ANSWERS TO REFEREE #1

We are very thankful to the referee for the numerous corrections and modifications he suggested. We have tried to carefully answer all the issues he/she raised. An additional analysis was included in the Supplement, and summarized in the manuscript. Note that following the suggestions of both referees, the manuscript has been substantially reformulated in order to improve its clarity.

For clarity, the suggestions/remarks of the referee are in italics, our direct answers in bold and our modifications in the manuscript in regular.

Summary

This paper presents an interesting analysis of the very extensive data set of vertical profiles of ozone and CO over the Frankfurt region collected by the MOZAIC-IAGOS program. The authors address: 1) Climatological vertical profiles of the mean O_3 and CO concentrations, 2) Seasonal variations, 3) Annual and seasonal O_3 trends, 4) Annual and seasonal CO trends, and 5) Changes in the O_3 seasonal cycle. The paper has been improved from the previous draft, but significant problems still remain that must be corrected before the paper is suitable for publication.

Although the data set considered is the densest record of vertical ozone and CO profiles in existence, the findings presented in this work are, on the whole, not new. In my opinion, the paper would be ready for publication when the first set of major issues and the minor and significant issues listed below have been addressed. However, I encourage the authors to address the second set of major issues, also listed below, which would make the paper more useful.

ANSWER : We took into account all the issues raised by the referee, see our answers below.

Major issues to be addressed before the paper is suitable for publication

1) p. 8, beginning on line 18. The authors write "The highest vertical gradients are found close to the surface all along the year (dry deposition and titration by NO) and at ...". In many continental locations over urban areas, high ozone concentrations are found near the surface, decreasing above the PBL, before increasing through the mid- to upper troposphere (sometimes described as a "C" shaped profile). Evidently the vertical profile over Frankfurt is different; this contrast should be discussed. One issue that should be included in that discussion is the influence of night-time titration of ozone. If the data are limited to daytime only profiles, is a "C" shaped profile seen?

ANSWER : Indeed, during the summer, high O_3 mixing ratios can be observed in the European boundary layer. We recently investigated the diurnal variations of ozone throughout the whole troposphere at Frankfurt (Petetin et al., Diurnal cycle of ozone throughout the troposphere over Frankfurt as measured by MOZAIC-IAGOS commercial aircraft, under review in Elementa Science of the Anthropocene). In the lower troposphere, the vertical gradients of seasonally-averaged O_3 are most of time positive, i.e. O_3 mixing ratios increase with the altitude. Close to the surface, the maximum O_3 mixing ratios are observed during the summer between 15 and 18 UTC. At this time, the vertical gradient is

greatly reduced compared to the vertical gradient in the early morning or at the end of the day (note that almost no observations are available between 0 and 3 UTC) but remains positive in the first 50 hPa above the surface. Between 950 and 700 hPa, the vertical gradient is close to zero and can reach slightly negative values, but the vertical profile of O_3 mixing ratios at this time does not show a clear "C" shaped profile. However, interestingly, such "C" shaped profile is very clearly observed if we consider the 95th percentile rather than the mean O_3 mixing ratios. We modified the text as follows : "During the summer, O_3 episodes are often observed in the European BL (van Loon et al., 2007; Meleux et al., 2007). High O_3 mixing ratios are also measured in urban environments, despite the presence of NO_x emitted locally by the anthropogenic activities (Vautard et al., 2007 reported a 95th percentile of daily O_3 maximum ranging between 70 and 100 ppb in 4 European megacities in 2010). Thus, one might have expected higher mixing ratios in the BL than in the lower free troposphere (sometimes described as a « C » shaped profile). However, observations do not show such a profile. One may suspect that this is due to the night-time titration of O_3 in the BL but limiting data to the afternoon does not highlight a clear « C » shaped profile. Actually, such « C » shaped profile is only observed when considering the 95th percentile rather than the mean O_3 mixing ratio (Petetin et al., Diurnal cycle of ozone throughout the troposphere over Frankfurt as measured by MOZAIC-IAGOS commercial aircraft, under review in Elementa Science of the Anthropocene). It means that the potentially high O_3 pollution in the BL during the summer can greatly modify the vertical profile of O_3 mixing ratios but only episodically. On average, the structure of the mean O_3 vertical profile in summer remains qualitatively the same (i.e. positive gradient through the whole troposphere) as during the rest of the year. ".

In the references :

van Loon, M., Vautard, R., Schaap, M., Bergström, R., Bessagnet, B., Brandt, J., Builtjes, P. J. H., Christensen, J. H., Cuvelier, C., Graff, a., Jonson, J. E., Krol, M., Langner, J., Roberts, P., Rouil, L., Stern, R., Tarrasón, L., Thunis, P., Vignati, E., White, L. and Wind, P.: Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble, Atmos. Environ., 41(10), 2083–2097, doi:10.1016/j.atmosenv.2006.10.073, 2007.

Meleux, F., Solmon, F. and Giorgi, F.: Increase in summer European ozone amounts due to climate change, Atmos. Environ., 41(35), 7577–7587, doi:10.1016/j.atmosenv.2007.05.048, 2007.

Vautard, R., Builtjes, P. H. J., Thunis, P., Cuvelier, C., Bedogni, M., Bessagnet, B., Honoré, C., Moussiopoulos, N., Pirovano, G. and Schaap, M.: Evaluation and intercomparison of Ozone and PM10 simulations by several chemistry transport models over four European cities within the CityDelta project, Atmos. Environ., 41(1), 173–188, doi:10.1016/j.atmosenv.2006.07.039, 2007.

2) p. 13, line 14 - The comparison with satellite data should be made more quantitative.

ANSWER : This part was modified as follows : "As shown in Fig. 5, the seasonal cycle of CO is characterized by maximum mixing ratios in late winter/early spring in the whole troposphere. Minimum mixing ratios are encountered in summer/early autumn in the LT and are slightly shifted to late summer/early autumn higher in altitude. Such a seasonal pattern is consistent with the seasonal variation observed in

background air masses arriving at the coastal site Mace Head (Derwent et al., 1998) or at a larger scale by satellite observations (Edwards et al., 2004; Worden et al., 2013). Averaged over the western Europe, the Terra/MOPITT CO tropospheric column maximizes at ~2.5 10^{18} molecules cm^{-2} in March-April and minimize at ~1.9 10^{18} molecules cm^{-2} in late summer, the ratio of the maximum over the minimum being 1.3 (Edwards et al., 2004). A very similar seasonal variation of tropospheric columns of CO has been observed by Zbinden et al. (2013) based on the MOZAIC data over the period 2002-2009. This is in good agreement with the amplitude of the seasonal cycle observed in the MT, the maximum CO mixing ratio being 1.35 higher than the minimum."

In the references :

Edwards, D. P., Emmons, L. K., Hauglustaine, D. A., Chu, D. A., Gille, J. C., Kaufman, Y. J., tron, G. P., Yurganov, L. N., Giglio, L., Deeter, M. N., Yudin, V., Ziskin, D. C., Warner, J., Lamarque, J.-F., Francis, G. L., Ho, S. P., Mao, D., Chen, J., Grechko, E. I. and Drummond, J. R.: Observations of carbon monoxide and aerosols from the Terra satellite: Northern Hemisphere variability, J. Geophys. Res., 109(D24), D24202, doi:10.1029/2004JD004727, 2004.

3) Section 3.3.1 - The discussion of ozone trends should be greatly simplified and reduced in length. Previous studies (Logan et al., 2012; Parrish et al., 2012) have quantified the longterm changes of ozone in the lower to mid-troposphere in the central European region. The lack of significant trends and the marginally significant trends found for ozone in Section 3.3.1 and Table 1 are completely consistent with previous work, and hence do not add any new information. For example, the only significant trend for the mean ozone in Table 1 is in the winter. The European winter shape factors given in Table 2 of Parrish et al. (2012) give an average trend of 0.61±0.25 %$O_{3,2000}$ yr^{-1}, which is statistically consistent with all of the O_3 trends found in this work. All of the derived ozone trends should be removed from Table 1 (perhaps included in the Supplement) and only a very short discussion of the ozone trends included. In this regard, picking a specific sub-period for analysis (here the 2000-2012 period) is statistically a very dangerous process, since relatively large trends can result from interannual variability over short periods (see for example, Barnes et al., 2016). The discussion of the trends over the sub-period should be eliminated.

ANSWER : We removed the trend results of O_3 in Table 1, as well as the discussion of trends over the period 2000-2012, and simplified this discussion as follows :

"All the annual and seasonal trends of the M(O_3) appear insignificant, except in winter for which a weakly significant increase is found in all three tropospheric layers (+0.83[+0.13;+1.67], +0.62[+0.05;+1.22] and +0.62[+0.02;+1.22]%$O_{3,2000}$ yr^{-1} in the LT, MT and UT, respectively). Previous trend analysis at the alpine sites (Zugspitze since 1978, Jungfraujoch and Sonnblick since 1990) have highlighted (i) a strong increase of O_3 during all seasons in the 1980s (around +0.6-0.9 ppb yr^{-1}), (ii) a persistent but lower increase in the 1990s during all seasons except summer where O_3 has levelled off, (iii) the extension of that levelling off in the 2000s to the other seasons and a slight decrease in summer (Logan et al., 2012; Parrish et al., 2012). This picture is in general agreement with our results in the lower part of the troposphere. More specifically, in winter, Parrish et al. (2012) found an average trend of +0.61±0.25 %$O_{3,2000}$ yr^{-1} at regional background sites in Europe over the 2-3 last decades, which is consistent with the trends found here over the period 1994-2012. At

low altitudes, this increase of O_3 in winter is mainly attributed to a reduced O_3 titration by NO due to decreasing NO_x emissions (e.g. Ordóñez et al., 2005). The persistent positive trends found higher in altitude suggest that wintertime O_3 has increased at a large scale (if not hemispheric) since air masses sampled by MOZAIC-IAGOS aircraft in both the MT and UT can be influenced by emissions from North America and Asia (as shown in Fig. 2).

Concerning the $P_5(O_3)$, a significant increase is found at the annual scale in all three tropospheric layers (+1.03[+0.36;+1.62], +0.42[+0.09;+0.68] and +0.63[+0.09;+0.99]$\%O_{3,2000}$ yr^{-1} in the LT, MT and UT, respectively). Conversely, trends of the $P_{95}(O_3)$ are all insignificant. Note that ignoring the autocorrelation of the data leads to some additional significant positive trends, including the $M(O_3)$ at the annual scale, the $P_5(O_3)$ in winter and autumn, and the $P_{95}(O_3)$ in winter, although not in all tropospheric layers (see Table S1 in the Supplement). It is beyond the scope of this study to investigate why the autocorrelation has a stronger effect on these specific seasons or layers, but this illustrates the strong influence of the serial dependence on the trend analysis and the necessity to take it into account."

4) Section 3.3.2 – In contrast to ozone, less is known about the trends of carbon monoxide through the depth of the troposphere, and statistically significant trends are derived in this work, so discussion of these trends is useful. It may also be useful to include the relative CO trends, now given in Table 1, with the absolute CO trends, now given in Table S1, in a single table in the body of the paper.

ANSWER : Following the recommendation of the referee, we added the absolute CO trends in Table 1.

5) Section 3.4.1 has been much improved but there are still problems that must be corrected. I do not believe that fitting of the seasonal cycle is correct. The problem is that the data are collected over the entire month, but the monthly average data are plotted at the start of the month (See Figure S7). Consequently, the derived date of maximum ozone is approximately one half month early. For example, the peak in the MT near the middle of the data record is approximately June 4. Parrish et al. (2013) found seasonal peaks of approximately June 20 at alpine peaks in Europe near 2003, which should be comparable to the MT data discussed here. The error identified above accounts for the difference in these findings. Fitting of the seasonal cycle must be corrected. I recommend the approach of Figure 1 of Parrish et al. (2013) where the monthly average data are plotted at the center of the month between 0 and 12, and t in Equation (1) of this paper in months, with values ranging between 0.5 and 11.5. The confidence limits on the results are given, and they indicate significant changes in both the magnitude and the phase of the ozone seasonal cycle. However, I am not yet convinced that the confidence limits are correct. I have examined the Press et al. (2007) reference that the authors cite for their determination of confidence limits. In that material, I do not see how the authors treated the very strong covariance of the individual data points in Figure 7; more information is required in the paper or in the Supplement. However, I suggest a second method be employed to calculate the confidence limits of the trends. The authors employ the excellent strategy of analyzing data from two separate, independent periods: 1995-2003 and 2004-2012. In effect, they obtain two determinations (with confidence limits) each of the magnitude and

phase of the seasonal cycle, which represent averages over the separate periods. To determine if the magnitude or phase is significantly different between the two periods, let us represent the quantity (with confidence limit) for the earlier and later periods by $A \pm \sigma_A$ and $B \pm \sigma_B$, respectively. A and B are significantly different if their difference (AB) is significantly different from zero. The confidence limit of the difference is approximated by $(\sigma_{A2} + \sigma_{B2})^{1/2}$. If the absolute value of A-B is greater than the confidence limit of the difference, then the trend is significant. Examining the 95% confidence intervals in Figure 7 suggests to me that not all of trends will be significant when evaluated in this manner. This suggested approach differs from that used by the authors, but the results should be approximately the same. These issues should be discussed. Note that even if the phase shifts in the upper troposphere are not statistically significant, the discussion in Section 3.4.2 regarding the altitude dependence of the seasonal shift is still valid, since the confidence limits provide constraints on how large the seasonal shift could be, and that constraint will be lower than the statistically significant shift found in the lower troposphere.

ANSWER : Two errors were found in this section: (i) the 15-day shift in the date of occurrence of the maximum O_3 (as noticed by the referees) and (ii) the calculation of the confidence intervals (that have not been converted from the "radian" space to the "month" space, leading to an underestimation of a factor $12/2\pi=1.9$). If we consider the monthly time series, the seasonal shift remains significant in the LT and MT. However, the confidence intervals on these shifts are greatly increased, leading to insignificant differences of seasonal change between the tropospheric layers. The confidence intervals are greatly reduced by fitting the sine function on the daily O_3 time series (rather than the monthly one). With this approach, the differences of seasonal shift between the tropospheric layers are significant. We thus decided to consider in this section the daily O_3 mixing ratios (and to briefly mention the larger confidence intervals obtained based on the monthly O_3 time series). Following the recommendations of the referee, we no longer use the approach of Press et al. (2007) and now simply calculate how significant is the difference of phase between the two 9-year periods (1995-2003 and 2004-2012). In addition, we decided to focus the discussion on the results obtained with the MOZAIC-IAGOS dataset, and thus removed the analysis of the seasonal change at ground and ozonesonde sites.
The text of Sect. 3.4 was modified as follows :
"**3.4.1 Evolution of the seasonal cycle at Frankfurt/Munich**

The seasonal variation of O_3 can be well approximated by a sine function fully characterized by three parameters: an offset value defined here as the average O_3 mixing ratio over the considered period, an amplitude, and a phase that determines at which period in the year the maximum of O_3 is reached. Following the approach of Parrish et al. (2013), one can fit a sine function over different periods of time and compare the results of the fit in order to highlight potential changes in the seasonal pattern of O_3. While Parrish et al. (2013) applied the sine fit to the monthly mean time series, we here consider the daily mean O_3 mixing ratio but the results from both approaches will be discussed . The equation of the fit is :

$$\tilde{y}(t) = y_0 + a \sin\left(\frac{2\pi t}{365} + \phi\right) \quad (1)$$

with t the time (in days, values ranging between 0.5 and 364.5), y_0 the offset mixing ratio (in ppb), a the amplitude (in ppb) and ϕ the phase. The date of the year of the

seasonal maximum of O_3 is then estimated as : $(\pi/2 - \phi) * 365/2\pi$ (Parrish et al., 2013). We apply the sine fit on the two 9-year time periods 1995-2003 and 2004-2012. As there is no overlap between these periods, the two datasets and the results of the sine fit are independent. The changes of amplitude and phase obtained with the sine fits are reported in Table 2.

