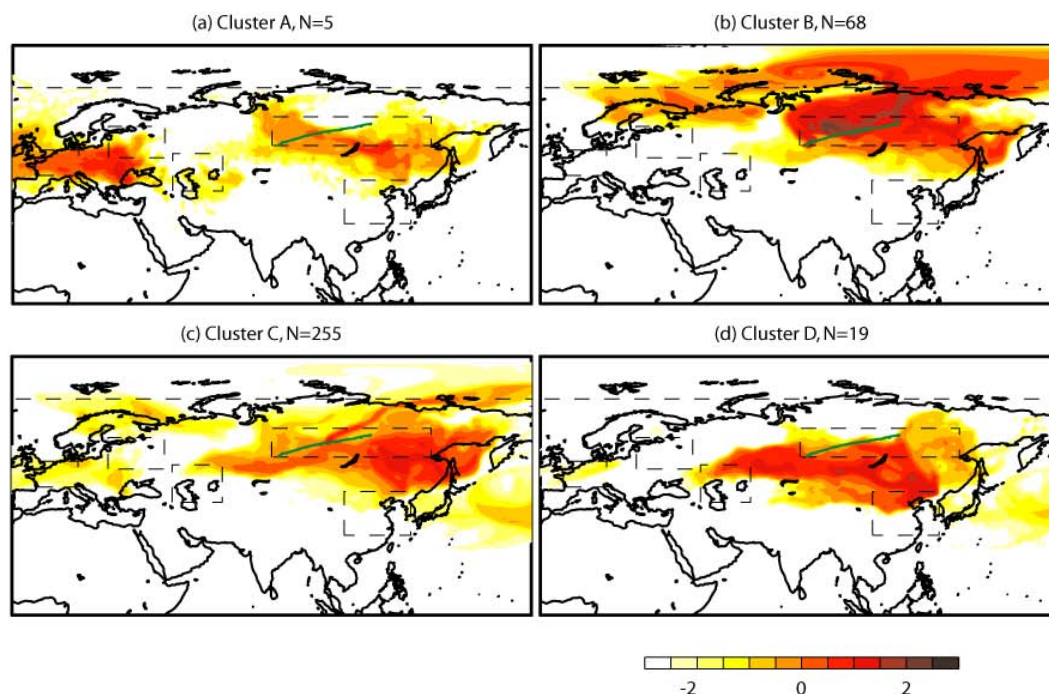
Fig. 5. Average footprint maps for data belonging to clusters A–D throughout Flight 1 (11 April 2006). Logarithmic color scale shows the log 10 of residence time, i.e. the PES below 300 m. The number of elements N in a particular cluster is given above each map.

4.1 Flight 1: pollution from North-eastern China

Figure 5 illustrate the ability of the clustering algorithm to separate transport patterns (and PES) between clusters identified throughout the 11 April 2006 flight. The panels show the average footprint (air mass residence time below 300 m) obtained from the population of backward plumes associated to each of the 4 clusters. Cluster A appears to gather footprints including air transported from the European boundary layer and, secondarily, local influence. Cluster B is related to strong local influence (brown and dark red in Fig. 5b) and to Arctic air masses.

Footprints classified as Cluster C (Fig. 5c) dominate the flight (73%). This cluster is characterized by a strong residence time in the free troposphere, and exhibits the highest stratospheric influence compared to the other clusters (Fig. 6f, green line). Potential emission sensitivity is typically stronger over South East and East Siberia. A marginal number (1%) of footprints are associated to Cluster A (Fig. 5a), where air masses are sensitive to European emissions. Cluster B (20%; Fig. 5b) gathers a significant amount of observations from air masses having resided over northern Siberia and the Arctic. Footprints in Cluster D (5%; Fig. 5d) are exposed to surface exchange in southern Siberia and the Kazakhstan. The four well-differentiated footprint maps illustrate the validity of cluster-based partitioning.

Figure 6 shows the distribution (quartiles and median) of CO_2, CO, O_3 and water vapour measurements associated to each cluster. Associated altitude is given in Fig. 6e, and the four (A–D) clusters' centroid (in std dev-normalized coordinates) position vector is given in Fig. 6f. Each cluster's centroid position vector can be related to the cluster average footprint map in Fig. 5.

The data associated to Cluster D (sensitive to NE China emissions) are encountered between 5 and 6.5 km altitude (Fig. 6e) and have a median CO concentration of 180 ppb (Fig. 6b), well differentiated from other clusters. On the opposite, CO_2 values (median 388.6 ppm; Fig. 6a) are not distinguishably higher. This cluster identifies the transport of North Eastern China emissions to eastern Siberia on 12 April 2006. The ability of the LPDM to resolve thin layers associated to synoptic features is a useful advantage over Eulerian models.

Cluster C is ubiquitous throughout Flight 1 and has a higher stratospheric influence reflected in its high O_3 concentrations (up to 85 ppb). Cluster B has high CO concentrations (median 174 ppb) corresponding to local emission sensitivity of the lower troposphere.

