

Anonymous Referee #2:

We greatly appreciate the positive feedback from the referee and the constructive comments. As described below, we have modified the manuscript according to suggestions and provided clarifications where necessary. We hope that the revised manuscript has improved in respect to the original paper. Please find a rebuttal against each point below.

Black, bold, italic: Referee's comments

Black: Author's reply

Green: Modifications in the manuscript

Note that based on the comments, two new figures were added to the manuscript, i.e. Fig. 14 and 19. References to figures in the author's reply refer to the new figure numbers.

The paper by Tack et al., describes mapping of NO₂ over Belgian cities using the APEX airborne spectral imager, an instrument that was not initially designed for the retrieval of trace gases. The authors present and their retrieval algorithm and discuss instrument performance during measurement flights. Airborne NO₂ results are compared with collocated car observations and yield good agreement.

Overall the paper is well written and arguments are mostly easy to follow. The manuscript presents a good summary of the data analysis and NO₂ results fulfills the in the introduction stated objective to assess the technical and operational capabilities of APEX NO₂ mapping. However, a more detailed analysis of the instrument in-flight performance, as well as a discussion on possible improvements is missing, which would be very fitting for the scope of AMT (see also specific comments).

1)

Abstract: Please state right away that Liege results are not further discussed or leave out.

We prefer to only mention Brussels and Antwerp in the abstract: "APEX data have been acquired under clear sky conditions over the two largest and most heavily polluted Belgian cities, i.e. Antwerp and Brussels on 15 April and 30 June 2015". However we mention the three sites in the introduction, but mentioning as well that the study focuses on Brussels and Antwerp: "In the framework of this study, dedicated APEX flights were performed above three of the largest and most heavily polluted urban areas in Belgium, namely Antwerp, Brussels and Liège. We report on the NO₂ retrieval scheme developed at BIRA-IASB and on the results of the flight campaigns performed in Belgium, with the focus on the cities of Antwerp and Brussels". Please note that we have also included the NO₂ map retrieved over Liège (Fig. 19) in Section 5 (see comment 12 of referee #1), however we keep the description short because of the mentioned reasons: "In the Liège data set, the highest NO₂ emissions are observed in the northeast, in the industrialised area of Herstal. The NO₂ levels range between 1 and 32 x 10¹⁵ molec cm⁻², with a mean VCD of 13.3 ± 3.1 x 10¹⁵ molec cm⁻². The overall error on the retrieved NO₂ VCDs is on average 23%".

2)

Section 1: Typically airborne experiments are cost intensive field measurements compared to e.g. car measurements. Please explain why APEX measurements are described as “cost effective”.

We fully agree that costs of an airborne campaign can be relatively high, of course depending to what you compare. We also agree that the statement “cost-effective” is not well-chosen here (please see also comment 1 of referee #1). Our intension was to say that we want to retrieve high-resolution NO₂ maps in an efficient way. For example, we could interpolate NO₂ maps from many car-DOAS measurements in a city but this wouldn't be efficient/cost-effective. As this statement is open for debate and out of the scope of this study, we prefer to remove it from the objectives.

3)

Section 2.1: p.4, line 21: Please define typical altitude, speed, integration time.

This is corrected and reformulated as: “The ground sample distance (GSD) is 3 (across-track) by 4 (along-track) m² assuming a typical altitude of 6.1 km AGL, ground speed of 72 m·s⁻¹ and integration time of 58 ms,…”

4)

Section 2.2: p.5, line 28: should say full NO2 columns below the plane.

Corrected.

5)

Section 3.1: p.6, line 14: How is the 58ms integration time chosen? Are there problems with saturated scans at times?

This is the typical integration time for APEX flights in spectral unbinned mode. It was not specifically chosen for the NO₂ inference application, but it was chosen several years ago, after various tests were made by the APEX operators. It was concluded that an integration time of 58 ms ensures the best balance between the signal-to-noise ratio and the occurrence of saturated pixels. We clarified this in the manuscript (not in Sect. 3.1 but at the first occurrence of the integration time in sect. 2.2 → see comment 3): “The ground sample distance (GSD) is 3 (across-track) by 4 (along-track) m² assuming a typical altitude of 6.1 km AGL, ground speed of 72 m·s⁻¹ and integration time of 58 ms, the latter being a good balance between the obtained signal-to-noise ratio and the occurrence of saturated scans.”

6)

Section 3.1: p.6, line 21: It is not clear which setting are used for the noise analysis, since Sect. 3.2 talks about calibration and analysis.

Thank you for this observation. This reference is indeed wrong and should refer to section 4.1 instead of section 3.2. We reformulated this as: “The applied DOAS settings for the noise analysis are provided in Sect. 4.1 and Table 4”

7)

Section 3.2: p.8: The strong variability of the instrument slit function is indeed a challenge for the DOAS analysis. Here further details on the instrument would be interesting, e.g., are there time traces of the instrument's pressure and temperature? Could these be correlated with the behavior of the slit function? Or even used to improve data analysis? The authors have optimized the DOAS analysis given the current state of the spectra, but recommendations of technical instrument improvements (if possible) or looking into exploiting technical in-situ data (if available) are missing, which would be well suited for the AMT audience.

We thank the reviewer for this pertinent remark. The instability of the spectral performance is indeed one of the limitations and a big challenge for the DOAS analysis. Changing environmental conditions, with pressure in particular, are to our opinion the most likely explanation. The APEX laboratory and in-flight spectral performance was intensively investigated during recent years by RSL (Remote Sensing Laboratories - University of Zürich) and EMPA (Swiss Federal Laboratories for Materials Science and Technology) and results are reported in a number of publications (e.g. D'Odorico et al., 2011; D'Odorico et al., 2012; Hueni et al., 2014; Kuhlmann et al., 2015; Kuhlmann et al., 2016).

The PhD dissertation of D'Odorico (2012), focusing on the APEX spectral performance, states in the conclusion: "APEX performance measured during laboratory characterisation cannot be assumed for the operational environment. Causes of deviations are to be sought in the airborne operational environment, the most significant of which are pressure and temperature excursions." This was also mentioned in the manuscript on p.7, starting from ll.27. In the conclusion of D'Odorico (2012) recommendations for technical instrument improvements were made and applied: "...These findings led to an instrument revision aimed at the stabilization of the system for a range of temperature and pressure conditions to be encountered during operation. The revision included the manufacturing of a pressure regulation mechanism for the automatic release and fill-in of nitrogen according to the change in flight altitude. However, experiments carried out in the following year (2010), revealed that the implemented design revisions did not fully solve the pressure/temperature dependency of the system. Spectral shifts in the range of one spectral pixel were yet again estimated in flight for both, VNIR and SWIR detector." The APEX spectral performance and related instabilities are also investigated in the very recent work of Kuhlmann et al. (2016). Also a new in-flight spectral calibration algorithm, based on a maximum a posteriori optimal estimation approach, is proposed in order to improve the quality of the fit.

We would like to refer to the related comments 4, 5 and 6 of reviewer #1. Based on these comments and comments 7 and 14 of reviewer #2, Sect. 3.2 has been completely revised and reformulated. We hope that the causes of the spectral shifts and increased in-flight FWHM are better explained now and that the aforementioned publications, focusing on the instrument calibration and spectral performance, have a larger visibility.

Technical changes to the APEX instrument to improve trace gas retrieval are not evident due to several reasons: as mentioned in the manuscript, the APEX instrument has been primarily designed for environmental remote sensing of the land surface and APEX is almost exclusively operated for this purpose. Nevertheless, it has been demonstrated that atmospheric trace gases, such as NO₂, can also be retrieved with APEX. This study has revealed some limitations related to the spectral performance in order to detect the fine trace gas absorption structures. Remote sensing of the land surface is, however, much less constraining and the in-flight FWHM and spectral shifts are well within the tolerances for these kind of applications. Most of the time, acquisitions are performed in spectrally binned mode (groups of spectral bands are binned in the VNIR spectral region) in order to increase SNR, while retaining the high spatial resolution of 3 by 4 m.

The BUMBA, AROMAT and AROMAPEX projects have increased the knowledge and experience for trace gas retrieval based on hyperspectral imagers and brought together several groups and communities. Now the official APEX lifetime is shortening, discussions are started for a possible follow-up instrument, taking into account the findings of the related studies. If the follow-up instruments needs to be suitable for both LULC applications and atmospheric research, the current limitations should be tackled, e.g. by designing a vacuum enclosure.

D'Odorico et al.: Performance assessment of onboard and scene-based methods for Airborne Prism Experiment spectral characterization, doi: <https://doi.org/10.1364/AO.50.004755>, 2011.

D'Odorico, P.: Monitoring the spectral performance of the APEX imaging spectrometer for inter-calibration of satellite missions, Remote Sensing Laboratories, Department of Geography, University of Zurich, 2012.

A. Hueni et al.: "APEX: Radiometry under spectral shift conditions," *IEEE Geoscience and Remote Sensing Symposium*, Quebec City, QC, 2014, pp. 1381-1384. doi: 10.1109/IGARSS.2014.6946692, 2014.

