Atmos. Chem. Phys., 12, 6757–6773, 2012
www.atmos-chem-phys.net/12/6757/2012/
doi:10.5194/acp-12-6757-2012

Atmospheric Chemistry and Physics

Impact of sampling frequency in the analysis of tropospheric ozone observations

M. Saunois[1,*], L. Emmons[1], J.-F. Lamarque[1], S. Tilmes[1], C. Wespes[1], V. Thouret[2,3], and M. Schultz[4]

[1]National Center for Atmospheric Research, Boulder, Colorado, USA
[2]Université de Toulouse, Toulouse, France
[3]Laboratoire d'Aérologie, Centre National de la Recherche Scientifique, Toulouse, France
[4]Forschungszentrum Jülich, 52425 Jülich, Germany
[*]now at: Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, Gif-Sur-Yvette, France

Correspondence to: M. Saunois (marielle.saunois@lsce.ipsl.fr)

Received: 8 July 2011 – Published in Atmos. Chem. Phys. Discuss.: 4 October 2011
Revised: 4 June 2012 – Accepted: 19 June 2012 – Published: 1 August 2012

Abstract. Measurements of ozone vertical profiles are valuable for the evaluation of atmospheric chemistry models and contribute to the understanding of the processes controlling the distribution of tropospheric ozone. The longest record of ozone vertical profiles is provided by ozone sondes, which have a typical frequency of 4 to 12 profiles a month. Here we quantify the uncertainty introduced by low frequency sampling in the determination of means and trends. To do this, the high frequency MOZAIC (Measurements of OZone, water vapor, carbon monoxide and nitrogen oxides by in-service AIrbus airCraft) profiles over airports, such as Frankfurt, have been subsampled at two typical ozone sonde frequencies of 4 and 12 profiles per month. We found the lowest sampling uncertainty on seasonal means at 700 hPa over Frankfurt, with around 5 % for a frequency of 12 profiles per month and 10 % for a 4 profile-a-month frequency. However the uncertainty can reach up to 15 and 29 % at the lowest altitude levels. As a consequence, the sampling uncertainty at the lowest frequency could be higher than the typical 10 % accuracy of the ozone sondes and should be carefully considered for observation comparison and model evaluation. We found that the 95 % confidence limit on the seasonal mean derived from the subsample created is similar to the sampling uncertainty and suggest to use it as an estimate of the sampling uncertainty. Similar results are found at six other Northern Hemisphere sites. We show that the sampling substantially impacts on the inter-annual variability and the trend derived over the period 1998–2008 both in magnitude and in sign throughout the troposphere. Also, a tropical case is discussed using the MOZAIC profiles taken over Windhoek, Namibia between 2005 and 2008. For this site, we found that the sampling uncertainty in the free troposphere is around 8 and 12 % at 12 and 4 profiles a month respectively.

1 Introduction

Tropospheric ozone is an important trace gas due to its role in the oxidative capacity of the global atmosphere, its effect on climate and its impact on air quality. This trace gas is monitored worldwide on various platforms (surface stations, balloons, aircraft, satellites) with diverse instruments (electronic cells, UV absorption instruments, Brewer-Dobson instruments, infrared spectrometers). After a continuous increase of ozone concentrations over Europe until the 1980s or 1990s (e.g. Logan, 1999; Naja et al., 2003; Ordóñez et al., 2005; Oltmans et al., 2006; Zbinden et al., 2006; Parrish et al., 2009), a leveling-off has been observed over the past decade (e.g. Ordóñez et al., 2005; Oltmans et al., 2006; Zbinden et al., 2006; Parrish et al., 2009). Since the 1980s, global anthropogenic emissions of ozone precursors have increased due to rapid economic development in Asia, while European and North American emissions have been decreasing (Vestreng et al., 2007; Monks et al., 2009). Tropospheric ozone variability is also influenced by biomass burning emissions (e.g. Simmonds et al., 2005; Koumoutsaris et al., 2008; Oltmans et al., 2010), atmospheric circulation (e.g. Rodriguez et al., 2004; Eckhardt et al., 2003),

changes in transport from the stratosphere (e.g. Fusco and Logan, 2003; Tarasick et al., 2005; Ordóñez et al., 2007) and residence time of air masses in the boundary layer (e.g. Naja et al., 2003; Solberg et al., 2008).

Due to the high temporal and spatial variability of ozone, long term measurements are necessary to determine changes in ozone concentrations with some degree of significance. While surface stations provide extensive datasets of ozone measurements, regular in-situ measurements of ozone in the free troposphere (i.e. not dedicated aircraft campaigns) were provided solely by balloon soundings, until the MOZAIC (Measurements of OZone, water vapor, carbon monoxide and nitrogen oxides by in-service AIrbus airCraft, Marenco et al., 1998) program was launched in 1994. Measurements of ozone vertical profiles are useful for the evaluation of numerical models (e.g. Logan, 1999; Emmons et al., 2000) and contribute to the understanding of the processes controlling the distribution of tropospheric ozone (e.g. Lamarque and Hess, 2004; Koumoutsaris et al., 2008).

Within the framework of international projects such as Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP, http://www.giss.nasa.gov/projects/accmip/), the Task Force on Hemispheric Transport of Air Pollution (HTAP; www.htap.org, Keating and Zuber, 2007) or the Chemistry-climate Model Validation Activity (CCMVal2, Eyring et al., 2010), it is necessary that observational data used for model evaluation and comparison to other observations be provided in a format comparable with model output. While the ideal way would be to output the model based on the sonde dates, the observational data are generally averaged on a monthly-mean time scale in order to facilitate the comparison of model to observation and to reduce the effort required in sharing data between different groups of research. However, the sampling frequency of the soundings is typically of 4 to 12 profiles per month. Thus, the monthly mean derived from those observational data will depend on how typical were the days sampled and thus, may be biased due to the sampling.

The ozone sonde data sets provide information about long-term changes in ozone concentration (e.g. Oltmans et al., 2006; Logan et al., 2012). However due to changes in techniques, the interpretation of the records may be difficult (e.g. Smit et al., 2007; Logan et al., 2012). Trends from ozone soundings and other platforms such as aircraft or surface sites are not always consistent with each other (Jonson et al., 2006; Oltmans et al., 2006; Chipperfield et al., 2007; Logan et al., 2012). To reconcile data from different platforms, a number of factors have to be accounted for. The aforementioned numerous sources of ozone variability complicate our understanding of ozone changes. Also, surface site measurements are made at high frequency (available on an hourly basis or less), but are often representative of local conditions, while soundings, giving vertical profiles, are limited spatially and are launched at low frequency. Cooper et al. (2010) used large data sets, with major contributions from MOZAIC, to discuss the springtime ozone increase over western North America. They suggest that weekly ozone sonde profiles were not sufficiently frequent to detect the positive ozone trend in the free troposphere.

The objective of this paper is to discuss and quantify the uncertainty in the analysis of low sampling frequency measurements such as ozone sondes. We aim to answer the main question: how significant are the signals measured from ozone sonde data sets? This question consists of several others: does a low sampling frequency influence the derived seasonal means? Does the time resolution impact the observed seasonal and inter-annual variabilities? Are trend estimates affected by low sampling frequency? Can we estimate the sampling uncertainty on the seasonal means, which could be used for model-observation comparisons or observation-observation comparisons?

For that purpose, we use the high frequency MOZAIC data set over Frankfurt and subsample these profiles at two typical sonde frequencies (4 and 12 profiles per month). This allows us to study to what extent time resolution can influence the observed seasonal mean and its variation, as if they were derived from different data sets. To our knowledge, this study is the first of the kind to assess the potential impact of sampling on the observed tropospheric ozone concentrations and variations using the high frequency MOZAIC data set.

The data and methodology are described in Sects. 2 and 3 respectively. Section 4 presents the effects of sampling derived from the ozone vertical profiles over Frankfurt on the seasonal means (Sect. 4.1), the annual and inter annual variabilities (Sect. 4.2), ozone trends (Sect. 4.4) and compares this sampling uncertainty with the instrument uncertainties (Sect. 4.3). A discussion of ozone trends is presented on the basis of MOZAIC, ozone sonde and surface measurements in Sect. 4.4. A generalization of the results for the Northern Hemisphere midlatitudes is presented in Sect. 5 and a tropical case study is discussed in Sect. 6. Conclusions are given in Sect. 7.

2 Observations

2.1 MOZAIC data

We use MOZAIC data covering the period 1995–2008 (http://mozaic.aero.obs-mip.fr, Marenco et al., 1998). The ozone measurements are made onboard MOZAIC aircraft with a dual beam UV absorption instrument having a detection limit of 2 ppbv and an overall precision of $\pm$(2 ppbv + 2 %) (Thouret et al., 1998). Special focus is given to the vertical profiles collected over Frankfurt, Germany. This airport is the most frequently sampled by MOZAIC, with a total of 12 676 vertical profiles between January 1995 and December 2008. On average, 75 profiles per month, i.e. more than two profiles a day, are provided. The MOZAIC measurements made over Vienna, Paris, New York, Boston, Osaka

and Tokyo are also used to generalise our results to the Northern Hemisphere midlatitudes in Sect. 5. The data collected in Namibia over Windhoek between 2005 and 2008 are used for the tropical case study presented in Sect. 6. The vertical profiles are binned by 100 hPa layers centered around the following mid-level pressures: 1000, 900, 800, 700, 600, 500, 400 and 300 hPa.

2.2 Ozone sondes and surface stations

The main area of interest, referenced hereafter as "Central Europe", is defined as the region between 44° N and 55° N latitude, and 3° E and 18° E longitude, which encompasses Frankfurt and several sounding stations. Figure 1 shows a map of this region and the measurement sites, where Frankfurt is denoted by a black star.

Six ozone sonde stations located near Frankfurt provided data over the period 1995–2008: Debilt, Hohenpeissenberg, Lindenberg, Payerne, Praha and Uccle (blue stars). The sounding data are available through the World Ozone and Ultraviolet Radiation Data Center (WOUDC, http://www.woudc.org). The ozone sonde data are treated as in Tilmes et al. (2011). The profiles already include the corrections performed by the data centers. For each profile a correction factor is suggested by the data center to scale the profile to ground-based ozone column measurements, for which the stratospheric fraction is dominant. This factor has not been applied here as it has little impact on the mean tropospheric profile, but we disregard profiles with a correction factor outside the range of 0.8 and 1.2. This filtering has only a small impact on the averaged profile between 1995 and 2009 (Tilmes et al., 2011). However, we filter out single profiles with column ozone values of more than 700 DU or of less than 50 DU, which would present unrealistic values of ozone profiles at the stratospheric maximum. The sonde profiles are binned the same way as MOZAIC profiles. For the elevated sites (Hohenpeissenberg and Payerne), the surface layer is 950–850 hPa instead of 1050–950 hPa.

More than 180 EMEP (European Monitoring and Evaluation Program) surface stations provide measurements of ozone concentrations. However, we keep only the EMEP stations located within our Central Europe region and having performed continuous measurements over the period 1998–2008. The data from the 48 remaining sites are filtered to retain only morning measurements so as to avoid a high diurnal variability and keep the same time window as the sondes. The surface stations appear as the green plus markers on the map (Fig. 1).

3 Methodology

As stated in the introduction, we aim to discuss and quantify the uncertainty associated with low sampling frequency measurements such as ozone sondes. The methodology detailed below explains how we subsampled the high frequency MOZAIC data set of ozone profiles over Frankfurt to create subsamples with sampling frequency similar to those of the European ozone sonde data sets.