Between 1995-2003 and 2004-2012, the amplitude of the O_3 seasonal cycle has significantly decreased in the whole troposphere, with a rate of decrease of -2.5±0.9, -1.1±0.5 and -2.1±1.0 ppb decade^{-1} in the LT, MT and UT, respectively. Reason for the decreasing amplitude is the significantly increased yearly O_3 minimum occurring in winter and to the same time constant O_3 maximum occurring in spring/summer (see Sect. 3.3.1). The differences of amplitude change between the different layers all remain statistically insignificant.

Table 2. Characteristics of the O_3 seasonal cycle over the periods 1995-2003 and 2004-2012 in all tropospheric layers. Amplitude and phase are obtained by fitting a sine function on the daily mean O_3 mixing ratios (see text).

Layer	Amplitude			Phase		
	Amplitude 1995-2003 (ppb)	Amplitude 2004-2012 (ppb)	Amplitude trend (ppb decade^{-1})	Date of seasonal maximum 1995-2003	Date of seasonal maximum 2004-2012	Shift (day decade^{-1})
UT	18.0±0.7	16.1±0.6	-2.1±1.0	23 June ± 2 days	20 June ± 2 days	-3.3±3.3
MT	11.5±0.3	10.5±0.3	-1.1±0.5	23 June ± 1 days	16 June ± 2 days	-7.8±2.5
LT	9.9±0.6	7.6±0.5	-2.5±0.9	18 June ± 3 days	2 June ± 4 days	-17.8±6.0

Over the period 1995-2003, the sine fit gives a seasonal maximum of O_3 the 18 June in the LT and the 23 June in the MT and UT. The date of seasonal maximum in the LT is in reasonable agreement with those obtained by Parrish et al. (2013) at two alpine sites (Jungfraujoch, Switzerland and Zugspitze, Germany) and at a lower elevation site (Hohenpeissenberg, Germany, ~50 km from Munich). Over the period 2004-2012, the seasonal maximum O_3 occurs the 2 June in the LT, the 16 June in the MT and the 20 June in the UT. Thus, the phase of the seasonal variations of O_3 shifted forward during the period 1995-2012. The seasonal shift between 1995-2003 and 2004-2012 is highly significant in the LT (-17.8±6.0 day decade^{-1}) and MT (-7.8±2.5 day decade^{-1}), and nearly insignificant in the UT (-3.3±3.3 day decade^{-1}). The differences of seasonal shift between the tropospheric layers are all significant, and the seasonal shift thus decreases with altitude. Note that applying the sine fit to the monthly O_3 mixing ratios give similar shift estimates but much larger uncertainties, leading to insignificant differences among the tropospheric layers (-13.3±11.6 and -6.7±6.5 day decade^{-1} in the LT and MT, respectively). Note that reducing the width of the time windows (to less than 9 years) does not give significantly different results.

At the three continental sites, Parrish et al. (2013) reported statistically significant rates of shift (at the 95% confidence level) ranging between -5 and -7 days decade^{-1} since 1970s while at the coastal site Mace Head, the rate was lower and insignificant (-3±3.7 days decade^{-1}). In comparison, the seasonal shift we obtained in the LT is significantly higher, but discrepancies are likely due to the fact that the studied

periods are different. As a faster change of phase is found between 2005 and 2008 (the 3 last years studied) (see Fig. 2 in Parrish et al. (2013)), restricting their analysis to our shorter period would likely lead to a higher seasonal shift (i.e., closer to our values).

3.4.2 Discussion

This previous analysis confirms that the ozone seasonal pattern in Central/Western Europe is changing, at least since the mid-1990s, moving toward a lower amplitude and an earlier O_3 maximum. It is worth noting that the MOZAIC-IAGOS observations above Frankfurt/Munich represent the worldwide densest dataset of O_3 vertical profiles, which gives robustness to our results. Thanks to vertical profile observations, it brings an interesting contribution by showing that this seasonal change of the phase above Frankfurt/Munich decreases with altitude. This may highlight that the O_3 seasonal pattern behave differently over the northern hemisphere continents (Europe, North America, Asia). Indeed, the FLEXPART-derived PES clearly shows that the air masses sampled by MOZAIC-IAGOS aircraft in the different tropospheric layers originate from different regions (see Fig. 2). The LT is predominantly influenced by the European emissions, the MT by both the European and Northern American emissions, the UT by both the Northern American and Asian emissions.

Parrish et al. (2013) exhaustively discussed several reasons that may explain this changing phase at surface in Europe, including changes in downward transport of stratospheric O_3, long-range transport, O_3 precursor's emissions and their spatial distribution, photochemical production and the potential influence of climate change. Our study does not provide an unambiguous explanation to either the seasonal trends discrepancies or the subsequent seasonal shifts (which would ideally require the use of global models able to correctly reproduce both O_3 seasonal patterns and trends throughout the troposphere). In terms of stratospheric contributions, the STE is known to peak in spring (Auvray and Bey, 2005; James et al., 2003; Tang et al., 2011) due to both enhanced downward transport (Appenzeller et al., 1996) and maximum mixing ratios in the lowermost stratosphere (e.g., Thouret et al., 2006). If the seasonal shift was induced by higher STT fluxes, one would expect stronger positive trends in spring close to the tropopause compared to the LT and a larger (and more significant) seasonal shift in the UT, which contradicts our observations. Thus, the STE is not likely the main reason explaining the shift of the O_3 seasonal pattern. The trend analysis (Sect. 3.3) has not highlighted any significant O_3 trend either in spring or summer, the large uncertainties being at least partly due to the strong interannual variability of O_3 mixing ratios."

Major issues to make the paper more useful:

1) Section 3.1.1 discusses the vertical profile of ozone. Previous work has also reported vertical profiles of ozone over nearby locations in Europe. It would be useful to quantitatively compare and contrast the profiles presented here with those measured at such locations (e.g., sondes from Hohenpeissenberg, Germany). If a comparison can be done with the sondes over the same period of time as the MOZAIC-IAGOS flights, then the accuracy of the sonde data could be evaluated. A comparison of the MOZAIC-IAGOS CV with the sonde CV would also be very informative.

ANSWER : Several previous studies have already extensively investigated the

consistency between MOZAIC-IAGOS O_3 dataset and the other in-situ observations (ozonesonde, surface sites), showing a reasonable agreement, in particular in Europe. Thus, to our opinion, it is not necessary to include such analysis in the paper. However, we do agree that more details on this point should be include in the paper, and we modified the text in Sect. 2.1 as follows :

"Note also that several studies have investigated the consistency of the MOZAIC-IAGOS O_3 dataset with other types of in-situ data (e.g., surface stations, ozonesonde) (Logan et al., 2012; Staufer et al., 2013, 2014; Tanimoto et al., 2015; Zbinden et al., 2013a). Focusing on O_3 changes in Europe, Logan et al. (2012) showed a reasonable agreement between aircraft and alpine sites but noticed the absence of O_3 increase in 1994-1998 in the sonde dataset (contrary to the two other types of data). Focusing on the pure tropospheric profiles, Zbinden et al. (2013) found a mean difference between MOZAIC-IAGOS and sondes of -2% in Germany, -8% in eastern United Sates and +1% over Japan. Tanimoto et al. (2015) obtained similar results, with differences between aircraft and sondes data around ±2% throughout the whole troposphere in Belgium, Germany and Japan, and ±5% at Hong Kong. The MOZAIC-IAGOS data at Munich were found to compare reasonably well with the surface observations at Hohenpeissenberg (slope of 0.97, correlation of 0.77). »

In the references:

Tanimoto, H., Zbinden, R. M., Thouret, V. and Nédélec, P.: Consistency of tropospheric ozone observations made by different platforms and techniques in the global databases, Tellus B, 67, doi:10.3402/tellusb.v67.27073, 2015.

Zbinden, R. M., Thouret, V., Ricaud, P., Carminati, F., Cammas, J.-P. and Nédélec, P.: Climatology of pure tropospheric profiles and column contents of ozone and carbon monoxide using MOZAIC in the mid-northern latitudes (24° N to 50° N) from 1994 to 2009, Atmos. Chem. Phys., 13(24), 12363–12388, doi:10.5194/acp-13-12363-2013, 2013.

2) Section 3.2.2 - I would expect the MT CO data above Frankfurt to compare well with surface CO concentrations at remote sites at the latitude of Frankfurt. A quantitative comparison with the zonal averages from Novelli and coworkers (Novelli et al., 1998; 2003) should be included. I suspect that this comparison will be more informative than the comparisons of the LT data with a few German surface sites.

ANSWER : The studies of Novelli et al. (1998; 2003) only include CO measurements performed before 2002. Thus, to our opinion, they are not very suitable for comparison with our MOZAIC-IAGOS observations over the 2002-2012 period. However, we agree with the referee that a more quantitative comparison with more surface stations can be useful. We added a study in the Supplement and modified the text in the paper as follows (note that these modifications also include the answer to the 5th significant issue raised by the referee) :

"As mentioned in Sect. 2.2 and 3.1.2, the data below 1 km were skipped in order to reduce the impact of the local emissions from both the neighbouring agglomeration and the other aircraft — on tarmac and/or during the take-off/landing phases (in case the MOZAIC-IAGOS aircraft closely follows other aircraft). Indeed, many studies have shown that airport activities can impact the air quality at the local scale (e.g. Hu et al., 2009; Pison and Menut, 2004; Yu et al., 2004). It is worth mentioning that in a

standard landing take-off cycle – comprising the approach, the taxi (plane on the tarmac), the take-off (acceleration phase on the tarmac) and the climb up to a standard atmospheric boundary layer of 915 m (Kesgin, 2006) – most of the CO emissions (85-95%) occur on the tarmac during the taxi phase (Kurniawan and Khardi, 2011). However, even above 1 km in the LT, one cannot exclude any influence of these emissions, or by the emissions of the neighbouring agglomeration.

To assess more precisely the spatial representativeness of the MOZAIC-IAGOS data in the LT and MT, a comparison is made with the CO mixing ratios measured at the World Meteorology Organisation (WMO) Global Atmosphere Watch (GAW) surface stations (see Sect. S.1 in the Supplement for details). Only the stations located between 45°N and 55°N (i.e. ±5° from the latitude of Frankfurt) and with at least 80% data capture for the period 2002-2012 (based on the monthly time series) are retained, which gives a set of 10 stations. The annual mean CO mixing ratio measured by MOZAIC-IAGOS aircraft in the LT (143 ppb) is in the lower range of the zonal average 155±28 ppb observed among the GAW surface stations. When considering only the stations above 1000 m (i.e. 3 stations all located in Europe), the zonal average is reduced to 145±19 ppb, which is very close to the mean CO observed in the LT. In the MT, the annual mean CO mixing ratio is 115 ppb, thus lower than the CO mixing ratios at the ground whatever the station, but the difference with the highest mountain station Jungfraujoch (3580 m elevation) is very small (-7%). Additionally, the annual MOZAIC-IAGOS CO data in both the LT and MT is strongly correlated with the CO observed at the ground (R between 0.61 and 0.94 at all stations except one at which R=0.41). Therefore, the comparison between the MOZAIC-IAGOS CO dataset at Frankfurt/Munich and the GAW dataset at the same latitude shows a good consistency, both in terms of mean annual CO mixing ratios and interannual variations. This ensures a satisfactory representativeness of the MOZAIC-IAGOS observations."

In the references :

Hu, S., Fruin, S., Kozawa, K., Mara, S., Winer, A. M. and Paulson, S. E.: Aircraft Emission Impacts in a Neighborhood Adjacent to a General Aviation Airport in Southern California, Environ. Sci. Technol., 43(21), 8039–8045, doi:10.1021/es900975f, 2009.

Kesgin, U.: Aircraft emissions at Turkish airports, Energy, 31(2-3), 372–384, doi:10.1016/j.energy.2005.01.012, 2006.

Kurniawan, J. S. and Khardi, S.: Comparison of methodologies estimating emissions of aircraft pollutants, environmental impact assessment around airports, Environ. Impact Assess. Rev., 31(3), 240–252, doi:10.1016/j.eiar.2010.09.001, 2011.

Pison, I. and Menut, L.: Quantification of the impact of aircraft traffic emissions on tropospheric ozone over Paris area, Atmos. Environ., 38(7), 971–983, doi:10.1016/j.atmosenv.2003.10.056, 2004.

Yu, K. N., Cheung, Y. P., Cheung, T. and Henry, R. C.: Identifying the impact of large urban airports on local air quality by nonparametric regression, Atmos. Environ., 38(27), 4501–4507, doi:10.1016/j.atmosenv.2004.05.034, 2004.

In the Supplement :

"**S.1 Annual mean CO from the GAW surface stations**

In this section, we compare the CO mixing ratios measured in the LT and MT by MOZAIC-IAGOS aircraft to the CO mixing ratios measured at the GAW surface stations. Monthly data from the GAW database are downloaded on the World Data

Centre for Greenhouse Gases (WDCGG) (http://ds.data.jma.go.jp/gmd/wdcgg/; data downloaded the 1st July 2016). We consider the period 2002-2012. Only the GAW surface stations located between 45°N and 55°N (i.e. ±5° from the latitude of Frankfurt) and with at least 80% of available data are retained. These criteria lead to a set of 10 stations described in Table S2.

Table S2 : Description of the GAW surface stations, annual mean CO mixing ratio over the periodf 2002-2012, and correlation (R) of the annual time series with the annual mean CO measured in the LT and MT by the MOZAIC-IAGOS aircraft.