4.2 Flight 5: fires and CO_2 uptake

Figure 7 shows the average footprint for each cluster for the 7 September 2006 flight. As in the previous section, clustering was performed on the whole campaign (here, the September 2006 campaign). Data within Cluster B (43%; Fig. 7b)
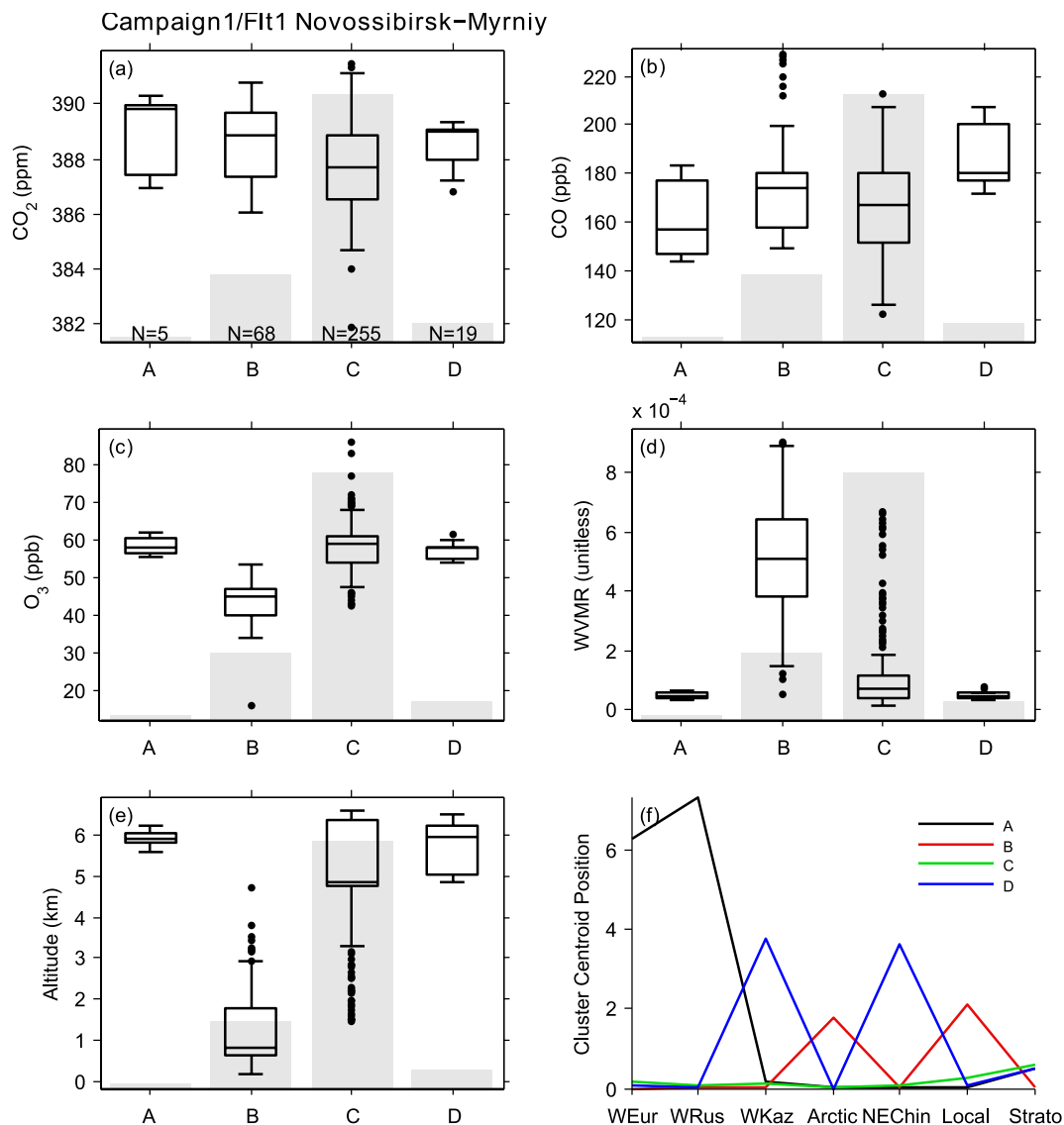
Fig. 6. **(a)** Box plots of median and inter-quartile range (IQR) for CO_2 concentrations measured during Flight 1 (11 April 2006) associated to each cluster (clusters A–D in abscissa, same denomination as in Fig. 5). Outliers within the IQR are included in the whiskers; outliers beyond the IQR are shown as dots. **(b)** Same for CO. **(c)** Same for O_3. **(d)** Same for water vapour mixing ratio. **(e)** Same for altitude. **(f)** Cluster centroid position vector for each cluster. Gray bars indicate the population of each cluster (number of samples).

were mostly associated to FT zonal flow, with limited sensitivity to potential surface emissions over European Russia and the Black Sea area. This cluster has the highest stratospheric signature of the dataset (Fig. 8f). Cluster D is also ubiquitous in this flight (39%; Fig. 7d). It gathers footprint with sensitivity highest over Kazakhstan, between Caspian and Aral Seas. A large number of fires detected by ATSR fire count in this region, probably of agricultural origin, were injected in the FT as the region was swept by a front (Paris et al., 2008). Cluster A (14%; Fig. 7a) is clearly associated to an Arctic air mass channelled southward in the BL. Cluster C (4%; Fig. 7c) reflects advection of air masses in contact with potential European emissions.

The European (Cluster C) air mass has a median CO concentration of 110 ppb (Fig. 8b) at a median altitude of 5 km, to be compared to 105 ppb for the more ubiquitous Cluster B. Although this is only a single case, it provides a case of eastward European pollution export in the mid-troposphere, suggesting that this export does not always follow the low level advection pathway described by modelling studies (Wild et al., 2003; Stohl et al., 2002; Duncan and Bey, 2004). Other characteristic of this air mass include relatively high CO_2 (median 376 ppm) and an O_3 median concentration of 53 ppb.
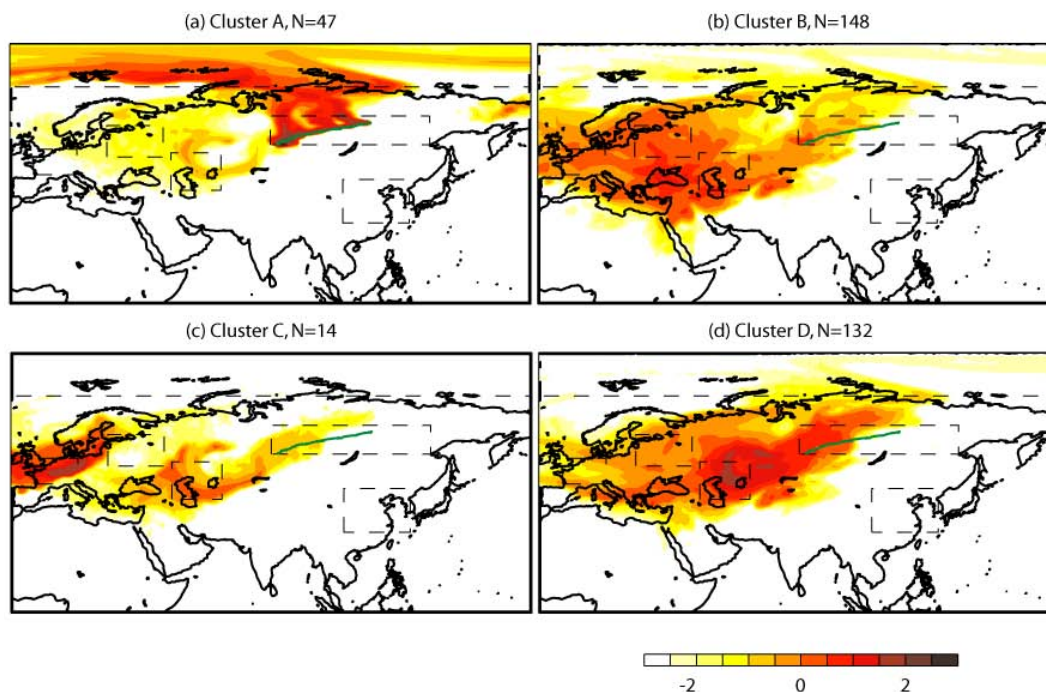
Fig. 7. Same as Fig. 5 for Flight 5 on 7 September 2006.

Figure 8b shows that Cluster D has the most elevated CO concentration, up to 158 ppb (median 121 ppb). It is also associated to high O_3 (median 55 ppb) but CO_2 mixing ratios are not particularly high (375 ppm). High CO and O_3 are likely to reflect the sensitivity of Cluster D (Fig. 7d) to the biomass burning detected in the Caspian region (especially in Northern Kazakhstan at $\sim$50° N, between 50 and 70° E). To illustrate our analysis we focus now on individual biomass burning plumes found through this flight. The plume with the highest CO concentration (144±14 ppb) exhibits a high CO-O_3 correlation (R^2=0.68, regression slope 0.24 mol mol^{-1}). This plume was determined to be a biomass burning plume using FLEXPART and satellite detected fire hot spots, with the fire hot spots located over Kazakhstan (http://zardoz.nilu.no/~andreas/YAK/), and hence is illustrative of Cluster D. The positive regression slope is usually associated to net O_3 production within the plume from combustion-generated precursors like CO (Pfister et al., 2006; Val Martin et al., 2006). In another biomass burning plume the regression slope was found to be negative (-0.10 mol mol^{-1}). Both of these plumes are encountered at 2–2.5 km altitude. The O_3-CO regression slopes are comparable to those found in the literature for boreal forest fire plumes (range 0.05–0.30 mol mol^{-1}; Val Martin et al., 2006; Real et al., 2007 and references therein).