Kuhlmann et al.: In-flight Spectral Calibration of the APEX Imaging Spectrometer using Fraunhofer Lines, ESA ATMOS Workshop, Heraklion, Greece, 2015

Kuhlmann et al.: An Algorithm for In-Flight Spectral Calibration of Imaging Spectrometers, doi: <https://doi.org/10.3390/rs8121017>, 2016.

8)

Section 4.3.1: Is there in-situ temperature and pressure data available or ground measurements that could be interpolated? Please discuss choice of US Standard Atmosphere.

Indeed, the used pressure and temperature profile is not correctly described in the manuscript, as "US Standard Atmosphere" should be "AFGL standard atmosphere for mid-latitude summer (Anderson et al., 1986)." This atmospheric profile is commonly used in AMF calculations and provides a good approximation. According to the study of Boersma et al. (2011), the effect of using a different atmospheric profile on the AMF calculations is very small: "using a midlatitude winter atmosphere profile (p,T) instead of a summer profile would change the tropospheric AMFs by 1 %." At the surface, the profiles have a different p of 5 hPa and a different T of 22 K.

Ground temperature and pressure are probably measured by the meteorological institute at different stations. However, the impact of including this in the AMF calculations would be very small.

Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO₂ column retrieval algorithm for the Ozone Monitoring Instrument, *Atmos. Meas. Tech.*, 4, 1905-1928, doi:10.5194/amt-4-1905-2011, 2011.

9)

Section 4.3.2: p.12, line 29: “similar statistics” here and elsewhere could go into an Appendix.

We prefer to add the sensitivity study results for Brussels as well to Table 5 instead of in an appendix. We refer to it in Section 4.3.2 as: “Similar results were obtained for the Brussels data set and are provided in Table 5.”

10)

Section 4.3.2: p.13 and Section 4.6, p.15, line 27: Based on Fig. 16, the station in Uccle is in a semipolluted area, so any AOD measured there will not be representative for heavily polluted areas. The error on the aerosol effect seems underestimated. Please include an assessment using higher AODs.

We thank the reviewer for pointing this out. We agree that the scarce data on the aerosol properties is a limitation and that indeed the Uccle station was in a semi-polluted area, taking into account the prevailing wind direction. During more recent flights over Berlin in the framework of the AROMAT/AROMAPEX campaign, we measured AODs in the same range but again the CIMEL station was outside the plume and thus in a semi-polluted area. Feedback on the campaign results were also focusing on more detailed data on the aerosol properties. This is something we should consider for future campaigns, e.g. by operating additional stations or a mobile system.

However we want to stress that flights were performed at clear-sky days with good visibility. We should also keep in mind that for example measurements in the plume (max. expected AODs) would also not be representative for the whole data set.

For the sensitivity tests, we assumed that all the aerosol extinction measured by the CIMEL takes place in the boundary layer, which leads to an upper estimate for the extinction in the BL, and thus the associated error. The difference of nadir AMF compared to a pure Rayleigh atmosphere is albedo-dependent and exhibits a maximum of 7%. This maximum estimated effect was further increased to 10% to account for higher aerosol loads. We have clarified this in Sect. 4.3.2: “Radiative transfer simulations with a corresponding well-mixed extinction at the surface yield an albedo-dependent aerosol effect of 7% or less when compared to the AMFs computed based on a Rayleigh atmosphere. However, a relative uncertainty of 10% is considered for all flights, as the AERONET station in Uccle was in a semi-polluted area and furthermore to take into account the AOT variability”.

11)

Section 5, p.18, line 4: dSCDs around 0 are not “well above detection limit”. Please be more specific with the respective statement (also in Fig. 15).

This statement can be indeed confusing. The DSCDs depend on the, partly arbitrary, choice of the reference SCD. In the case there is much NO₂ in the reference, the DSCDs can be close to zero or even negative. We could state that the observed DSCD signal or DSCD variability is clearly popping out of the error bars/detection limit (1-sigma standard error), but maybe this is also confusing. Therefore we suggest to use “detection limit” only in the context of the VCDs. This is adapted in the text: “The 1-sigma slant error of APEX retrievals has a typical value between 3.4 and 4.4 x 10¹⁵ molec cm⁻² on the DSCD, corresponding to a detection limit of approximately 1.8 and 2.3 x 10¹⁵ molec cm⁻² on the VCD, assuming a typical AMF of 1.9 (See Sect. 4.3). In order for an absorber to be clearly identified, the retrieved column needs to be larger than this threshold.” And in the caption of Fig. 16:

“The retrieved NO₂ VCDs are well above the detection limit of approximately 1.8 to 2.3 x 10¹⁵ molec cm⁻²”.

12)

All Tables: Tables typically have a short caption. Please move explanations to text.

It is often required that tables/figures and their captions are self-explanatory to the largest extent possible. Nevertheless, the caption of Table 5 was much too long and unclear. Therefore we splitted the original table in 2 parts: sensitivity study of 1) albedo/viewing/sun geometry (Table 5) and 2) a priori NO₂ profile (Table 6). Table 6 is restructured and the captions are reformulated. We would also like to refer to comment 13 of referee #1.

13)

Table 1: Please choose detectors or pixel consistently. What does the plane speed relate to? The section “other” seems very random. I suggest integration in text where fitting.

We thank the reviewer for these observations. “Detectors” is replaced by “pixels”. Please see also comment 21 of referee #1, related to this. We replaced “plane speed” by “ground speed”. The section “other” is indeed a bit random and is removed accordingly. However we kept “ground speed” and “integration time” in the table as this is strongly linked to the along-track spatial resolution.

14)

Table 3: Can you comment on why the in-flight FWHM is significantly larger than nominal?

This is a very pertinent but difficult question to answer. This was already observed in previous studies, e.g. Popp et al. (2012), without providing an explanation. D’Odorico (2012) states in the conclusion: “APEX performance measured during laboratory characterisation cannot be assumed for the operational environment. Causes of deviations are to be sought in the airborne operational environment, the most significant of which are pressure and temperature excursions.”

In the very recent study of Kuhlmann et al. (2016), huge efforts were done to better understand the larger in-flight FWHMs. According to their study, there is no significant sensitivity of the FWHM to pressure changes (in contrast to the spectral shifts). The impact was studied of 1) filling-in of the Fraunhofer lines by the ring effect or underestimation of the O₄ absorption; 2) filling-in because of spatial binning of radiance spectra with slightly varying central wavelengths; 3) filling-in of dark lines because of radiometric calibration errors. However all of these effects seem too small to explain the larger in-flight FWHM. The authors conclude that probably a combination of not fully corrected CCD readout smear (resulting in spectral smoothing) and filling-in of the Fraunhofer lines causes the larger in-flight FWHMs. In sect. 3.2 we have added an explanation based on the analysis and conclusion of Kuhlmann et al. (2016):” **Based on an intensive analysis, the study of Kuhlmann et al. (2016) concludes that the larger in-flight FWHMs are likely explained by a combination of (1) not fully corrected CCD readout smear, resulting in a spectral smoothing, and (2) filling-in of the Fraunhofer lines.**”

15)

Table 4: Please explain “resol”

This cross-section is based on a linearisation of the dependence of spectral Fraunhofer lines on the slit function (or instrumental line shape ISRF) width. In fact, it corresponds to the first derivative of solar reference with respect to the slit function width. This cross-section allows to compensate for the impact of small changes of the instrumental resolution. Please see also Danckaert et al. (2015) and Beirle et al. (2017).

We have changed the paragraph in the manuscript to make this more clear: “Beside relevant high-pass filtered trace gas laboratory cross-sections (NO₂ and O₄), a low-order polynomial term, a synthetic Ring spectrum and a synthetic resolution cross-section are fitted to the logarithm of the ratio of the observed spectrum, and a reference spectrum. They account for respectively (1) smooth broad-band variations (e.g. reflection at the Earth’s surface) and Rayleigh and Mie scattering; (2) the Ring effect (Grainger and Ring, 1962), i.e. the filling-in of Fraunhofer lines by rotational Raman scattering on air molecules, and (3) the impact of small changes of the instrumental resolution. The synthetic resolution cross-section is based on a linearisation of the dependence of spectral Fraunhofer lines on the slit function width and corresponds to the first derivative of solar reference with respect to the slit function width (Danckaert et al., 2015; Beirle et al., 2017).”

16)

Fig. 3: Is the reference noise taken into account here?

For each binning level (i.e. the raw APEX spectra are spatially binned in the along- and across-track direction according to a power of 2, 2^n , with n ranging from 0 to 8), a clean reference area was binned in the same way. Making use of one “fixed” reference for all the different binning levels would be impossible because the reference spectrum should always be binned across-track in the same way as the spectra to be analysed. This way, all spectra from a “track” (including the reference spectrum) share the same optical configuration, taking into account that each detector is assumed to be a 1D-instrument with its own optical aberrations. This has been clarified in Sect. 3.1: “For a test area, the raw APEX spectra are spatially binned in the along- and across-track direction according to a power of 2, 2^n , with n ranging from 0 to 8. For each binning level, a clean reference area is binned in the same way”.