Station name (contributor*)	Location	Altitude	Annual mean CO (ppb)	Correlation with the LT (MT) CO
Fraserdale (EC)	81.57°W, 49.88°N	210 m	135	0.68 (0.77)
Hohenpeissenberg (DWD)	11.02°E, 47.8°N	985 m	165	0.74 (0.57)
Hegyhatsal (NOAA/ESRL)	16.65°E, 46.95°N	248 m	209	0.61 (0.59)
Jungfraujoch (Empa)	7.987°E, 46.548°N	3580 m	123	0.94 (0.96)
Kollumerwaard (RIVM)	6.28°E, 53.33°N	0 m	196	0.71 (0.72)
Park Falls (NOAA/ESRL)	90.27°W, 45.92°N	868 m	141	0.74 (0.83)
Mace Head (AGAGE)	9.9°W, 53.33°N	8 m	135	0.62 (0.62)
Ochsenkopf (NOAA/ESRL)	11.8°E, 50.03°N	1185 m	160	0.41 (0.54)
Shemya Island (NOAA/ESRL)	174.08°E, 52.72°N	40 m	132	0.78 (0.85)
Schauinsland (UBA)	7.92°E, 47.92°N	1205 m	151	0.79 (0.78)

*Contributors : EC, Environment Canada; DWD, Deutscher Wetterdienst; NOAA/ESRL, National Oceanic & Atmospheric Administration / Earth System Research Laboratory; Empa, Swiss Federal Institute for Materials Science and Technology; RIVM, Dutch National Institute for Public Health and the Environment; AGAGE, Advanced Global Atmospheric Gases Experiment; UBA, Umwelt Bundesamt

The annual mean CO mixing ratio is first calculated for each station individually, and then averaged for all stations. It gives an annual mean CO mixing ratio of 155 ppb over the period 2002-2012 (the standard deviation among the stations is 28 ppb). It is 15% higher than the zonal average of ~135 ppb given by Novelli et al. (1998) in the 1990s at this latitude. As shown in Table S2, large differences among the individual stations are found, mean CO mixing ratios ranging from 132 ppb at Shemya Island (in the Pacific) to 209 ppb at Hegyhatsal (in Hungary). The standard deviation inferred from the annual mean CO of all stations is 28 ppb. The annual mean CO mixing ratio given by the MOZAIC-IAGOS data in the LT is 143 ppb for the same period. It is in the lower range of the zonal average CO inferred from the GAW surface stations at this latitude. When considering only the surface stations above 1000 m above sea level (a.s.l.) (i.e. 3 stations all located in Europe : Jungfraujoch, Ochsenkopf and Schauinsland), the zonal average is reduced to 145 ppb, thus very close to the annual mean CO observed in the LT by MOZAIC-IAGOS aircraft. The annual mean CO mixing ratio in the MT (115 ppb) is substantially lower than the zonal average given by surface stations, but the difference with the highest mountain station Jungfraujoch (located at 3580 m) is very small (-7%).

To further assess the representativeness of the MOZAIC-IAGOS observations, the correlations between the annual mean CO mixing ratios at the GAW surface stations and the CO mixing ratios observed in the LT and MT are given in Table S2. Except at

Ochsenkopf where the correlation is only 0.40, all GAW stations show correlations with the MOZAIC-IAGOS data in the LT exceeding 0.60 (up to 0.94 at the mountain site Jungfraujoch, close to Frankfurt/Munich). Except at Hohenpeissenberg, higher correlations are found with the MOZAIC-IAGOS data in the MT. Thus, the interannual variation obtained in the MOZAIC-IAGOS data in the LT and MT is consistent with the interannual variation observed at regional and global surface sites at this latitude.
Therefore, although the measurements performed by the MOZAIC-IAGOS aircraft in the LT may still be influenced by some local emissions, these comparisons with the GAW surface stations highlight a good consistency, both in terms of mean annual CO mixing ratios and interannual variations. This gives confidence on the representativeness of the MOZAIC-IAGOS observations."

Significant issues

1) p. 10, lines 15-16 - This sentence is confusing. As shown by the seasonal cycle of CO illustrated in Fig. 5, CO in the free troposphere does indeed change significantly from one season to the other; please clarify.

ANSWER : Indeed, this sentence was not well formulated. The paragraph was modified as follows : "The Figure 4 shows that the daily variability is lower for CO than for O_3, in particular at the surface and close to the tropopause. Over the entire tropospheric column, the annual daily variability of CO is 20%. It ranges between 44% close to the surface and 17% in the free troposphere where it remains almost constant with altitude. A very similar picture is drawn for the different seasons. The highest values at the surface (in the second half of the troposphere) are encountered in winter/autumn (summer)."

2) p. 10, line 19 and p. 13, line 15 - It is not clear to me that a maximum of primary CO emissions occurs in winter (at northern mid-latitudes), especially if biomass burning emissions are included. This issue should be discussed in more detail with references.

ANSWER : The seasonal variation of the total (anthropogenic and biomass burning) CO emissions strongly depends on which regions are considered. We investigated the seasonal variation of the combination of the MACCity anthropogenic emissions with the GFASv1.2 biomass burning emissions, averaged over the period 2003-2012. If we consider only the northern mid-latitudes (i.e. Europe, United States, Central-East Asia), there is a broad maximum from late autumn to winter. If we add the boreal regions (i.e. Alaska, Canada, Siberia), the emissions maximize in spring-summer (with a secondary maximum in October). If we add the regions from the northern tropics (i.e. Central America, Middle-East, South-East Asia), the maximum emissions occur in spring. At the global scale, the maximum emissions occur in late summer.
However, to our opinion, that discussion about the seasonal variation of CO emissions is not worth being included in the paper. The seasonal pattern observed in this study is quite common. The sentence on page 10 was removed, and the paragraph on page 13 modified as follows :
"As shown in Fig. 5, the seasonal cycle of CO is characterized by maximum mixing ratios in late winter/early spring in the whole troposphere. Minimum mixing ratios are

encountered in summer/early autumn in the LT and are slightly shifted to late summer/early autumn higher in altitude. Such a seasonal pattern is consistent with the seasonal variation observed in background air masses arriving at the coastal site Mace Head (Derwent et al., 1998) or at a larger scale by satellite observations (Edwards et al., 2004; Worden et al., 2013). Averaged over the western Europe, the Terra/MOPITT CO tropospheric column maximizes at ~2.5 10^{18} molecules cm^{-2} in March-April and minimize at ~1.9 10^{18} molecules cm^{-2} in late summer, the ratio of the maximum over the minimum being 1.3 (Edwards et al., 2004). A very similar seasonal variation of tropospheric columns of CO has been observed by Zbinden et al. (2013) based on the MOZAIC data over the period 2002-2009. This is in good agreement with the amplitude of the seasonal cycle observed in the MT, the maximum CO mixing ratio being 1.35 higher than the minimum. The winter-time maximum results from the accumulation of the primary CO emissions at northern mid-latitudes when the photolysis is limited. In summer, CO mixing ratios minimize due to a more effective photolytic destruction, despite an enhanced secondary formation from biogenic compounds and additional emissions from biomass burning (in particular in boreal regions). A rather similar seasonal pattern is observed with the CO 5th and 95th percentiles except that the peak is sharper (February-March) in the LT for the 95th percentile."

3) p. 12, beginning on line 14 - If the spring 1998 anomaly is indeed related to enhanced STE, then one might expect a stronger anomaly in the MT and UT, which are closer to the source of the STE, than in the LT; however that does not seem to be the case. This should be discussed further.

ANSWER : The Fig. 7 shows that in spring 1998, the anomaly is strongest in the UT, and roughly similar in the MT and LT. This discussion was not perfectly clear, and was thus modified as follows :

"The highest monthly mean mixing ratios in the LT (above 60 ppb, the 99th percentile) are observed in August 1995, May 1998, August 2003, July 2006 (Fig. 6). The spring 1998 anomaly is related to the 1997 El Niño that enhanced the pollution export from Asia (due to a higher convective activity and a strengthening of the subtropical jet stream) and North America and may have increased the STE (Koumoutsaris et al., 2008; Zeng and Pyle, 2005). This anomaly is visible in the whole troposphere, and is the strongest in the UT. The anomalies in August 2003 and July 2006 are related to the severe heat waves that struck a large part of Europe (Ordóñez et al., 2005; Solberg et al., 2008; Struzewska and Kaminski, 2008; see also Tressol et al. (2008) for a detailed analysis of the 2003 heat wave with the MOZAIC measurements). They are the strongest in the LT but remain visible in the MT.

The annual mean O_3 mixing ratios are highly correlated between the three tropospheric layers (R=0.87, 0.75 and 0.94 between the LT/MT, LT/UT and MT/UT, respectively). As the sources and sinks of O_3 in the troposphere strongly vary with altitude, such high correlations were not expected. This may be explained by the fact that both the first kilometre and the tropopause layer are not taken into account in this study, which likely greatly reduces the differences of interannual variation among the tropospheric layers as defined in this study. In addition, as shown in Sect. 2.2, the altitude of the UT (defined here based on the PV values) is biased low compared to the UT derived based on the chemical tropopause. Thus, the UT may be less influenced by the stratosphere and more by the free troposphere, which may increase the correlation between the MT and UT. Similar correlations are obtained at the

seasonal scale. An interesting exception is the low correlation found between the LT and UT in summer (R=0.26) and spring (0.46). This may be due to the fact that the BL is deeper during these seasons, which keeps away the 1-2 km LT from the influence of the free troposphere."

4) p. 12, lines 18-20 - Exactly what is being correlated must be clarified. Are these correlations of annual averages? What can be the cause of this correlation? Has anyone noted this before? More discussion is required.

ANSWER : Yes, we are talking about the correlations of the annual averages. To our knowledge, no previous study have investigated such a phenomenon. We discussed in more detail this point, see our answer to the previous significant issue.

5) p. 13, lines 22-25 - If I understand correctly, aircraft do not emit significant amounts of CO. A reference should be given if my understanding is incorrect; otherwise the discussion should be revised.

ANSWER : Here, we are not saying that aircraft do not emit significant amounts of CO. Actually, they do. We added several references in the text (see our answer to the second major issue) and made some modification in order to make the discussion clearer.

6) Throughout the discussion of trends, the authors should keep clearly in mind that the lack of a statistically significant trend derived from their data does NOT mean that there is no trend. Rather a trend may exist, but their data are too variable to extract that trend. For example, on p. 18, the sentence "Conversely, all trends in autumn are insignificant." is correct, but the previous phrase "...while the P_5(CO) is decreasing only in winter (in the MT and UT) and summer (only in the LT)" is not worded correctly. A similar problem exists in line 11 on p. 20; the present analysis does not establish a significant difference in the trends between winter and spring/summer.

ANSWER : We modified this Sect. 3.3.2 as follows : "As previously mentioned in the beginning of Sect. 3.3, CO trends are here investigated relatively to the 2004 reference year. Over the period 2002-2012, the M(CO) at the annual scale significantly decreases in the whole troposphere, with trends of -1.51[-2.42;-0.44], -1.55[-2.34;-0.72] and -1.36[-2.05;-0.74]%CO_{2004} yr^{-1} in the LT, MT and UT, respectively. Similar negative trends are also obtained for the P_5(CO) and P_{95}(CO) in all the tropospheric layers. At the seasonal scale, the M(CO) and P_{95}(CO) show negative trends in winter, spring and summer, although not always in all the tropospheric layers, while the P_5(CO) is significantly decreasing in winter (in the MT and UT) and summer (in the LT). Conversely, all trends in autumn are insignificant. Note that the results without taking into consideration the autocorrelation of the data show significant negative trends of the P_5(CO) in most layers and during all the seasons, except autumn (see Table S1 in the Supplement). These results are in general agreement with previous studies in Europe (e.g., Karlsdóttir et al., 2000; Novelli et al., 2003; Dils et al., 2009; Worden et al., 2013). Based on satellite observations, Worden et al. (2013) highlighted over Europe a decrease of the CO total columns, around -1.44+0.22% yr^{-1} with MOPITT over 2001-2011 and -1.00±0.33% yr^{-1} with AIRS over 2003-2011, thus in the range of our results over Frankfurt. Over the period 1995-2007, Gilge et al. (2010) found trends of -3.36±1.08 and -1.51±0.64 ppb yr^{-1}

(reduced to -2.65±0.04 ppb yr^{-1} by filtering the background values (Zellweger et al., 2009)) at two alpine sites from the WMO GAW network, in reasonable agreement with our absolute Theil-Sen slope estimates at Frankfurt/Munich in the LT and MT (-2.24[-3.59;-0.65] and -1.85[-2.79;-0.85] ppb yr^{-1}). ".

7) p. 18, lines 22-24 - The wording of this introductory sentence is awkward. It should be clearly stated that if the trends in ozone are significantly different between seasons, then there is necessarily a change in the seasonal cycle.
ANSWER : Actually, our first sentence is not correct. We found that only a few trends are significant, but the differences of trend between the seasons remain insignificant (due to either the absence of trend or a high interannual variability hiding any trend). Thus, to our opinion, the trend results obtained in Sect. 3.3.1 do not allow by themselves to conclude on the presence or absence of a significant change in the seasonal cycle. We modified the text as follows : "In Sect. 3.3.1, we highlighted that only a few O_3 trends are statistically significant. However, the differences of trends between the seasons remain insignificant. It is worth noting that an insignificant trend does not imply the absence of trend since a trend can be hidden by a strong variability. In this section, we investigate if these trends come along with a change of the O_3 seasonal cycle above Frankfurt/Munich (Sect. 3.4.1). Results are discussed in Sect. 3.4.2."

8) Figure 7 has a trace of "Fit error" in each graph. This quantity is not defined, and not discussed in the manuscript. I suggest these traces be removed from the graphs, or at least defined and discussed in the paper.
ANSWER : We removed this figure and greatly modified this section.

9) p. 22, lines 18-20: This discussion of ozone anomalies is not clear. I cannot see any relevance to Figure 7. This discussion must be improved.
ANSWER : We moved this Figure in the Supplement and greatly modified the Sect. 3.4, see our answer to the major issue 5.

Minor issues

1) p. 3, line 10 - The authors write "... the confidence in the results remains limited by the high uncertainties at stake in both models and ...". This is not proper English usage.
ANSWER : This part of the text was removed, following the recommendations of the other referee.

2) p. 3, lines 30-31 - The authors write "... limited representativeness of measurements in the boundary layer measurements...". Eliminate one of the "measurements".
ANSWER : Modification applied.

3) The caption to Figure 5 should define the meaning of the shaded regions.
ANSWER : We added the following sentence in the caption : "The shaded areas show the ±2 standard error (i.e. the uncertainty in the average at a 95% confidence level)."

4) p. 11, lines 13-14 - The sentence should read: "In the MT and UT, maximum concentrations occur between May and August (highest concentrations in July).
ANSWER : Modification applied.

5) p. 12, line 4 - Start new paragraph.
ANSWER : We removed this sentence as the variability is already discussed in Sect. 3.1.1, and we started a new paragraph after.

6) p. 12, line 9 - Remove "have".
ANSWER : Modification applied.

7) p. 19, line 5 – The last time period should be (2004-2012) not (2003-2012).
ANSWER : Modification applied.

ANSWERS TO REFEREE #2

For clarity, the suggestions/remarks of the referee are in italics, our direct answers in bold and our modifications in the manuscript in regular.

The authors have extensively answered the concerns raised by the reviewers and the manuscript has changed substantially. In doing so, most critical issues were solved and many parts were improved. So far so good. However, one major flaw I have, namely that the wording is largely not acceptable. It's not so much the English spelling itself, but that the sentences and the argumentation are partially so awkward that I could decipher what is meant only after reading the relevant parts many times. Often the real information was hided and even sometimes not written. Under the line, I often didn't catch the information of a section and at the end I couldn't repeat what I have read, that is the take-home message was missing.
One main reason in my opinion is that you frequently try to condense an argumentation in one convoluted sentence, in which a native speaker would do it in three short sentences and in a logically straight-forward way. Second main reason is the lax and imprecise wording, see the many many examples listed below. Please, read each sentence carefully and think what you have read. Often the sense can only be understood with considerable creativity and own interpretation. In addition, you steadily give twisted phrases, e.g. "the 1994-2012 period" (also in the title), "the Frankfurt and Munich airports", ... This is slang and inadequate for a scientific paper ... and to my knowledge also in French not used. Below, I give more detailed suggestions in the abstract only, but thereafter also many ... as improving the wording is not the task of a reviewer. Please work on it! Consider that the impact of a (linguisticly) badly written paper is poor. I have to say that this was also the reason why the review took so long. After the first reading, I only said "oh gosh" and the fun to work on it was quite limited.
ANSWER : When reading the suggested corrections, we understand the concerns of the referee about the wording of our paper. We deeply thank him/her for his/her efforts (and patience) to solve this problem. We took into account all of his/her recommendations and tried to improve the wording in the paragraphs added or modified in this revised manuscript.