Both of these plumes have a high CO-CO_2 correlation (respectively R^2=0.80 and 0.78) with negative CO-CO_2 regression slopes of -0.06 and -0.09 mol mol^{-1} respectively. These negative slopes reflect competing loss processes such

as CO oxidation within the plume and CO_2 assimilation by intact vegetation, and reflect the difficulty to retrieve fire emission ratio with the airmass integrating sources and sinks upstream of the burning area.

Cluster A gathers observations affected by Arctic air (Fig. 8f) with low CO_2, CO and O_3 concentrations values (median values 373 ppm CO_2, 103 ppb CO and 34 ppb O_3). It is typically encountered at low altitude (1 km) within the BL and consists of moist air (Fig. 8d). Advection of Arctic air in the campaign domain sampled during September 2006 occurred in the boundary layer. FLEXPART results indicate that total PES over the Arctic was 5 days or more in the last 10 days for the majority of data in Cluster A. The low CO and CO_2 concentrations suggest a strong isolation from pollution sources in the air mass history. During the five days of meridian transport from the Arctic, exposure to uptake by high latitude vegetation with permanent daylight period and to surface deposition must also have contributed to low CO_2 and low O_3 values respectively.

4.3 Flight 9: CO_2 uptake, forest fire and stratospheric input

Figure 9 shows the average footprint for each of the 4 clusters for the 14 August 2007 flight. Cluster B indicates the dominant footprint pattern for this flight (65%; Fig. 9b). It has a strong stratospheric component (Fig. 10f) and the weak surface footprint density is spread over Western Siberia. Cluster C (17%; Fig. 9c) has a high sensitivity north of Lake Baikal.
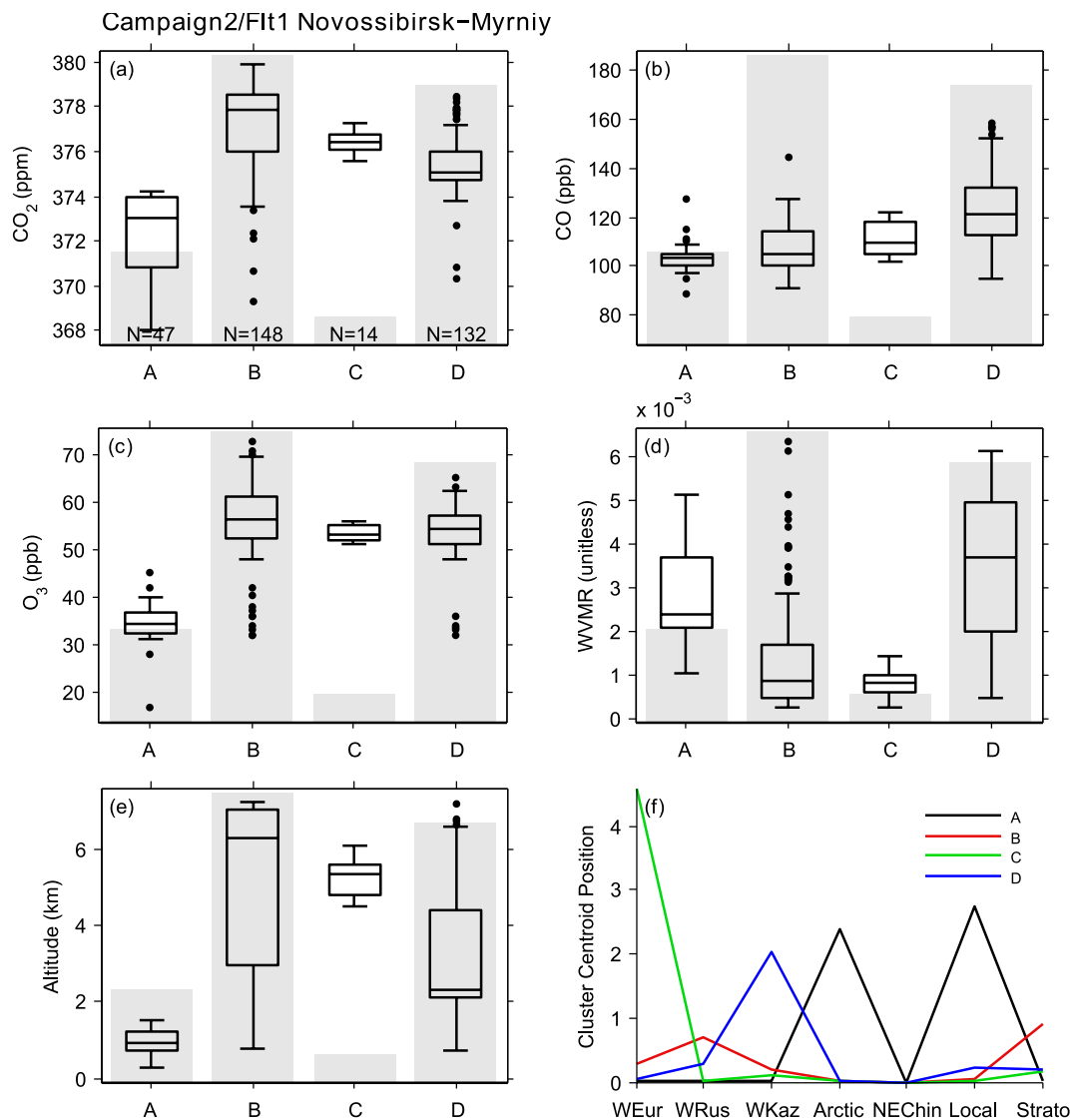
Fig. 8. Same as Fig. 6 for Flight 5 on 7 September 2006.

Cluster A (9%; Fig. 9a) is centred over north European Russia and Scandinavia while Cluster D (9%; Fig. 9d) reflects air of Arctic origin, comparable to Cluster A in the September 2006 case.

Figure 10b–c shows that Cluster C has the highest CO values (median 122 ppb, upper quartile 191 ppb). Such high CO values point to regional fire influence, as seen on ATSR fire atlas for the period of the campaign (not shown). The altitude range is very large (1–6 km). The highest CO concentrations within this cluster (300 ppb) were observed in two distinct layers in the lower FT (see Fig. 11).

Comparing CO_2 concentrations in the low-altitude Arctic clusters of September 2006 (Cluster A) and August 2007 (Cluster D), the latter (median 369 ppm) is lower by $\sim$4 ppm (despite the anthropogenic inter-annual increase of $\sim$1.4 ppm yr^{-1}). The same is found for O_3 with 25 ppb in

August 2007 vs. 34 ppb in September 2006. In Siberia FT CO_2 concentrations are minimum in August (Ramonet et al., 2002), and September concentrations increase by about 2 ppm in the FT and 5 ppm in the BL relative to August.