17)

Fig. 6: include description of lines in legend or caption.

Ok, we have reformulated the caption as follows: “In-flight spectral calibration: **a)** the spectral resolution (FWHM) and **b)** the spectral shift and its dependency on the across-track scanline pixel position (spectral smile), plotted for 490 nm, i.e. the middle of the analysis window, for different flightlines. A second order polynomial has been fitted to each calibration set”.

18)

Please review manuscript for rules on spelling out digits.

Thank you for noticing this. The whole manuscript was scanned to check and was corrected accordingly. The following rules were not always applied consequently:

"1) use words for cardinal numbers less than 10; use numerals for 10 and above (e.g. three flasks, seven trees, 6m, 9 days, 10 desks).

2) use numerals for cardinal number less than 10 for items that are unit of time or when used in scientific sense or mathematical sense."

19)

p.3, line 1: Figs. 1 and 2

Corrected. Also corrected on p.16, line 24; p.19, line 29 and p.20, line25

20)

p.9, line 1 (SCDi) is already introduced

Corrected.

21)

p.9 Eqs. 2 and 3: "ref" should be subscript

Corrected.

22)

p.9, line 19: exchange word order of "wavelength" and "lower"

Corrected.

23)

p.11ff: Please be consistent with either Box-AMF or box-AMF.

Corrected. Should be box-AMF.

24)

p.15, end of line 22: "the" is missing

I'm sorry but we don't find back where "the" is missing in one of these sentences.

25)

p.17, line 20: 11:30 UTC

Corrected.

26)

Fig. 4: labels a) and b) are missing

Corrected.

27)

Fig. 6: Check rules on including units: “()” or “[]”, cursive or not?

We thank the reviewer for this observation. Square brackets should be used, as well as italic text. This was indeed not applied in a consistent way in the manuscript and is corrected accordingly in Fig. 6, 8, 16 and 20.

High resolution mapping of the NO₂ spatial distribution over Belgian urban areas based on airborne APEX remote sensing

Frederik Tack¹, Alexis Merlaud¹, Marian-Daniel Iordache², Thomas Danckaert¹, Huan Yu¹, Caroline Fayt¹, Koen Meuleman², Felix Deutsch², Frans Fierens³, and Michel Van Roozendael¹

¹BIRA-IASB, Royal Belgian Institute for Space Aeronomy, Brussels, 1180, Belgium

²VITO, Flemish Institute for Technological Research, Mol, 2400, Belgium

³IRCEL-CELINE, Belgian Interregional Environment Agency, Brussels, 1210, Belgium

Correspondence to: Frederik Tack (frederik.tack@aeronomie.be)

Abstract. We present retrieval results of tropospheric nitrogen dioxide (NO₂) vertical column densities (VCDs), mapped at high spatial resolution over three Belgian cities, based on the DOAS analysis of Airborne Prism EXperiment (APEX) observations. APEX, developed by a Swiss-Belgian consortium on behalf of ESA (European Space Agency), is a pushbroom hyperspectral imager characterised by a high spatial resolution and high spectral performance. APEX data have been acquired under clear sky conditions over the two largest and most heavily polluted Belgian cities, i.e. Antwerp and Brussels on 15 April and 30 June 2015. Additionally, a number of background sites have been covered for the reference spectra. The APEX instrument was mounted in a Dornier DO-228 airplane, operated by Deutsches Zentrum für Luft- und Raumfahrt (DLR). NO₂ VCDs were retrieved from spatially aggregated radiance spectra allowing to resolve urban plumes at the resolution of 60x80 m². Main sources in the Antwerp area appear to be related to (petro)chemical industry while traffic-related emissions dominate in Brussels. The NO₂ levels observed in Antwerp range between 3 and 35 x 10¹⁵ molec cm⁻², with a mean VCD of 17.4 ± 3.7 x 10¹⁵ molec cm⁻². In the Brussels area, smaller levels are found, ranging between 1 and 20 x 10¹⁵ molec cm⁻² and a mean VCD of 7.7 ± 2.1 x 10¹⁵ molec cm⁻². The overall error on the retrieved NO₂ VCDs is on average 21% and 28% for the Antwerp and Brussels data set, respectively. Low VCD retrievals are mainly limited by noise (1-sigma slant error), while high retrievals are mainly limited by systematic errors. Compared to coincident car mobile-DOAS measurements performed in Antwerp and Brussels, both data sets are in good agreement with correlation coefficients around 0.85 and slopes close to unity. APEX retrievals tend to be on average 12% and 6% higher for Antwerp and Brussels, respectively. Results demonstrate that the NO₂ distribution in an urban environment, and its fine scale variability, can be mapped accurately with high spatial resolution and in a relatively short time frame, and the contributing emission sources can be resolved. High resolution quantitative information about the atmospheric NO₂ horizontal variability is currently rare, but can be very valuable for (air quality) studies at the urban scale.

1 Introduction

Nitrogen dioxide (NO₂) is an atmospheric trace gas and a key pollutant that attracts considerable attention due to several reasons: (1) NO₂ is a proxy for air pollution in general, as its abundance mostly coincides with a range of other pollutants; (2) recent studies by the World Health Organization (WHO, 2013) have shown that NO₂ exposure can have a direct health impact; (3) it is a precursor in the formation of aerosols (Chan et al., 2010) and tropospheric ozone (Crutzen, 1970) and therefore it contributes locally to radiative forcing (Solomon et al., 1999), through which it indirectly affects the climate system. In an urban environment, NO₂ mainly originates from anthropogenic sources such as the burning of fossil fuels, related to industrial activities and traffic. NO₂ is a pollutant with a strong local character and exhibits concentrations that are highly variable in space (horizontal and vertical) and time, often exceeding limits set by the European legislation, i.e. hourly average limit of 200 µg/m³ and an annual mean limit value of 40 µg/m³ (EU Directive 2008/50/EC). For the reasons stated, the accurate monitoring and mapping of the NO₂ variability at high spatial resolution is of great relevance.

In this paper, a method is presented to retrieve tropospheric NO₂ vertical column densities (VCDs) from hyperspectral Airborne Prism Experiment (APEX) observations. A major objective of the study is to assess the technical and operational capabilities of APEX to map the NO₂ field at city scale and at high spatial resolution, in a relatively short time ~~and cost~~ effective way, and furthermore to characterise all aspects of the retrieval approach. APEX, developed by a Swiss-Belgian consortium on behalf of ESA (European Space Agency), is a pushbroom hyperspectral imager, integrating spectroscopy and 2-dimensional (2D) spatial mapping in one single system. The well-established differential optical absorption spectroscopy technique (DOAS; Platt and Stutz, 2008) is applied to the observed backscattered solar radiation, in order to quantify the abundance of NO₂ in the atmosphere, based on its fine molecular absorption structures.

For atmospheric NO₂ detection, a high spectral resolution is necessary in order to resolve the fast varying spectral signatures of this molecule. High spatial resolution, i.e. higher than 100 m, is also required to identify small scale gradients in the NO₂ amounts and to resolve individual emission sources. Although the APEX instrument has been primarily designed for environmental remote sensing of the land surface, it also allows observing atmospheric trace gases as demonstrated in the precursor study of Popp et al. (2012) on APEX tropospheric NO₂ mapping over Zurich, Switzerland. In the framework of this study, dedicated APEX flights were performed above three of the largest and most heavily polluted urban areas in Belgium, namely Antwerp, Brussels and Liège. We report on the NO₂ retrieval scheme developed at BIRA-IASB and on the results of the flight campaigns performed in Belgium, with the focus on the cities of Antwerp and Brussels.

For about 2 decades, systematic mapping of the horizontal distribution of tropospheric trace gases, such as NO₂, has been performed at global scale using spaceborne sensors like SCIAMACHY (Scanning Imaging Absorption Chartography), GOME (Global Ozone Monitoring Experiment), GOME-2, and OMI (Ozone Monitoring Experiment) (see e.g. Richter and Burrows, 2002; Beirle et al., 2010; Boersma et al., 2011; Hilboll et al., 2011; Valks et al., 2011; Bucselá et al., 2013). Such observations are well suited for global monitoring, however, the coarse spatial resolution of typically several tens of

kilometers make them inadequate to detect city-scale NO₂ variability and to resolve individual emission sources (see Figs. 1 and 2). Several studies discuss instruments and experiments for atmospheric trace gas retrieval from airborne platforms. Most of these works focus on the retrieval of the vertical distribution of trace gases, based on along-track Multi-AXis (MAX-) DOAS observations (see e.g. Petritoli et al., 2002; Melamed et al., 2003; Wang et al., 2005; Bruns et al., 2006; Dix et al., 2009; Merlaud et al., 2012; Baidar et al., 2013). Only a few recent studies report on the high resolution 2D spatial mapping of the NO₂ horizontal distribution from an airborne platform. The discussed hyperspectral imaging systems are based on a whiskbroom (Merlaud et al., 2013b; Liu et al., 2015) or pushbroom setup (Heue et al., 2008; Popp et al., 2012; General et al., 2014; Lawrence et al., 2015; Schönhardt et al., 2015; Meier et al., 2016; Nowlan et al., 2016).