Minor remarks

P.1, l.12. "vertical profiles of ozone (O3) and carbon monoxide (CO)" instead of "ozone and carbon monoxide vertical profiles"
ANSWER : Modification applied.

P.1, l.12. Exchange "at several" with "in three"
ANSWER : Modification applied.

P.1, l.12. Exchange "the Frankfurt and Munich airports" with "the German airports Frankfurt and Munich" and improve it later in the text
ANSWER : Modification applied.

P.1, l.13. Exchange "the densest dataset in the world (about 96)" with "the worldwide densest vertical in-situ dataset of O3 and CO (with ~96)"
ANSWER : Modification applied.

P.1, l.14. "mean vertical profile of ozone" instead of "mean ozone vertical profile"
ANSWER : Modification applied.

P.1, l.15. Skip "vertical"
ANSWER : Modification applied.

P.1, l.15. Exchange "kilometre (deposition, titration by NO) during the whole year and close to the tropopause (stratosphere-troposphere exchanges)" with "kilometre (due to dry deposition at ground and titration by NO) during the whole year and close to the tropopause (due to stratosphere-to-troposphere in-mixing)"
ANSWER : Modification applied.

P.1, l.17. "mean vertical profile of CO"
ANSWER : Modification applied.

P.1, l.17. Skip "in the concentrations"
ANSWER : Modification applied.

P.1, l.18. free troposphere
ANSWER : Modification applied.

P.1, l.18. Exchange "In terms of seasonal variations, the mean O3 has a minimum in November-December in the whole troposphere, a broad spring/summer maximum …" with "O3 minimizes in November-December, shows a broad spring/summer maximum…"
ANSWER : Modification applied.

P.1, l.21. Exchange "mean CO seasonal profile" with "seasonal variation (or course) of CO"
ANSWER : Modification applied.

P.1, l.22. Exchange "surface, refined to September-October higher in the troposphere, while maximum concentrations occur" with "surface and in September-October higher in the troposphere, while the maximum occurs"
ANSWER : Modification applied.

P.1, l.23. Exchange "the 1994-2012 period" with "the period 1994-2012" and later in the text
ANSWER : Modification applied here and elsewhere.

P.1, l.23. "O3 has changed insignificantly" instead of "the mean O3 trends are mostly insignificant"
ANSWER : Modification applied.

P.1, l.27. all three tropospheric
ANSWER : Modification applied.

P.1, l.28. "In contrast, for CO the mean" instead of "Conversely, the mean CO"
ANSWER : Modification applied.

P.1, l.29. all three …
ANSWER : Modification applied.

P.1, l.32. Exchange "The changes in the O3 seasonal cycle are also investigated, with a focus on the phase. Ozone maxima occur earlier and earlier with a shift around -12.1…" with "The phase of the seasonal variation of O3 was found to change in the entire troposphere. The O3 maxima moves forward in time with a rate of -12.1…"
ANSWER : Modification applied.

P.1, l.32ff. In my opinion, the abstract is too long. You can skip e.g. the first 2-3 lines of p.2 This were improvements for the abstract only.
ANSWER : We removed this sentence.

P.2, l.20. The term "stratosphere-troposphere exchanges" doesn't exist, "stratosphere-troposphere exchange (STE)" does. Change this everywhere in the text
ANSWER : Modification applied here and elsewhere.

P.2, l.25. Exchange "ozonesonde long-term observations" with "long-term ozonesonde observations"
ANSWER : Modification applied.

P.3, l.8. Exchange "quantify all the terms of the ozone budget" with "quantify all factors influencing the budget of ozone"
ANSWER : Modification applied.

P.3, l.11. The sentence "An alternative but more qualitative approach consists in taking benefit from the different seasonal patterns of the various ozone budget terms (e.g., precursors emissions, photochemical production, stratospheric intrusions, transport regimes) and aims at linking the evolution of the ozone seasonality to changes in the contribution of its various sources and sinks." is really awful and exemplary demonstrate my major concern. By the way "ozone budget terms" doesn't exist. Moreover this "alternative" will not work. You can't change and follow one single of the listed process (like in a laboratory) without influencing the other. Thus skip the sentence or change it (strongly).
ANSWER : We removed this sentence.

P.3, l.14. Exchange "Mace Head coastal site" with "coastal site Mac Head"
ANSWER : Modification applied.

P.3, l.19. „Derwent et al. (2013) noticed that the still increasing baseline ozone levels do not extend to the European ozone load…" I don't understand this sentence. Do you mean "the increasing baseline ozone levels do not impact the European load of O3? … which I also would not understand.
ANSWER : Here, we mean that O_3 mixing ratios in European air masses have increased much less than the baseline O_3 mixing ratios, which thus suggests a possible compensation between (increasing) O_3 imports and (decreasing) local production. We modified the paragraph as follows: "Observations at the coastal site Mace Head have shown an average annual increase of +0.25±0.09 ppb $year^{-1}$ of the baseline (i.e. originating from the Northern Hemispheric marine boundary layer)

O_3 mixing ratios during the period 1988-2012 (Simmonds et al., 2004, Derwent et al., 2013). This increase has been the strongest in winter and spring and the lowest in summer, and has slowed down during the 2000s (Derwent et al., 2013). In contrast, the annual O_3 mixing ratios in European air masses have shown a much lower increase (Derwent et al., 2013), which suggests a possible compensation between a decrease of O_3 local formation in Europe and an increase of O_3 imports."

P.3, l.25. "Such a shift may reflect some changes in the contributions of the various ozone sources and sinks, e.g.,…" O3 is controlled by its sources and sinks. Thus "such a shift does reflect…", but the second half of the sentence then makes little sense.
ANSWER : Here, we simply give the several possible reasons for this shift, as discussed in Parrish et al. (2013). We modified the sentence as follows : "Such a shift may reflect some changes in the transport pathways, precursors emissions and photochemistry of O_3, possibly due to climate change (Parrish et al., 2013)."

P.3, l.27. After Parrish et al (2013) a line break. You should use much more line breaks!
ANSWER : Modification applied.

P.3, l.32. Skip "measurements (close to precursors emissions and/or deposition sink)"
ANSWER : Modification applied.

P.3, l.35. Change "long" to "moderate"
ANSWER : Modification applied.

P.4, l.1. Change "interesting" to "powerful"; skip "useful"
ANSWER : Modification applied.

P.4, l.4. Skip "used in the paper"
ANSWER : Modification applied.

P.4, l.12. Change "since 1994 and 2002" to "since 1994 (O3) and 2002 (CO)"
ANSWER : Modification applied.

P.4, l.18. … the precision, but more important what is the accuracy?
ANSWER : For both O3 and CO, the accuracy is ±2 and ±5 ppb, and the precision ±2 and ±5%, respectively. We modified the text as follows : "Ozone measurements were performed using a dual-beam UV-absorption monitor (time resolution of 4 seconds) with an accuracy/precision estimated at about ±2 ppbv / ±2% (Thouret et al., 1998). Carbon monoxide was measured by an improved infrared filter correlation instrument (time resolution of 30 seconds) with an accuracy/precision estimated at ±5 ppbv / ±5% (Nédélec et al., 2003)."

P.4, l.27. Skip "briefly"
ANSWER : Modification applied.

P.4, l.29. Change "systems operations" with "instrumentation"
ANSWER : Modification applied.

P.4, l.32. the airport Frankfurt
ANSWER : Modification applied (here and elsewhere).

P.5, l.7. Change "As tropospheric air masses are subject to very different constraints depending on their altitude (e.g., distance from surface emissions or stratosphere)" with "As tropospheric ozone shows strongly varying sources, sinks and lifetimes with height"
ANSWER : Modification applied.

P.5, l.7. Change "The UT is defined here as the 60 hPa-width layer below tropopause plus 15 hPa" with "The UT is defined here as the layer having its top at the tropopause plus 15 hPa and spanning a pressure of 60 hPa, that is a layer ~1.6 km thick and starting/ending ~2.1/~0.5 km below the tropopause"
ANSWER : Modification applied.

P.5, l.12. Change "Data collected in the 1-2 km layer" with "Data collected below"
ANSWER : Modification applied.

P.5, l.12. Change "The p PV=2 parameter is used to determine the DT pressure at the top of the selected tropospheric vertical profile." with "The pressure at the DT (pPV=2) minus 15 hPa defines the top of the UT applied here."
ANSWER : Modification applied.

P.6, l.4. "For instance, the ozone criteria may give a lower dynamical tropopause (DT) compared to the thermal method (Bethan et al., 1996)." In this sentence sticks some mistakes. First, O3 can define the chemical tropopause (CT), which is in the mid-latitudes mostly close to the DT, but it can't define the DT". Secondly, a thermal method doesn't exist; you mean the thermal tropopause (TT, which is well defined). Thirdly, the CT (and DT) is only on average below the TT, in the mid-latitude over the UK by ~800m, but can also be above (at anticyclonic flow) or well below by up to 4-5 km (at cyclonic flow). Here, read the papers by V. Wirth (Thermal versus dynamical tropopause in upper tropospheric balanced flow anomalies. Quart. J. Roy. Met. Soc., 126, 299-317., and the ones thereafter).
ANSWER : We modified the sentence as follows : "It is worth noting that the determination of the tropopause altitude is associated to several uncertainties. Some uncertainties arise from the choice of the method used to locate the tropopause. For instance, the chemical tropopause (defined in terms of both O_3 mixing ratio and vertical gradient of O_3 mixing ratio, also referred as the ozone tropopause) is on average below the thermal tropopause (Bethan et al., 1996)."

P.7, l.10. ... which indicates that the ECMWF-derived 2 PVU doesn't work for defining the tropopause. ECMWF estimates it far too low. How often this appears. Later you have to discuss the consequences of such an ill-defined tropopause. The closer you get to the tropopause, the more questionable your O3 trends etc. get.
ANSWER : Indeed, the ECMWF-derived PV is a source of error that leads to an uncertainty on the UT O_3. On the flights where it was possible, we compared the DT pressure derived from PV values with the pressure at which O_3 reaches (and remains above) 150 ppb (taken here as a simplified estimate of the chemical tropopause). There is a systematic bias of +21 hPa, the DT being located below

the chemical tropopause. The UT O_3 may thus be biased low.
The text was modified as follows : "A good example is given in Fig. 1 where the abrupt O_3 increase (corresponding to the tropopause) occurs 2 km above the DT derived from PV values. However, our approach allows to assess in which layer (MT or UT) observations belong even when the tropopause is not reached (within the 400 km around the airport). In order to assess the uncertainties introduced by an erroneous DT pressure, we compared it with the pressure at which the O_3 mixing ratio reaches 150 ppb (taken here a simplified estimate of the chemical tropopause) and remains above at higher altitude (in order to avoid stratospheric intrusions in the troposphere). This was done on all vertical profiles where it was possible, which represents 46% of the dataset. On average over the period 1994-2012, the mean bias of the DT pressure compared to the 150 ppb O_3-isopleth is +21 hPa, while the 5^{th}, 10^{th}, 90^{th} and 95^{th} percentiles of this bias are -72, -32, +78 and +99 hPa, respectively. Therefore, the DT derived from PV values tends to be located below the 150 ppb O_3-isopleth, which may bias low the O_3 mixing ratios in the UT. However, the discrepancy remains moderate on most profiles. It is beyond the scope of this study to investigate in more details the influence of the method used to locate the tropopause."

We also mentioned this point in Sect. 3.2.1 as a possible explanation of the high correlation of O_3 mixing ratios between the UT and MT: "The annual mean O_3 mixing ratios are highly correlated between the three tropospheric layers (R=0.87, 0.75 and 0.94 between the LT/MT, LT/UT and MT/UT, respectively). As the sources and sinks of O_3 in the troposphere strongly vary with altitude, such high correlations were not expected. This may be explained by the fact that both the first kilometre and the tropopause layer are not taken into account in this study, which likely greatly reduces the differences of interannual variation among the tropospheric layers as defined in this study. In addition, as shown in Sect. 2.2, the altitude of the UT (defined here based on the PV values) is biased low compared to the UT derived based on the chemical tropopause. Thus, the UT may be less influenced by the stratosphere and more by the free troposphere, which may increase the correlation between the MT and UT."

P.7, l.19. Skip "by the aircraft"
ANSWER : Modification applied.

P.7, l.26. "sometimes referred as the potential emissions sensitivity (PES), that is the potential to catch up emissions from certain regions"
ANSWER : Modification applied.

P.8, Figure caption. Can't combine "Average residence time" with the units given in the picture. You show this figure, but later it is (at least I believe) not once used, e.g. for interpreting the vertically differing trends etc. Do it!

ANSWER : As indicated in the text, the residence times are normalized by the air density. We added the information in the caption as follows : "Figure 2: Average residence time in the first kilometre (normalized by the air density) of air masses sampled in all three tropospheric layers around Frankfurt. The average is calculated based on all the FLEXPART simulations over the period 1994-2012. Note the irregular scale. " **We also added a reference to this figure in Sect. 3.3.1** « The persistent positive trends found higher in altitude suggest that wintertime O_3 has increased at a large scale (if not hemispheric) since air masses sampled by MOZAIC-

IAGOS aircraft in both the MT and UT can be influenced by emissions from North America and Asia (as shown in Fig. 2)." **and 3.4. "**Thanks to vertical profile observations, it brings an interesting contribution by showing that this seasonal change of the phase above Frankfurt/Munich decreases with altitude. This may highlight that the O_3 seasonal pattern behave differently over the northern hemisphere continents (Europe, North America, Asia). Indeed, the FLEXPART-derived PES clearly shows that the air masses sampled by MOZAIC-IAGOS aircraft in the different tropospheric layers originate from different regions (see Fig. 2). The LT is predominantly influenced by the European emissions, the MT by both the European and Northern American emissions, the UT by both the Northern American and Asian emissions.**".**

P.8, l.13. You show data between ground and 12 km, but not to the tropopause. Correct!
ANSWER : The sentence was modified as follows : "The annual mean O_3 mixing ratio increases with altitude, from 21 ppb at ground to 81 ppb at 12 km (47 ppb at 2 km)."

P.8, l.17. What is an O3 abundance? I would say an O3 density (in e.g. O3 molecules per cm^3) … which may decrease with height. Change e.g. to "mixing ratio"
ANSWER : Indeed, we mean "mixing ratio" here, we corrected the text.

P.8, l.18. You can't simply write "(dry deposition …)". "(due to dry deposition at the ground and enhanced titration by NO in the PBL)" would be possible. Please, be precise and clear. This lax wording is terrible. I think you are scientist.
ANSWER : Modification applied.

P.9, l.1. No "tropopause", but "12 km"
ANSWER : Modification applied.

