

Author comment: We would like to thank reviewer #1 for the positive assessment of our manuscript and the very constructive comments. We followed them as described in detail below. In the following, the reviewer's comments will be in normal typeface, our responses in italic and bold typeface.

In this paper, the authors report on a new and exciting application of the airborne APEX instrument, namely imaging of tropospheric NO₂ columns at unprecedented spatial resolution over an extended area around Zurich, Switzerland. They briefly describe the instrument, measurements, data analysis, and uncertainties, and then compare measurements taken at two different times of the day over the Zurich area. The observations provide a highly structured spatial map of tropospheric NO₂, with good consistency between the two observation periods but lower values and less structures in the afternoon. The APEX data are then compared to a high resolution emission inventory, an annual model average and surface in-situ observations with overall good consistency.

The paper reports on a very interesting new measurement method and fits well into the scope of AMT. It is well written, provides all the information necessary and shows some interesting results. I therefore recommend publication in AMT after minor revisions and consideration of the points made below.

Referee comment 1. One aspect of the paper which is not convincing to me is the comparison with in situ data. I have two concerns about this comparison:

First, as the authors point out several times, the two data sets represent different quantities, the surface concentration and the column. While there often is a linear relation between the two, we should not expect this relation to be constant over time, in particular if the BL depth changes as much as the authors report for this case. In very first approximation, one would expect the ratio between column and surface concentration to increase by a factor of 2 – 4 if the BL height increases by a factor of 4. Therefore, it doesn't make much sense to include both the morning and afternoon measurements in the same graph (Figure 12) or linear fit.

Answer: We fully agree. The slope between surface in-situ and column NO₂ critically depends on vertical mixing which changes over the course of the day with the evolution of the boundary layer. In the revised manuscript we will replace the single linear fit with separate fits for the morning and afternoon overpasses.

Secondly, if we look at the two data sets independently, the correlation looks acceptable in the afternoon but in the morning, 8 out of 9 APEX measurements give the same value within uncertainties while the surface observations vary by a factor of 4. To me, this is not a strong confirmation of the ability of APEX to "observe spatial NO₂ gradients".

I'm actually not too concerned about this mismatch as the spatial pattern of the measurements looks very reasonable and there are many uncertainties in the in-situ measurements and the comparison of these two different quantities. However, I think the authors should re-consider the discussion and interpretation of this comparison.

Answer: Right, the correlation between columns and in situ measurements is indeed rather poor for the morning overpasses. The uncertainty in the in situ measurements is typically of the order of 10% for this type of NO₂ measurements in an urban environment where the concentrations of potentially interfering NO_y species are much lower than the concentrations of NO₂. Thus, the uncertainty in the in situ measurements is probably not an important factor contributing to the scatter. Rather, this scatter is likely related to the problem of comparing point-measurements at the surface, which are strongly influenced by local sources (note that several of the sites are located directly at a road) with column integrals, which are influenced not only by local but also by more distant upstream sources. The discussion of the comparison with in situ data in Section 4.4 will be changed accordingly. Please see also our response to referee #2.

Referee comment 2. The comparison between morning and afternoon data is also affected by the interpretation of the in situ data. The authors use these as confirmation for strong NO₂ decrease between the two time periods. However, as in the same time the BL depth increases so much, the situation is not as straight forward. Also, while higher wind speeds and stronger vertical mixing will reduce spatial gradients, the reduced lifetime (more OH) will counteract this effect and in general I would expect larger spatial gradients at high sun.

Answer: We agree that the decrease in the in situ values between morning and afternoon could potentially be due to an increase in the height of the PBL only, which in turn would not affect the columns. An increase in OH would indeed sharpen the gradients (because of a reduced lifetime of NO₂) rather than smoothing them out. An increase in OH from morning to afternoon would thus be compatible with a decrease in NO₂ but not with the smoother pattern observed in the afternoon. The main reason for the lower values in the afternoon must be a stronger dilution due to the much stronger winds. Note that since the solar zenith angle was approximately the same during the morning and afternoon overpasses and since OH-levels depend on NO_x concentrations in a strongly non-linear way (e.g. Jaeglé et al., Geophys. Res. Lett. 1998), it is not clear whether OH levels were on average higher or lower in the afternoon.

The discussion of the spatial distribution of NO₂ in Section 4.2 and the conclusions will be adapted accordingly.