Cluster B shows the highest stratospheric component for this flight. The highest O_3 values (up to 91 ppb) associated to it are consistent with the "stratospheric" classification. High O_3 concentrations (up to 90 ppb; Fig. 11) are associated to more than 50% of freshly (2 days) exported stratospheric particles.
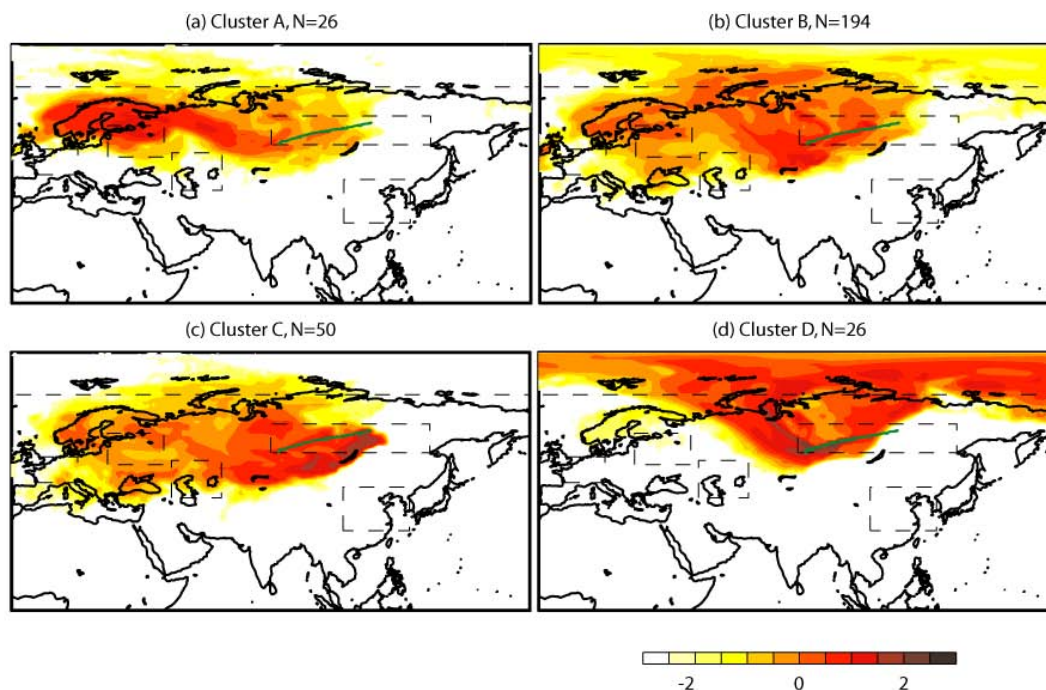
Fig. 9. Same as Fig. 5 for Flight 9 on 17 August 2007.

5 Statistical relationships between source regions and species across campaigns

In order to compare between campaigns and seasons, and to discuss the cluster analysis we now resort to pair-wise correlations statistics (correlations and regression slopes) directly associating observed trace gas mixing ratios and individual retroplume's residence times over each region (Fig. 2). Pearson correlation coefficients were calculated after excluding data associated to residence time <100 s. This data selection rejects less than 1% of the data range. We report also the Kendall τ parameter which is a "rank" correlation coefficient and therefore describes better increasing or decreasing but non-linear relationships (Wilks, 2006).

5.1 Variations of CO_2 associated to source regions

Figure 12 shows the influence that airmasses' residence time in the lowest 300 m in the Arctic and Siberian regions exerts on CO_2 measurements. Since CO_2 fluxes are highly variable at the sub-day timescale, whereas simulated exposure to fluxes is averaged over 10 days, the relations between CO_2 concentrations and PES over source or sink regions are not expected to be strongly linear, making the Kendall τ a useful complement to the Pearson correlation coefficient to describe the relationship between residence time and mixing ratios. Linear regression slopes and correlation coefficients are reported in Table 1. CO_2 mixing ratios are significantly ($p < 0.01$) anti-correlated with PES

over local ecosystems both in September 2006 ($R=-0.20$, $\tau=-0.37$, Table 1) and August 2007 ($R=-0.41$, $\tau=-0.43$). The regression slope is steeper in August than in September (-3.9×10^{-4} ppm s^{-1} and -1.0×10^{-4} ppm s^{-1} respectively), documenting a stronger uptake in August. Similarly, for the Arctic region in summer, significant correlation is observed between CO_2 concentrations and the PES, also with stronger correlations in August 2007 than in September 2006 (Fig. 12). Therefore, the more air travelled through the Arctic region before reaching our aircraft, the more it has been depleted in CO_2 by sub-arctic ecosystems. As discussed in Sect. 4, air having resided in the Arctic surface layer and eventually reaching our aircraft is typically associated to advection in the BL, and therefore integrates between 1 and 10 days of exposure to uptake by Siberian sub-arctic ecosystems. Together, these correlation coefficients indicate that the 10-day cumulated exposure to local and Arctic fluxes in August 2007 explains ~50% of the total CO_2 variance, when boreal and sub-arctic ecosystem uptake is at its annual maximum.