P.9, l.1."(exchanges with the stratospheric reservoir)" … boah … change
ANSWER : We replaced this text by : "(due to STE)"

P.9, l.3. "very low in winter and to a lesser extent in autumn" … likewise boah, you mean "very low in winter and even more very low in autumn"? … what is lower than very low. Once again, you are scientist! "low" is a poor description anyway, but lower then very low is already quite silly … but what means now "quite" … and thereafter you write "substantially enhanced" another time … boah
ANSWER : We modified the text as follows : "Above 3 km, the mean vertical gradients are +1.1, +1.5, +3.0 and +5.1 ppb km^{-1} in winter, autumn, spring and summer, respectively."

P.9, Figure caption. How the standard deviation is inferred, from the individual profiles or daily averaged? Change "overall climatological profile considering all seasons" to "annual mean profile"
ANSWER : The standard deviation is inferred from the daily averaged profiles. We added this information in the caption and applied the correction suggested by the referee.

P.9, l.10. Start with "To further characterize the variability of O3 and CO above

Frankfurt/Munich, we…" In my opinion "daily variability" is a better word than "coefficient of variation"! P.9, l.21. Change "close to the tropopause" to "In the UT". Exchange "exchanges" with "in-mixing" … citations are missing

ANSWER : We modified the paragraph as follows :

"To further characterize the variability of O_3 and CO above Frankfurt/Munich, we now investigate the daily variability at both the annual and seasonal scales. The daily variability is here defined as the coefficient of variation (CV) of the daily-averaged mixing ratios, that is the standard deviation normalized by the corresponding (i.e. annual or seasonal) mean mixing ratio. Vertical profiles of the daily variability for O_3 and CO are shown in Fig. 4. Results about CO will be discussed in Sect. 3.1.2. The daily variability of O_3 at the annual scale ranges between 20 and 73% depending on the altitude, with a mean value of 32%. The maximum daily variability of O_3 is found at ground (73%) and at 12 km (53%) where it is likely driven by intense shallow and transient exchanges between the stratosphere and the troposphere (Stohl et al., 2003b). Conversely, the minimum daily variability is found at about 3.4 km. Such daily variability is lower at the seasonal scale, at most altitudes and during most seasons, but the shape of the vertical profiles remains similar. The seasonal daily variability minimizes at 4.4 km in autumn and between 3.1-3.4 km during the other seasons, thus close to the minimum annual diurnal variability. Similarly, it maximizes at the surface and close to the tropopause. Interestingly, the daily variability above 11 km is noticeably higher in spring than during the other seasons, which again may be due to the day-to-day variability of STE that peaks during that season.

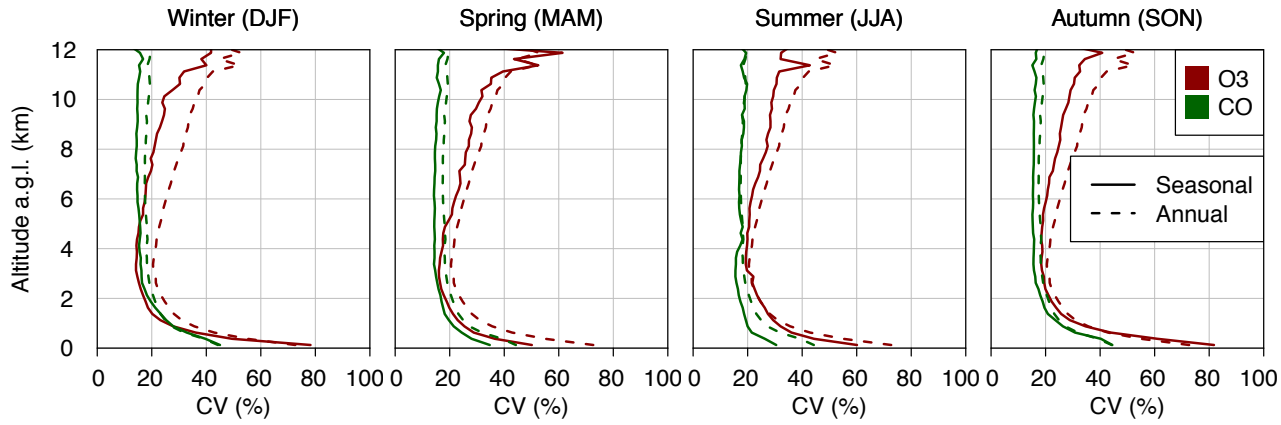

Figure 4: Annual (dotted lines, the same for all panels) and seasonal (continuous lines) daily variability of O_3 and CO mixing ratios above Frankfurt/Munich. The daily variability is here defined as the coefficient of variation (CV) of the daily-averaged mixing ratios (i.e. the standard deviation normalized by the mean)."

P.10, l.8. "high in altitude"? be precise!

ANSWER : We modified the sentence as follows : "The annual mean vertical profile of CO (Fig. 3) shows mixing ratios ranging from 150 ppb at 1 km to 80 ppb at 12 km. ».

P.10, l.8. "Over the entire tropospheric column, the mean CO mixing ratio reaches 117 ppb." No clue what you want to say, e.g. "is 117 ppb", "reaches to 117 ppb", …?

ANSWER : We mean "is 117 ppb", modification applied.

P.10, l.10. Change "the CO mixing ratios in the first kilometre strongly increase as one moves closer to surface emissions (up to 243 ppb at the surface)" to "At the ground and close to surface emissions, CO maximizes with 243 ppb on the annual average"
ANSWER : Modification applied.

P.10, l.13. You write so often "concentrations" which is "molecules cm^-3", but you mean "mixing ratios" ... be precise
ANSWER : Indeed, in all the paper, we are talking about mixing ratios, we applied the correction here and elsewhere.

P.10, l.15. Of course, CO clearly and significantly does change over the year. But the daily variability exceeds the seasonal amplitude. Change! P.10, l.15. Also "The seasonal climatological profiles always remain at less than one sigma from the annual climatology" is not correct and the English ... boah ... this sentence has a (or no) sense like "in the night it is darker than outside" and the rest of the para is likewise not ok. P.10, l.21. "Concerning the CV (Fig. 4), one can see that CO is less variable than O3" ... again, what a strange wording. Why not simply "As shown in Fig. 4, the daily variability (DV) of CO is lower than the one of O3, in particular at altitude above 4 km…" P.10, l.22. Also formulations like "The annual CV of CO shows values ranging from 44% …" are misleading and makes it so difficult to follow. Simply write "The annually averaged variability of CO ranges from 44%", but don't speak of "the annual CV (curriculum vitae) shows values …" In our research field "value" is basically an "item non grata" ... it can be everything

ANSWER : We modified the text as follows : "The Figure 4 shows that the daily variability is lower for CO than for O_3, in particular at the surface and close to the tropopause. Over the entire tropospheric column, the annual daily variability of CO is 20%. It ranges between 44% close to the surface and 17% in the free troposphere where it remains almost constant with altitude. A very similar picture is drawn for the different seasons. The highest values at the surface (in the second half of the troposphere) are encountered in winter/autumn (summer)."

P.11, to P.14, bottom. The entire chapter 3.2 is hard to digest. It is 'only' a description of ups and downs, fairly boring and at the end I had nothing in my mind, basically didn't learn anything. I would shorten it and focus and the clearest features.
ANSWER : Following several recommendations of the first referee, we largely modified this section :

"The average seasonal variations of O_3 and CO in all three tropospheric layers around Frankfurt/Munich are given in Fig. 5, and their long-term time series are shown in Fig. 6.

3.2.1 Ozone

As noted in Sect. 3.1.1, the mean tropospheric O_3 increases with altitude, with average mixing ratios (over the whole period) of 44, 53 and 63 ppb in the LT, MT and UT, respectively. A clear seasonal pattern is emphasized in the entire tropospheric column. In the LT, the seasonal variation shows a broad spring/summer maximum and a minimum in winter, in good accordance with the seasonal variations observed at

surface in Europe (Wilson et al., 2012). In the MT and UT, mixing ratios maximize between May and August. The O_3 5th percentile shows higher mixing ratios in April-May in the LT and MT while the seasonal variations in the UT remains similar than for the mean O_3. The 95th percentile shows a maximum in spring/summer in all tropospheric layers, this maximum being sharper in the UT than in the LT and MT.

The highest monthly mean mixing ratios in the LT (above 60 ppb, the 99th percentile) are observed in August 1995, May 1998, August 2003, July 2006 (Fig. 6). The spring 1998 anomaly is related to the 1997 El Niño that enhanced the pollution export from Asia (due to a higher convective activity and a strengthening of the subtropical jet stream) and North America and may have increased the STE (Koumoutsaris et al., 2008; Zeng and Pyle, 2005). This anomaly is visible in the whole troposphere, and is the strongest in the UT. The anomalies in August 2003 and July 2006 are related to the severe heat waves that struck a large part of Europe (Ordóñez et al., 2005; Solberg et al., 2008; Struzewska and Kaminski, 2008; see also Tressol et al. (2008) for a detailed analysis of the 2003 heat wave with the MOZAIC measurements). They are the strongest in the LT but remain visible in the MT.

The annual mean O_3 mixing ratios are highly correlated between the three tropospheric layers (R=0.87, 0.75 and 0.94 between the LT/MT, LT/UT and MT/UT, respectively). As the sources and sinks of O_3 in the troposphere strongly vary with altitude, such high correlations were not expected. This may be explained by the fact that both the first kilometre and the tropopause layer are not taken into account in this study, which likely greatly reduces the differences of interannual variation among the tropospheric layers as defined in this study. In addition, as shown in Sect. 2.2, the altitude of the UT (defined here based on the PV values) is biased low compared to the UT derived based on the chemical tropopause. Thus, the UT may be less influenced by the stratosphere and more by the free troposphere, which may increase the correlation between the MT and UT. Similar correlations are obtained at the seasonal scale. An interesting exception is the low correlation found between the LT and UT in summer (R=0.26) and spring (0.46). This may be due to the fact that the BL is deeper during these seasons, which keeps away the 1-2 km LT from the influence of the free troposphere.

3.2.2 Carbon monoxide

On average, CO mixing ratios of 143, 115 and 101 ppb are found in the LT, MT and UT, respectively. Mixing ratios in the UT are thus only 29% lower than in the LT close to local emissions. In comparison with O_3, the monthly mean of the daily variability of CO is lower and similar in all three tropospheric layers (around 14-16%). Such a result is expected due to the longer lifetime of CO in comparison with O_3 in most of the troposphere, which leads to a higher regional and hemispheric background (Junge, 1974). As shown in Fig. 5, the seasonal cycle of CO is characterized by maximum mixing ratios in late winter/early spring in the whole troposphere. Minimum mixing ratios are encountered in summer/early autumn in the LT and are slightly shifted to late summer/early autumn higher in altitude. Such a seasonal pattern is consistent with the seasonal variation observed in background air masses arriving at the coastal site Mace Head (Derwent et al., 1998) or at a larger scale by satellite observations (Edwards et al., 2004; Worden et al., 2013). Averaged over the western Europe, the Terra/MOPITT CO tropospheric column maximizes at ~2.5 10^{18} molecules cm^{-2} in March-April and minimize at ~1.9 10^{18} molecules cm^{-2} in late summer, the ratio of the maximum over the minimum being 1.3 (Edwards et al.,

2004). A very similar seasonal variation of tropospheric columns of CO has been observed by Zbinden et al. (2013) based on the MOZAIC data over the period 2002-2009. This is in good agreement with the amplitude of the seasonal cycle observed in the MT, the maximum CO mixing ratio being 1.35 higher than the minimum. The winter-time maximum results from the accumulation of the primary CO emissions at northern mid-latitudes when the photolysis is limited. In summer, CO mixing ratios minimize due to a more effective photolytic destruction, despite an enhanced secondary formation from biogenic compounds and additional emissions from biomass burning (in particular in boreal regions). A rather similar seasonal pattern is observed with the CO 5^{th} and 95^{th} percentiles except that the peak is sharper (February-March) in the LT for the 95^{th} percentile.

As mentioned in Sect. 2.2 and 3.1.2, the data below 1 km were skipped in order to reduce the impact of the local emissions from both the neighbouring agglomeration and the other aircraft — on tarmac and/or during the take-off/landing phases (in case the MOZAIC-IAGOS aircraft closely follows other aircraft). Indeed, many studies have shown that airport activities can impact the air quality at the local scale (e.g. Hu et al., 2009; Pison and Menut, 2004; Yu et al., 2004). It is worth mentioning that in a standard landing take-off cycle – comprising the approach, the taxi (plane on the tarmac), the take-off (acceleration phase on the tarmac) and the climb up to a standard atmospheric boundary layer of 915 m (Kesgin, 2006) – most of the CO emissions (85-95%) occur on the tarmac during the taxi phase (Kurniawan and Khardi, 2011). However, even above 1 km in the LT, one cannot exclude any influence of these emissions, or by the emissions of the neighbouring agglomeration.

To assess more precisely the spatial representativeness of the MOZAIC-IAGOS data in the LT and MT, a comparison is made with the CO mixing ratios measured at the World Meteorology Organisation (WMO) Global Atmosphere Watch (GAW) surface stations (see Sect. S.1 in the Supplement for details). Only the stations located between 45°N and 55°N (i.e. ±5° from the latitude of Frankfurt) and with at least 80% data capture for the period 2002-2012 (based on the monthly time series) are retained, which gives a set of 10 stations. The annual mean CO mixing ratio measured by MOZAIC-IAGOS aircraft in the LT (143 ppb) is in the lower range of the zonal average 155±28 ppb observed among the GAW surface stations. When considering only the stations above 1000 m (i.e. 3 stations all located in Europe), the zonal average is reduced to 145±19 ppb, which is very close to the mean CO observed in the LT. In the MT, the annual mean CO mixing ratio is 115 ppb, thus lower than the CO mixing ratios at the ground whatever the station, but the difference with the highest mountain station Jungfraujoch (3580 m elevation) is very small (-7%). Additionally, the annual MOZAIC-IAGOS CO data in both the LT and MT is strongly correlated with the CO observed at the ground (R between 0.61 and 0.94 at all stations except one at which R=0.41). Therefore, the comparison between the MOZAIC-IAGOS CO dataset at Frankfurt/Munich and the GAW dataset at the same latitude shows a good consistency, both in terms of mean annual CO mixing ratios and interannual variations. This ensures a satisfactory representativeness of the MOZAIC-IAGOS observations.

Along the period 2002-2012, the highest CO annual mixing ratios are encountered in 2003. This is in agreement with the satellite measurements that show on this year a high positive anomaly on CO total columns in Europe and more generally in North Hemisphere, notably due to intense boreal fires (Worden et al., 2013). High mixing ratios in the LT are also observed during the winter 2010, concomitantly with a cold

snap over Europe that may have induced higher CO emissions (for the residential heating) (Cattiaux et al., 2010)."

P.11, l.3. Change "annual and monthly" with "long-term"
ANSWER : Modification applied.

P.11, l.9. Skip "previously" and write instead "As noted in section …" Change this at other locations, too!!!
ANSWER : Modification applied here and elsewhere.

P.11, l.11. Change "in the LT" to "in the entire tropospheric column"
ANSWER : Modification applied.

P.11, l.13. Change "in May and August" to "between May and August"
ANSWER : Modification applied.