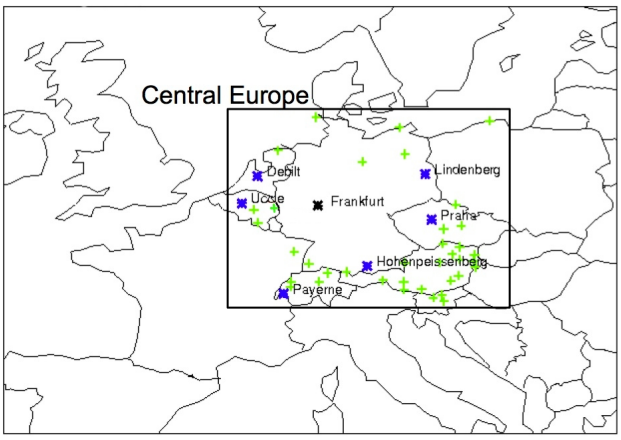

Fig. 1. Map of Europe. The Central Europe region, defined by the area between 44° N and 55° N latitude, and 3° E and 18° E longitude, is shown in the black rectangle. The MOZAIC city, Frankfurt, is shown with a black star, the six ozone sounding sites with blue stars and the EMEP surface stations with green plus symbols.

3.1 Time autocorrelation and effective sample size

Temporal autocorrelation in time series can significantly reduce the amount of information that would be available from the same number of independent data points, and thus increase the error estimates. We have tested the temporal correlation in the MOZAIC daily time series of ozone profiles over Frankfurt. Ozone profiles over Frankfurt from the MOZAIC aircraft are numerous but not regular, leading to individual or several days within a month that are not documented. Missing values in the daily time series of a month will have an effect on the estimation of autocorrelation. In order to avoid any misrepresentation of the temporal autocorrelation, we calculate time correlation based on months that have at least one profile a day. There are sixty-one months of the kind over the period 1995–2008. When more than one profile a day is available, the profile for this day is randomly selected. Sixty-one daily time series of one month are analyzed for each of the seven pressure layers. We estimate the first order autoregressive coefficient (r_1) for each month and then take the average of these estimates across the months. We found that the estimated autoregressive coefficient $\overline{r_1}$ (between adjacent days) is about 0.10–0.26 with a maximum value at 900 hPa and a minimum value at 600 hPa (Table 1). Using all the available months ($168 = 12 \cdot 14$) to estimate $\overline{r_1}$ leads to a range of 0.17–0.35.

To test the significance level of r_1, we use the one-sided test recommend by WMO (1966) and compute the 95 % significance level for r_1 with

Table 1. First order autoregressive coefficients ($\overline{r_1}$) and scaling factors for effective sample size derived from the 61 full months at rate 1/1 (sample every day), rate 1/2 (sample every other day) and rate 1/4 (sample every fourth day).

Pressure level	Rate 1/1		Rate 1/2		Rate 1/4
	$\overline{r_1}$	$f_{1/1} = \frac{1-\overline{r_1}}{1+\overline{r_1}}$	$\overline{r_1}$	$f_{1/2} = \frac{1-\overline{r_1}}{1+\overline{r_1}}$	$\overline{r_1}$
1000 hPa	0.15	0.74	0.03	0.94	0.08
900 hPa	0.26	0.59	0.17	0.71	0.03
800 hPa	0.21	0.65	0.14	0.75	0.02
700 hPa	0.14	0.75	0.09	0.83	0.03
600 hPa	0.10	0.82	0.05	0.90	0.05
500 hPa	0.11	0.81	0.00	1.00	0.05
400 hPa	0.13	0.77	0.03	0.94	0.08
300 hPa	0.16	0.73	0.03	0.94	0.04

$$r_{1,.95} = \frac{-1 + 1.645\sqrt{(N_\mathrm{d} - 2)}}{N_\mathrm{d} - 1} \quad (1)$$

where N_d is the number of daily observations (i.e. 29 to 31). For $N_\mathrm{d} = 30$, $r_{1,.95} = 0.26$. As a result, $\overline{r_1}$ is below or at the limit of the threshold, which means that the null hypothesis of no relationship between adjacent days ($\rho_1 = 0$) cannot be rejected. In other words, adjacent day observations may be uncorrelated, especially in the free troposphere above 800 hPa. We also determine the minimum time lag necessary to reach independence between observations by screening the correlogram of each month. We estimate that independence is reached when the autocorrelation coefficient is lower than the 95 % confidence limit $\frac{-1}{N_\mathrm{d}} + \frac{2}{\sqrt{N_\mathrm{d}}}$ (equal to 0.30 for $N_\mathrm{d} = 30$). Using the 61 full months only, we found that independency is reached for a time lag of one day in about 62–90 % of the cases and two days in another 7–31 % of the cases (the percentage depends on the pressure level). Less than 7 % of the correlograms show significant correlation for time lag equal to or higher than 3 days. These first results suggest that ozone measurements made every other day are generally independent. We further subsampled the 61 full months at two sampling rates: rate 1/2 (sample every other day) and rate 1/4 (sample every fourth day), creating 122 ($61 \cdot 2$) time series at rate 1/2 and 244 ($61 \cdot 4$) at rate 1/4. We estimate the first order autoregressive coefficient for each time series and take the average of these estimates across the time series. The results are reported in Table 1. The autocorrelation between observations made every other day is lower than 0.17, supporting an independence of observations made every other day. As expected, we found that the samples at rate 1/4 do not present significant temporal autocorrelation.

As the first order autoregressive coefficient is found to be significant for these Frankfurt daily time series of tropospheric ozone, the sample size needs to be adjusted for autocorrelation in time series. However, as previously stated, the MOZAIC measurements are irregular and this makes it difficult to estimate r_1 for each month. As a result, we have used the same average estimate $\overline{r_1}$ for all months. The effective sample size for one month is given by

$$N_{\mathrm{d,eff}} = N_\mathrm{d} \cdot f = N_\mathrm{d} \frac{1 - \overline{r_1}}{1 + \overline{r_1}}. \quad (2)$$

A first order correlation of 0.26 leads to a scaling to about 59 % of the original sample size. The scaling factors (f) are given in Table 1 ($f_{1/1}$ for every day sample and $f_{1/2}$ for every other day sample, for every fourth day sample we will consider $f_{1/4} = 1$). The effective sample size will be used to estimate the standard error and confidence interval on the seasonal means in Sects. 3.5 and 4 for Frankfurt and in Sect. 5 for the other northern midlatitude sites. We also estimate the proper scaling factors for Windhoek and use them in Sect. 6 (factors not shown).

3.2 Morning subset of the Frankfurt MOZAIC data set

The methodology used in this paper aims to mimic the ozone sonde sampling. Sondes are launched generally at 11:30 UT or 12:00 UT in five of the six ozone sonde stations located in Central Europe (Debilt, Lindenberg, Payerne, Praha and Uccle) and around 05:30 UT at the Hohenpeissenberg station. To avoid the effect of a strong diurnal cycle in the lowest levels and to match the time window of the balloon launch in Central Europe, we retain the MOZAIC profiles taken between 05:00 and 13:00 UT. Figure 2 shows the number of MOZAIC profiles (black line) per season for each year as well as the number of profiles per season available within this morning time window (blue line). As most of the MOZAIC flights were transatlantic, they took off and landed in the morning in Frankfurt. As a result, the morning subset of ozone profiles over Frankfurt represents 79 % of the entire dataset and often includes more than 100 profiles per season (Fig. 2). The morning subset of this data set includes on average more than one profile per day, allowing us to subsample each month at two typical sonde frequencies: 4 and 12 profiles a month (i.e. 12 and 36 profiles per season) as described hereafter. However, the reader should keep in mind that the profiles are taken irregularly, meaning that there are

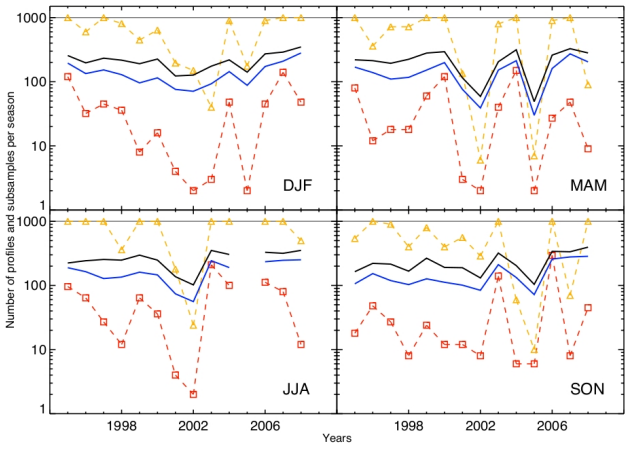

Fig. 2. Time record of the number of vertical profiles taken per season over Frankfurt between 1995 and 2008, in the whole data set (black) and in the morning (blue) subset in solid lines. The number of seasonal subsamples created with a frequency of 4 (orange) and 12 (red) profiles a month using the "regular" sampling method is shown by the dashed lines.

days with more than two profiles, days with one or two profiles, and even days without observations.

3.3 Monthly subsampling

To better mimic the regular sampling of the soundings and to create subsamples, we use a "regular" sampling method, which is illustrated in Fig. 3. The subsampling is done for each month; this is why we consider two frequencies (N_f) equal to 4 or 12 profiles per month. Consider a theoretical month documented with 60 regularly-spaced profiles (twice a day, which is close to the reality at Frankfurt). If we subsample this data set to 4 profiles, we can create $60/4 = 15$ subsamples by taking every 15th profile. Considering there were two profiles a day, the first subsample corresponds to day 1 (1st profile), day 8 (2nd profile), day 16 (1st profile) and day 23 (2nd profile) (Fig. 3a), i.e. one profile every week. The second subsample corresponds to day 1 (2nd profile), day 9 (1st profile), day 16 (2nd profile) and day 24 (1st profile); and so on for the other subsamples. If we want subsamples of 12 profiles using the same month, then we create $60/12 = 5$ subsamples by taking every 5th profile. As a result, the first subsample corresponds to one profile of days 1, 3, 6, 8, 11, 13, 16, 18, 21, 23, 26, 28; the second subsample corresponds to one profile of days 1, 4, 6, 9, 11, 14, 16, 19, 21, 24, 26, 29; the third subsample corresponds to one profile of days 2, 4, 7, 9, 12, 14, 17, 19, 22, 24, 27, 29; and so on (Fig. 3b). In this example, we take a profile every 2 or 3 days, which is close to the reality of the thrice-weekly sampling of soundings. In the case of a number of profiles a month which is not divisible by 4 or 12, some profiles are not used as we do not allow multiple uses of profiles.

This method avoids selecting sequential days, which is consistent with the sampling frequency of ozone sondes, even though some MOZAIC profiles are discarded in this way. Except for a few months which are documented with fewer than 40 profiles, we were able to create more than 10 subsamples of 4 profiles. On the contrary, there were less than 10 subsamples created with 12 profiles for each month because there are fewer than 120 profiles per month. As a compromise between representativity and data availability, we limit the number of monthly subsamples to 10 for both frequencies. As a consequence, for the example of 60 profiles we would use only the first 10 subsamples of 4 profiles. The reader should note here that the irregularity of the MOZAIC measurements makes it difficult to sample exactly every 2 or 3 days (or every week) as presented in the example in Fig. 3. The sampling chosen here leads to slightly different sampling days than those of the ozone sondes. However, considering the high number of profiles over Frankfurt (except in 2002 and 2005), the sampling frequencies are similar (weekly or thrice weekly) to those of the ozone sonde frequencies and allows to produce more samples.

We also tried a "random" sampling method in which the profiles are randomly picked within the month. A random sampling allows eventually to consider any profiles and to create 10 subsamples, whatever the number of profiles available and their frequency. Despite the fact that profiles from sequential days might be selected, potentially giving more weight to a particular time/event in the monthly mean, this method provides similar results.