Besides the increased characterisation of horizontal trace gas distribution and related chemical processes in urban areas, high resolution NO₂ maps can be valuable for the calibration and validation of satellite products, chemical transport models (CTM) and high resolution air quality models, such as e.g. the RIO-IFDM high resolution air quality model (Lefebvre et al., 2013), developed by the Belgian Interregional Environment Agency (IRCEL-CELINE) in cooperation with the Flemish Institute for Technological Research (VITO). There is also a general growing interest from the remote sensing community for airborne hyperspectral imagers as they can complement and link ground-based measurements, spaceborne observations, and model data.

The paper is organised as follows: Section 2 introduces the optical characteristics of the APEX instrument, as well as the hyperspectral data sets acquired at three major Belgian cities. Section 3 discusses post-flight, pre-processing steps in order to improve the signal-to-noise ratio (SNR) and the spectral performance. Section 4 characterises the key steps of the developed methodology to derive NO₂ VCDs from APEX spectra and to produce high resolution NO₂ maps, with proper error characterisation. The following section presents and discusses the retrieved NO₂ field above Antwerp and Brussels. In Section 6, the retrieved VCDs are quantitatively assessed by intercomparison with correlative data sets (Mobile-DOAS and mini-MAX-DOAS).

2 APEX instrument and data sets

2.1 APEX optical unit and spatial performance

The APEX instrument was designed and developed by a Swiss-Belgian consortium on behalf of ESA. Since 2011, APEX has been fully operational for data acquisition flights. A brief overview of the optical unit, the core element of the instrument, is given here. Solar radiation, backscattered by the Earth's surface or atmosphere within the field of view (FOV) of the instrument enters the optical unit through a curved slit. A collimator groups and redirects the light towards a beamsplitter, separating the VNIR (Visible Near Infrared; 370 - 970 nm) from the SWIR (Shortwave Infrared; 950 - 2500 nm) channels. The VNIR radiance is spectrally dispersed by a prism, resulting in a greater light throughput and the absence

of overlapping orders when compared to a diffraction grating, but with the drawback of a limited resolution and a non-linear wavelength scale (Platt and Stutz, 2008). The dispersed radiation is then projected on a 2D CCD (Charged Coupled Device) 14 bit depth area detector, recording the intensity in a series of narrow, contiguous spectral bands. The VNIR CCD can record up to 335 unbinned spectral bands. The SWIR channel is not further taken into account in this study as the DOAS analysis is applied on a small part of the visible wavelength region, i.e. 470-510 nm. ~~The sealed optical unit is enclosed by a thermo regulated box in order to be temperature ($19^{\circ}\text{C} \pm 1^{\circ}\text{C}$) and pressure (dry nitrogen atmosphere with partial differential pressure control) stabilised.~~ The sealed optical unit is enclosed by a thermo-regulated box in order to be temperature ($19^{\circ}\text{C} \pm 1^{\circ}\text{C}$) stabilised, while pressure in the spectrometer is kept at 200 hPa above ambient pressure (dry nitrogen atmosphere with partial differential pressure control). To obtain a good compensation of the aircraft movement, the whole instrument is mounted and operated on a Leica PAV-30 stabilised platform. A more complete description of the APEX instrument, its optical unit and its calibration concept can be found in Itten et al. (2008), Jehle et al. (2010), D’Odorico (2012), Schaepman et al. (2015) and Kuhlmann et al. (2016).

The spatial performance and relevant specifications of the APEX sensor are given in Table 1. The pushbroom imaging spectrometer consists of 1000 detector pixels across-track (spatial dimension x), which are illuminated simultaneously, and 335 pixels in the spectral dimension $I(\lambda)$. The plane formed by the spatial dimension x and spectral dimension $I(\lambda)$ is called a frame. Mapping of the NO_2 distribution below the sensor is performed by the swath imaging of the pushbroom scanner (spatial dimension x) and the forward motion of the airborne platform (spatial dimension y). The resulting 3D hyperspectral image cubes, built by acquiring consecutive frames, consist of two planimetric dimensions (x , y) and a third spectral dimension $I(\lambda)$. With a field of view (FOV) of 28° and a flight altitude of 6.1 km AGL during data acquisition, which is much higher than the sampled air masses containing the majority of NO_2 , the swath covers an area of approximately 3 km across-track. The across-track spatial resolution is determined by the instrument’s instantaneous field of view (IFOV) and the platform altitude, while the along-track resolution is determined by the platform motion and the integration time. The ground sample distance (GSD) is 3 (across-track) by 4 (along-track) m^2 assuming a typical altitude of 6.1 km AGL, ground speed of $72\text{ m}\cdot\text{s}^{-1}$ and integration time of 58 ms, the latter being a good balance between the obtained signal-to-noise ratio and the occurrence of saturated scans.

2.2 APEX acquisition and data sets

Air pollution levels in the northern part of Belgium are among the highest in Europe. The sources are mainly related to the high population density and related traffic, and regular transport from industrial sources in the Rhine-Ruhr area in Germany. The annual mean NO_2 VCDs for 2015, retrieved based on OMI observations (OMNO2d, Giovanni, NASA GES DISC), range from 1 to 1.2×10^{16} molec cm^{-2} in the northern part of Belgium and are well above the background levels, as can be observed in Fig. 1. The NO_2 levels are around $7 - 8 \times 10^{15}$ molec cm^{-2} for the southern part.

The coarse spatial resolution of current global monitoring spaceborne instruments makes them inappropriate for studies of the NO₂ field at the scale of cities. This is illustrated in Fig. 2 where the typical spatial extent of a nadir observation from the OMI and the future TROPOMI (TROPOspheric Monitoring Instrument) spaceborne sensors is illustrated. The latter is the spectrometer payload of the ESA Sentinel-5 Precursor satellite, to be launched in 2017. For the sake of comparison, details of the APEX flightplan over the Antwerp area on 15 April 2015 are provided as well. Whereas the Antwerp city center is nearly fully covered by one TROPOMI pixel, it is sampled by approximately 5000 APEX pixels, which is needed to detect small-scale NO₂ horizontal variability and to resolve individual emission sources. Note that the original APEX pixels of 3 by 4 m² are spatially binned to pixels of 60 by 80 m², as will be explained in Sect. 3.1.

The APEX instrument has been installed in the Dornier DO-228 D-CODE, based at the Braunschweig research airport and operated by Deutsches Zentrum für Luft- und Raumfahrt (DLR). APEX data has been acquired over the three largest and most heavily polluted Belgian cities, i.e. Brussels, Antwerp and Liège, on 14, 15 April and 30 June 2015. The three acquisition sites are indicated on Fig. 1. Main characteristics of the acquired data sets and covered sites are given in Table 2.

The flight campaign was carefully planned (Vreys et al., 2016a) in order to optimise the acquisition for trace gas retrieval purposes. However, availability of the plane and crew, technical constraints, flight permissions granted by air traffic control (ATC), and most of all the meteorological conditions seemed to be the main factors constraining and determining the actual flights. APEX was installed in the Dornier DO-228 on 16 March 2015 and starting from that date, a 1 month window was scheduled to perform the necessary flights. In order to get an optimal NO₂ signal, early spring was considered to be a good trade-off: NO₂ exhibits a peak in the seasonal cycle around winter in the northern hemisphere. On the other hand, from the perspective of the solar zenith angle (SZA), flights in summer are considered to be more suitable. High sun maximises the signal or the light backscattered to the sensor and reduces the smoothing of the signal due to shallow sun elevation angle, as discussed in Lawrence et al. (2015). In summer, there are also more clear-sky days. Due to poor weather conditions, actual flights could only be performed at the end of the foreseen time window. The flights were conducted from a small commercial airport, southeast of the Antwerp city center. There were in total three flight days. A flight over Brussels was performed on the first campaign day, 14 April 2015. However, due to presence of a substantial amount of scattered clouds, seriously disrupting the analysis of the data, Brussels was re-flown on 30 June 2015. All other flights were performed under clear-sky conditions and good visibility.

In general, the flightpath consists of adjacent straight flightlines, flown systematically in opposite directions, with overlapping footprints (approximately 20% overlap in the across-track direction between consecutive flightlines), in order to have a gap-free coverage. For example the flightlines above Antwerp and Brussels were alternately flown from south to north, and from north to south, with the first flightline in the west. The spectra, acquired during the banking of the plane in order to prepare the acquisition of the next adjacent flightline, are not taken into account in the processing, due to the large roll angles. Flights are performed at a constant altitude of 6.1 km AGL, which is well above the planetary boundary layer (PBL) in order to capture the full NO₂ column below the plane. The flights were performed as close as possible to local noon

when the sun reaches its highest position, as far this was made possible by ATC and the weather conditions. The **three** cities were fully covered and also a number of flightlines were extended to rural background sites with decreased levels of NO₂. Additionally, a number of background sites were covered when the plane was flying from or to the airport or when flying from one city to another. These reference spectra, containing low NO₂ absorption, are used in the DOAS analysis.