Referee comment 3. There is one aspect which the authors only touch upon but which in my opinion is very relevant for their study and future applications of it. As these data represent a large step forward in terms of spatial resolution, spatial representativeness problems already discussed for satellite obser-

variations will become a significant issue. All the a priori data used (topography, surface spectral reflectance, NO₂ profiles, aerosol profiles) are needed at the resolution of the measurements as otherwise large errors can be introduced to individual data points. The authors did a good job on topography and surface reflectance but could not use equally well resolved profiles of NO₂ and aerosols. As all quantities, surface reflectance, NO₂ emission strength, NO₂ profile, and aerosol profile are expected to be highly correlated on small spatial scales, this is a relevant aspect which should be discussed in the paper. Further complication is added by the fact that at a flight altitude of 5 km and a spatial resolution of the order of 100 m, the independent pixel approximation used in the data analysis might no longer hold and three-dimensional effects of the radiative transfer might have to be considered.

Answer: We thank Referee #1 for raising these important points. Highly detailed aerosol extinction and NO₂ a priori data are currently not available at the resolution required for the APEX NO₂ processing. We will enhance the discussion on these aspects in the final version of this manuscript as follows:

In last paragraph subsection 4.2.:

"An additional influence on the accuracy of the AMF can be expected from retrieval input parameters with insufficient spatial resolution not matching the high resolution of APEX. Heckel et al. (2011), for example, studied the impact of coarse resolution retrieval input parameters (a priori NO₂ profile, surface reflectance, and aerosol information) on satellite retrievals of tropospheric NO₂ VCD with significantly smaller pixel size. They identified the a priori profile and surface albedo to have the largest impact on the retrieval uncertainty. In our approach we therefore derive the surface reflectance directly from the APEX data at high resolution. Profiles of NO₂ and aerosols, on the other hand, are taken from coarse resolution data sets which represents an important remaining error source. For an improved retrieval and error quantification, future flight campaigns should therefore aim at flying vertical NO₂ and aerosol profiles with complementary in situ instrumentation. Aerosol information (AOD) can potentially be derived from APEX data itself in the future (Seidel et al., 2011). The high spatial resolution of the APEX data also has implications on the radiative transfer calculations. For example, photons from neighboring pixels can be scattered into the instantaneous field-of-view. Considering such three dimensional effects in the radiative transfer computations would potentially improve future APEX NO₂ retrievals. "

Referee comment 4. Because the AMF is varying with location, it would be good to add a figure like Fig. 7 showing for either the morning or the afternoon overpasses the DSC, the AMF and the VC to show which of the variability is due to surface reflectance and which is in the raw data.

Answer: We will provide a Figure similar to Fig. 7. The discussion of this new Figure (c.f. Fig 1 below) is included in section 4.2:

"Differential SCD, AMF as well as the VCD are illustrated in Fig. XX exemplarily for the central morning flight stripe. As already underlined by Fig. YY, the surface reflectance has the largest impact on the AMF. For example, the lowest AMFs can be found over the dark forested areas and the highest over man-made structures such as buildings and roads. In our study region therefore significant correlation between bright surfaces and enhanced NO_x due to emissions from motorways and residential or industrialized areas can be found. Darker surfaces, in contrast, corresponded to comparatively clean vegetated areas. Without accounting for the varying surface reflectance, the contrast between polluted and clean areas would therefore be clearly overestimated. Fig. XX further demonstrates that spatially varying AMF also has an impact on small scale NO_x features like e.g. in the lower (southern) part of the flight stripe around the in-situ site Wettswil Filderen (c.f. map in Fig. XX). However, overall the largest part of the VCD variability can clearly be linked to the variability of the dSCD. "