In August 2007 North Eastern China is an influent source of CO_2, as denoted by the strong positive correlation ($R=0.54$, slope$=5.0\times10^{-2}$ ppm s^{-1}; Table 1). We note that the clustering analysis did not identify a cluster with dominant NE China signal (Sect. 4.3., Fig. 10). The strong correlation combined with this absence of a distinct cluster suggests that NE Chinese emissions enhanced CO_2 concentrations over Siberia in an ubiquitous and diffuse way. On the opposite, in April 2006 both a highly polluted cluster of
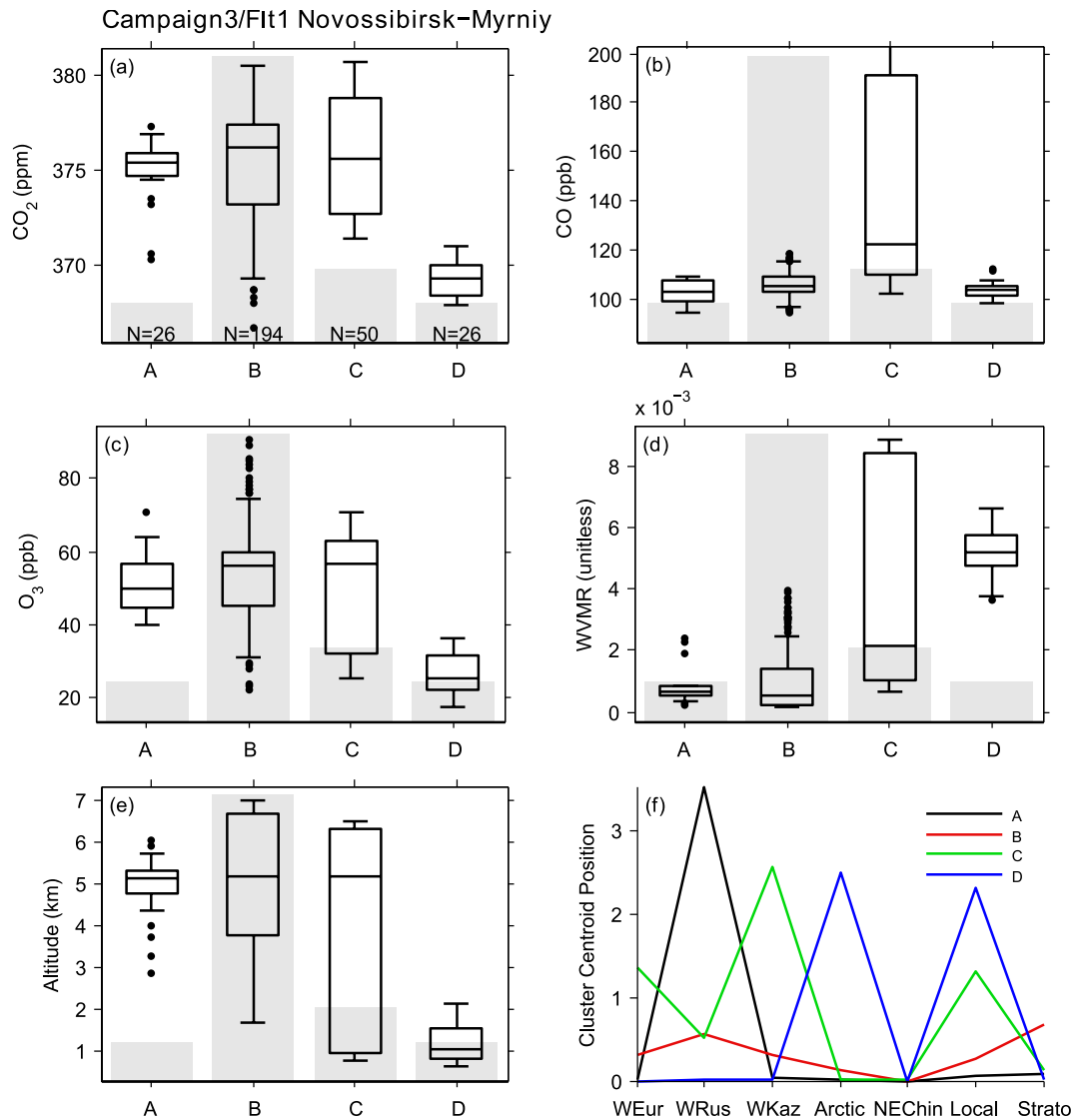
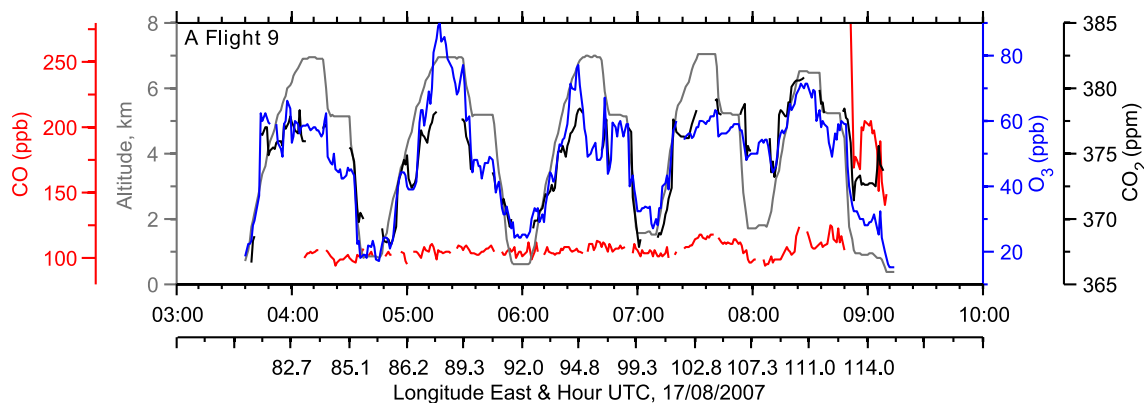**Fig. 10.** Same as Fig. 6 for Flight 9 on 17 August 2007.

Table 1. Regional PES-concentration pairwise statistics for CO_2.

	April 2006			September 2006			August 2007		
	τ^a	R^b	s^c	τ^a	R^b	s^c	τ^a	R^b	s^c
W Europe	(−0.07)	(0.03)	–	0.31	(0.03)	–	0.42	(0.19)	–
W Russia	(0.05)	−0.29	−3.7e-4	0.15	0.11	4.9e-4	0.11	−0.12	−1.9e-3
W Kazak	0.13	(−0.04)	–	(0.01)	−0.15	−2.1e-4	0.25	−0.13	−3.0e-3
Arctic	(0.06)	−0.25	−1.3e-4	−0.29	−0.35	−5.0e-4	−0.40	−0.37	−4.7e-4
NE China	0.31	(−0.01)	–	(0.00)	(0.68)	–	0.14	0.54	5.0e-2
Local	0.40	0.14	5.0e-5	−0.37	−0.20	−1.0e-4	−0.43	−0.41	−3.9e-4
Strato	(−0.01)	−0.35	−4.0e-2	0.37	(0.00)	–	0.37	(−0.21)	−1.4e-2

[a] Kendall's tau. Values are given between parentheses if not statistically robust ($p > 0.01$)

[b] Pearson correlation coefficient. Values are given between parentheses if not statistically robust ($p > 0.01$)

[c] Robust linear regression slope ($ppm\,s^{-1}$). No value is given if both Pearson and Kendall correlations are not robust, or if only Pearson correlation is not robust with $|R| < 0.2$.

Fig. 11. Time series of measured CO_2 (black), CO (red) and O_3 (blue) concentrations as well as altitude (grey) during Flight 9 on 17 August 2007. Aircraft longitude is also given every half hour.

Table 2. Regional PES-concentration pair-wise statistics for CO.

	April 2006			September 2006			August 2007		
	τ[a]	R[b]	s[c]	τ[a]	R[b]	s[c]	τ[a]	R[b]	s[c]
W Europe	−0.08	(−0.04)	–	(0.02)	(−0.01)	–	0.18	−0.28	−7.9e-02
W Russia	(0.03)	−0.25	−5.3e-03	0.09	(0.02)	–	(0.03)	(−0.07)	–
W Kazak	0.17	(−0.11)	–	0.29	0.21	3.6e-03	0.26	0.17	3.6e-03
Arctic	(0.03)	(0.04)	–	0.15	−0.18	−4.1e-03	−0.06	−0.13	–
NE China	0.32	(0.00)	–	(−0.03)	(−0.86)	–	0.17	−0.22	−1.7e-01
Local	0.29	(0.01)	–	0.27	(0.06)	–	0.14	0.36	−3.8e-05
Strato	(−0.04)	−0.36	−6.3e-01	−0.15	(−0.42)	−1.1e+00	(0.03)	0.26	–

[a] Kendall's tau. Values are given between parentheses if not statistically robust ($p > 0.01$)
[b] Pearson correlation coefficient. Values are given between parentheses if not statistically robust ($p > 0.01$)
[c] Robust linear regression slope ($ppm\,s^{-1}$). No value is given if both Pearson and Kendall correlations are not robust, or if only Pearson correlation is not robust with $|R| < 0.2$.

dominant NE China footprint, and a consistent positive relationship ($\tau = 0.31$) between CO_2 and NE China PES were found, indicating that spring contribution from NE China to CO_2 over Siberia occurred in specific plumes through specific transport event, rather than contributing to the background concentration.

5.2 Variations of CO and O_3 associated to source regions

Tables 2 and 3 show the pair-wise statistics associating potential source regions and CO and O_3 concentrations respectively. The CO burden over Siberia is partly due to emissions from remote anthropogenic sources such as in the NE China region in April 2006 ($\tau = 0.32$; Table 2), where uplift of polluted air and advection toward Siberia was documented. Western Kazakhstan sources in September 2006 ($\tau = 0.29$) and August 2007 ($\tau = 0.26$) also contributed to the CO burden. In the case of Kazakhstan CO emissions are most likely due to biomass burning occurring in these regions. More local sources also have an impact through all campaigns.