3.4 Creation of the seasonal subsamples

The seasonal subsamples are derived from the monthly subsamples. If there are $n1$, $n2$ and $n3$ subsamples for month 1, month 2 and month 3 respectively, then we derive $n^{N_f}_{seas,yr} = n1 \times n2 \times n3$ subsamples for a given season and year. Consequently, a monthly subsample may be used in several seasonal subsamples. The number of subsamples created per season and per year is given in Fig. 2. As there are up to 10 monthly subsamples, the maximum number of seasonal subsamples is 1000. This value is often reached for the frequency of 4 profiles a month. On the contrary, since generally less than 10 monthly subsamples of 12 profiles could be created, the maximum of seasonal subsamples at this frequency is around 120. Using this "regular" sampling method, the number of subsamples is highly dependent on the number of profiles available. In particular, fewer profiles were available for the years 2002 and 2005, especially in the spring and summer (Fig. 2). In order to keep a minimum of two subsamples per season, per year and for each frequency, we discard the spring of 2005 and the summers of 2002 and 2005 from our discussion. Discarding these years does not significantly affect the results regarding the trends presented in Sect. 4.4.

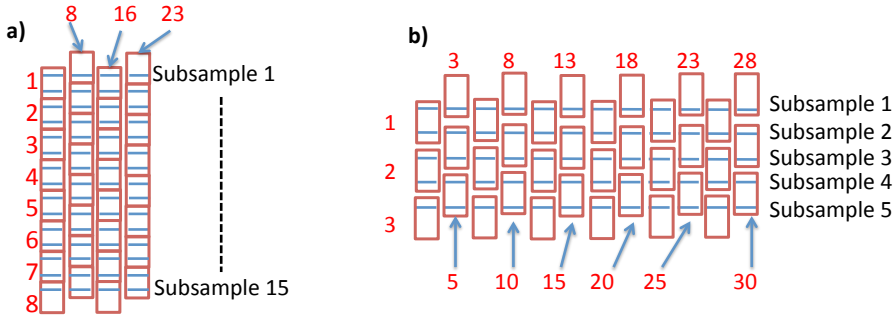

Fig. 3. Illustration of the "regular" sampling method based on a theoretical month, for which 60 profiles are available (shown as small blue lines). We consider here that two profiles per day are available. Each day is represented as a red box and the red number is the day of the month from 1 to 30. In **(a)** we can create 15 (60 divided by 4) subsamples of four profiles each. The first subsample corresponds to day 1 (1st profile), day 8 (2nd profile), day 16 (1st profile) and day 23 (2nd profile). In **(a)**, day 8 is repeated in the second column to include two profiles (same for day 23). The second subsample corresponds to day 1 (2nd profile), day 9 (1st profile), day 16 (2nd profile) and day 24 (1st profile); and so on for the other subsamples. In **(b)** we create 5 (60 divided by 12) subsamples of 12 profiles each. The first subsample corresponds to one profile of days 1, 3, 6, 8, 11, 13, 16, 18, 21, 23, 26, 28. Days 3, 8, 13, 18, 23 and 28 are repeated from the bottom of a column to the top of the next column. This method allows the selection of regularly spaced profiles.

3.5 Definition of the metrics

In this study, we aim to give a quantitative estimate of the uncertainty that arises from low time resolution, depending on the season, altitude level and sampling frequency. We explain here how we define the sampling uncertainty and other metrics presented in the text and figures. A summary of the metrics is presented in Table 2.

3.5.1 Seasonal mean, standard error and confidence limit

First, we calculate the seasonal mean concentrations of ozone at each pressure level. For each year, we define by $\overline{x}_{\mathrm{seas,yr}}$ the seasonal mean derived from the $N_{\mathrm{seas,yr}}$ MOZAIC morning profiles; the subscript "seas,yr" means that this value is derived for each season and year. The sample standard deviation associated with the seasonal sample of $N_{\mathrm{seas,yr}}$ profiles is called $s_{\mathrm{seas,yr}}$. The standard error of the mean $\overline{x}_{\mathrm{seas,yr}}$ is defined by

$$\overline{se}_{\mathrm{seas,yr}} = \frac{s_{\mathrm{seas,yr}}}{\sqrt{N_{\mathrm{eff}}}} \quad (3)$$

with

$$N_{\mathrm{eff}} = f_{1/1}\frac{N_{\mathrm{seas,yr}}}{2} \quad (4)$$

N_{eff} is the effective sample size accounting for autocorrelation in the daily observations. The scaling factor $f_{1/1}$ is that derived in Sect. 3.1. However, there are on average two profiles per day in the Frankfurt morning data set. Assuming that a second profile taken the same day does not provide more information than the first one, we divide $N_{\mathrm{seas,yr}}$ by 2. Using an effective sample size instead of the real size of the sample increases the standard error estimate. The seasonal distributions of the MOZAIC measurements are found to be close to normal except for the lowest levels, for which the distributions are cut off at zero due to ozone titration by NO in the vicinity of the airport. In these cases, part of the theoretical normal distribution would be in the negative side, which is not realistic for ozone concentrations. The 95 % confidence interval for $\overline{x}_{\mathrm{seas,yr}}$ is defined as

$$\mathrm{CI}_{95\%} = [\overline{x}_{\mathrm{seas,yr}} - t_{0.05}\cdot\overline{se}_{\mathrm{seas,yr}}, \overline{x}_{\mathrm{seas,yr}} + t_{0.05}\cdot\overline{se}_{\mathrm{seas,yr}}] \quad (5)$$

where $t_{0.05}$ is the 95th percentile of the Student's t-distribution with $N_{\mathrm{eff}} - 1$ degree of freedom. This confidence interval is represented in Fig. 4 by the blue vertical error bars. For the lowest levels where the distributions are skewed, the confidence interval may not represent a 95 % confidence interval, however, for clarity, we keep the same metrics for the lowest levels.

Similar definitions are used for a subsample of the MOZAIC morning data set. Considering there are $N^{i,N_{\mathrm{f}}}_{\mathrm{seas,yr}}$ profiles in a subsample i, whose seasonal mean is $\overline{x}^{i,N_{\mathrm{f}}}_{\mathrm{seas,yr}}$, then the standard error of the mean is defined as

$$\overline{se}^{i,N_{\mathrm{f}}}_{\mathrm{seas,yr}} = \frac{s^{i,N_{\mathrm{f}}}_{\mathrm{seas,yr}}}{\sqrt{N_{\mathrm{eff}}}} \quad (6)$$

where $s^{i,N_{\mathrm{f}}}_{\mathrm{seas,yr}}$ is the sample standard deviation of the subsample i and with

$$N_{\mathrm{eff}} = f_{1/2}N^{i,N_{\mathrm{f}}}_{\mathrm{seas,yr}} \text{ for } N_{\mathrm{f}} = 12, \text{ and} \quad (7)$$

$$N_{\mathrm{eff}} = N^{i,N_{\mathrm{f}}}_{\mathrm{seas,yr}} \text{ for } N_{\mathrm{f}} = 4. \quad (8)$$

Table 2. List of the metrics used and their definition. $t_{0.05}$ is the 95th percentile of the Student's t-distribution with $N_{\text{eff}} - 1$ degrees of freedom. N_{eff} is defined in Eq. (4) for the morning data set and in Eqs. (7) and (8) for the subsamples.

Metric	Definition
N_{f}	subsampling frequency, here 4 or 12 profiles per month
$N_{\text{seas,yr}}$	number of MOZAIC morning profiles per season and per year
$\overline{x}_{\text{seas,yr}}$	seasonal mean using all morning profiles (for one year)
$s_{\text{seas,yr}}$	sample standard deviation of the $N_{\text{seas,yr}}$ morning profiles
$\overline{se}_{\text{seas,yr}}$	standard error of the mean $\overline{x}_{\text{seas,yr}}$ defined by Eq. (3)
N_{eff}	effective sample size ($= f_{1/1}\frac{N_{\text{seas,yr}}}{2}$, Eq. 4)
$\pm t_{0.05} \cdot \overline{se}_{\text{seas,yr}}$	95 % confidence limits on $\overline{x}_{\text{seas,yr}}$
$n_{\text{seas,yr}}^{N_{\text{f}}}$	number of created subsamples per season and per year
$N_{\text{seas,yr}}^{i,N_{\text{f}}}$	number of profiles in the subsample i for one season and year
$\overline{x}_{\text{seas,yr}}^{i,N_{\text{f}}}$	seasonal mean using the $N_{\text{seas,yr}}^{i,N_{\text{f}}}$ profiles of subsample i
$s_{\text{seas,yr}}^{i,N_{\text{f}}}$	subsample standard deviation of the $N_{\text{seas,yr}}^{i,N_{\text{f}}}$ profiles of subsample i
$\overline{se}_{\text{seas,yr}}^{i,N_{\text{f}}}$	standard error of the mean $\overline{x}_{\text{seas,yr}}^{i,N_{\text{f}}}$ defined by Eq. (6)
N_{eff}	effective sample size for the subsamples (see Eqs. 7 and 8)
$\pm t_{0.05} \cdot \overline{se}_{\text{seas,yr}}^{i,N_{\text{f}}}$	95% confidence limits on $\overline{x}_{\text{seas,yr}}^{i,N_{\text{f}}}$
$y_{\text{seas,yr}}^{i,N_{\text{f}}}$	difference $\overline{x}_{\text{seas,yr}}^{i,N_{\text{f}}} - \overline{x}_{\text{seas,yr}}$
$s_{\text{seas,yr}}^{\text{samp,N}_{\text{f}}}$	sample standard deviation of the distribution $f_{\text{seas,yr}}^{N_{\text{f}}}(y_{\text{seas,yr}}^{i,N_{\text{f}}})$
$s_{\text{seas,clim}}^{\text{samp,N}_{\text{f}}}$	sample standard deviation of the distribution $g_{\text{seas}}^{N_{\text{f}}}(y_{\text{seas,yr}}^{i,N_{\text{f}}})$
$2.6 \cdot s_{\text{seas,yr}}^{\text{samp,N}_{\text{f}}}$	sampling uncertainty (one year of data)
$2.6 \cdot s_{\text{seas,clim}}^{\text{samp,N}_{\text{f}}}$	climatological sampling uncertainty

Here the superscript "i, N_{f}" denotes reference to a subsample i at the frequency N_{f}. The number of profiles $N_{\text{seas,yr}}^{i,N_{\text{f}}}$ per season and per year is equal to 12 or 36 for a frequency $N_{\text{f}} = 4$ or 12 respectively. For a frequency of $N_{\text{f}} = 12$, which leads to a sampling every other day or every three days on average, we use the first order autoregressive coefficient derived for the rate $1/2$ in Sect. 3.1 as an estimate of the autocorrelation in the monthly subsamples. Thus the subsample size ($N_{\text{seas,yr}}^{i,N_{\text{f}}}$) is scaled with the factor $f_{1/2}$. For the frequency $N_{\text{f}} = 4$, the results found in Sect. 3.1 suggest insignificant autocorrelation between profiles taken every week. Thus we use the real subsample size to estimate the standard error. The confidence interval for each of the subsample means is calculated as

$$\text{CI}_{95\%} = t_{0.05} \cdot \overline{se}_{\text{seas,yr}}^{i,N_{\text{f}}} \tag{9}$$

where $t_{0.05}$ is the 95th percentile of the Student's t-distribution with $N_{\text{eff}} - 1$ degree of freedom.