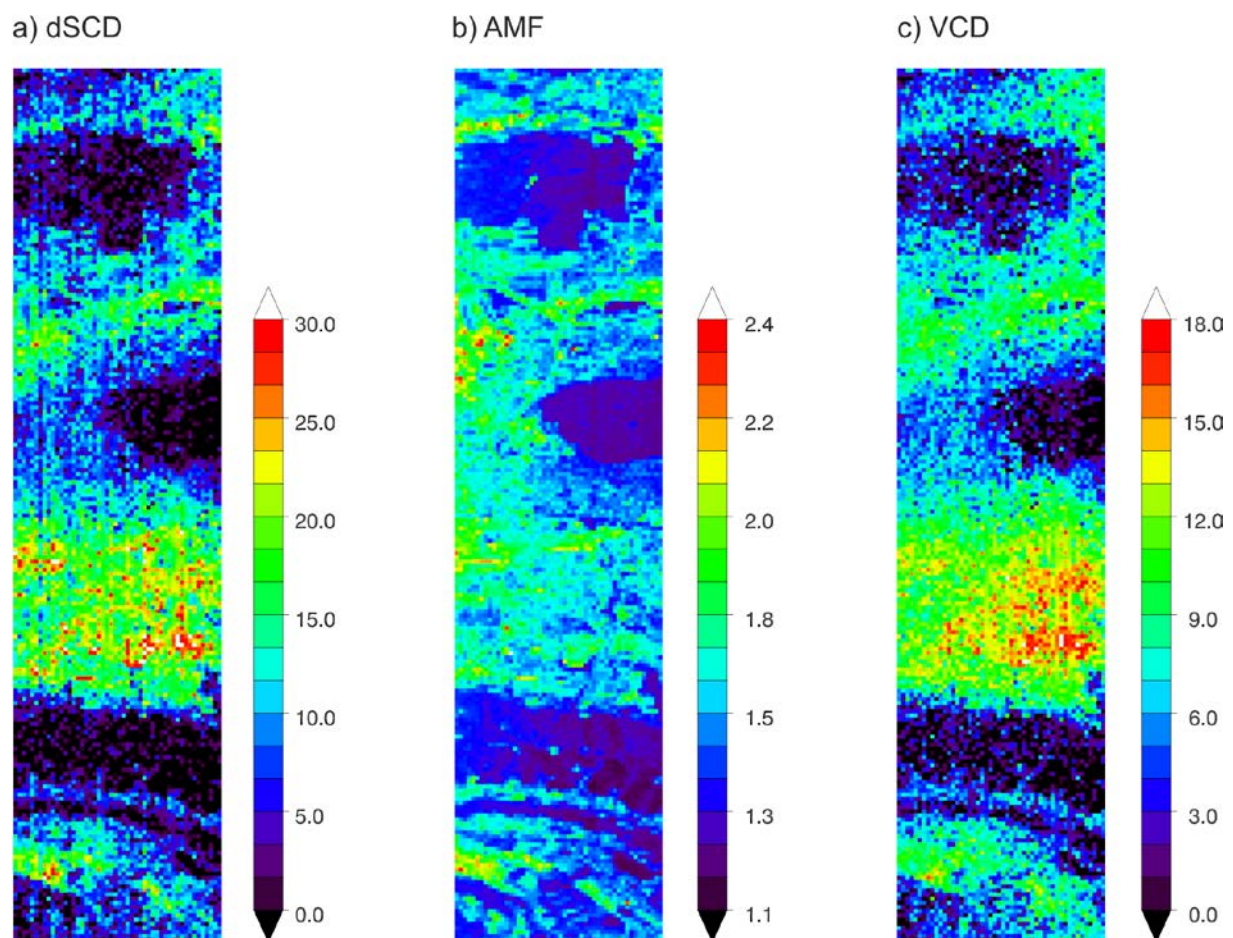


Figure 1 Differential SCD, AMF as well as the finally retrieved VCD for the central morning flight stripe.

Referee comment 5. The fitting window selected is affected by O₃ and H₂O absorption features, but neither of the two interfering species appears to be included in the fit. While O₃ might cancel as spatial gradients are weak and the tropospheric absorption is small, this is not true for H₂O. Please discuss the reasons for excluding these absorbers and why that's acceptable.