P.12, l.8. "… is related to the 1997 El Niño that have enhanced stratospheric-tropospheric exchanges (Ordóñez et al., 2007)" … STE, but also the claim and citation are in my opinion not correct. May be in the meantime there are some first and vague speculations about the impact of ENSO on STE.
ANSWER : The referee is right about this reference of Ordoñez et al. (2007), there is an error, it is not the good one. The relevant reference here is Zeng and Pyle (2005). Over the period 1990-2001, they highlighted a good anti-correlation between the Southern Oscillation Index (SOI) and the STE simulated the UM/CHEM model (UK Met Office Unified Model combined with a tropospheric chemistry module) (with a delay of one year). In particular, they found a strong anomaly of the STE in 1998, after the strong El Niño of 1997-1998. We modified the sentence as follows : "The spring 1998 anomaly is related to the 1997 El Niño that enhanced the pollution export from Asia (due to a higher convective activity and a strengthening of the subtropical jet stream) and North America and may have increased the STE (Koumoutsaris et al., 2008; Zeng and Pyle, 2005)."
In the references :
Zeng, G. and Pyle, J. A.: Influence of El Niño Southern Oscillation on stratosphere/troposphere exchange and the global tropospheric ozone budget, Geophys. Res. Lett., 32(1), L01814, doi:10.1029/2004GL021353, 2005.

P.12, l.19. "variation" instead of "variability"
ANSWER : Modification applied.

P.13, l.7. "which illustrates the high contribution of the CO background at the hemispheric scale" … makes no sense
ANSWER : We removed this part of the sentence: « Mixing ratios in the UT are thus only 29% lower than in the LT close to local emissions. »

P.13, l.9. At the tropopause O3 is longer lived than CO
ANSWER : We modified the sentence as follows : "Such a result is expected due to the longer lifetime of CO in comparison with O_3 in most of the troposphere, which leads to a higher regional and hemispheric background (Junge, 1974)."

P.13, l.22. "in section …" instead of "previously"

ANSWER : Modification applied.

P.13, l.23. I thought you skip the ground-based data. So "on tarmac, take-off/landing" should not count

ANSWER : Indeed, we skipped the data below 1 km in order to minimize the impact of these emissions. As aircraft approach the airport the last phase of their landing or move away quickly from the airport during their takeoff, the emissions on tarmac likely does not influence the observations performed above 1 km. However, even above 1 km, one cannot exclude any influence of the emissions from the Frankfurt (or Munich) city and the other planes flying in close corridors. This is the reason why we compared the mixing ratios in the LT to those measured at regional surface stations in Germany. We modified the sentence as follows : "As mentioned in Sect. 2.2 and 3.1.2, the data below 1 km were skipped in order to reduce the impact of the local emissions from both the neighbouring agglomeration and the other aircraft — on tarmac and/or during the take-off/landing phases (in case the MOZAIC-IAGOS aircraft closely follows other aircraft). Indeed, many studies have shown that airport activities can impact the air quality at the local scale (e.g. Hu et al., 2009; Pison and Menut, 2004; Yu et al., 2004). It is worth mentioning that in a standard landing take-off cycle – comprising the approach, the taxi (plane on the tarmac), the take-off (acceleration phase on the tarmac) and the climb up to a standard atmospheric boundary layer of 915 m (Kesgin, 2006) – most of the CO emissions (85-95%) occur on the tarmac during the taxi phase (Kurniawan and Khardi, 2011). However, even above 1 km in the LT, one cannot exclude any influence of these emissions, or by the emissions of the neighbouring agglomeration. "

P.13, l.34. "Considering the monthly time series, correlations are improved due to the seasonal variations (from 0.61 to 0.90)." doesn't work. P.14, l.5. R = -0.21 means no correlation

ANSWER : This paragraph was greatly modified, and these sentences removed.

P.15, Table 1. Add vertical lines after the 2. and 5. row to enhance the visibility. The sense and significance of the different time periods (1994-2012 and 2000-2012) don't come clear to me. You short motivate this a bit. As written in the introduction, there was a positive O3 trend until ~2000 and thereafter no trend or at certain sites even a weak decrease. The logical way to go would be to analyze the trend between 1994 and ~2004 and between ~2002 and 2012, right? Why you didn't do it? As written, here you have to motivate and explain your approach.

ANSWER : We investigated the trends over the sub-period 2000-2012 precisely because of a levelling-off of the O_3 mixing ratios in the 2000s has been noticed in many studies. We agree that the explanation of this choice was insufficient in the text. However, following the recommendations of the first referee, we decided to avoid analysing the trend over sub-periods because of the strong interannual variability of O_3 in the troposphere makes such analysis tricky. We thus removed the trend results over the period 2000-2012.

P.16, l.6. "in all three" instead of "in the three"

ANSWER : Modification applied here and elsewhere.

P.17, l.4. "Most of the few positive trends found here over the whole period are due to an increase of O3 in the 1990s." makes no sense to me. Elevated values/trend at the beginning of the considered time period should cause decreasing O3 and a negative trend.

ANSWER : We mean here that there has been an increase in the 1990s and a levelling-off during the 2000s. But as explained in our previous answer, we removed the analysis of trends over the sub-period 2000-2012.

P.17, l.4. "Interestingly, at regional background sites in Europe over the 2-3 last decades, Parrish et al. (2012) highlighted that O3 trends, when they are expressed relatively to the concentration in 2000, are quite similar (around +1% O3,2000 yr -1) whatever the site and the season." ... quite awkward English. P.17. The entire para / page, without any line break is very difficult to understand. For me there is no structure and even after 3 times reading, I do not see the take-home message.

ANSWER : Following the recommendations of both referees, we greatly simplified this discussion of O_3 trends:

"All the annual and seasonal trends of the $M(O_3)$ appear insignificant, except in winter for which a weakly significant increase is found in all three tropospheric layers (+0.83[+0.13;+1.67], +0.62[+0.05;+1.22] and +0.62[+0.02;+1.22]$\%O_{3,2000}$ yr^{-1} in the LT, MT and UT, respectively). Previous trend analysis at the alpine sites (Zugspitze since 1978, Jungfraujoch and Sonnblick since 1990) have highlighted (i) a strong increase of O_3 during all seasons in the 1980s (around +0.6-0.9 ppb yr^{-1}), (ii) a persistent but lower increase in the 1990s during all seasons except summer where O_3 has levelled off, (iii) the extension of that levelling off in the 2000s to the other seasons and a slight decrease in summer (Logan et al., 2012; Parrish et al., 2012). This picture is in general agreement with our results in the lower part of the troposphere. More specifically, in winter, Parrish et al. (2012) found an average trend of +0.61±0.25 $\%O_{3,2000}$ yr^{-1} at regional background sites in Europe over the 2-3 last decades, which is consistent with the trends found here over the period 1994-2012. At low altitudes, this increase of O_3 in winter is mainly attributed to a reduced O_3 titration by NO due to decreasing NO_x emissions (e.g. Ordóñez et al., 2005). The persistent positive trends found higher in altitude suggest that wintertime O_3 has increased at a large scale (if not hemispheric) since air masses sampled by MOZAIC-IAGOS aircraft in both the MT and UT can be influenced by emissions from North America and Asia (as shown in Fig. 2).

Concerning the $P_5(O_3)$, a significant increase is found at the annual scale in all three tropospheric layers (+1.03[+0.36;+1.62], +0.42[+0.09;+0.68] and +0.63[+0.09;+0.99]$\%O_{3,2000}$ yr^{-1} in the LT, MT and UT, respectively). Conversely, trends of the $P_{95}(O_3)$ are all insignificant. Note that ignoring the autocorrelation of the data leads to some additional significant positive trends, including the $M(O_3)$ at the annual scale, the $P_5(O_3)$ in winter and autumn, and the $P_{95}(O_3)$ in winter, although not in all tropospheric layers (see Table S1 in the Supplement). It is beyond the scope of this study to investigate why the autocorrelation has a stronger effect on these specific seasons or layers, but this illustrates the strong influence of the serial dependence on the trend analysis and the necessity to take it into account. "

P.17, l.35. "...This is likely due to the fact that only the troposphere is considered in this present study". The variability in the lowermost stratosphere (LMS) is much higher and depends strongly on the actual flight route/statistics. You will thus need

much more data in the LMS to infer significant trends.
ANSWER : Indeed, a higher variability is expected in the LMS. However, in our study, we filtered the observations performed in the LMS (as well as in the tropopause layer). Of course, there is an uncertainty in this approach since we notably rely on the ECMWF PV fields. According to the sensitivity test we added in Sect. 2.2 (i.e. the comparison of our estimation of the tropopause based on the ECMWF-derived 2 pvu with the pressure of the 150 ppb O_3 isopleth), the altitude of UT may be biased low. Anyway, following the recommendation of the first referee, we greatly simplified this Sect. 3.3.1 and removed this part of the discussion.

P.18, eq. 1 and theta_month. If I enter theta = 0, I end up at a maximum on 15. March. Or if t=1 corresponds to 1.Jan, then 31.Dec is t=12.97. More logical would be, if t=1 corresponds to 15. Jan (middle of month) and t=12 to 15. Dec
ANSWER : Indeed, there were some persistent errors in this section, including the one mentioned by the referee. We greatly modified this analysis, following the recommendations of the first referee:

"The seasonal variation of O_3 can be well approximated by a sine function fully characterized by three parameters: an offset value defined here as the average O_3 mixing ratio over the considered period, an amplitude, and a phase that determines at which period in the year the maximum of O_3 is reached. Following the approach of Parrish et al. (2013), one can fit a sine function over different periods of time and compare the results of the fit in order to highlight potential changes in the seasonal pattern of O_3. While Parrish et al. (2013) applied the sine fit to the monthly mean time series, we here consider the daily mean O_3 mixing ratio but the results from both approaches will be discussed . The equation of the fit is :

$$\tilde{y}(t) = y_0 + a \sin\left(\frac{2\pi t}{365} + \phi\right) \quad (1)$$

with t the time (in days, values ranging between 0.5 and 364.5), y_0 the offset mixing ratio (in ppb), a the amplitude (in ppb) and ϕ the phase. The date of the year of the seasonal maximum of O_3 is then estimated as : $(\pi/2 - \phi) * 365/2\pi$ (Parrish et al., 2013). We apply the sine fit on the two 9-year time periods 1995-2003 and 2004-2012. As there is no overlap between these periods, the two datasets and the results of the sine fit are independent. The changes of amplitude and phase obtained with the sine fits are reported in Table 2.

Between 1995-2003 and 2004-2012, the amplitude of the O_3 seasonal cycle has significantly decreased in the whole troposphere, with a rate of decrease of -2.5±0.9, -1.1±0.5 and -2.1±1.0 ppb decade^{-1} in the LT, MT and UT, respectively. Reason for the decreasing amplitude is the significantly increased yearly O_3 minimum occurring in winter and to the same time constant O_3 maximum occurring in spring/summer (see Sect. 3.3.1). The differences of amplitude change between the different layers all remain statistically insignificant.

Table 2. Characteristics of the O_3 seasonal cycle over the periods 1995-2003 and 2004-2012 in all tropospheric layers. Amplitude and phase are obtained by fitting a sine function on the daily mean O_3 mixing ratios (see text).

Layer	Amplitude	Phase

	Amplitude 1995-2003 (ppb)	Amplitude 2004-2012 (ppb)	Amplitude trend (ppb decade^{-1})	Date of seasonal maximum 1995-2003	Date of seasonal maximum 2004-2012	Shift (day decade^{-1})
UT	18.0±0.7	16.1±0.6	-2.1±1.0	23 June ± 2 days	20 June ± 2 days	-3.3±3.3
MT	11.5±0.3	10.5±0.3	-1.1±0.5	23 June ± 1 days	16 June ± 2 days	-7.8±2.5
LT	9.9±0.6	7.6±0.5	-2.5±0.9	18 June ± 3 days	2 June ± 4 days	-17.8±6.0

Over the period 1995-2003, the sine fit gives a seasonal maximum of O_3 the 18 June in the LT and the 23 June in the MT and UT. The date of seasonal maximum in the LT is in reasonable agreement with those obtained by Parrish et al. (2013) at two alpine sites (Jungfraujoch, Switzerland and Zugspitze, Germany) and at a lower elevation site (Hohenpeissenberg, Germany, ~50 km from Munich). Over the period 2004-2012, the seasonal maximum O_3 occurs the 2 June in the LT, the 16 June in the MT and the 20 June in the UT. Thus, the phase of the seasonal variations of O_3 shifted forward during the period 1995-2012. The seasonal shift between 1995-2003 and 2004-2012 is highly significant in the LT (-17.8±6.0 day decade^{-1}) and MT (-7.8±2.5 day decade^{-1}), and nearly insignificant in the UT (-3.3±3.3 day decade^{-1}). The differences of seasonal shift between the tropospheric layers are all significant, and the seasonal shift thus decreases with altitude. Note that applying the sine fit to the monthly O_3 mixing ratios give similar shift estimates but much larger uncertainties, leading to insignificant differences among the tropospheric layers (-13.3±11.6 and -6.7±6.5 day decade^{-1} in the LT and MT, respectively). Note that reducing the width of the time windows (to less than 9 years) does not give significantly different results.

At the three continental sites, Parrish et al. (2013) reported statistically significant rates of shift (at the 95% confidence level) ranging between -5 and -7 days decade^{-1} since 1970s while at the coastal site Mace Head, the rate was lower and insignificant (-3±3.7 days decade^{-1}). In comparison, the seasonal shift we obtained in the LT is significantly higher, but discrepancies are likely due to the fact that the studied periods are different. As a faster change of phase is found between 2005 and 2008 (the 3 last years studied) (see Fig. 2 in Parrish et al. (2013)), restricting their analysis to our shorter period would likely lead to a higher seasonal shift (i.e., closer to our values)."

Now 3 examples for the strange (and difficult to follow) type of formulation:
P.20, l.7. "Results show a decrease around 1-2 ppb of the amplitude" instead of simply "The amplitude decreases by 1-2 ppb"
ANSWER : This paragraph was modified, see our previous answer.

P.20, l.10. "… This is consistent with the fact that O3 has increased significantly in the whole troposphere during the winter (the season of minimum O3) but not during spring/summer (the season of maximum O3)" instead of "Reason for the decreasing amplitude is the significantly increased yearly O3 minimum occurring in winter and to the same time constant O3 maximum occurring in spring/summer"
ANSWER : Modification applied.

P.20, l.13. "Concerning the phase of the O3 seasonal cycle, results clearly highlight a

shift toward earlier O3 maximum along the 1995-2012 period" instead of simply "The phase of the seasonal variation of O3 was found to shift forward during the period 1995-2012"
ANSWER : We modified the sentence as follows : "Thus, the phase of the seasonal variations of O_3 shifted forward during the period 1995-2012."

P.21, l.14. coastal site Mace Head
ANSWER : Modification applied here and elsewhere.

P.21, l.16. 3 last years studied
ANSWER : Modification applied.

P.21, l.31. "much lower ones in the UT and to a lesser extent in the MT." ... boah
ANSWER : We modified this section and this sentence was removed.

P.21, l.32. ... worldwide densest
ANSWER : Modification applied.

P.22, l.1. "shift extent" ... boah
ANSWER : We modified this section and this text was removed.

P.22, l.6. "an unambiguous explanation". If so, you give some vague hints
ANSWER : In this sentence, we say that our study does <u>not</u> provide an unambiguous explanation to the different trends and seasonal shifts.