3.5.2 Sampling uncertainty

In order to assess the effect of low sampling frequency, we analyze the differences between the overall seasonal mean $\overline{x}_{\text{seas,yr}}$, derived from the morning MOZAIC data set, defining our true value, and the biased seasonal means derived from the generated low frequency subsamples $\overline{x}_{\text{seas,yr}}^{i,N_{\text{f}}}$ ($i = 1$, $n_{\text{seas,yr}}^{N_{\text{f}}}$ with $n_{\text{seas,yr}}^{N_{\text{f}}}$ the number of subsamples of 4 or 12 profiles a month created for a given season and year). We note $y_{\text{seas,yr}}^{i,N_{\text{f}}} = \overline{x}_{\text{seas,yr}}^{i,N_{\text{f}}} - \overline{x}_{\text{seas,yr}}$ these differences. At a frequency $N_{\text{f}} = 4$, there are generally $n_{\text{seas,yr}}^{N_{\text{f}}} = 1000$ subsamples per season and per year, while at the frequency $N_{\text{f}} = 12$, $n_{\text{seas,yr}}^{N_{\text{f}}}$ is lower than 120 and around 50–70 (Fig. 2). For each year and each season, we consider the discrete probability distribution of these differences $f_{\text{seas,yr}}^{N_{\text{f}}}(y_{\text{seas,yr}}^{i,N_{\text{f}}})$ composed of $n_{\text{seas,yr}}^{N_{\text{f}}}$ subsamples. The sample standard deviation associated with the distribution $f_{\text{seas,yr}}^{N_{\text{f}}}$ is called $s_{\text{seas,yr}}^{\text{samp,N}_{\text{f}}}$, the superscript "samp" referring to "sampling uncertainty" and the subscript "seas,yr" recalling that this value is derived for a given year and season. The number of seasonal subsamples in each year is generally large enough to make the distributions $f_{\text{seas,yr}}^{N_{\text{f}}}$ close to normal. This is true for a frequency of $N_{\text{f}} = 4$, but less frequent at $N_{\text{f}} = 12$, especially for particular years such as 2002, 2003 or 2005, when $n_{\text{seas,yr}}^{N_{\text{f}}}$ is below a few dozens (Fig. 2). We defined, for each year and each season, the sampling uncertainty on the observed means as $2.6 \cdot s_{\text{seas,yr}}^{\text{samp,N}_{\text{f}}}$. In the case of a normal distribution, this value corresponds to the 99 % sampling uncertainty and ensures that 99 % of the biased seasonal means are within $\pm 2.6 \cdot s_{\text{seas,yr}}^{\text{samp,N}_{\text{f}}}$. When the number of subsamples is below this limit, we keep the same estimate; even though it does not correspond to 99 % of the probability distribution, it still gives a range for the biased seasonal mean spread.

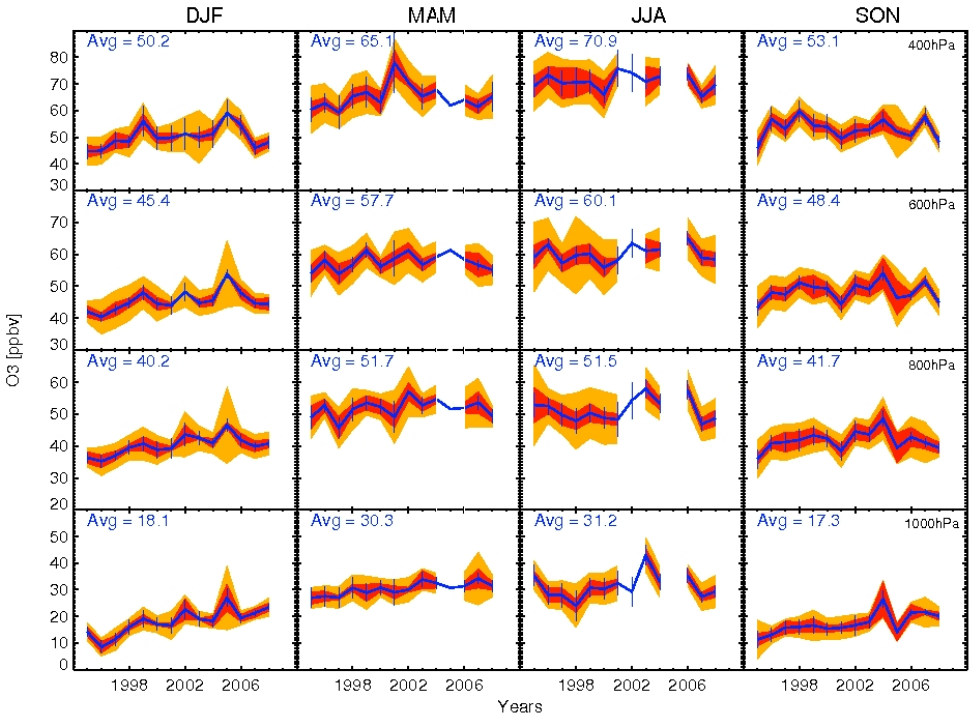

Fig. 4. Time evolution of the ozone seasonal means over Frankfurt between 1995 and 2008 as observed by MOZAIC aircraft at four pressure levels (1000, 800, 600, 400 hPa, from bottom to top). The seasonal mean ($\overline{x}_{\mathrm{seas,yr}}$) from the morning data set is shown in blue with the 95 % confidence limit error bars ($\mathrm{CI}_{95\%} = t_{0.05} \cdot \overline{se}_{\mathrm{seas,yr}}$). The shaded areas represent the sampling uncertainty (range $[\overline{x}_{\mathrm{seas,yr}} - 2.6 \cdot s^{\mathrm{samp,N_f}}_{\mathrm{seas,yr}}$, $\overline{x}_{\mathrm{seas,yr}} + 2.6 \cdot s^{\mathrm{samp,N_f}}_{\mathrm{seas,yr}}]$, see Sect. 3.5.2 for details) derived from the ensemble of subsamples of 4 (orange) and 12 (red) profiles a month for each year and season. At the top left corner of each rectangle, the average seasonal means over the entire period are shown.

In order to obtain a single value per season (and per altitude level), we aggregated the 14 distributions $f^{N_f}_{\mathrm{seas,yr}}$ from each year in the period 1995–2008 into a single distribution, which we call $g^{N_f}_{\mathrm{seas}}$. For each season, the resulting distribution $g^{N_f}_{\mathrm{seas}}(y^{i,N_f}_{\mathrm{seas,yr}})$ is then composed of approximately 9000 (800) seasonal subsamples for a frequency of $N_f = 4$ (12) profiles a month. These distributions are found to be normal and are fitted with a gaussian distribution. We call sampling uncertainty the value given by $2.6 \cdot s^{\mathrm{samp,N_f}}_{\mathrm{seas,clim}}$ (the subscript "clim" refers here to "climatology"), where $s^{\mathrm{samp,N_f}}_{\mathrm{seas,clim}}$ is the sample standard deviation of the fitted distribution. We consider the fitted standard deviation of the distribution in order to exclude outliers, which appear mainly in winter and spring. The fitted sample standard deviation is similar to the sample standard deviation for distributions free of outlier. This sampling uncertainty estimate ensures that 99 % of the seasonal means derived from the subsamples are within $\pm 2.6 \cdot s^{\mathrm{samp,N_f}}_{\mathrm{seas,clim}}$ of their true value. This time, $s^{\mathrm{samp,}N_f}_{\mathrm{seas,clim}}$ refers to a seasonal estimate derived using the whole morning MOZAIC period, and we have called it climatological sampling uncertainty.

4 Effect of the sampling over Frankfurt

In this section, we discuss to what extent the sampling impacts on the observed seasonal means and the annual and inter annual variabilities using the ensemble of subsamples created from MOZAIC morning profiles over Frankfurt.

4.1 Sampling uncertainty on the ozone seasonal mean

Figure 4 presents the variations of ozone concentrations from the MOZAIC morning subset over Frankfurt (in blue) between 1995 and 2008 at four pressure levels (1000, 800, 600 and 400 hPa) and for the four seasons (winter, spring, summer and fall). The error bars on the seasonal means are shown in blue in Fig. 4 and correspond to the 95 % confidence limits, $\mathrm{CI}_{95\%}$ calculated from Eq. (5).

The shaded areas represent the 95 % sampling uncertainty (range of the biased seasonal means, defined as $\pm 2.6 \cdot s^{\mathrm{samp,}N_f}_{\mathrm{seas,yr}}$ in Sect. 3.5.2) at the two frequencies $N_f = 4$ and 12 (in orange and red respectively). As expected, the sampling uncertainty on the seasonal mean increases when the sampling frequency decreases. The results show that the sampling uncertainty is larger than the 95 % confidence limit on the"true" seasonal mean, especially for the lower frequency $N_f = 4$.

We use the climatological sampling uncertainty ($2.6 \cdot s^{\mathrm{samp},N_\mathrm{f}}_{\mathrm{seas,clim}}$ defined in Sect. 3.5.2) to draw a more general picture of the results for each season. Table 2 summarizes the values of these climatological sampling uncertainties in percentages relative to the true value of the seasonal mean for four pressure levels. In addition, Fig. 5 shows the vertical profiles of the climatological sampling uncertainty in orange and red solid lines for the frequencies $N_\mathrm{f} = 4$ and 12 respectively. The sampling uncertainty on the seasonal mean ranges between 9 to 29 % for the 4 profile-a-month data sets as compared to 5–15 % for the 12 profile-a-month data sets. For surface ozone, the narrowest ranges in ppb are observed in the winter and fall. However as these months have lower ozone concentrations, the uncertainty on the seasonal mean represents up to 15 and 29 % of the true value for a frequency of 4 and 12 profiles a month respectively. In the free troposphere, the lowest uncertainty is found in winter. The sampling uncertainty on the seasonal mean, as calculated in our study, is higher than 10 % at the lowest time resolution, except between 700 and 500 hPa for most of the seasons. For a 12 profile-a-month frequency, the sampling uncertainty generally drops below 10 % in the free troposphere. The lowest sampling uncertainty is observed in the free troposphere at 700 hPa. This result suggests that over Frankfurt the 700 hPa level is the best candidate for comparing observations to other observations or to models and limiting the bias due to different sampling frequencies. At this level, the sampling uncertainty is 4.6, 4.2, 5.5, 5.6 % for winter, spring, summer and fall respectively for a 12 profile-a-month dataset (8.6, 9.0, 10.8, 8.7 % for 4 profile-a-month).

The sampling uncertainty on the seasonal means derived from the "random" sampling method is generally similar but slightly higher than the values derived from the "regular" sampling method (not shown). The estimates of the sampling uncertainty from this method are generally within 4 percentage points (unit of an arithmetic difference of two percentages) of the values presented for the "regular" method in Table 3, with exceptions in the lowest levels where the differences may reach 17 percentage points.

In Fig. 5 we also compare the sample standard deviation at different frequencies. For the entire morning data set, the sample standard deviation ($s_{\mathrm{seas,yr}}$) is estimated for each year (for each season and pressure level) and we plot the average of these estimates across the years in dot-dashed blue line. For the frequency $N_\mathrm{f} = 4$ or 12, the sample standard deviation ($s^{i,N_\mathrm{f}}_{\mathrm{seas,yr}}$) is estimated for each subsample and each year (for each season and pressure level), and we plot the average of these estimates across the subsamples and the years in a dot-dashed orange or red lines. The results show that the average sample standard deviation is similar whatever the sampling frequency, although a bit higher when considering the high frequency data set. Also, the sample standard deviation is always higher than the sampling uncertainty on the seasonal means. Both metrics (sample standard deviation and sampling uncertainty) have a well-marked C-shape, showing higher variability of ozone concentrations in the boundary layer (air masses affected by fresh emissions, subject to dry deposition of ozone, turbulence) and in the upper troposphere (potential impact of stratosphere-troposphere exchanges). Higher variability between the profiles enhances the potential differences between subsamples and then makes the distributions $f^{N_\mathrm{f}}_{\mathrm{seas,yr}}(y^{i,N_\mathrm{f}}_{\mathrm{seas,yr}})$ broader at these levels compared to those in the middle troposphere.