Answer: In our first attempts to fit NO₂ SCD from APEX measurements, we also included O₃ and H₂O. However, our analysis subsequently showed that NO₂ SCD and O₃ SCD were highly correlated and likewise that NO₂ SCD and H₂O SCD were highly anti-correlated which cannot be expected from atmospheric chemistry considerations. In addition, O₃ SCD and H₂O SCD were unrealistically high and histogram plots revealed a lack of small NO₂ SCD. It was found that these undesired effects were caused by too many cross-sections being fitted in a too small fitting interval. The radiometric noise and large (wavelength-dependent) offsets in the APEX data did not allow extending the fitting interval or moving it to shorter wavelengths. For these reasons, and considering the expected weak O₃ contribution and the fact that the APEX flights took place under dry clear-sky conditions, O₃ and H₂O cross-sections were both omitted in the fit which resulted in a much more realistic NO₂ SCD histogram and spatial distribution and reduced noise. We will integrate this in the revised manuscript, section 3.1.:

"We also tested integrating O₃ and H₂O absorption in the DOAS fit which led to distinctly worse results probably due to a too small fitting window relative to the number of cross sections. Retrieved slant O₃ and H₂O columns reached unrealistic values and were correlated (O₃) or anticorrelated (H₂O) with NO₂. These interfering gases were therefore omitted hereinafter."

Referee comment 6. What is the rationale of using a polynomial in the destriping? I don't see any justification for this and would suggest using the assumption that NO₂ on average does not depend on scan line.

Figs. 7 and 8 suggest that the NO₂ VCDs averaged along track are unlikely to be constant in the across track direction but are rather varying smoothly, for example caused by the inhomogeneous distribution of NO_x sources. After trying different options, including one mentioned by the Referee, we found that a polynomial leads to the best result in terms of destriping. We rephrased the corresponding part to clarify this as follows:

"Assuming that NO₂ VCD averaged in along-track (column) direction varies smoothly across-track (row), a fifth degree polynomial was fitted to the column averages. The residuals per column were finally subtracted from the initially retrieved NO₂ field."

Referee comment 7. Another aspect related to destriping is the question related to the consistency between different overpasses of the same location. The authors state that the consistency is good and that differences can be explained by transport. While I agree that there are no large discontinuities, there seems to be a gradient in each of the overpasses and I think it would be very interesting to show a direct comparison of the NO₂ results for the overlapping parts of the measurements (which now are hidden in the figure), for example as a line plot.

Answer: This is a good suggestion. We will add a plot illustrating the north-south transect of the overlapping region between the central and eastern flight overpasses (c.f. Fig. 1 of the original manuscript). The following text will be added to the revised manuscript (sec 4.2).

"The three different flight lines per mosaic generally superimpose well. This is underlined by Fig. XX which shows the N-S transects of the APEX NO₂ VCD in the overlapping region (c.f. map in Fig. 1) of the central (green line) and eastern (black line) morning overpasses. The two curves are in very good agreement with a correlation coefficient of around 0.95. However, the values from the central line are biased against the eastern line. The mean difference between these two curves is around $1.9 \cdot 10^{15}$ (or 28%) which might be due to several reasons. For example, the NO₂ columns in the "pollution-free" areas where the reference spectra are selected can slightly differ due to terrain variations such that using reference spectra from lower altitudes (higher NO₂ columns) lead to relatively lower NO₂ VCD than those from higher altitudes. In addition, the two flight lines do not observe exactly the same air mass at the same location due to the varying observation geometry and the time lag between the two measurements."