The absence of a strong correlation between O_3 mixing ratio and airmass residence time over biomass-burning-affected areas of Kazakhstan does not support a strong average O_3 production during advection (Table 3), contrarily to our observations on CO-O_3 correlation in individual plumes (Sect. 4.2).

O_3 statistics exhibit significant correlations with the simulated stratosphere influence in April 2006. This strong correlation ($R = 0.43$) suggests that stratosphere-troposphere exchange is a strong source term to the free tropospheric O_3 spring budget. However, stratosphere Pearson correlation coefficients are weaker and not significant in the summer campaigns, indicating that other processes dominate summer tropospheric O_3 budget, such as local photochemistry and surface deposition.

Exposure to surface influence in the Siberian BL consistently reduces O_3 concentrations in all campaigns, with patterns similar to that of CO_2. This effect, attributed to deposition on forests, is stronger in summer months (Table 3). With positive correlations, European emissions (in Western

Table 3. Regional PES-concentration pair-wise statistics for O_3.

	April 2006			September 2006			August 2007		
	τ^a	R^b	s^c	τ^a	R^b	s^c	τ^a	R^b	s^c
W Europe	0.28	0.35	8.5e-03	0.31	(0.01)	–	0.39	0.34	1.2e-01
W Russia	0.18	0.24	1.1e-03	0.07	(−0.06)	–	0.13	−0.10	−8.2e-03
W Kazak	0.12	−0.20	−3.7e-04	0.07	(0.06)	–	0.17	(−0.05)	–
Arctic	−0.44	−0.36	−6.3e-04	−0.33	−0.47	−3.8e-03	−0.45	−0.45	−1.8e-03
NE China	0.08	−0.36	−1.7e-03	−0.10	(0.15)	–	0.16	0.44	2.1e-01
Local	−0.47	−0.53	−9.8e-04	−0.44	−0.76	−2.5e-03	−0.50	−0.69	−2.3e-03
Strato	0.45	0.43	1.0e-01	0.41	(0.30)	5.8e-01	0.52	(−0.02)	–

[a] Kendall's tau. Values are given between parentheses if not statistically robust ($p > 0.01$)

[b] Pearson correlation coefficient. Values are given between parentheses if not statistically robust ($p > 0.01$)

[c] Robust linear regression slope ($ppm\,s^{-1}$). No value is given if both Pearson and Kendall correlations are not robust, or if only Pearson correlation is not robust with $|R| < 0.2$.
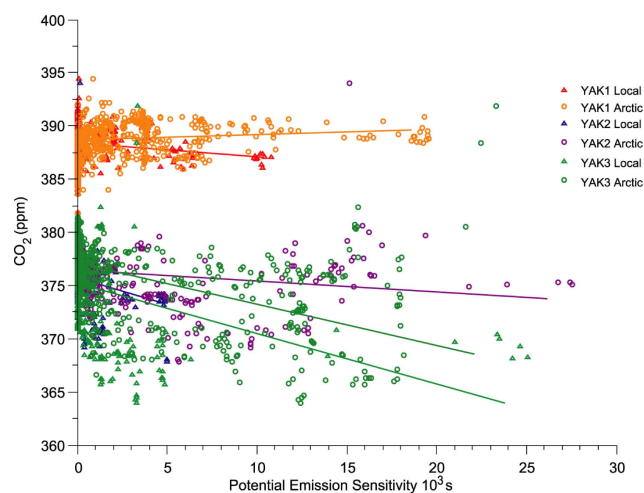
Fig. 12. Scatter plot of CO_2 measurements against PES in the Arctic (circles) and Siberia (triangles) regions, with robust linear regression fit for each of the two regions and each of the three campaigns. YAK-1: April 2006 (red and orange for Siberia and Arctic, respectively), YAK-2: September 2006 (blue and purple), YAK-3: August 2007 (light green and green).

Europe and European Russia) seem to be a source of O_3 over Siberia in all seasons, confirming the cluster studies of Sect. 4. This influence is significant for Western Europe in April 2006 ($R=0.35$) and August 2007 ($R=0.34$), and for Western Russia in April 2006 only ($R=0.24$). Our results indicate that emissions in North Eastern China also play an important role in the O_3 budget. In August 2007 they contribute positively to O_3 mixing ratios ($R=0.44$), whereas this influence is negative in April 2006 ($R=-0.36$). This negative spring influence might be due to heterogeneous chemistry in the presence of heavy aerosol load in the air mass following the dust storm that occurred in the region prior to and during the campaign (Paris et al., 2008).

6 Conclusion

We analysed three intensive campaigns above Siberia resulting in a total of ∼70 h of CO_2, CO and O_3 measurements. The sampling strategy was to perform consecutive ascents and descents to obtain a tomography of the atmosphere above Siberia. Our data analysis used a novel technique based on the clustering of LPDM footprints and correlation analysis. The clustering technique was found to be able to separate tracers' concentrations although it was based solely on atmospheric transport modelling. CO concentrations of ∼109 ppb were observed in clusters associated to a dominant European emissions sensitivity in September 2006 and ∼105 ppb in August 2007. Dominant NE China PES was associated to concentrations of 180 ppb in April 2006. High CO and O_3 concentrations (median values 121 ppb and 55 ppb respectively) were found in clusters associated with fires in Kazakhstan in September 2006. High correlation ($R^2=0.68$) and robust linear relationships with regression slopes ranging between −0.10 and 0.24 $ppb\,ppb^{-1}$ where found in individual plumes.

Summer (August 2007, September 2006) tropospheric variability of CO_2 was found to be largely explained (∼50%) by uptake due to exposure to boreal and sub-arctic ecosystems. It is likely that permanent sunlight exposure of vegetation above the Arctic Circle contributes to a more constant uptake in this region, therefore strongly depleting CO_2 in slowly transiting air masses.

A discernable influence of air exposure to European emissions was observed, associated with higher O_3 concentrations. O_3 was also affected by large sink processes in the Siberian BL and constant stratospheric contribution. North Eastern Chinese emissions contribute to higher O_3 mixing ratio in summer, but tend to decrease them in spring.

Airborne measurements across Siberia during three YAK campaigns have been capable to detect concentrations originating from sources located in remote regions, thus

validating the sampling strategy. The cluster-based method proposed here was successful at separating air masses with different chemical compositions, although the classification scheme is based only on simulated transport properties and eventually matched independently against air chemical composition measurements. Another advantage of our clustering method is that it is independent from our limited knowledge of sources and sinks, especially for the case of Siberia. Single-flight cluster analysis exhibited a variety of air masses origin, related to a large variability in CO_2, CO and O_3. The technique used here, although it has sufficient resolving power, would benefit from more refinements. The cluster-based method should be able to ensure comparability between different campaigns at different locations and/or periods (given no strong concentrations trend between campaigns). Another possible application of this technique to surface observatories or regular aircraft profiles is to contribute to the definition of background conditions and to provide a systematic classification of air masses origin.

Edited by: M. Heimann

The publication of this article is financed by CNRS-INSU.

References

Cooper, O. R., Moody, J. L., Parrish, D. D., Trainer, M., Ryerson, T. B., Holloway, J. S., Hubler, G., Fehsenfeld, F, C., Oltmans, S. J., and Evans, M. J.: Trace gas signatures of the airstreams within North Atlantic cyclones: Case studies from the North Atlantic Regional Experiment (NARE '97) aircraft intensive, J. Geophys. Res., 106(D6), 5437–5456, 2001.