5 **3 Post-flight pre-processing of APEX data for trace gas retrieval**

Hyperspectral observations from the APEX instrument are organised in 3D data cubes per flightline. Different standard APEX products are available: raw level 0 data representing the photon counts on the detector and level 1 data, being the at-sensor radiance after radiometric correction by application of a gain and offset to each detector, based on absolute radiometric calibration with a certified integrating sphere, among others (Hueni et al., 2008). For the purpose of trace gas retrieval in this study a customised product, named level 0-DC, is generated, i.e. a level 0 data product which is spectrally calibrated and corrected for dark current. The dark current is measured at start and end of each acquired flightline, based on a shutter mechanism. This product provides the most suitable spectral performance and the lowest DOAS fit root mean square (RMS) error that can be obtained from the APEX observations.

To reduce the data load and processing time, a spectral subset is taken on the level 0-DC product. The VNIR channel observes radiation in the 370 – 970 wavelength range while only the subset 370 – 600 nm is taken into account for the spectral calibration and the DOAS analysis.

3.1 Spatial binning and signal enhancement

The raw, spectrally and spatially unbinned APEX spectra, acquired with an integration time of 58 ms, have a typical signal-to-noise ratio (**SNR**) around 150. In order to reduce the noise level and increase the sensitivity of the instrument to NO₂, the raw pixels are spatially aggregated in postprocessing in the along- and across-track direction. If only shot noise is assumed (which is a good assumption for UV-Vis detectors), the noise should decrease with \sqrt{N} according to photon statistics, with N the number of binned spectra.

For a test area, the raw APEX spectra are spatially binned in the along- and across-track direction according to a power of 2, 2^{*n*}, with *n* ranging from 0 to 8. **For each binning level, a clean reference area is binned in the same way.** The number of binned spectra and the corresponding RMS of the noise, from the DOAS fit, are plotted on a double logarithmic scale in Fig. 3. The applied DOAS settings for the noise analysis are provided **in Sect. 4.1 and Table 4.** The red curve in the Allan plot represents the statistical shot noise \sqrt{N} , scaled to the noise of the unbinned raw spectra. High noise, around 7 x 10⁻³, can be observed for the unbinned APEX spectra. The measured noise follows photon statistics well until approximately 200 binned spectra, where a deviation of the noise slope occurs. Due to dominant instrumental noise, such as dark current and read-out noise, and systematic errors in the DOAS fit, saturation occurs and the noise cannot be reduced significantly anymore. The

noise is reduced to 4×10^{-4} or a favorable SNR of 2500, after binning of 20 pixels in along- and across-track direction. As the curve is slightly deviating from the statistical shot noise \sqrt{N} , the improvement of the SNR, compared to the unbinned spectra, is slightly less than 20 times.

The obtained higher SNR improves the detection limit of the instrument at the cost of a reduction of the original spatial resolution. A binning of 20 by 20 pixels is found to be an appropriate trade-off between the obtained instrument sensitivity and the spatial detail, leading to an effective GSD of approximately $60 \times 80 \text{ m}^2$.

The impact of a different spatial binning on the NO_2 VCD retrievals is illustrated in Fig. 4. The VCDs retrieved from a co-located column on flightline eight of the Antwerp data set are plotted from north to south for three different binning levels, i.e. 8^2 , 20^2 and 50^2 pixels. This results in an effective GSD of 24×32 , 60×80 , and $150 \times 200 \text{ m}^2$, respectively. Flightline eight of the Antwerp data set is binned according to three different binning levels, i.e. 8^2 , 20^2 and 50^2 pixels, resulting in an effective GSD of 24×32 , 60×80 , and $150 \times 200 \text{ m}^2$, respectively. Then, the same north-south along-track profile is taken from the three different sets of retrieved VCDs. As can be seen, the three NO_2 VCD along-track profiles show the same patterns of enhanced NO_2 , consisting of two major plumes related to industrial activities in the Antwerp harbor. However, the 8^2 binned data contains a lot of noise. On the other hand, the 50^2 binned data smooths out effective NO_2 signals, e.g. at pixel 2050. The smoothing effect can also be observed in Fig. 5. An orthogonal horizontal profile was taken through the second plume and a Gaussian was fitted on the obtained profile NO_2 VCDs for the three different binning levels. A broadening of the plume can be observed for the higher binning levels. The FWHM increases from 1685, 2129 to 2331 m for 8^2 , 20^2 and 50^2 pixels, respectively. The smoothing effect can also be observed in Fig. 5. A horizontal profile of approximately 2500 m was taken perpendicular to the second major plume and is indicated by a red dotted line in Fig. 15. Then, a Gaussian model was fitted to the obtained profile NO_2 VCDs for the three different binning levels. A broadening of the plume can be observed for the higher binning levels. The width of the fitted Gaussians is expressed as FWHM and broadens from 1685, 2129 to 2331 m for a binning of 8^2 , 20^2 and 50^2 pixels, respectively.

3.2 APEX spectral performance and wavelength calibration

A key characteristic of the spectral performance is the instrument spectral response function (ISRF or slit function), being the response of the instrument to a signal as a function of wavelength. The ISRF can be determined by a peak response, i.e. central wavelength (CW) and response shape, i.e. full width at half maximum (FWHM). Typically for pushbroom sensors, each across-track detector element should be considered as a 1-D instrument, due to optical aberrations and misalignments, with an intrinsic spectral response which is slightly different from the others.

In order to determine the spectral response of each pixel and to obtain a good alignment in the DOAS fit between the analysed spectra, reference and the absorption cross-sections, an in-flight spectral calibration is performed with the QDOAS software (Danckaert et al., 2015) prior to further processing. The accurate wavelength calibration aligns the Fraunhofer lines in the in-flight APEX spectra with a high resolution solar reference (Chance and Kurucz, 2010). The latter is iteratively

convoluted with an adjusted APEX slit function until a best match with the spectra is found, characterising the effective shift and FWHM. The APEX nominal wavelength-pixel relation, determined under laboratory conditions, is used as initial values in the calibration procedure in order to converge to the best solution. We determine FWHM and shift values in five sub-windows between 370 and 600 nm by minimizing the chi-square of the differences between the observed spectrum and the solar reference in each window. Then, we fit polynomials through the five resulting shift and FWHM values. As the pushbroom sensor consists of 1000 pixels across-track, which are binned by 20, the output consists of 50 different calibration sets in total.

During operation, airborne instruments are typically exposed to changes of environmental conditions (changes in pressure, humidity and temperature, vibrations, mechanical stress, etc.) which can affect the instrument characteristics and degrade its performance. Even though the optical unit of APEX is sealed in a nitrogen atmosphere and temperature stabilised, deviations can be significant between the nominal and effective in flight spectral performance. This has been extensively investigated in D'Odorico (2012) and confirmed by our study. Nominal parameters of the spectral performance, determined by the laboratory calibration, are given in Table 3. As a prism dispersion element is used, the FWHM is a non linear function and broadening with wavelength. The nominal FWHM has a Gaussian shape and is approximately 1.5 nm at 490 nm, i.e. the center of the fitting window, for the nadir detector element. Spectral shift of the CW from an absorption line should be smaller than 0.2 nm according to Schaepman et al. (2015).

The spectral calibration on the in flight spectra points out a broadening of the ISRF and bigger shifts than specified by the nominal calibration (see Table 3). In Fig. 6, the dependency of the FWHM and the shift on the scanline pixel position are plotted for 490 nm, and subsequently a second order polynomial has been fitted. The spectral calibration output is provided for several flightlines, i.e. 2 flightlines from the Antwerp data set (day 105), 1 flightline from the Liège data set (day 105) and 1 flightline from the Brussels data set (day 181). The FWHM ranges between 2.4 nm, for the outer detectors, and 3.3 nm, for the nadir detector, which is more than twice the nominal value and at the limits to detect subtle spectral variations. Also the spectral shift, exhibiting the typical smile curvature across the detector array (Richter et al., 2011), can be up to 4 times larger than the nominal value, e.g. 0.8 nm in case of the nadir detector on the Brussels flightline 10. Significant changes in FWHM and spectral shift can occur between different flightlines, largely attributed to changing pressure, temperature and environmental conditions, according to D'Odorico and Schaepman (2012) and Schaepman et al. (2015). In Fig. 6, minor changes of the slit function are detected between observations acquired on the same flightline. Bigger changes can, however, be observed between Antwerp flightline 8 and Liège flightline 3, acquired during the same flight (day 105) but with a time interval of approximately 1 hour. Substantial changes of both the FWHM and the spectral shift occur between the Antwerp (day 105) and Brussels (day 181) flight. This time dependent variability of the APEX slit function in operational conditions can be critical for the analysis of the spectra, as discussed in Sect. 4.2.