P.22, l.11. (e.g., Thouret et al., 2006). Sorry, but there are hundreds of publications which are better suited for STE processes and the springtime maximum of downward stratospheric transport
ANSWER : The reference of Thouret et al. (2006) is related to the second part of the sentence (i.e. the maximum of O_3 in spring in the LS). We added the reference of Appenzeller et al. (1996) to justify the first part (i.e. the enhanced downward transport during spring) :
"In terms of stratospheric contributions, the STE is known to peak in spring (Auvray and Bey, 2005; James et al., 2003; Tang et al., 2011) due to both enhanced downward transport (Appenzeller et al., 1996) and maximum mixing ratios in the lowermost stratosphere (e.g., Thouret et al., 2006)."
In the references :
Appenzeller, C., Holton, J. R. and Rosenlof, K. H.: Seasonal variation of mass transport across the tropopause, J. Geophys. Res. Atmos., 101(D10), 15071–15078, doi:10.1029/96JD00821, 1996.

P.22, l.14. which contradicts our observations
ANSWER : Modification applied.

P.22, l.18. our period ?
ANSWER : We replaced this text by : "the period 1994-2012"**.**

P.22, l.31. "... based on the potential vorticity extracted from ECMWF meteorological data" No, only for defining the upper border of the UT
ANSWER : We modified the text as follows : "We focus on the troposphere, each

vertical profile being subdivided in three tropospheric layers: the lower, mid- and upper troposphere (LT, MT and UT, respectively). The UT is defined relative to the dynamical tropopause, based on the potential vorticity extracted from ECMWF meteorological data. »

List of changes in the manuscript.
For clarity, the removed sentences are in red, the added/modified sentences in green. The sign « ➔ » means « is replaced by ».
:

- **Page 3, lines 8-14 :** « A full understanding […] various sources and sinks. » **is removed.**
- **Page 3, lines 14-22 :** « At the Mace Head coastal site […] and an increase of ozone imports » ➔ « Observations at the coastal site Mace Head have shown an average annual increase of +0.25±0.09 ppb $year^{-1}$ of the baseline (i.e. originating from the Northern Hemispheric marine boundary layer) O_3 mixing ratios during the period 1988-2012 (Simmonds et al., 2004, Derwent et al., 2013). This increase has been the strongest in winter and spring and the lowest in summer, and has slowed down during the 2000s (Derwent et al., 2013). In contrast, the annual O_3 mixing ratios in European air masses have shown a much lower increase (Derwent et al., 2013), which suggests a possible compensation between a decrease of O_3 local formation in Europe and an increase of O_3 imports."
- **Page 3, line 32 :** « (close to precursors emissions and/or deposition sink) » **is removed**
- **Page 3, line 35 :** « As a long lifetime […] » ➔« As a moderate lifetime […] »
- **Page 4, line 1 :** « interesting pollution tracer » ➔ « powerful pollution tracer »
- **Page 4, lines 19-21** : « Note also […] Staufer et al., 2013, 2014). » **is removed (actually, moved to another paragraph, see below)**
- **Page 4, line 32** : « Note also that several studies have investigated the consistency of the MOZAIC-IAGOS O_3 dataset with other types of in-situ data (e.g., surface stations, ozonesonde) (Logan et al., 2012; Staufer et al., 2013, 2014; Tanimoto et al., 2015; Zbinden et al., 2013). Focusing on O_3 changes in Europe, Logan et al. (2012) showed a reasonable agreement between aircraft and alpine sites but noticed the absence of O_3 increase in 1994-1998 in the sonde dataset (contrary to the two other types of data). Focusing on the pure tropospheric profiles, Zbinden et al. (2013) found a mean difference between MOZAIC-IAGOS and sondes of -2% in Germany, -8% in eastern United Sates and +1% over Japan. Tanimoto et al. (2015) obtained similar results, with differences between aircraft and sondes data around ±2% throughout the whole troposphere in Belgium, Germany and Japan, and ±5% at Hong Kong. The MOZAIC-IAGOS data at Munich were found to compare reasonably well with the surface observations at Hohenpeissenberg (slope of 0.97, correlation of 0.77)." **is added.**
- **Page 5, line 15 :** « The influence of the local emissions on the observations in the LT will be briefly discussed in Sect. 3.2.2. » **is added.**
- **Page 5, lines 23-24 : «** The $p_{PV=2}$ parameter is used to determine the DT pressure at the top of the selected tropospheric vertical profile. **»** ➔ **«** The pressure at the DT ($p_{PV=2}$) plus 15 hPa defined the top of the UT applied here. **»**
- **Page 7, line 12 : «** In order to assess the uncertainties introduced by an erroneous DT pressure, we compared it with the pressure at which the O_3 mixing ratio reaches 150 ppb (taken here a simplified estimate of the

chemical tropopause) and remains above at higher altitude (in order to avoid stratospheric intrusions in the troposphere). This was done on all vertical profiles where it was possible, which represents 46% of the dataset. On average over the period 1994-2012, the mean bias of the DT pressure compared to the 150 ppb O_3-isopleth is +21 hPa, while the 5th, 10th, 90th and 95th percentiles of this bias are -72, -32, +78 and +99 hPa, respectively. Therefore, the DT derived from PV values tends to be located below the 150 ppb O_3-isopleth, which may bias low the O_3 mixing ratios in the UT. However, the discrepancy remains moderate on most profiles. **» is added.**

- **Page 8, lines 12 – page 9, line 5 : «** Between the surface (2 km) and the tropopause, the annually-averaged O_3 mixing ratios range between 21-81 (47-81) ppb. Averaged over the entire tropospheric column, the annually-averaged O_3 mixing ratio is 56 ppb. The seasonal variability is strong, with minimum annually-averaged O_3 (over the tropospheric column) in winter (44 ppb) and autumn (48 ppb), and maximum ones in summer (67 ppb) and spring (61 ppb). The O_3 abundance clearly increases with altitude. The highest vertical gradients are found close to the surface all along the year (dry deposition and titration by NO) and at the vicinity of the tropopause during spring and summer (exchanges with the stratospheric reservoir). The inflexion of vertical gradients at about 1 km a.g.l. has already been mentioned in Chevalier et al. (2007). In the free troposphere, the vertical gradients are very low in winter and to a lesser extent in autumn, but substantially enhanced during spring and summer where concentrations quickly increase with altitude. **» ➔ «** The standard deviation shown in Fig. 3 is inferred from the daily mean vertical profiles, and thus represents the daily variability at the seasonal scale. Over the entire tropospheric column, the annual mean O_3 mixing ratio is 56 ppb. The mean O_3 over the tropospheric column shows the minimum mixing ratios in winter (44 ppb) and autumn (48 ppb), and the maximum ones in summer (67 ppb) and spring (61 ppb). The annual mean O_3 mixing ratio increases with altitude, from 21 ppb at ground to 81 ppb at 12 km (47 ppb at 2 km). The highest vertical gradients are found close to the surface (due to dry deposition and enhanced titration by NO in the BL) during the whole year and close to 12 km during spring and summer (due to STE). The inflexion of vertical gradients at about 1 km a.g.l. has already been mentioned in Chevalier et al. (2007). Above 3 km, the mean vertical gradients are +1.1, +1.5, +3.0 and +5.1 ppb km^{-1} in winter, autumn, spring and summer, respectively. **»**
- **Page 9, line 5 : «** During the summer, O_3 episodes are often observed in the European BL (van Loon et al., 2007; Meleux et al., 2007). High O_3 mixing ratios are also measured in urban environments, despite the presence of NO_x emitted locally by the anthropogenic activities (Vautard et al., 2007 reported a 95th percentile of daily O_3 maximum ranging between 70 and 100 ppb in 4 European megacities in 2010). Thus, one might have expected higher mixing ratios in the BL than in the lower free troposphere (sometimes described as a « C » shaped profile). However, observations do not show such a profile. One may suspect that this is due to the night-time titration of O_3 in the BL but limiting data to the

afternoon does not highlight a clear « C » shaped profile. Actually, such « C » shaped profile is only observed when considering the 95th percentile rather than the mean O_3 mixing ratio (Petetin et al., Diurnal cycle of ozone throughout the troposphere over Frankfurt as measured by MOZAIC-IAGOS commercial aircraft, under review in Elementa Science of the Anthropocene). It means that the potentially high O_3 pollution in the BL during the summer can greatly modify the vertical profile of O_3 mixing ratios but only episodically. On average, the structure of the mean O_3 vertical profile in summer remains qualitatively the same (i.e. positive gradient through the whole troposphere) as during the rest of the year. **» is added.**

- **Page 9, lines 10-23 : The text is slightly reformulated as «** To further characterize the variability of O_3 and CO above Frankfurt/Munich, we now investigate the daily variability at both the annual and seasonal scales. The daily variability is here defined as the coefficient of variation (CV) of the daily-averaged mixing ratios, that is the standard deviation normalized by the corresponding (i.e. annual or seasonal) mean mixing ratio. Vertical profiles of the daily variability for O_3 and CO are shown in Fig. 4. Results about CO will be discussed in Sect. 3.1.2. The daily variability of O_3 at the annual scale ranges between 20 and 73% depending on the altitude, with a mean value of 32%. The maximum daily variability of O_3 is found at ground (73%) and at 12 km (53%) where it is likely driven by intense shallow and transient exchanges between the stratosphere and the troposphere (Stohl et al., 2003b). Conversely, the minimum daily variability is found at about 3.4 km. Such daily variability is lower at the seasonal scale, at most altitudes and during most seasons, but the shape of the vertical profiles remains similar. The seasonal daily variability minimizes at 4.4 km in autumn and between 3.1-3.4 km during the other seasons, thus close to the minimum annual diurnal variability. Similarly, it maximizes at the surface and close to the tropopause. Interestingly, the daily variability above 11 km is noticeably higher in spring than during the other seasons, which again may be due to the day-to-day variability of STE that peaks during that season. **»**
- **Page 10, lines 15-26 : «** Contrary to O_3 [...] are encountered in winter/autumn (summer). **» ➔ «** The Figure 4 shows that the daily variability is lower for CO than for O_3, in particular at the surface and close to the tropopause. Over the entire tropospheric column, the annual daily variability of CO is 20%. It ranges between 44% close to the surface and 17% in the free troposphere where it remains almost constant with altitude. A very similar picture is drawn for the different seasons. The highest values at the surface (in the second half of the troposphere) are encountered in winter/autumn (summer). **»**
- **Page 11, line 4 : «** [...] (although the profile is sharper in the UT). **» ➔ «** [...] , this maximum being sharper in the UT than in the LT and MT. **»**
- **Page 11, lines 4-6 : «** On average, the monthly mean of the daily O_3 variability (represented by monthly standard deviations in Fig. 6) represents about 20% of the mean in the LT and MT, and 25% in the UT. **» is removed.**

- **Page 12, lines 8-11 : «** The spring 1998 anomaly has been discussed by Koumoutsaris et al. (2008) and is related to the 1997 El Niño that have enhanced stratospheric-tropospheric exchanges (Ordóñez et al., 2007) and pollution export from Asia (higher convective activity and strengthening of the subtropical jet stream) and North America. **»** ➔ **«** The spring 1998 anomaly is related to the 1997 El Niño that enhanced the pollution export from Asia (due to a higher convective activity and a strengthening of the subtropical jet stream) and North America and may have increased the STE (Koumoutsaris et al., 2008; Zeng and Pyle, 2005). This anomaly is visible in the whole troposphere, and is the strongest in the UT.**»**
- **Page 12, lines 14-18 : «** Most of these O_3 anomalies in the LT are not always distinguishable in the MT, where the highest concentrations (above 68 ppb, the 99th percentile) are encountered in August 2004 and July 2002, 2006 and 2008. Similarly, the highest monthly concentrations do not always coincide between the MT and UT (see for instance the high O_3 mixing ratio observed in the UT during 2011 and 2012 summers). **» is removed.**
- **Page 12, lines 18-20 : «** Nevertheless, on a yearly average, a very similar interannual variability is found between the three tropospheric layers as illustrated by high correlations (R=0.87, 0.75 and 0.94 between the LT/MT, LT/UT and MT/UT, respectively). **»** ➔ **«** The annual mean O_3 mixing ratios are highly correlated between the three tropospheric layers (R=0.87, 0.75 and 0.94 between the LT/MT, LT/UT and MT/UT, respectively). As the sources and sinks of O_3 in the troposphere strongly vary with altitude, such high correlations were not expected. This may be explained by the fact that both the first kilometre and the tropopause layer are not taken into account in this study, which likely greatly reduces the differences of interannual variation among the tropospheric layers as defined in this study. In addition, as shown in Sect. 2.2, the altitude of the UT (defined here based on the PV values) is biased low compared to the UT derived based on the chemical tropopause. Thus, the UT may be less influenced by the stratosphere and more by the free troposphere, which may increase the correlation between the MT and UT. Similar correlations are obtained at the seasonal scale. An interesting exception is the low correlation found between the LT and UT in summer (R=0.26) and spring (0.46). This may be due to the fact that the BL is deeper during these seasons, which keeps away the 1-2 km LT from the influence of the free troposphere. **»**
- **Page 13, line 7 : «** , which illustrates the high contribution of the CO background at the hemispheric scale **» is removed**
- **Page 13, line 10 : «** in most of the troposphere **» is added**
- **Page 13, lines 14-19 : «** Such a seasonal pattern is rather consistent with satellite observations (Worden et al., 2013). It results from a maximum of primary CO emissions in winter (at northern mid-latitudes) associated to a limited photolysis which increases the CO lifetime and allows its accumulation in the atmosphere. This delays the maximum of CO concentrations to late winter/early spring, after which concentrations start to decrease due to a more effective photolytic destruction in

summer, despite an enhanced secondary formation from biogenic compounds. » ➔ « Such a seasonal pattern is consistent with the seasonal variation observed in background air masses arriving at the coastal site Mace Head (Derwent et al., 1998) or at a larger scale by satellite observations (Edwards et al., 2004; Worden et al., 2013). Averaged over the western Europe, the Terra/MOPITT CO tropospheric column maximizes at ~2.5 10^{18} molecules cm^{-2} in March-April and minimize at ~1.9 10^{18} molecules cm^{-2} in late summer, the ratio of the maximum over the minimum being 1.3 (Edwards et al., 2004). A very similar seasonal variation of tropospheric columns of CO has been observed by Zbinden et al. (2013) based on the MOZAIC data over the period 2002-2009. This is in good agreement with the amplitude of the seasonal cycle observed in the MT, the maximum CO mixing ratio being 1.35 higher than the minimum. The winter-time maximum results from the accumulation of the primary CO emissions at northern mid-latitudes when the photolysis is limited. In summer, CO mixing ratios minimize due to a more effective photolytic destruction, despite an enhanced secondary formation from biogenic compounds and additional emissions from biomass burning (in particular in boreal regions). »

- **Page 13, line 22 – page 14, line 6 : «** As previously mentioned, the local emissions both from the neighbouring agglomeration and from the other aircraft — on tarmac and/or during the take-off/landing phases (in case the MOZAIC-IAGOS aircraft closely follows other aircraft) — may add some variability depending on the local dispersion conditions and thus influence the CO measurements in the LT. To assess more precisely the spatial representativeness of these MOZAIC-IAGOS LT data, some surface measurements are available at four German stations from the World Meteorology Organisation (WMO) Global Atmosphere Watch (GAW) database: Hohenpeissenberg (47.8°N, 11.0°W; at 50 km South-West from Munich), Neuglobsow (53.1°N, 13.0°W), Schauinsland (47.9°N, 7.9°W), Ochsenkopf site (50.0°N, 11.8°W). We investigate the correlations of both the seasonal and annual mean CO concentrations between the MOZAIC-IAGOS measurements (in the LT) and these surface observations (see Fig. S4 in the Supplement). At the annual scale, a reasonable agreement is found, with correlations (R) of annual mean CO between 0.56 (at Neuglobsow) and 0.81 (at Schauinsland). Considering the monthly time series, correlations are improved due to the seasonal variations (from 0.61 to 0.90). At the seasonal scale, correlations remain satisfactory in winter, with values between 0.65 and 0.87. However, results appear more contrasted among the GAW stations during the other seasons. In spring, all correlations are above 0.69 except Hohenpeissenberg (0.52). In summer, both Hohenpeissenberg and Ochsenkopf have a low correlation with MOZAIC-IAGOS (0.34 and 0.52, respectively), while very high correlations (above 0.96) are found for the two other stations. In autumn, the correlation is very low for Neuglobsow (-0.21), moderate for Ochsenkopf (0.51) and satisfactory for Hohenpeissenberg and Schauinsland (0.73 and 0.85). » ➔ « As mentioned in Sect. 2.2 and 3.1.2, the data below 1 km were skipped in order to reduce the impact of the local emissions from both the neighbouring agglomeration and the other

aircraft — on tarmac and/or during the take-off/landing phases (in case the MOZAIC-IAGOS aircraft closely follows other aircraft). Indeed, many studies have shown that airport activities can impact the air quality at the local scale (e.g. Hu et al., 2009; Pison and Menut, 2004; Yu et al., 2004). It is worth mentioning that in a standard landing take-off cycle – comprising the approach, the taxi (plane on the tarmac), the take-off (acceleration phase on the tarmac) and the climb up to a standard atmospheric boundary layer of 915 m (Kesgin, 2006) – most of the CO emissions (85-95%) occur on the tarmac during the taxi phase (Kurniawan and Khardi, 2011). However, even above 1 km in the LT, one cannot exclude any influence of these emissions, or by the emissions of the neighbouring agglomeration.