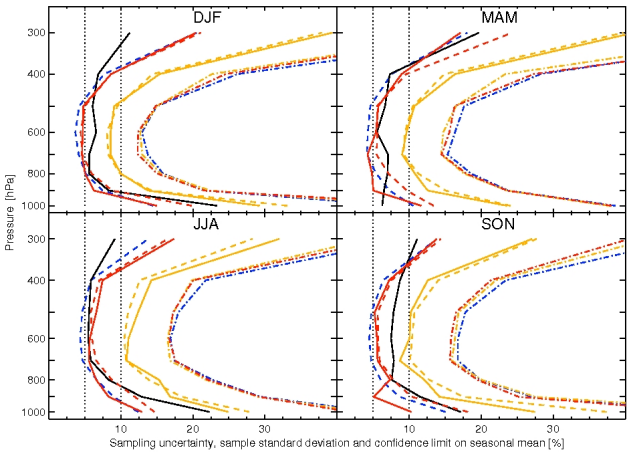

Fig. 5. Vertical profiles of the climatological sampling uncertainty ($2.6 \cdot s^{\mathrm{samp,N_f}}_{\mathrm{seas,clim}}$, see Sect. 3.5.2) derived from Frankfurt MOZAIC data for the four seasons (orange and red solid lines for the frequency of $N_\mathrm{f} = 4$ and 12, respectively). The blue dot-dashed line represents the average value of the sample standard deviation $s_{\mathrm{seas,yr}}$ derived from the whole morning data set (average across years for each season and level), and the orange and red dot-dashed line the average values of the sample standard deviation $s^{i,N_\mathrm{f}}_{\mathrm{seas,yr}}$ derived from the subsamples (average across years and across subsamples for each season, level and frequency). Similarly, the dashed lines represents the average estimate of the 95 % confidence limits on the seasonal means ($t_{0.05} \cdot \overline{se}_{\mathrm{seas,yr}}$ in blue and $t_{0.05} \cdot \overline{se}^{i,N_\mathrm{f}}_{\mathrm{seas,yr}}$ in orange and red). The black solid line shows the profiles of ozone inter annual variability as calculated in Sect. 4.2. The vertical dotted lines are the 5 and 10 % value lines, uncertainty values commonly quoted for measurement error of ozone sondes.

Assessment of the sampling uncertainty for any site requires high frequency data set, which is not feasible. In the following we suggest an easy-to-calculate estimate of the sampling uncertainty suitable for any tropospheric ozone data set. As presented in the Methodology, the 95 % confidence limit on the seasonal mean has been calculated for the MOZAIC morning data set as well as for any subsample, this confidence interval being easy to calculate. We take the average of these values across the subsamples and across the years to derive an estimate of the mean 95 % confidence limit on the seasonal mean for both frequencies. These estimates are plotted as orange and red dashed lines in Fig. 5. The results show that the average 95 % confidence limit on

the biased seasonal mean is close to the sampling uncertainty we derived in this study (an absolute difference of less than 3 percentage points). This also means that the 95 % confidence interval (as defined in this study) of the seasonal mean produced by the subsample most probably contains the true mean value. As expected, this average 95 % confidence limit on the seasonal mean using the entire morning data set of profiles (blue dashed line) is much lower than the confidence limit using fewer profiles.

4.2 Sampling effect on observed annual and inter-annual variabilities

Our results show that different low time frequency samples may show substantially different seasonal means and suggest that the derived annual and inter annual variabilities may be biased by the sampling.

In Fig. 4, we observe that the seasonal cycle is well marked in the entire morning data set. The seasonal differences of the long-term means between the cold and the warm months (DJF vs. JJA) are 41, 33, 28 and 72 % of the cold month concentrations (22 to 42 % of the warm months), respectively for the four pressure levels considered (from the top to the surface, see Fig. 4). These differences are higher than the sampling uncertainty on the seasonal means (Table 3), meaning that the seasonal cycle can be distinguished even when using the low frequency measurements.

The variability of ozone concentrations from one year to another (for each season) is calculated from the morning MOZAIC subset as $(\overline{x}_{\mathrm{seas,yr+1}} - \overline{x}_{\mathrm{seas,yr}})/\overline{x}_{\mathrm{seas,yr}}$. On average, over the four seasons, the inter-annual variability (IAV) is below 8 % in the free troposphere and ranges between 7 and 20 % in the two lowest levels (black solid lines in Fig. 5). As a result, the observed IAV signal is generally higher than the 95 % confidence limit of the seasonal means derived from the MOZAIC morning data set (dashed blue line), except at the highest altitude levels. This suggests that a high frequency data set may be used to disclose inter-annual variability in the tropospheric ozone. When using a data set at a frequency of 12 profiles a month, this capacity is reduced. For the frequency $N_{\mathrm{f}} = 4$, the sampling uncertainty is much higher than the inter-annual variability, leading to an uncertain IAV in such low frequency data sets. Consequently, except for extreme events, the IAV signal might possibly be masked by the sampling effect and the observed IAV signal will be highly dependent on the sampling, especially at the lowest time resolution over this region.

4.3 Sampling uncertainty versus measurement uncertainties

First, it is worth noting that the MOZAIC instrument uncertainty is typically 2–3 ppb for a concentration lower than 50 ppb (5 %), which is lower than the sampling uncertainty on the seasonal mean.

The accuracy of ozone sonde measurements is often quoted as ±5 % (Smit and Kley, 1998). A series of experiments evaluated the sonde performance and indicated a precision of better than ±(3–5) % and an accuracy of about ±(5–10) % up to 30 km altitude if standard operating procedures for ECC sondes are used (Smit et al., 2007). These values are represented on Fig. 5 in dotted lines. The sampling uncertainty as estimated in our study is always higher than 5 % accuracy. Between 900 and 400 hPa, the sampling uncertainty at the frequency of $N_{\mathrm{f}} = 12$ is within the accuracy range of 5–10 %. At a frequency of $N_{\mathrm{f}} = 4$, the sampling uncertainty is generally higher than 10 %, except around 700 hPa.

4.4 Sampling effect on ozone trends

To assess the effect of sampling frequency on ozone trends, we calculated the linear trend over the period 1998–2008. This time period is shorter than the MOZAIC period, but corresponds to a period over which the sonde and MOZAIC measurements agree the best (Logan et al., 2012; Tilmes et al., 2011). Seasonal ozone trends over the period 1998–2008 are derived from the whole morning MOZAIC data set using a weighted linear regression. For each seasonal mean, the standard error on the mean ($\overline{se}_{\mathrm{seas,yr}}$) is used as an error measurement in the linear regression; the weight put on a seasonal mean is then the inverse of the square root of the standard error. Using the same approach, linear trends are also derived from the measurements made at the six European sonde sites and the 48 EMEP surface stations. Weighting the seasonal means with the standard error allows us to take into account the uncertainty for each of them. The weighting greatly raises the uncertainty estimate of the trend, but the trend magnitude remains unchanged. As a consequence, the 1-sigma uncertainty of the trend is highly dependent on the standard error of the mean $\overline{se}_{\mathrm{seas,yr}}$ used for the weighting, and therefore dependent on the number of data. Figure 6 displays the distribution of the trends and the 1-sigma uncertainty estimates of these trends for the whole morning MOZAIC data set (black diamond), the European sondes (blue stars) and the surface station (black plus). In order to visualize the significance of the trend, the dashed lines corresponding to $y = x$ and $y = x/2$ (i.e. 1-sigma = slope and 2-sigma = slope) are added; this means that markers falling between the dashed lines represent trends that are not statistically significant.

The MOZAIC morning subset shows positive trends in winter and fall in the lowest level, while the surface ozone trend in summer is negative ($-0.3\,\mathrm{ppbv\,yr^{-1}}$). The trends derived using all MOZAIC profiles (and not only morning profiles) show a negative summer trend of around $-0.7\,\mathrm{ppbv\,yr^{-1}}$. These results for surface ozone are in agreement with previous studies (Ordóñez et al., 2005; Zbinden et al., 2006; Jonson et al., 2006; Oltmans et al., 2006; Jeannet et al., 2007; Gilge et al., 2010). They most probably result from the decrease in ozone precursor emissions during

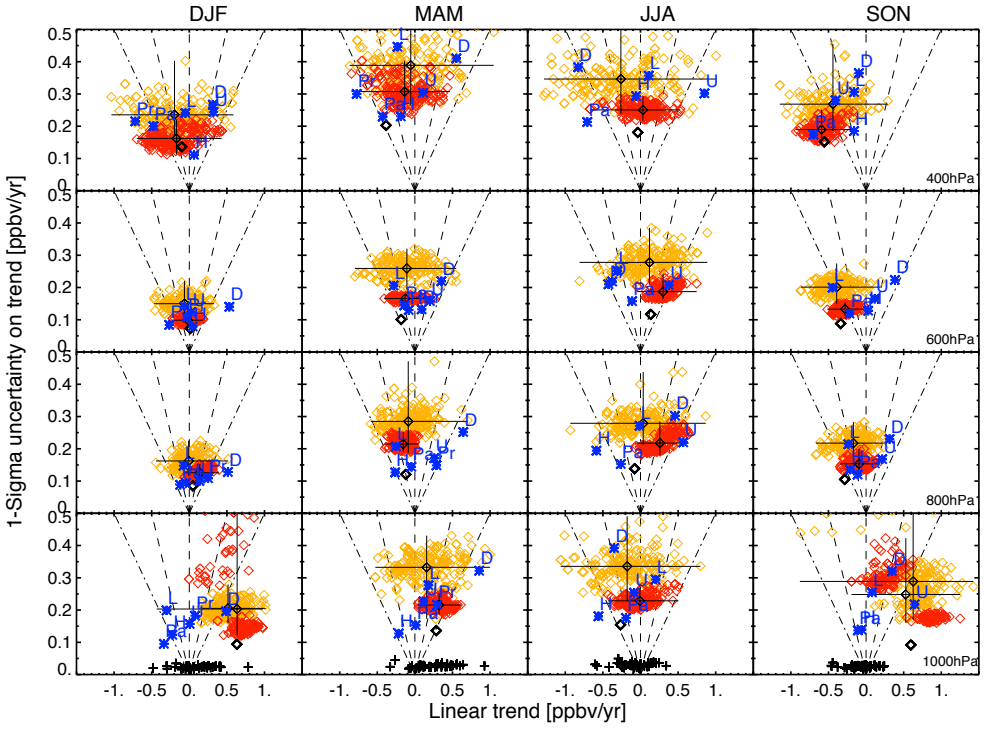

Fig. 6. 1-sigma uncertainty estimate of the linear trend against the linear trend for the four seasons and at four pressure levels (1000, 800, 600, 400 hPa, from bottom to top). Black plus symbols give the trends at 48 surface stations. The linear trend is calculated over the period 1998–2008. Blue stars show the trends derived from ozone sondes at the six stations (D = Debilt, H = Hohenpeissenberg, L = Lindenberg, Pa = Payerne, Pr = Praha, U = Uccle) at 1000 hPa except for H and Pa, value at 900h Pa. The black diamond corresponds to the trends derived from the whole morning MOZAIC data set in Frankfurt. For each frequency, 200 random time series were created from the ensemble of Frankfurt subsamples. The resulting trends are given by the cloud of diamonds in red (frequency of 12 profiles a month) and orange (4 profiles a month). The mean, minimum and maximum of these distributions are shown with the vertical and horizontal black thin lines. The dashed and dot-dashed lines show the 67 % and 95 % confidence limit lines for the slope.

this period. In the cold months, the trend probably results from a reduced ozone titration by nitrogen monoxide. During summer, the decrease in ozone precursors probably leads to a weaker photo-chemical ozone production during pollution episodes. Surface stations give the lowest uncertainty in the slope due to their large amount of data. Most suggest a positive trend in winter and spring, whereas in summer and fall, trends appear more scattered around zero. The seasonal trends vary with the altitude of the stations (not shown). Above 1 km, the results suggest a negative trend in summer, positive in winter and spring, and a near-zero trend during the fall season, in agreement with MOZAIC measurements in Frankfurt. Trends derived from the sondes are also scattered around zero, in a range similar to the surface stations. Local effects in the boundary layer prevent us from drawing any conclusions without proper analysis of the vicinity of each site (which is beyond the scope of this study). However, this result highlights the range of the surface ozone trend in Europe.