During operation, airborne instruments are typically exposed to changes of environmental conditions (changes in pressure, humidity and temperature, vibrations, mechanical stress, etc.) which can affect the instrument characteristics and degrade its spectral performance. Even though the optical unit of APEX is temperature stabilized and sealed in a nitrogen atmosphere, kept at 200 hPa above ambient pressure, deviations can occur between the nominal and effective in-flight spectral performance. This has been extensively investigated by D’Odorico et al. (2011) and Kuhlmann et al. (2016) and confirmed by our study. Nominal parameters of the spectral performance, determined by the laboratory calibration, are given in Table 3. Based on laboratory measurements, the CW spectral shift should be smaller than 0.2 nm according to Schaepman et al. (2015). However, based on a recent re-analysis, a larger uncertainty between 0.4 and 1 nm was observed in real spectra (Kuhlmann et al., 2016). As a prism dispersion element is used, the FWHM is a non-linear function and broadening with wavelength. The slit function is assumed to have a Gaussian shape, with a nominal FWHM of approximately 1.5 nm at 490 nm, i.e. the center of the fitting window, for the nadir detector element.

The spectral calibration on the in-flight spectra points out bigger CW shifts than specified by the nominal calibration as well as a broadening of the ISRF (see Table 3). In Fig. 6 a) the FWHM and b) the spectral shift and its dependency on the across-track scanline pixel position (spectral smile), are plotted for 490 nm for different across-track scanlines and flightlines: the mean of 20 scanlines 1) at the start and 2) at the end of flightline eight of the Antwerp data set (day 105), 3) at the start of flightline three of the Liège data set (day 105), and 4) at the start of flightline one of the Brussels data set (day 181). A second order polynomial has been fitted to each calibration set. In Fig. 6, minor changes of the slit function are detected between observations acquired on the same flightline. A spectral shift of approximately 0.18 nm can be observed between Antwerp flightline eight and Liège flightline three, acquired during a single flight (day 105) but with a time interval of approximately 1 hour. Large spectral shifts up to 0.8 nm can be observed between the Antwerp (day 105) and Brussels (day 181) flights. This is largely attributed to changing environmental conditions, with pressure changes in particular (D’Odorico et al., 2012). As mentioned before, the pressure of the nitrogen gas in the spectrometer is kept at 200 hPa above ambient pressure. Pressure changes, due to changing ambient pressure or inaccuracies of the pressure regulation, can affect the index of refraction of the nitrogen gas and, subsequently, the dispersion at the prism, resulting in spectral shifts (Kuhlmann et al., 2016). This time-dependent variability of the APEX slit function in operational conditions can be critical for the analysis of the spectra, as discussed in Sect. 4.2.

The range of wavelength shifts in the across-track direction, exhibiting the typical smile curvature (Richter et al., 2011) in Fig. 6.b, is approximately 0.4 nm, which is well within the given tolerance of 0.35 pixels (Schaepman et al., 2015). The FWHM ranges between 2.4 nm, for the outer detectors, and 3.3 nm, for the nadir detector, which is more than twice the nominal value and at the limits to detect subtle spectral variations. According to Kuhlmann et al. (2016), there is no significant sensitivity of the FWHM to pressure changes (in contrast to the spectral shifts). Based on an intensive analysis, the study of Kuhlmann et al. (2016) concludes that the larger in-flight FWHMs are likely explained by a combination of (1) not fully corrected CCD readout smear, resulting in a spectral smoothing, and (2) filling-in of the Fraunhofer lines.

Beside the spectral resolution and shift, the spectral sampling interval (SSI) and sampling ratio are important characteristics of the spectral performance. As specified in Table 3, the SSI is approximately 0.9 nm. According to Chance (2005), the Nyquist rate is just met in case of a sampling ratio of 3.1 to 3.6 pixels per FWHM, assuming a Gaussian ISRF. However, being at the limits, APEX is slightly undersampling the spectra, which can complicate the DOAS analysis due to reduced spectral information.

4 NO₂ vertical column density retrieval algorithm

The flowchart in Fig. 7 illustrates the key steps of the applied NO₂ VCD retrieval algorithm. For each of the modules, the main input and output data sets and their respective file formats are specified. A NO₂ vertical column can be derived for each APEX pixel or measured spectrum i , according to:

$$VCD_i = \frac{SCD_i}{AMF_i} \quad (1)$$

In Eq. (1), the NO₂ VCD_i or the integrated amount of molecules per cm² expected for a single, vertical transect of the atmosphere is defined as the ratio of the measured slant column density or the number of molecules per cm² detected in an observation (SCD_i) and a corresponding air mass factor (AMF_i) (Solomon et al., 1987). The direct output of the DOAS analysis (see Sect. 4.1) is not SCD_i , but a differential slant column density ($DSCD_i$), being the difference of the concentration of NO₂ integrated along the effective light path and the NO₂ concentration in a reference spectrum (SCD_{ref}). The residual NO₂ amount in the background spectrum needs to be determined in order to convert $DSCD_i$ to SCD_i (see Sect. 4.2). Therefore Eq. (1) needs to be rewritten as:

$$VCD_i = \frac{DSCD_i + SCD_{ref}}{AMF_i} \quad (2)$$

Or

$$VCD_i = \frac{DSCD_i + (VCD_{ref} * AMF_{ref})}{AMF_i} \quad (3)$$

AMFs account for enhancements in the optical path length of the slant column, due to viewing and sun geometry, albedo, aerosol scattering and the NO₂ vertical profile (see Sect. 4.3).

4.1 DOAS analysis of the measured spectra

The obtained binned APEX spectra are analysed in the 470-510 nm visible wavelength region by application of an adapted version of the QDOAS non-linear least-squares spectral fitting tool, developed at BIRA-IASB (Danckaert et al., 2015). The adapted version allows processing the spectra of full APEX flightlines simultaneously, with both the input and output in a

network common data form (netCDF) file format. Note that the small 470-510 nm fitting window is not considered to provide the highest sensitivity to NO₂. Instruments specifically designed for NO₂ trace gas retrieval usually operate in the 425-490 nm interval. This broad window contains more spectral information and strongly structured NO₂ absorption features, while there is low interference with absorption structures of other trace gases, providing optimal sensitivity to NO₂.
5 However, interference with unidentified instrumental artefacts or features prevents us from extending the fitting window to lower-wavelengths lower than 470 nm. Currently, the chosen wavelength interval is considered to be the best trade-off between sensitivity on the one hand and minimum interference with other absorbers and instrumental structures on the other hand.

The basic idea of the DOAS approach is to separate broadband and narrow band signals in the spectra and to isolate the narrow molecular absorption structures (widths usually smaller than a few nm).
10 Beside relevant high-pass filtered trace gas laboratory cross sections (NO₂ and O₄), a synthetic Ring spectrum, a synthetic resolution cross section and a low order polynomial term are fitted to the logarithm of the ratio of the observed spectrum, and a reference spectrum. They account for respectively (1) the Ring effect (Grainger and Ring, 1962), i.e. the filling in of Fraunhofer lines by rotational Raman scattering on air molecules (2) small differences in spectral resolution between the reference and the spectrum to analyze,
15 due to the unstable slit function, and (3) smooth broad band variations (e.g. reflection at the Earth's surface) and Rayleigh and Mie scattering. O₃ and H₂O cross sections were not fitted due to cross-correlations and overparameterisation of the small fitting interval. Further details about the main DOAS settings can be found in Table 4. Beside relevant high-pass filtered trace gas laboratory cross-sections (NO₂ and O₄), a low-order polynomial term, a synthetic Ring spectrum and a synthetic resolution cross-section are fitted to the logarithm of the ratio of the observed spectrum, and a reference spectrum. They
20 account for respectively (1) smooth broad-band variations (e.g. reflection at the Earth's surface) and Rayleigh and Mie scattering; (2) the Ring effect (Grainger and Ring, 1962), i.e. the filling-in of Fraunhofer lines by rotational Raman scattering on air molecules, and (3) the impact of small changes of the instrumental resolution. The synthetic resolution cross-section is based on a linearisation of the dependence of spectral Fraunhofer lines on the slit function width and corresponds to the first derivative of solar reference with respect to the slit function width (Danckaert et al., 2015; Beirle et al., 2017). O₃ and H₂O
25 cross-sections were not fitted due to cross-correlations and overparameterisation of the small fitting interval. Further details about the main DOAS settings can be found in Table 4.

A typical DOAS fit of an APEX spectrum is illustrated in Fig. 8. The direct output or resulting coefficients of the applied fitting algorithm are NO₂ differential slant columns. A NO₂ DSCD of 2.61×10^{16} molec cm⁻² is retrieved with an RMS on the residuals of 4.03×10^{-4} . In the absence of systematic structures, the RMS is the standard deviation of a Poisson distribution,
30 corresponding to the measured photons. The RMS corresponds with a favorable SNR of 2500, and is in line with the obtained signal enhancement after spatial binning, as discussed in Sect. 3.1.