Dorling, S. R., Davies, T. D., and Pierce, C. E.: Cluster-Analysis – a Technique for Estimating the Synoptic Meteorological Controls on Air and Precipitation Chemistry – Method and Applications, Atmos. Environ., 26(14), 2575–2581, 1992.

Duncan, B. N. and Bey, I.: A modeling study of the export pathways of pollution from Europe: Seasonal and interannual variations (1987–1997), J. Geophys. Res., 109(D8), D08301, doi:10.1029/2003JD004079, 2004.

Emanuel, K. A. and Zivkovic-Rothman, M.: Development and evaluation of a convection scheme for use in climate models, J. Atmos. Sci., 56(11), 1766–1782, 1999.

Eneroth, K., Kjellstrom, E., and Holmen, K.: Interannual and seasonal variations in transport to a measuring site in western Siberia and their impact on the observed atmospheric CO_2 mixing ratio, J. Geophys. Res., 108(D21), 4660, doi:10.1029/2002JD002730, 2003.

Fehsenfeld, F. C., Ancellet, G., Bates, T. S., Goldstein, A. H., Hardesty, R. M., Honrath, R., Law, K. S., Lewis, A. C., Leaitch, R., McKeen, S., Meagher, J., Parrish, D. D., Pszenny, A. A. P., Russell, P. B., Schlager, H., Seinfeld, J., Talbot, R., and Zbinden, R.: International Consortium for Atmospheric Research on Transport and Transformation (ICARTT): North America to Europe – Overview of the 2004 summer field study, J. Geophys. Res., 111, D23S01, doi:10.1029/2006JD007829, 2006.

Forster, C., Stohl, A., and Seibert, P.: Parameterization of convective transport in a Lagrangian particle dispersion model and its evaluation, J. Appl. Meteorol. Clim., 46(4), 403–422, 2007

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO_2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework, J. Geophys. Res., 108(D24), 4757, doi:10.1029/2003JD003770, 2003.

Han, Y. J., Holsen, T. A., Hopke, P. K., and Yi, S. M.: Comparison between back-trajectory based modeling and Lagrangian backward dispersion modeling for locating sources of reactive gaseous mercury, Environ. Sci. Technol., 39(10), 3887–3887, 2005.

Jacob, D. J., Crawford, J. H., Kleb, M. M., Connors, V. S., Bendura, R. J., Raper, J. L., Sachse, G. W., Gille, J. C., Emmons, L., and Heald, C. L.: Transport and Chemical Evolution over the Pacific (TRACE-P) aircraft mission: Design, execution, and first results, J. Geophys. Res., 108(D20), 1–19, 2003.

Kozlova, E. A., Manning, A. C., Kisilyakhov, Y., Seifert, T., and Heimann, M.: Seasonal, synoptic, and diurnal-scale variability of biogeochemical trace gases and O_2 from a 300-m tall tower in central Siberia, Global Biogeochem. Cy., 22, GB4020, doi:10.1029/2008GB003209, 2008.

Lauvaux, T., Uliasz, M., Sarrat, C., Chevallier, F., Bousquet, P., Lac, C., Davis, K. J., Ciais, P., Denning, A. S., and Rayner, P. J.: Mesoscale inversion: first results from the CERES campaign with synthetic data, Atmos. Chem. Phys., 8, 3459–3471, 2008, http://www.atmos-chem-phys.net/8/3459/2008/.

Law, K. S. and Stohl, A.: Arctic air pollution: Origins and impacts, Science, 315(5818), 1537–1540, 2007.

Lelieveld, J., Berresheim, H., Borrmann, S., Crutzen, P. J., Dentener, F. J., Fischer, H., Feichter, J., Flatau, P. J., Heland, J., Holzinger, R., Korrmann, R., Lawrence, M. G., Levin, Z., Markowicz, K. M., Mihalopoulos, N., Minikin, A., Ramanathan, V., de Reus, M., Roelofs, G. J., Scheeren, H. A., Sciare, J., Schlager, H., Schultz, M., Siegmund, P., Steil, B., Stephanou, E. G., Stier, P., Traub, M., Warneke, C., Williams, J., and Ziereis, H.: Global air pollution crossroads over the Mediterranean, Science, 298(5594), 794–799, 2002.

Levin, I., Ciais, P., Langenfelds, R., Schmidt, M., Ramonet, M., et al.: Three years of trace gas observations over the EuroSiberian domain derived from aircraft sampling – a concerted action, Tellus B, 54(5), 696–712, 2002.

Liang, Q., Jaegle, L., Jaffe, D. A., Weiss-Penzias, P., Heckman, A., and Snow, J. A.: Long-range transport of Asian pollution to the northeast Pacific: Seasonal variations and transport pathways of carbon monoxide, J. Geophys. Res., 109, D23S07, doi:10.1029/2003JD004402, 2004.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, J. Geophys. Res., 108(D16), 4493, doi:10.1029/2002JD003161, 2003.

Liu, H., Jacob, D. J., Chan, L. Y., Oltmans, S. J., Bey, I., Yantosca, R. M., Harris, J. M., Duncan, B. N., and Martin, R. V.:, Sources of tropospheric ozone along the Asian Pacific Rim: An analysis of ozonesonde observations, J. Geophys. Res., 107(D21), 4573, doi:10.1029/2001JD002005, 2002.

Liu, H. Y., Jacob, D. J., Bey, I., Yantosca, R. M., Duncan, B. N., and Sachse, G. W.: Transport pathways for Asian pollution outflow over the Pacific: Interannual and seasonal variations, J. Geophys. Res., 108(D20), 8786, doi:10.1029/2002JD003102, 2003.

Lloyd, J., Langenfelds, R. L., Francey, R. J., Gloor, M., Tchebakova, N. M., et al.: A trace-gas climatology above Zotino, central Siberia, Tellus B, 54(5), 749–767, 2002.

Methven, J., Arnold, S. R., Stohl, A., Evans, M. J., Avery, M., Law, K., Lewis, A. C., Monks, P. S., Parrish, D. D., Reeves, C. E., Schlager, H., Atlas, E., Blake, D. R., Coe, H., Crosier, J., Flocke, F. M., Holloway, J. S., Hopkins, J. R., McQuaid, J., Purvis, R., Rappengluck, B., Singh, H. B., Watson, N. M., Whalley, L. K., and Williams, P. I.: Establishing Lagrangian connections between observations within air masses crossing the Atlantic during the International Consortium for Atmospheric Research on Transport and Transformation experiment, J. Geophys. Res.-Atmos., 111(D23), D23S62, doi:10.1029/2006JD007540, 2006.

Moody, J. L. and Galloway, J. N.: Quantifying the relationship between atmospheric transport and the chemical composition of precipitation on Bermuda, Tellus B, 40(5), 463–479, 1988.

Nakazawa, T., Sugawara, S., Inoue, G., Machida, T., Makshyutov, S., and Mukai, H.: Aircraft measurements of the concentrations of CO_2, CH_4, N_2O, and CO and the carbon and oxygen isotopic ratios of CO_2 in the troposphere over Russia, J. Geophys. Res.-Atmos., 102(D3), 3843–3859, 1997.