In the free troposphere at 600 hPa, the seasonal trends derived from MOZAIC are weaker, and not always statistically significant, with negative trends in fall and spring and positive trends in summer and winter. These results are in agreement with the recent study of Logan et al. (2012). For the sondes (blue stars), as the number of profiles is lower, the uncertainty in the estimate of the trend is larger, leading generally to insignificant trends. Some sonde measurements are in general agreement (same sign) with MOZAIC (e.g. Lindenberg, Hohenpeisenberg, Payerne) while others are not (e.g. Debilt and Uccle in the free troposphere). Other studies have already highlighted such discrepancies between European sites (Oltmans et al., 2006; Logan et al., 2012), and Logan et al. (2012) suggest that unusually high or low ozone concentrations measured at these sites in some years are responsible for these differences. In the upper level, the trends are more scattered around zero, except in fall when all present a negative trend (although of little significance for most of the sonde sites).

However, we have showed in the previous section that the observed seasonal means may be significantly impacted by the sampling frequency, especially at a frequency of $N_f = 4$. Thus there could be a potential effect of sampling on the observed ozone trend. To quantify this potential effect, we have created 200 random time series at each frequency ($N_f = 4$

or 12). Each time series is built as follows: for each of the 11 yr, a seasonal subsample is randomly selected among the $n_{\mathrm{seas,yr}}^{N_\mathrm{f}}$ available subsamples. As a result, we calculate an ensemble of 200 linear trends for each sampling frequency. The linear trends are calculated in the same manner as described above for the MOZAIC morning data set. These 200 low frequency MOZAIC trends are over-plotted in Fig. 6 in red and orange diamonds for the 12 and 4 profile-a-month frequencies, respectively. The mean, maximum and minimum values of these ensemble of trends and their estimates are represented by the thin black lines over the cloud of points. The mean values of the trends derived from the ensembles are generally similar to the trends derived from the full MOZAIC morning subset (black diamond) (within 0.2 ppbv yr^{-1}). As expected, the scattering of the trends is greater at a frequency of 4 profiles a month than at 12 profiles a month. Winter and fall trends from the MOZAIC morning subset being well pronounced in the lowest level, the distribution of the subsample trends remains mainly in the positive quadrant. The narrowest scattering in the cold months is due to the lower variability of ozone (see Sect. 4.1). In summer and spring, the higher variability and the less marked trends result in a larger scattering around a null trend. Comparing the sonde trends (blue stars) to the Frankfurt subsample trends (clouds of red and orange diamonds), we observe that the uncertainty estimate of the trend of a given sounding station is close to that of the ensemble at a similar sampling frequency. Obviously, this results from the measurement frequency at each station (close to 4–7 profiles a month for Debilt, Lindenberg and Praha and around 12 a month over Hohenpeissenberg, Payerne and Uccle). Also, the sonde trend markers fall surprisingly well within the red and orange clouds of Frankfurt subsamples, except for winter near the surface. If we consider the free troposphere only (altitude above 800 hPa), our results show that the linear trend derived from a subsample of Frankfurt data set could yield a value similar to those derived from any of the sondes, either positive or negative, in agreement or not with the MOZAIC morning data set. The trends extracted from low frequency data sets over the 1998–2008 period can be highly biased and not representative even if apparently significative. Thus our study suggests that the apparent aforementioned discrepancies observed between sondes, as well as between sondes and MOZAIC, may be attributed to sampling frequency, even though geophysical variations or differences in measurement strategy cannot be ruled out. However, it is worth noting that our results apply for this specific time period (1998–2008), which presents small variations of ozone concentrations. We applied the same approach using the subsamples created with the "random" sampling method. The main characteristics of the distributions of trends obtained from the 200 random time series are generally similar to those using the "regular" sampling method (not shown).

5 Generalization to the Northern Hemisphere midlatitudes

In this section, we aim to generalize our results to the Northern Hemisphere midlatitudes. The Frankfurt data set was the best candidate to start this study, since more than two profiles per day were collected. However, other cities in the Northern Hemisphere are well documented, such as Vienna, Paris, New York, Boston, Tokyo and Osaka. The number of profiles collected per season over these cities is summarized in Table 4.

For Frankfurt, Vienna, Paris and New York, the average number of profiles collected per season and per year allows subsampling these data sets at the two typical ozone sonde frequencies (more than 60 profiles per season on average). The data sets over Boston, Tokyo and Osaka have a lower frequency and thus can be subsampled only at the 4 profile-a-month frequency. In this section, all the profiles are kept without time filtering to retain the greatest number of profiles. A test performed for Frankfurt shows that the time filtering affects only the lowest levels and has little influence in the free troposphere (not shown). Also, we use the "random" sampling method in order to produce the maximum number of subsamples. The choice of sampling method reveals no significant impact on the results obtained for Frankfurt in Sect. 4; therefore we argue that applying this method here is appropriate. We use the same methodology as for the Frankfurt morning data set. We assume similar autocorrelation and use the ratio derived in Sect. 3.1 to calculate the effective sample size for the seven cities.

For these seven cities, we derive estimates of the climatological sampling uncertainty (defined as $2.6 \cdot s_{\mathrm{seas,clim}}^{\mathrm{samp},N_\mathrm{f}}$ in Sect. 3.5.2) in the same way as for the Frankfurt morning data set and then plotted these values against pressure levels in Fig. 7, color coded by cities. The vertical profiles for Vienna, Paris, New York and Boston are similar to those for Frankfurt in regards to the shape and the order of magnitude. As expected, the sampling uncertainty is higher at the lowest frequency for all these sites. They all show that the sampling uncertainty at a 12 profile-a-month frequency is lower than the 10 % measurement accuracy in the free troposphere (around 8 % between 500 and 800 hPa), while the sampling uncertainty at 4 profile-a-month frequency is systematically greater than the measurement accuracy of 10 %. For a frequency of 4 profiles a month, the sampling uncertainty is around 10–18 % between 500 and 800 hPa. These results also suggest a greater sampling uncertainty for Tokyo and Osaka (20 to 30 % in JJA and SON) than for Europe and North America in the summer and fall. This is in agreement with larger ozone distributions observed over the sonde stations at Tateno and Kagoshima in Japan for the same seasons (Tilmes et al., 2011). In addition, the mean sample standard deviation (average across subsamples and years) is much higher in summer and fall at the two Japanese sites than for the other five MOZAIC sites, which present a standard deviation

Table 3. Climatological sampling uncertainty, defined as the $2.6 \cdot s^{\mathrm{samp,N_f}}_{\mathrm{seas,clim}}$ (see Sect. 3.5.2 for details). Values are given for the "regular" sampling method in percentages relative to the overall seasonal mean at both sampling frequencies ($N_\mathrm{f} = 4$ or 12 profiles a month) and at four pressure levels.

	Frequency N_f	Winter	Spring	Summer	Fall
400 hPa	4	15.4	16.4	14.2	12.6
	12	8.6	9.0	7.5	7.2
600 hPa	4	8.6	9.6	11.1	10.1
	12	4.7	5.4	5.7	5.5
800 hPa	4	10.0	10.8	15.3	12.1
	12	5.0	4.9	6.6	7.4
1000 hPa	4	29.2	23.6	24.9	27.5
	12	15.0	11.7	12.9	10.3

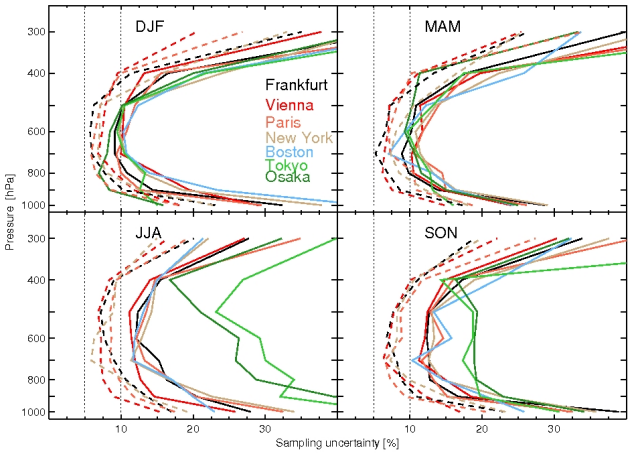

Fig. 7. Vertical profiles of the climatological sampling uncertainty in percentages ($2.6 \cdot s^{\mathrm{samp,N_f}}_{\mathrm{seas,clim}}$) for the seven northern mid-latitude sites, color-coded by cities. The sampling uncertainty at the frequency $N_\mathrm{f} = 4$ (12) is shown in solid (dashed) lines. As stated in the text, only Frankfurt, Vienna, Paris and New York are shown for the frequency $N_\mathrm{f} = 12$. The vertical dotted lines are the 5 and 10 % value lines, uncertainty values commonly quoted for measurement error of ozone sondes.

similar in magnitude to that of Frankfurt (not shown). This Japanese region is influenced either by the pollution emitted by biomass burning in Siberia and China (Streets et al., 2003) or by poor ozone air masses transported from the tropics (Logan, 1985). Moreover, there are large latitudinal gradients in ozone over Japan in the summer and autumn (Logan, 1999), so that ozone concentrations measured by the aircraft depend on their routes into and out of these airports. As a result, the variability sampled by the aircraft over Japan depends on the dynamic regime under which the site is at the time of the sampling (influence of monsoon circulation and convective systems, transport of midlatitude air masses). This leads to a greater variability of ozone concentrations in the Japanese free troposphere which largely impacts the distributions $g^{N_\mathrm{f}}_{\mathrm{seas}}(y^{i,N_\mathrm{f}}_{\mathrm{seas,yr}})$. Our results might also be biased by the smaller number of profiles available for the two Japanese sites as compared to the North American and European sites. However, the results found for Boston, similar to Frankfurt even though even fewer profiles were available, tend to corroborate our findings for Tokyo and Osaka.

We also compared the vertical profiles of the average sampling uncertainty ($2.6 \cdot s^{\mathrm{samp},N_\mathrm{f}}_{\mathrm{seas,clim}}$) with the vertical profiles of the average 95 % confidence limit on the seasonal mean ($\mathrm{CI}_{95\,\%}$) for each of these cities (as in Fig. 5). For the sake of clarity, we do not show these profiles in Fig. 7. As found for Frankfurt in Sect. 4.3, the sampling uncertainty for both frequencies is higher or similar to the 95 % confidence interval on the biased seasonal mean for all stations in the Northern Hemisphere (difference of 1–3 percentage points on average between 800 and 500 hPa). The 95 % confidence interval on the biased seasonal mean could be used as an estimate of the sampling uncertainty, although it may underestimate the sampling uncertainty slightly.