4.2 Background spectrum

In the DOAS analysis, the concentration of NO₂ is determined with respect to an unknown amount of NO₂ in a selected reference spectrum (SCD_{ref}). This differential approach largely reduces the impact of both systematic instrumental instabilities and the prominent Fraunhofer lines in the spectra, which blur out the much finer trace gas absorption structures. The approach also cancels out the stratospheric NO₂ contribution to the signal, making the measurements only sensitive to tropospheric absorption, under the assumption that the stratospheric NO₂ field has a negligible spatial and temporal variability in the time between the acquisition of the reference spectrum and the measurements. At daytime, the stratospheric NO₂ column is characterised by a near-linear slow increase due to the photolysis of N₂O₅ at mid-latitudes. The diurnal increase of the NO₂ stratospheric column between 80° SZA sunrise and sunset is estimated to be approximately 1×10^{14} molec cm⁻² per hour (Tack et al., 2015) and thus much smaller than the retrieved VCDs. The latter have a mean value of 1.7×10^{16} and 7.7×10^{15} molec cm⁻² for the Antwerp and Brussels data set, respectively. All flights were performed close to local noon, and in general there was a minor time interval between the acquisition of the spectra and the reference of less than 1 hour.

The residual NO₂ amount in the reference (SCD_{ref}) is an unknown that needs to be estimated, which is a general shortcoming of all airborne DOAS pushbroom imagers. Some studies assume that there is no (Schönhardt et al., 2015) or very little, e.g. 1×10^{15} molec cm⁻² (Popp et al., 2012), tropospheric NO₂ in a background area. Due to a combination of the nature of the APEX instrument and the study area, these assumptions were not valid, as explained here:

~~The effective slit function is affected by environmental conditions during operation of APEX, despite the fact that the instrument is sealed, pressure and temperature stabilised. Fig. 6 shows that changes of the slit function can be significant for different flightlines within a same flight, and certainly for different flight days.~~ The in-flight slit function is affected by environmental conditions during operation of APEX, despite the fact that the instrument is sealed, pressure- and temperature-regulated. Fig. 6 shows that changes of the slit function occur between different flightlines of a single flight. Large spectral shifts can be observed between different flights/campaigns. The unstable slit function can cause misregistrations and spurious residuals in the DOAS fit. Even when the latter are small, they can impact the retrievals, considering that very fine absorption structures are analysed. As a consequence, a reference cannot be used to analyse a certain spectrum if the spectral performance deviates too much. Quality parameters of the fit, such as the RMS, were carefully checked in order to detect significant changes of the effective slit function between the analysed spectra and reference. In case of the Brussels data set (day 181), reference data were selected per individual flightline due to the larger instability affecting the slit function.

The nature of the reference area further complicates the estimation of SCD_{ref} . Based on a priori information, a number of candidate background areas around the cities were covered during the taxi flight between airport and survey area, far from emission sources and with decreased levels of NO₂. Due to the unstable slit function and its time-dependency, unfortunately these reference flightlines could not be used in the analysis in most cases. Instead, the references needed to be selected closer to the city, where the NO₂ levels are relatively high and generally have a strong spatial variability. Both for the Antwerp (day

105) and Brussels (day 181) data set, the references were selected in the south part of the flightlines, upwind of the main sources in the city. The SCD_{ref} was estimated based on co-located mobile-DOAS measurements. Following Eq.(2), the retrieved VCD will be overestimated in case the effective SCD_{ref} is lower than the estimated amount.

4.3 Air mass factor calculations

5 A slant column depends on multiple light paths of backscattered solar radiation, contributing to the spectrum observed by the instrument. In order to derive the effective optical path length through the atmosphere and thus to be able to interpret and compare observations, transfer of radiation in the atmosphere needs to be modelled and appropriate enhancement factors need to be calculated. In this study, NO₂ box-AMFs ($BAMF_i$) have been calculated with the linearised radiative transfer model (RTM) LIDORT 2.6 (Spurr, 2008). The box-AMFs describe the sensitivity of the measurements as a function of
 10 altitude, resulting in a height-dependent assessment of the instruments sensitivity (Wagner et al., 2007). The radiative transfer equation is solved in a multi-layer, multiple scattering atmosphere using the discrete ordinate method. RTM simulations are performed at 490 nm, i.e. the middle of the NO₂ fitting window. 93 atmospheric height layers j are defined from the ground surface to the top of the atmosphere (TOA), the latter determined at 120 km altitude, and for each layer a box-AMF is retrieved. The vertical discretisation consists of 40 layers of 50 m thickness, until 2 km and 12 layers of 500 m
 15 between 2 and 8 km. Above 8 km, the altitude grid of the US Standard Atmosphere is adopted. A total AMF ($TAMF_i$) can be derived for each APEX spectrum i by integration of the $BAMF_j$ along an a priori NO₂ vertical profile:

$$TAMF_i = \frac{\sum_{j=0}^{TOA} BAMF_j * VCD_j}{\sum_{j=0}^{TOA} VCD_j} \quad (4)$$

In Eq. (4), $BAMF_j$ and VCD_j refer to the box-AMF and the a priori partial NO₂ VCD of atmospheric layer j . A mean TAMF
 20 of 1.9 is obtained for the Antwerp and Brussels data set.

4.3.1 RTM parameters

LIDORT numerically reproduces the state of the atmosphere and transfer of the solar radiation through the atmosphere based on a priori information on all parameters that affect the lightpath, e.g. the surface albedo, sun and observation geometry, and atmospheric properties (pressure, temperature, cloud cover, absorber and aerosol vertical profiles). (1) The surface albedo, as
 25 well as (2) the sun and viewing geometry, i.e. SZA, viewing zenith angle (VZA) and relative azimuth angle (RAA), can be extracted from the observations. These additional (meta)data sets are provided for each observed spectrum and are spatially binned accordingly. The surface albedo is approximated by the APEX reflectance value (Level 2 product), which is obtained for each pixel after application of an atmospheric correction algorithm with MODTRAN4 (Berk et al., 1999) on the radiometrically calibrated level 1 at-sensor radiance product (Biesemans et al. 2007, Sterckx et al. 2016). The viewing and
 30 sun geometry are output products of the APEX orthorectification module, as described in Sect. 4.4. (3) **Pressure and**

temperature profiles are provided by the US Standard Atmosphere at mid-latitude. Pressure and temperature profiles are taken from the AFGL standard atmosphere for mid-latitude summer (Anderson et al., 1986). (4) Clouds usually introduce a major uncertainty in atmospheric modelling. However, all flights were performed under clear-sky conditions avoiding the necessity of a cloud retrieval scheme. (5) A priori NO₂ vertical profile shapes are obtained from the regional-scale air quality model AURORA (Air quality modelling in Urban Regions using an Optimal Resolution Approach; Lauwaet et al., 2014) at a high resolution of 1 x 1 km². The AURORA model consists of a fine vertical grid with 40 layers of 50 m thickness and 37 layers of 100 m thickness to a maximum altitude of 5700 m. Model output for the time of flight and overpass location was computed. Interpolation on the coupled 3D model grid provides a local NO₂ vertical profile for each APEX pixel. The obtained NO₂ profile shapes are assessed in the next section. (6) Aerosols can both enhance or reduce the AMF, depending on their vertical distribution. Since the latter information was not available and since all flights took place on clear spring/summer days with good visibility, a pure Rayleigh atmosphere was considered to compute the VCDs. The uncertainty related to this assumption is, however, discussed and quantified in the next section.

4.3.2 RTM sensitivity study

To study the impact of the input parameters on the AMF computations, sensitivity tests were performed based on different scenarios with varying input in the radiative transfer modelling. First, the mean and 1-sigma standard deviation were calculated for the albedo, RAA, VZA and SZA based on 73000 pixels/observations of the Antwerp data set. In Table 5, each row corresponds to an RTM parameter for which two scenarios are provided: mean +/- 1-sigma level (68%). For the study of a certain parameter, the other RTM parameters are assigned a fixed value μ , being 5%, 94.1°, 7°, and 54.6° for the albedo, RAA, VZA and SZA, respectively. For each scenario, corresponding TAMFs are computed.

For the sun and viewing geometries, the TAMF variability is low, i.e. less than 6% for input within $\mu \pm 1\sigma$. From the studied input parameters, the surface albedo has clearly the most significant impact on the AMF computations, which is consistent with previous investigations such as Boersma et al. (2004) and Lawrence et al. (2015). The TAMF variability can be up to 65% for albedo input within $\mu \pm 1\sigma$, illustrating the importance of having accurate knowledge of the surface properties in case of airborne imaging applications. Similar results were obtained for the Brussels data set and are provided in Table 5.

The strong dependency of the TAMF computations to the albedo is illustrated in Fig. 9. Similar statistics were obtained for the Brussels data set.

In order to obtain a better understanding of the instrument vertical sensitivity and the impact of the albedo on the radiative transfer, BAMFs are plotted as a function of the altitude in Fig. 10 for five different surface albedo scenarios. The sensitivity of the instrument is strongly height-dependent. The sensor is mostly sensitive to NO₂ observed directly below the airplane, with its sensitivity decreasing towards the ground surface, except for the scenario with very high albedo (45%). In case of high albedo, much of the incident radiance is reflected towards the sensor, increasing its sensitivity to NO₂ and thus significantly impacting the air mass factor. The sensitivity is about 2.9 directly under the plane and decreases to 0.3 to 2.1 at

the ground surface for an albedo of 0.3 to 8%, respectively. The sensitivity to the atmospheric layers above the sensor is almost constant and relatively low, around 1.7. This is very close to the geometrical AMF ($1/\cos(\text{SZA})$).