Nedelec, P., Cammas, J.-P., Thouret, V., Athier, G., Cousin, J.-M., Legrand, C., Abonnel, C., Lecoeur, F., Cayez, G., and Marizy, C.: An improved infrared carbon monoxide analyser for routine measurements aboard commercial Airbus aircraft: technical validation and first scientific results of the MOZAIC III programme, Atmos. Chem. Phys., 3, 1551–1564, 2003,
http://www.atmos-chem-phys.net/3/1551/2003/.

Newell, R. E. and Evans, M. J.: Seasonal changes in pollutant transport to the North Pacific: the relative importance of Asian and European sources, Geophys. Res. Lett., 27(16), 2509–2512, 2000.

Oberlander E. A., Brenninkmeijer, C. A. M., Crutzen, P. J., Elansky, N. F., Golitsyn, G. S., Granberg, I. G., Scharffe, D. H., Hofmann, R., Belikov, I. B., Paretzke, H. G., and van Velthoven, P. F. J.: Trace gas measurements along the Trans-Siberian railroad: The TROICA 5 expedition, J. Geophys. Res.-Atmos., 107(D14), 4206, doi:10.1029/2001JD000953, 2002.

Owen, R. C., Cooper, O. R., Stohl, A., and Honrath, R. E.: An analysis of the mechanisms of North American pollutant transport to the central North Atlantic lower free troposphere, J. Geophys. Res., 111, D23S58, doi:10.1029/2006JD007062, 2006.

Paris, J.-D., Ciais, P., Nédélec, P., Ramonet, M., Belan, B. D., Yu. Arshinov, M., Golitsyn, G. S., Granberg, I., Stohl, A., Cayez, G., Athier, G., Boumard, F., and Cousin, J.-M.: The YAK-AEROSIB transcontinental aircraft campaigns: new insights on the transport of CO_2, CO and O_3 across Siberia, Tellus B, 60(4), 551–568, doi:10.1111/j.1600-0889.2008.00369.x, 2008.

Pfister, G. G., Emmons, L. K., Hess, P. G., Honrath, R., Lamarque, J. F., Martin, M. V., Owen, R. C., Avery, M. A., Browell, E. V., Holloway, J. S., Nedelec, P., Purvis, R., Ryerson, T. B., Sachse, G. W., and Schlager, H.: Ozone production from the 2004 North American boreal fires, J. Geophys. Res., 111 (D24), 2006.

Pochanart, P., Akimoto, H., Kajii, Y., Potemkin, V. M., and Khodzher, T. V.: Regional background ozone and carbon monoxide variations in remote Siberia/East Asia, J. Geophys. Res., 108(D1), 4028, doi:10.1029/2001JD001412, 2003.

Ramonet, M., Ciais, P., Nepomniachii, I., Sidorov, K., Neubert, R. E. M., Langendorfer, U., Picard, D., Kazan, V., Biraud, S., Gusti, M., Kolle, O., Schulze, E. D., and Lloyd, J.: Three years of aircraft-based trace gas measurements over the Fyodorovskoye southern taiga forest, 300 km north-west of Moscow, Tellus B, 54(5), 713–734, 2002.

Real, E., Law, K. S., Weinzierl, B., Fiebig, M., Petzold, A., Wild, O., Methven, J., Arnold, S., Stohl, A., Huntrieser, H., Roiger, A., Schlager, H., Stewart, D., Avery, M., Sachse, G., Browell, E., Ferrare, R., and Blake, D.: Processes influencing ozone levels in Alaskan forest fire plumes during long-range transport over the North Atlantic, J. Geophys. Res., 112, D10S41, doi:10.1029/2006JD007576, 2007.

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/.

Sirois, A. and Bottenheim, J. W.: Use of Backward Trajectories to Interpret the 5-Year Record of Pan and O-3 Ambient Air Concentrations at Kejimkujik National-Park, Nova-Scotia, J. Geophys. Res., 100(D2), 2867–2881, 1995.

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

Stohl, A. and Thomson, D. J.: A density correction for Lagrangian particle dispersion models, Bound.-Lay. Meteorol., 90(1), 155–167, 1999.

Stohl, A., Eckhardt, S., Forster, C., James, P., and Spichtinger, N.: On the pathways and timescales of intercontinental air pollution transport, J. Geophys. Res., 107(D23), 4684, doi:10.1029/2001JD001396, 2002.

Stohl, A., Berg, T., Burkhart, J. F., Fjaeéraa, A. M., Forster, C., Herber, A., Hov, Ø., Lunder, C., McMillan, W. W., Oltmans, S., Shiobara, M., Simpson, D., Solberg, S., Stebel, K., Ström, J., Trseth, K., Treffeisen, R., Virkkunen, K., and Yttri, K. E.: Arctic smoke – record high air pollution levels in the European Arctic due to agricultural fires in Eastern Europe in spring 2006, Atmos. Chem. Phys., 7, 511–534, 2007a,
http://www.atmos-chem-phys.net/7/511/2007/.

Stohl, A., Forster, C., Huntrieser, H., Mannstein, H., McMillan, W. W., Petzold, A., Schlager, H., and Weinzierl, B.: Aircraft measurements over Europe of an air pollution plume from Southeast Asia aerosol and chemical characterization, Atmos. Chem. Phys., 7, 913–937, 2007b, http://www.atmos-chem-phys.net/7/913/2007/.

Stohl, A., Hittenberger, M., and Wotawa, G.: Validation of the Lagrangian particle dispersion model FLEXPART against large-scale tracer experiment data, Atmos. Env., 32(24), 4245–4264, 1998.

Takegawa, N., Kondo, Y., Koike, M., Chen, G., Machida, T., Watai, T., Blake, D. R., Streets, D. G., Woo, J. H., Carmichael, G. R., Kita, K., Miyazaki, Y., Shirai, T., Liley, J. B., and Ogawa, T.: Removal of NO_x and NO_y in Asian outflow plumes: Aircraft measurements over the western Pacific in January 2002, J. Geophys. Res., 109(D23), D23S04, doi:10.1029/2004JD004866, 2004.

Traub, M., Fischer, H., de Reus, M., Kormann, R., Heland, H., Ziereis, H., Schlager, H., Holzinger, R., Williams, J., Warneke, C., de Gouw, J., and Lelieveld, J.: Chemical characteristics assigned to trajectory clusters during the MINOS campaign, Atmos. Chem. Phys., 3, 459–468, 2003, http://www.atmos-chem-phys.net/3/459/2003/.

Val Martin, M., Honrath, R. E., Owen, R. C., Pfister, G., Fialho, P., and Barata, F.: Significant enhancements of nitrogen oxides, black carbon, and ozone in the North Atlantic lower free troposphere resulting from North American boreal wildfires, J. Geophys. Res., 111, D23S60, doi:10.1029/2006JD007530, 2006.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., and Arellano Jr., A. F.: Interannual variability in global biomass burning emissions from 1997 to 2004, Atmos. Chem. Phys., 6, 3423–3441, 2006, http://www.atmos-chem-phys.net/6/3423/2006/.

Wild, O., Pochanart, P., and Akimoto, H.: Trans-Eurasian transport of ozone and its precursors, J. Geophys. Res., 109, D11302, doi:10.1029/2003JD004501, 2004.

Wilks, D. S.: Statistical methods in the atmospheric sciences, Academic Press, Amsterdam, Boston, 627 pp., 2006.