To conclude this section, the results derived from the detailed study undertaken for Frankfurt in Sect. 4 can be extended to other northern midlatitude sites in Europe and North America, which are not influenced by strong changes in air mass composition (such as tropical air masses). For sites more similar to Tokyo and Osaka in terms of ozone variability, the sampling uncertainty is much higher during the summer and fall seasons. As a consequence, we suggest a careful interpretation of the observed ozone means (and ozone variations) over Japan.

6 Tropical case: Windhoek, Namibia

To extend our discussion to the tropics, we use the daily data collected over Windhoek, Namibia (22° S, 17° E). Windhoek is located in the Khomas Highland plateau area (around 1700 m a.s.l.). Its international airport have been visited by the MOZAIC aircraft under the carrier Air Namibia since December 2005. We use here the measurements collected between December 2005 and November 2008. During this three year period, there were 250, 262, 267 and 263 profiles collected over Windhoek in winter, spring, summer and fall respectively (leading to around one profile per day). The random sampling method was applied to the Windhoek data set. The time period recorded over Windhoek is shorter than for Frankfurt, but the frequency is high. Thus the results presented should be representative of the ozone variability in this region (except for the inter annual variability).

The results are presented in Fig. 8, similar to Fig. 5 for Frankfurt. Over Windhoek, the vertical profiles of the sampling uncertainty and the sample standard deviation do not present the well-marked C-shape profiles seen in Figs. 5 and 7. At the highest levels, the ozone variability observed

Table 4. Number of profiles collected per season over the 14 yr period 1995–2008 by the MOZAIC aircraft over Frankfurt, Vienna, Paris, New York, Boston, Tokyo and Osaka. Numbers in parenthesis are the average numbers of profiles collected per season per year.

City (Period)	DJF	MAM	JJA	SON	Total
Frankfurt (1995–2008)	3013 (215)	3026 (216)	3406 (243)	3231 (231)	12 676
Vienna (1995–2006)	867 (62)	1136 (81)	1517 (108)	1245(89)	4765
Paris (1995–2004)	1040 (74)	961 (69)	1062 (76)	1090 (78)	4153
New York (1995–2006)	762 (54)	778 (56)	846 (60)	863 (62)	3249
Boston (1995–2006)	198 (14)	190 (14)	332 (24)	298 (21)	1018
Tokyo (1995–2006)	307 (22)	410 (29)	455 (33)	346 (25)	1518
Osaka (1995–2006)	293 (21)	349 (25)	400 (29)	409 (29)	1451

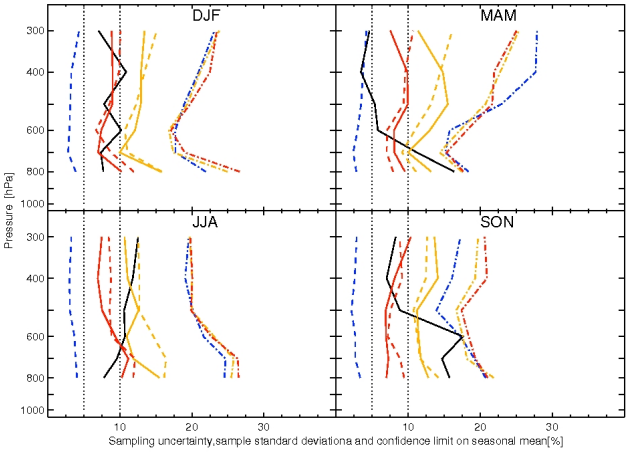

Fig. 8. Same analysis as in Fig. 5 (Frankfurt) but for Windhoek.

from the Windhoek profiles is much lower than those from the Northern Hemisphere sites. This could be the result of a lesser influence of stratospheric mixing at these altitudes, as the tropopause is much higher in the tropics (around 100 hPa). At the lowest levels, the ozone levels and variability are also lower than those seen over the Northern Hemisphere cities.

Regarding the sample standard deviation (dot-dashed lines), the profiles are similar whatever the frequency. The seasons DJF and MAM show different shapes than in JJA and SON. These differences could be linked to the migration of the inter tropical convergence zone (ITCZ), leading to meteorological and chemical differences between the wet and dry seasons respectively. In the 600–300 hPa layer, the sample standard deviation is enhanced (up to 30 %) during the winter and fall. During these periods of the year, the ITCZ, located in the Southern Hemisphere, is associated with deep convection, resulting in significant emissions of NO_x from lightning (e.g. Bond et al., 2002). These irregular convective systems contribute to the modulation of ozone production in the upper troposphere (e.g. Edwards et al., 2003; Sauvage et al., 2007) and hence lead to higher ozone variability compared to dry months (such as JJA) over Windhoek (J.-P. Cammas, personal communication, 2011). From the surface to 300 hPa, the sampling uncertainty calculated for Windhoek is around 8 % and 12 % for the 12 and 4 profile-a-month frequencies respectively. These values are similar to what was found in the free troposphere (between 800 and 500 hPa) at the Northern Hemisphere midlatitudes.

For this tropical site, the sampling uncertainty and the 95 % confidence interval of the subsample seasonal mean are close (a difference of less than 5 percentage points). Here, the IAV is of the same magnitude or lower than the sampling uncertainty for both frequencies, except in the lowest levels in SON. Fall is the burning season in this region (e.g. Sauvage et al., 2007), and Windhoek is under the influence of important sources of ozone precursors from biomass burning, the magnitude of which may vary from one year to another. However, further study would be needed to better understand the processes controlling the ozone vertical distribution in this area, which is beyond the scope of this study.

7 Conclusions

We have used high frequency MOZAIC data sets to discuss the effect of sampling in the analysis of ozone vertical profiles in order to estimate the uncertainty that arises when using low time resolution data sets such as ozone sondes. We subsampled the MOZAIC profiles at two typical ozone sonde frequencies, 4 and 12 profiles a month. We performed a detailed analysis using the Frankfurt data set, as this is the best documented airport. In addition, we used other northern midlatitude sites to generalise our findings, and the Windhoek, Namibia data set to discuss a tropical case.

We defined the climatological sampling uncertainty as $2.5 \cdot s^{\mathrm{samp},N_f}_{\mathrm{seas,clim}}$, where $s^{\mathrm{samp},N_f}_{\mathrm{seas,clim}}$ is the standard deviation of the distribution of the differences between the subsample seasonal means and the overall mean. This metric has been derived per season and per pressure levels. As expected, the sampling uncertainty is higher at the lower time resolution.

The vertical profiles of the average sample standard deviation have a well-marked C-shape for all the Northern Hemisphere sites, which suggests higher variability of ozone in the lowest and highest levels, probably due to local anthropogenic pollution events and the potential impact of

stratosphere-troposphere exchange, respectively. As a result, the sampling uncertainty presents a similar shape. The lowest uncertainty is found in the free troposphere at 700 hPa, with values around 5 and 10 % for the 12 and 4 profile-a-month frequencies respectively over Frankfurt. As a consequence, this level is the best candidate for observation comparison and model evaluation purposes in the northern mid-latitudes. For the tropical case (Windhoek), the sampling uncertainty in the free troposphere is around 8 and 12 % at 12 and 4 profiles a month respectively.

We found that: (1) at a 12 profile-a-month frequency, the sampling uncertainty drops below the measurement accuracy (5–10 %) in the free troposphere, while at 4 profile-a-month the sampling uncertainty is generally higher than the measurement accuracy and should be considered, (2) the sample standard deviation remains the same whatever the sampling frequency, (3) the climatological sampling uncertainty is similar to the average 95 % confidence limit on the subsample seasonal mean detected by a subsample, and (4) the 95 % confidence interval of the seasonal mean produced by a sample (as derived in this study) will most probably contain the true mean value and should be used for observation to model or observation comparisons.

We discussed the accuracy of low time resolution measurements to detect ozone variations at different time scales. Over Frankfurt, we concluded that: (1) the seasonal cycle is well observed even at the lowest frequency, (2) the IAV signal is generally too low, and consequently masked by the sampling effect, (3) the trend derived over the 11 yr period 1998–2008 varies significantly in magnitude and even in sign with the samples. As a consequence, apparent discrepancies between sites might be attributed to a low frequency sampling in some cases (even though geophysical variations or differences in measurement strategy may also interfere).

These results are valid for the European region, which is under the influence of similar air masses as those of Frankfurt. They might also apply to other northern midlatitude sites such as in North America.

Our results suggest that the sampling frequency should be taken into account for further discussion of observed variations and for observation comparisons or model evaluation. This also strengthens the need to have model outputs on the dates of the sondes for model evaluation, as monthly mean comparisons may be biased by the sampling.

To conclude, this study highlights the significant effect of sampling when using low time resolution measurements. We provide estimates of the sampling uncertainty that arises from such data sets, and we believe that these estimates should be considered for observation comparison and model evaluation. This study strengthens the need for regular and high frequency measurements of tropospheric ozone (at least every two or three days). These high frequency ozone observations are the key to obtaining accurate observations of the inter annual variability and decadal changes in ozone concentrations and to better understand changes in ozone concentrations.

Acknowledgements. The National Center for Atmospheric Research is operated by the University Corporation for Atmospheric Research under the sponsorship of the National Science Foundation. The authors acknowledge the strong support of the European Commission, Airbus, and the airlines (Lufthansa, Austrian, Air France) who carry free of charge the MOZAIC equipment and perform the maintenance since 1994. MOZAIC is presently funded by INSU-CNRS (France), Météo-France, and Forschungszentrum (FZJ, Jülich, Germany). The MOZAIC data base is supported by ETHER (CNES and INSU-CNRS). The authors are grateful to all the agencies that provide ozone sonde data to the World Ozone and Ultraviolet Radiation Data Centre (WOUDC): German Weather Service – Meteorological Observatory at Hohenpeissenberg and Lindenberg, the National Meteorological Institute of the Netherlands, MeteoSwiss, the Royal Meteorological Institute of Belgium and the Czech HydroMeteorological Institute. The authors are grateful to the two anonymous referees who helped to improve the manuscript.

Edited by: D. Shindell

References

Bond, D. W., Steiger, S., Zhang, R., Tie, X., and Orville, R.: The importance of NOx production by lightning in the tropics, Atmos. Environ., 36(9), 1509–1519, 2002.

Chipperfield, M. P., Fioletov, V. E., Bregman, B., Burrows, J., Connor, B. J., Haigh, J. D., Harris, N. R. P., Hauchecorne, A., Hood, L. L., Kawa, S. R., Krzyscin, J. W., Logan, J. A., Muthama, N. J., Polvani, L., Randel, W. J., Sasaki, T., Stahelin, J., Storlarski, R. S., Thomason, L. W., and Zawodny, J. M.: Global Ozone: Past and Present, Chapter 3 in Scientific 25 assessment of ozone depletion: 2006, Global Ozone Research and Monitoring Project-Report No. 50, World Meteorological Organization, Geneva, Switzerland, 2007.

Cooper, O.R., Parrish, D. D., Stohl, A., Trainer, M., Nédélec, P., Thouret, V., Cammas, J.-P., Oltmans, S. J., Johnson, B. J., Tarasick, D., Leblanc, T., McDermid, I. S., Jaffe, D., Gao, R., Stith, J., Ryerson, T., Aikin, K., Campos, T., Weinheimer, A., and Avery, M. A.: Increasing springtime ozone mixing ratios in the free troposphere over western North America, Nature, 463, 344–348, doi:10.1038/nature08708, 2010.

Eckhardt, S., Stohl, A., Beirle, S., Spichtinger, N., James, P., Forster, C., Junker, C., Wagner, T., Platt, U., and Jennings, S. G.: The North Atlantic Oscillation controls air pollution transport to the Arctic, Atmos. Chem. Phys., 3, 1769–1778, doi:10.5194/acp-3-1769-2003, 2003.