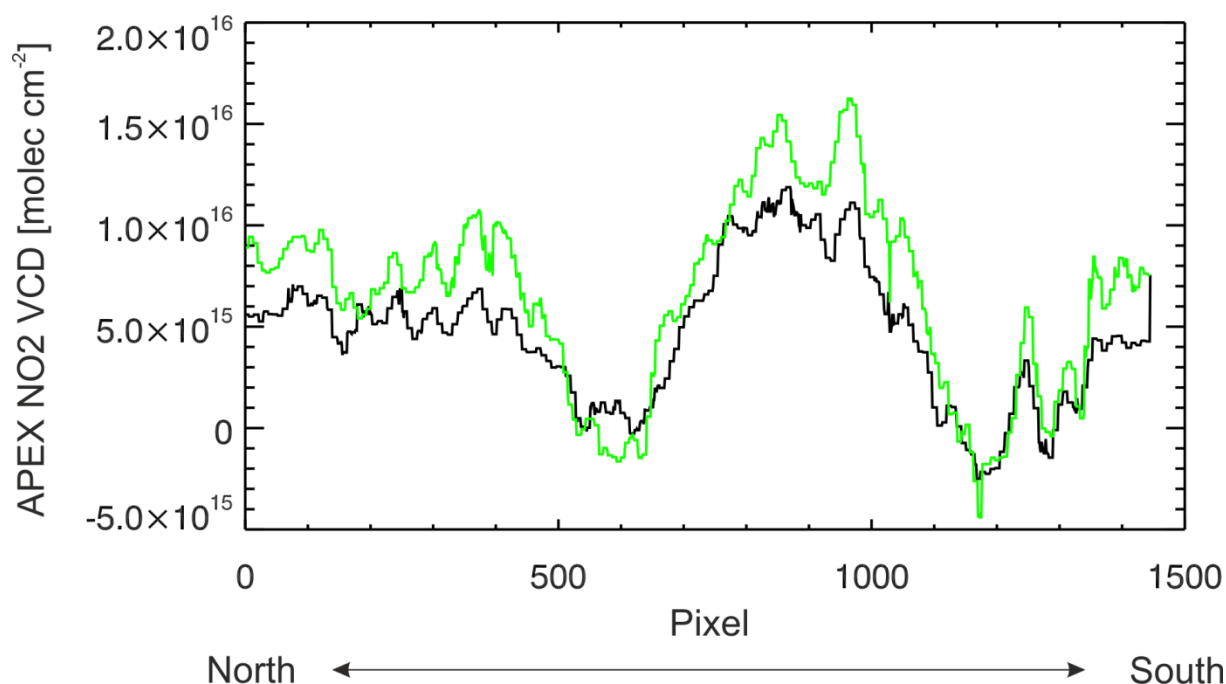


Figure 2 N-S transects of the APEX NO₂ VCD in the overlapping region (c.f. map in Fig. 1) of the central (green line) and eastern (black line) morning overpasses.

Referee comment 8. The discussion of the effect of the reference spectra taken over a “clean” region is in part misleading. While it is true that the stratospheric contribution will cancel if time is short enough as suggested in eqs. 4 and 5, the same holds for the tropospheric part above the aircraft. Only variations in the free troposphere will contribute to the retrieved DSCDs, and considering the lifetime of NO₂ in these altitudes as well as average wind speeds, there should be relatively little variability. This has an impact on the interpretation of the data (they are more weighted to the BL) and on comparison of measurements taken at different times of the day.

Answer: We completely agree with Referee #1 that also the free-troposphere part above the aircraft largely cancels out because only very small (spatial and temporal) variability can be expected. We already considered this point in the original version of the manuscript (p.2459, l19):

“Note that dSCD varies primarily due to different NO₂ below the aircraft, mainly in the boundary layer where NO₂ profiles peak in contrast to the usually low concentrations in the free troposphere.”

However, we make this point more clearly by adding the following in Sec. 3.2:

“Note that dSCD varies primarily due to different NO₂ below the aircraft, mainly in the boundary layer where NO₂ profiles usually peak close to the source. Free tropospheric NO₂ and particu-

larly the tropospheric NO₂ above the aircraft is expected to contribute only very little as it is probably similar for the reference and sample observations and in addition the NO₂ concentrations and above-aircraft AMF values are low."

Referee comment 9. In Fig. 8, it would have been nicer to have more similar colour scales – as the figures are now, the extreme values in the two panels are represented opposing ends of the colour scale.

Answer: Unfortunately, we do not have control of the color scale of the modeled surface concentration map.

Referee comment 10. It would be good if Fig. 11 could have better image quality – it is difficult to see the plume and other details in the RGB images

Answer: We will make sure that the figure will be displayed sufficiently large in the final version.

References:

Jaeglé, L., D.J. Jacob, W.H. Brune, D. Tan, I. Faloona, A.J. Weinheimer, B.A. Ridley, T.L. Campos, and G.W. Sachse (1998), Sources of HO_x and production of ozone in the upper troposphere over the United States, Geophys. Res. Lett., 25, 1705-1708

Heckel, A., Kim, S.-W., Frost, G. J., Richter, A., Trainer, M., and Burrows, J. P.: Influence of low spatial resolution a priori data on tropospheric NO₂ satellite retrievals, Atmos. Meas. Tech., 4, 1805-1820, doi:10.5194/amt-4-1805-2011, 2011.

Seidel, F.C., A.A. Kokhanovsky, M.E. Schaepman: Fast retrieval of aerosol optical depth and its sensitivity to surface albedo using remote sensing data, Atmos. Res. (2011), doi:10.1016/j.atmosres.2011.03.006