To assess more precisely the spatial representativeness of the MOZAIC-IAGOS data in the LT and MT, a comparison is made with the CO mixing ratios measured at the World Meteorology Organisation (WMO) Global Atmosphere Watch (GAW) surface stations (see Sect. S.1 in the Supplement for details). Only the stations located between 45°N and 55°N (i.e. ±5° from the latitude of Frankfurt) and with at least 80% data capture for the period 2002-2012 (based on the monthly time series) are retained, which gives a set of 10 stations. The annual mean CO mixing ratio measured by MOZAIC-IAGOS aircraft in the LT (143 ppb) is in the lower range of the zonal average 155±28 ppb observed among the GAW surface stations. When considering only the stations above 1000 m (i.e. 3 stations all located in Europe), the zonal average is reduced to 145±19 ppb, which is very close to the mean CO observed in the LT. In the MT, the annual mean CO mixing ratio is 115 ppb, thus lower than the CO mixing ratios at the ground whatever the station, but the difference with the highest mountain station Jungfraujoch (3580 m elevation) is very small (-7%). Additionally, the annual MOZAIC-IAGOS CO data in both the LT and MT is strongly correlated with the CO observed at the ground (R between 0.61 and 0.94 at all stations except one at which R=0.41). Therefore, the comparison between the MOZAIC-IAGOS CO dataset at Frankfurt/Munich and the GAW dataset at the same latitude shows a good consistency, both in terms of mean annual CO mixing ratios and interannual variations. This ensures a satisfactory representativeness of the MOZAIC-IAGOS observations. **»**

- **Figure 7 is removed**
- **In Table 1, the trend results of O3 are removed, and replaced by the absolute trends of CO.**
- **Page 16, line 5 – page 17, line 36 : All the discussion of Sect. 3.3.1 is simplified as follows : «** All the annual and seasonal trends of the $M(O_3)$ appear insignificant, except in winter for which a weakly significant increase is found in all three tropospheric layers (+0.83[+0.13;+1.67], +0.62[+0.05;+1.22] and +0.62[+0.02;+1.22]$\%O_{3,2000}$ yr^{-1} in the LT, MT and UT, respectively). Previous trend analysis at the alpine sites (Zugspitze since 1978, Jungfraujoch and Sonnblick since 1990) have highlighted (i) a strong increase of O_3 during all seasons in the 1980s (around +0.6-0.9 ppb yr^{-1}), (ii) a persistent but lower increase in the 1990s during all seasons except summer where O_3 has levelled off, (iii) the extension of

that levelling off in the 2000s to the other seasons and a slight decrease in summer (Logan et al., 2012; Parrish et al., 2012). This picture is in general agreement with our results in the lower part of the troposphere. More specifically, in winter, Parrish et al. (2012) found an average trend of $+0.61\pm0.25$ $\%O_{3,2000}$ yr^{-1} at regional background sites in Europe over the 2-3 last decades, which is consistent with the trends found here over the period 1994-2012. At low altitudes, this increase of O_3 in winter is mainly attributed to a reduced O_3 titration by NO due to decreasing NO_x emissions (e.g. Ordóñez et al., 2005). The persistent positive trends found higher in altitude suggest that wintertime O_3 has increased at a large scale (if not hemispheric) since air masses sampled by MOZAIC-IAGOS aircraft in both the MT and UT can be influenced by emissions from North America and Asia (as shown in Fig. 2).

Concerning the $P_5(O_3)$, a significant increase is found at the annual scale in all three tropospheric layers (+1.03[+0.36;+1.62], +0.42[+0.09;+0.68] and +0.63[+0.09;+0.99]$\%O_{3,2000}$ yr^{-1} in the LT, MT and UT, respectively). Conversely, trends of the $P_{95}(O_3)$ are all insignificant. Note that ignoring the autocorrelation of the data leads to some additional significant positive trends, including the $M(O_3)$ at the annual scale, the $P_5(O_3)$ in winter and autumn, and the $P_{95}(O_3)$ in winter, although not in all tropospheric layers (see Table S1 in the Supplement). It is beyond the scope of this study to investigate why the autocorrelation has a stronger effect on these specific seasons or layers, but this illustrates the strong influence of the serial dependence on the trend analysis and the necessity to take it into account. **»**

- **Page 18, lines 22-24 : «** In the previous section, we highlighted differences in the O_3 trends depending on the season and the tropospheric layer. Here, we investigate if these contrasted trends come along a change of the O_3 seasonal cycle above Frankfurt/Munich (Sect. 3.4.1). » ➔ « In Sect. 3.3.1, we highlighted that only a few O_3 trends are statistically significant. However, the differences of trends between the seasons remain insignificant. It is worth noting that an insignificant trend does not imply the absence of trend since a trend can be hidden by a strong variability. In this section, we investigate if these trends come along with a change of the O_3 seasonal cycle above Frankfurt/Munich (Sect. 3.4.1). »

- **Page 18, line 26 – page 21, line 25 : The whole paragraph is modified and the figure removed. The new text is : «** The seasonal variation of O_3 can be well approximated by a sine function fully characterized by three parameters: an offset value defined here as the average O_3 mixing ratio over the considered period, an amplitude, and a phase that determines at which period in the year the maximum of O_3 is reached. Following the approach of Parrish et al. (2013), one can fit a sine function over different periods of time and compare the results of the fit in order to highlight potential changes in the seasonal pattern of O_3. While Parrish et al. (2013) applied the sine fit to the monthly mean time series, we here consider the daily mean O_3 mixing ratio but the results from both approaches will be discussed . The equation of the fit is :

$$\tilde{y}(t) = y_0 + a \sin\left(\frac{2\pi t}{365} + \phi\right) \quad (1)$$

with t the time (in days, values ranging between 0.5 and 364.5), y_0 the offset mixing ratio (in ppb), a the amplitude (in ppb) and ϕ the phase. The date of the year of the seasonal maximum of O_3 is then estimated as : $(\pi/2 - \phi) * 365/2\pi$ (Parrish et al., 2013). We apply the sine fit on the two 9-year time periods 1995-2003 and 2004-2012. As there is no overlap between these periods, the two datasets and the results of the sine fit are independent. The changes of amplitude and phase obtained with the sine fits are reported in Table 2.

Between 1995-2003 and 2004-2012, the amplitude of the O_3 seasonal cycle has significantly decreased in the whole troposphere, with a rate of decrease of -2.5±0.9, -1.1±0.5 and -2.1±1.0 ppb decade^{-1} in the LT, MT and UT, respectively. Reason for the decreasing amplitude is the significantly increased yearly O_3 minimum occurring in winter and to the same time constant O_3 maximum occurring in spring/summer (see Sect. 3.3.1). The differences of amplitude change between the different layers all remain statistically insignificant.

Table 2. Characteristics of the O_3 seasonal cycle over the periods 1995-2003 and 2004-2012 in all tropospheric layers. Amplitude and phase are obtained by fitting a sine function on the daily mean O_3 mixing ratios (see text).

Layer	Amplitude			Phase		
	Amplitude 1995-2003 (ppb)	Amplitude 2004-2012 (ppb)	Amplitude trend (ppb decade^{-1})	Date of seasonal maximum 1995-2003	Date of seasonal maximum 2004-2012	Shift (day decade^{-1})
UT	18.0±0.7	16.1±0.6	-2.1±1.0	23 June ± 2 days	20 June ± 2 days	-3.3±3.3
MT	11.5±0.3	10.5±0.3	-1.1±0.5	23 June ± 1 days	16 June ± 2 days	-7.8±2.5
LT	9.9±0.6	7.6±0.5	-2.5±0.9	18 June ± 3 days	2 June ± 4 days	-17.8±6.0

Over the period 1995-2003, the sine fit gives a seasonal maximum of O_3 the 18 June in the LT and the 23 June in the MT and UT. The date of seasonal maximum in the LT is in reasonable agreement with those obtained by Parrish et al. (2013) at two alpine sites (Jungfraujoch,

Switzerland and Zugspitze, Germany) and at a lower elevation site (Hohenpeissenberg, Germany, ~50 km from Munich). Over the period 2004-2012, the seasonal maximum O_3 occurs the 2 June in the LT, the 16 June in the MT and the 20 June in the UT. Thus, the phase of the seasonal variations of O_3 shifted forward during the period 1995-2012. The seasonal shift between 1995-2003 and 2004-2012 is highly significant in the LT (-17.8±6.0 day decade^{-1}) and MT (-7.8±2.5 day decade^{-1}), and nearly insignificant in the UT (-3.3±3.3 day decade^{-1}). The differences of seasonal shift between the tropospheric layers are all significant, and the seasonal shift thus decreases with altitude. Note that applying the sine fit to the monthly O_3 mixing ratios give similar shift estimates but much larger uncertainties, leading to insignificant differences among the tropospheric layers (-13.3±11.6 and -6.7±6.5 day decade^{-1} in the LT and MT, respectively). Note that reducing the width of the time windows (to less than 9 years) does not give significantly different results.

At the three continental sites, Parrish et al. (2013) reported statistically significant rates of shift (at the 95% confidence level) ranging between -5 and -7 days decade^{-1} since 1970s while at the coastal site Mace Head, the rate was lower and insignificant (-3±3.7 days decade^{-1}). In comparison, the seasonal shift we obtained in the LT is significantly higher, but discrepancies are likely due to the fact that the studied periods are different. As a faster change of phase is found between 2005 and 2008 (the 3 last years studied) (see Fig. 2 in Parrish et al. (2013)), restricting their analysis to our shorter period would likely lead to a higher seasonal shift (i.e., closer to our values). **»**

- **Page 21, line 29 – page 22, line 2 : «** Thanks to vertical profile observations, it brings an interesting contribution by highlighting that these seasonal changes above Frankfurt/Munich depend on altitude, with the highest shift in the LT and much lower ones in the UT and to a lesser extent in the MT. It is worth noting that the MOZAIC-IAGOS observations above Frankfurt/Munich represent the densest dataset of O_3 vertical profiles in the world, which gives robustness to our results. Qualitatively, a similar behaviour is observed in ozonesonde observations but the reliability of these results is limited by a much lower number of vertical profiles. However, a high variability from one site to the other in Europe affects both the rate of decrease with altitude and the shift extent itself, as confirmed by results obtained at surface stations. **»** ➔ **«** It is worth noting that the MOZAIC-IAGOS observations above Frankfurt/Munich represent the worldwide densest dataset of O_3 vertical profiles, which gives robustness to our results. Thanks to vertical profile observations, it brings an interesting contribution by showing that this seasonal change of the phase above Frankfurt/Munich decreases with altitude. This may highlight that the O_3 seasonal pattern behave differently over the northern hemisphere continents (Europe, North America, Asia). Indeed, the FLEXPART-derived PES clearly shows that the air masses sampled by MOZAIC-IAGOS aircraft in the different tropospheric layers originate from different regions (see Fig. 2). The LT is predominantly influenced by the European emissions, the MT by both the European and Northern

American emissions, the UT by both the Northern American and Asian emissions. **»**

- **Page 22, lines 18-24 : «** However, one can qualitatively notice on Fig. 7 that over our period, the O_3 anomalies in summer tend to be more and more negative in the LT, contrary to the UT where they tend to be slightly positive at the end of the period. This is probably mainly due to the decrease of O_3 precursors emissions in Europe (Derwent et al., 2003; Solberg et al., 2005; Jonson et al., 2006). As there are no such differences of O_3 anomalies between the tropospheric layers during the spring, this may at least partly explain the higher seasonal shift found close to the surface. **» is removed.**
- **Page 24, lines 1-15 : «** This study also investigates the changes in the O_3 seasonal cycle (by fitting sinusoids over 9-years moving time periods) with a focus on the phase. Results highlight a statistically significant change of the phase in the LT, ozone maxima occurring earlier by -12.1±4.1 days decade^{-1} on average (at a 95% confidence level), in general agreement with previous results from the literature (Parrish et al., 2013). Observations at most surface stations in Central Europe show seasonal shifts in the same direction (toward earlier maxima), but with a strong variability from one station to the other. A major contribution of this study concerns the dependence on altitude of this seasonal shift, as it is found to decrease by a factor of two in the mid-troposphere (-5.2±2.3 days decade^{-1}) and five in the upper troposphere (-2.3±2.1 days decade^{-1}). Qualitatively, a similar dependence on altitude is obtained with ozonesonde observations at most sites in Europe. The occurrence of negative O_3 anomalies during the summer in the LT but not in the other layers, probably induced by the reduction of O_3 precursors emissions in Europe, may at least partly explain the higher seasonal shift observed at low altitudes. **»** ➔ **«** This study also investigates the changes in the O_3 seasonal cycle (by fitting sinusoids over the 9-years periods 1995-2003 and 2004-2012) with a focus on the phase. Results highlight a statistically significant change of the phase in the LT, ozone maxima occurring earlier by -17.8±6.0 days decade^{-1} on average (at a 95% confidence level), in general agreement with previous results from the literature (Parrish et al., 2013). A major contribution of this study concerns the dependence on altitude of this seasonal shift, as it is found to decrease with altitude, with -7.8±2.5 days decade^{-1} in the mid-troposphere and -3.3±3.3 days decade^{-1} in the upper troposphere (i.e. nearly insignificant). **»**