Edwards, D. P., Lamarque, J. F., Attié, J. L., Emmons, L. K., Richter, A., Cammas, J. P., Gille, J. C., Francis, G. L., Deeter, M. N., Warner, J., Ziskin, D. C., Lyjak, L. V., Drummond, J. R., and Burrows, J. P.: Tropospheric Ozone Over the Tropical Atlantic: A Satellite Perspective, J. Geophys. Res., 108, 4237, doi:10.1029/2002JD002927, 2003.

Emmons, L. K., Hauglustaine, D., Muller, J.-F., Caroll, M., Brasseur, G. P., Brunner, D., Staehelin, J., Thouret, V., and

Marenco, A.: Data composites of airborne observations of tropospheric ozone and its precursors, J. Geophys. Res., 105, 20497–20538, 2000.

Eyring, V., Shepherd, T. G., and Waugh, D. W.: SPARC Report on the Evaluation of Chemistry-Climate Models, SPARC Report No. 5, WCRP-132, WMO/TD-No.1526, 2010.

Fusco, A. C. and Logan, J. A.: Analysis of 1970–1995 trends in tropospheric ozone at Northern Hemisphere midlatitudes with the GEOS-CHEM model, J. Geophys. Res., 108, 4449, doi:10.1029/2002JD002742, 2003.

Gilge, S., Plass-Duelmer, C., Fricke, W., Kaiser, A., Ries, L., Buchmann, B., and Steinbacher, M.: Ozone, carbon monoxide and nitrogen oxides time series at four alpine GAW mountain stations in central Europe, Atmos. Chem. Phys., 10, 12295–12316, doi:10.5194/acp-10-12295-2010, 2010.

Jeannet, P., Stübi, R., Levrat, G., Viatte, P., and Staehelin, J.: Ozone balloon soundings at Payerne (Switzerland): Reevaluation of the time series 1967–2002 and trend analysis, J. Geophys. Res., 112, D11302, doi:10.1029/2005JD006862, 2007.

Jonson, J. E., Simpson, D., Fagerli, H., and Solberg, S.: Can we explain the trends in European ozone levels?, Atmos. Chem. Phys., 6, 51–66, doi:10.5194/acp-6-51-2006, 2006.

Keating, T. J. and Zuber, A.: Task Force on Hemispheric Transport of Air Pollution, Air Pollut. Stud. 16, U.N. Econ. Comm. for Europe, New York, hemispheric transport of air pollution 2007 interim report, 2007.

Koumoutsaris, S., Bey, I., Generoso, S., and Thouret, V.: Influence of El Niño Southern Oscillation on the interannual variability of tropospheric ozone in the northern midlatitudes, J. Geophys. Res., 113, D19301, doi:10.1029/2007JD009753, 2008.

Lamarque, J.-F. and Hess, P. G.: Arctic Oscillation modulation of the Northern Hemisphere spring tropospheric ozone, Geophys. Res. Lett., 31, 6127, doi:10.1029/2003GL019116, 2004.

Logan, J.: Tropospheric ozone – seasonal behavior, trends, and anthropogenic influence, J. Geophys. Res., 90, 10463–10482, 1985.

Logan, J.: An Analysis of Ozone sonde Data for the Lower Stratosphere: Recommendations for Testing Models, J. Geophys. Res., 104, 16151–16170, 1999.

Logan, J. A., Staehelin, J., Megretskaia, I. A., Cammas, J.-P., Thouret, V., Claude, H., De Backer, H., Steinbacher, M., Scheel, H. E., Stübi, R., Fröhlich, M., and Derwent, M.: Changes in ozone over Europe since 1990: analysis of ozone measurements from sondes, regular aircraft (MOZAIC) and alpine surface sites, J. Geophys. Res., 117, D09301, doi:10.1029/2011JD016952, 2012.

Marenco, A., Thouret, V., Nédélec, P., Smit, H., Helten, M., Kley, D., Karcher, F., Simon, P., Law, K., Pyle, J., Poschmann, G., Von Wrede, R., Hume, C., and Cook, T.: Measurement of ozone and water vapor by Airbus in-service aircraft: The MOZAIC airborne program, An overview, J. Geophys. Res., 103, 25631–25642, 1998.

Monks, P. S., Granier, C., Fuzzi, S., Stohl, A., Williams, M. L., Akimoto, H., Amann, M., Baklanov, A., Baltensperger, U., Bey, I., Blake, N., Blake, R. S., Carslaw, K., Cooper, O. R., Dentener, F., Fowler, D., Fragkou, E., Frost, G. J., Generoso, S., Ginoux, P., Grewe, V., Guenther, A., Hansson, H. C., Henne, S., Hjorth, J., Hofzumahaus, A., Huntrieser, H., Isaksen, I. S. A, Jenkin, M. E., Kaiser, J., Kanakidou, M., Klimont, Z., Kulmala, M., Laj, P., Lawrence, M. G., Lee, J. D., Liousse, C., Maione, M., McFiggans, G., Metzger, A., Mieville, A., Moussiopoulos, N., Orlando, J. J., O'Dowd, C. D., Palmer, P. I., Parrish, D. D., Petzold, A., Platt, U., Pöschl, U., Prévôt, A. S. H., Reeves, C. E., Reimann, S., Rudich, Y., Sellegri, K., Steinbrecher, R., Simpson, D., ten Brink, H., Theloke, J., van der Werf, G. R., Vautard, R., Vestreng, V., Vlachokostas, Ch., and von Glasow, R.: Atmospheric composition change – global and regional air quality, Atmos. Environ., 43, 5268–5350, doi:10.1016/j.atmosenv.2009.08.021, 2009.

Naja, M., Akimoto, H., and Staehelin, J.: Ozone in background and photochemically aged air over central Europe: Analysis of long-term ozone sonde data from Hohenpeissenberg and Payerne, J. Geophys. Res., 108, 4063, doi:10.1029/2002JD002477, 2003.

Oltmans, S., Lefohn, A., Harris, J., Galbally, I., Scheel, H., Bodeker, G., Brunke, E., Claude, H., Tarasick, D., Johnson, B., Simmonds, P., Shadwick, D., Anlauf, K., Hayden, K., Schmidlin, F., Fujimoto, T., Akagi, K., Meyer, C., Nichol, S., Davies, J., Redondas, A., and Cuevas, E.: Long-term changes in tropospheric ozone, Atmos. Environ., 40, 3156–3173, doi:10.1016/j.atmosenv.2006.01.029, 2006.

Oltmans, S. J, Lefohn, A. S, Harris, J. M., Tarasick, D. W., Thompson, A. M., Wernli, H., Johnson, B. J., Novelli, P. C., Montzka, S. A., Ray, J. D., Patrick, L. C., Sweeney, C., Jefferson, A., Dann, T., Davies, J., Shapiro, M., and Holben, B. N.: Enhanced ozone over western North America from biomass burning in Eurasia during April 2008 as seen in surface and profile observations, Atmos. Environ.,44, 4497–4509, 2010.

Ordóñez, C., Mathis, H., Furger, M., Henne, S., Hüglin, C., Staehelin, J., and Prévôt, A. S. H.: Changes of daily surface ozone maxima in Switzerland in all seasons from 1992 to 2002 and discussion of summer 2003, Atmos. Chem. Phys., 5, 1187–1203, doi:10.5194/acp-5-1187-2005, 2005.

Ordóñez, C., Brunner, D., Staehelin, J., Hajinicolaou, P., Pyle, J. A., Jonas, M., Wernli, H., and Prévôt, A. S. H.: Strong influence of lowermost stratospheric ozone on lower tropospheric background ozone changes over Europe, Geophys. Res. Lett., 34, L07805, doi:10.1029/2006GL029113, 2007.

Parrish, D. D., Millet, D. B., and Goldstein, A. H.: Increasing ozone in marine boundary layer inflow at the west coasts of North America and Europe, Atmos. Chem. Phys., 9, 1303–1323, doi:10.5194/acp-9-1303-2009, 2009.

Rodriguez, S., Torres, C., Guerra, J.-C., and Cuevas, E.: Transport pathways of ozone to marine and free-troposphere sites in Tenerife, Canary Islands, Atmos. Environ., 38, 4733–4747, 2004.

Sauvage, B., Thouret, V., Cammas, J.-P., Brioude, J., Nédélec, P., and Mari, C.: Meridional ozone gradients in the African upper troposphere, Geophys. Res. Lett., 34, L03817, doi:10.1029/2006GL028542, 2007.

Simmonds, P. G., Manning, A. J., Derwent, R. G., Ciais, P., Ramonet, M., Kazan, V., and Ryall, D.: A burning question. Can recent growth rate anomalies in the greenhouse gases be attributed to large-sacle biomass burning event, Atmos. Environ., 39, 2513–2517, 2005.

Smit, H. G. J. and Kley, D.: The 1996 WMO International intercomparison ozone sondes under quasi flight conditions in the environmental chamber at Jülich, report, World Meteorol. Organ., Geneva, Switzerland, 1998.

Smit, H. G. J., Straeter, W., Johnson, B. J., Oltmans, S. J., Davies, J., Tarasick, D. W., Hoegger, B., Stubi, R., Schmidlin, F. J., Northam, T., Thompson, A. M., Witte, J. C., Boyd, I., and

Posny, F.: Assessment of the performance of ECC ozone sondes under quasiflight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE), J. Geophys. Res., 112, D19306, doi:10.1029/2006JD007308, 2007.

Solberg, S.,Hov, Ø., Søvde, A., Isaksen, I. S. A., Coddeville, P., De Backer, H., Forster, C., Orsolini, Y., and Uhse, K.: European surface ozone in the extreme summer 2003, J. Geophys. Res., 113, D07307, doi:10.1029/2007JD009098, 2008.

Streets, D. G., Yarber, K. F., Woo, J., and Carmichael, G.: Biomass burning in Asia: Annual and seasonal estimates and atmospheric emissions, Global Biogeochem. Cy., 17, 1099, doi:10.1029/2003GB002040, 2003.

Tarasick, D. W., Fioletov, V. E., Wardle, D. I., Kerr, J. B., and Davies, J.: Changes in the vertical distribution of ozone over Canada from ozone sondes: 1980–2001, J. Geophys. Res., 110, D02304, doi:10.1029/2004JD004643, 2005.

Thouret, V., Marenco, A., Logan, J. A., Nédélec, P., and Grouhel, C.: Comparisons of ozone measurements from the MOZAIC airborne program and the ozone sounding network at eight locations, J. Geophys. Res., 103, 25695–25720, doi:10.1029/98JD02243, 1998.

Tilmes, S., Lamarque, J.-F., Emmons, L. K., Conley, A., Schultz, M. G., Saunois, M., Thouret, V., Thompson, A. M., Oltmans, S. J., Johnson, B., and Tarasick, D.: Ozonesonde climatology between 1995 and 2009: description, evaluation and applications, Atmos. Chem. Phys. Discuss., 11, 28747–28796, doi:10.5194/acpd-11-28747-2011, 2011.

Vestreng, V., Mareckova, K., Kakareka, S., Malchykhina, A., and Kukharchyk, T.: Inventory Review 2007; Emission Data reported to LRTAP Convention and NEC Directive, Tech. Rep. MSC-W Technical Report 1/07, Norw. Meteorol. Inst., Oslo, 2007.

World Meterorological Organization: Technical Note No. 79: Climatic Change, WMO-No, 195.TP.100, Geneva, 80 pp., 1966.

Zbinden, R. M., Cammas, J.-P., Thouret, V., Nédélec, P., Karcher, F., and Simon, P.: Mid-latitude tropospheric ozone columns from the MOZAIC program: climatology and interannual variability, Atmos. Chem. Phys., 6, 1053–1073, doi:10.5194/acp-6-1053-2006, 2006.