While the NO_2 horizontal distribution can be mapped, based on airborne APEX hyperspectral data, the details of the vertical distribution of NO_2 in the atmosphere are not well known. As discussed in Sect. 4.3.1, in this work, a time and space dependent NO_2 profile is interpolated on the high resolution AURORA 3D model grid for each APEX pixel. Then, TAMFs are computed based on integration along the obtained a priori NO_2 profile. In Fig. 11, the mean AURORA profile is plotted for the Antwerp study area. **Three** extreme cases of local AURORA profiles are plotted as well, i.e. for an urban, industrial and semi-rural site. Over the harbor, with many industrial activities, large NO_2 concentrations can be observed for the surface layer. Also the higher atmospheric layers contain relatively more NO_2 , probably due to emitting stacks which can be as high as 70 m. Over the city center, the bulk of the NO_2 can be related to traffic emissions at the surface.

The impact of the NO_2 profile shape on the TAMF computation is reported in **Table 6** by comparison of **two** scenarios: integration of the BAMFs along (1) local NO_2 vertical profiles A_{interp} , interpolated on the AURORA model grid and (2) a fixed AURORA NO_2 profile A_{harbor} , with very high NO_2 concentrations for the surface layers, and which can be assumed to be an extreme case. A scatter plot and linear regression analysis for the TAMFs, obtained by both scenarios, are shown in Fig. 12. The higher surface concentrations of the A_{harbor} profile give a larger weight to the relatively low BAMFs at the surface (see Fig. 10) resulting in a mean TAMF decrease of 7.5%, when compared to the first scenario. This impact reduces to zero at higher TAMF values, which can be related to higher albedo values, which have a much smaller decrease in sensitivity to the surface. In a second sensitivity test, the TAMFs obtained by integration along local AURORA NO_2 vertical profiles A_{interp} are compared with TAMFs computed based on a simple NO_2 box profile of 0.5 and 1 km height and well-mixed in the boundary layer. In case of a NO_2 box profile of 0.5 km, TAMFs are on average 7% smaller when compared to local AURORA NO_2 profiles, while TAMFs are slightly higher (~1%) in case of a box profile of 1 km.

Previous studies (Leitao et al., 2010; Meier et al., 2016) indicate that aerosols can enhance or reduce the AMF, depending on their position with respect to the NO_2 layer, the optical thickness, and the absorption of the aerosol layer. An aerosol optical thickness (AOT) lower than 0.15, at 500 nm, was measured during all flights by a CIMEL sun photometer at the AERONET station (Holben et al., 1998) in Uccle (50.78° N, 4.35° E, 100 m ASL). ~~Radiative transfer simulations with a corresponding well mixed extinction at the surface yield an albedo dependent aerosol effect of 10% or less when compared to the AMFs computed based on a Rayleigh atmosphere. This relative uncertainty is considered for all flights.~~ Radiative transfer simulations with a corresponding well-mixed extinction at the surface yield an albedo-dependent aerosol effect of 7% or less when compared to the AMFs computed based on a Rayleigh atmosphere. However, a relative uncertainty of 10% is considered for all flights, as the AERONET station in Uccle was in a semi-polluted area and furthermore to take into account the AOT variability.

4.4 Postprocessing: destriping

After application of the retrieval equation, i.e. Eq. (3), a bias correction is applied on the retrieved NO₂ VCDs to cope with the across-track stripe-like pattern in the generated maps. Striping is inherent to pushbroom sensors due to the intrinsic spectral response of each detector which is slightly different from the others (see Sect.3.2). The applied bias correction is based on the algorithms presented in Boersma et al. (2011), Popp et al. (2012) and Lawrence et al. (2015). For each flightline, the column NO₂ values are averaged and a third order polynomial is fitted to the column averages. The deviation from the polynomial is treated as a detector dependent bias and used as a correction factor to subtract from the retrieved NO₂ columns. The procedure removes the across-track striping to a great extent, while retaining the NO₂ spatial patterns. For visualisation purposes, the retrieved VCDs are smoothed by convolving a fifth order Savitzky-Golay least-squares polynomial filter (Savitzky and Golay, 1964; Schafer, 2011) in the across track direction. For visualisation purposes, the retrieved NO₂ VCD map is convolved by a low-pass filter in the across-track direction, reducing the high frequencies or short-scale variability. The applied Savitzky-Golay filter is based on a least-squares fitting of a second order polynomial over a span of five pixels (Savitzky and Golay, 1964; Schafer, 2011).

4.5 Spectra geolocation and georeferencing

The APEX sensor is equipped with a high-grade Applanix POS/AV 410 navigation system, which records sensor position, i.e. latitude, longitude and elevation, and orientation, i.e. pitch, roll and heading. Concurrently, global positioning system (GPS) base station data, for differential correction, and data originating from an inertial measurement unit (IMU) are recorded. All telemetry is blended in real-time and logged for post-processing to allow proper georeferencing of the spectra (Mostafa and Hutton, 2001). On 14 April also a boresight calibration flight (Mostafa, 2001) was performed over Oostend, Belgium for accurate georeferencing purposes. The boresight angles account for misalignments between the IMU axis and the sensor axis and they are computed every time APEX is mounted in the aircraft. Ground control points (GCPs) selected from orthophotos are identified in the APEX data and, following a Monte Carlo procedure, optimal parameters to compensate for roll, pitch and yaw errors are inferred, greatly reducing geolocation errors which are usually lower than 1 unbinned spatial pixel (Vreys et al., 2016b). The orthorectification and georeferencing module receives the sensor interior and exterior orientation, boresight calibration data and digital elevation model (DEM) data as input. The module outputs the position and the complete viewing geometry for each pixel or measured spectrum, allowing a proper mapping of the retrieved NO₂ spatial distribution. In Vreys et al. (2016b), the georeferencing module and its qualitative and quantitative assessment is discussed more deeply.

The georeferenced VCDs and intermediate products, e.g. DSCDs and AMFs, are eventually gridded and overlaid on Google Maps layers in an open-source geographic information system (GIS) environment, QGIS 2.10.1 (QGIS development team, 2009).

4.6 NO₂ VCD error budget

5 The overall error on the retrieved NO₂ VCDs originates from uncertainties in the calculated DSCDs, SCD_{ref}, and AMFs. One assumes that the contributing uncertainties are sufficiently uncorrelated as they arise from nearly independent steps. Based on Eq. (2), the overall error of the NO₂ VCD retrieval algorithm can be quantified based on the following error propagation method:

$$\sigma_{VCD_i}^2 = \left(\frac{\sigma_{DSCD_i}}{AMF_i}\right)^2 + \left(\frac{\sigma_{SCD_{ref}}}{AMF_i}\right)^2 + \left(\frac{SCD_i}{AMF_i^2} \times \sigma_{AMF_i}\right)^2 \quad (5)$$

10

(i) **The error** on the DOAS fit (1-sigma standard deviation), σ_{DSCD_i} , is a direct output of QDOAS for each fit and has a typical value between 3.4 and 4.4 x 10¹⁵ molec cm⁻² on the APEX DSCD retrievals, when spectra are binned 20 by 20 pixels. Note that this is approximately **1** order of magnitude higher than for the fixed ground-based stations, e.g. the BIRA-IASB MAX-DOAS instrument (Tack et al., 2015), having a higher SNR and better spectral performance. Whereas σ_{DSCD_i} is a rather minor error source in case of ground-based stations, it becomes a significant contributor to the total error in case of an airborne imager.

15

(ii) The second error source originates from the estimation of the NO₂ residual amount in the reference spectra, $\sigma_{SCD_{ref}}$. As SCD_{ref} is determined from co-located mobile-DOAS measurements, the overall error on the mobile-DOAS retrievals is taken into account. The mobile-DOAS error estimation is discussed in Merlaud (2013a) and Constantin et al. (2013). The mean overall error for the Antwerp and Brussels data set is 1.8 x 10¹⁵ molec cm⁻².

20

(iii) The error in the calculation of the air mass factor, σ_{AMF_i} , is caused by the uncertainties in the assumptions made for the radiative transfer model parameters (See Sect. 4.3.1). The contributing uncertainties can be summed in quadrature to obtain an overall error estimate σ_{AMF_i} . According to Boersma et al. (2004), the error budget associated with the computation of the AMF is dominated by the cloud fraction, surface albedo, and NO₂ profile shape: (1) As flights were performed under clear-sky conditions, cloud fraction is not considered an error source in this case. (2) Sensitivity tests, performed in Sect. 4.3.2, indicate that the surface albedo has the most significant impact on the effective lightpath, thus on the AMF. Within the albedo 1-sigma interval, the AMF variability can be up to 65%. However, as absolute radiances can be directly derived from the APEX instrument, the albedo can be determined with relatively high accuracy. For a realistic estimate of the uncertainty, following study was performed: several albedo types were measured in the field with an ASD FieldSpec-4 spectrometer

25

(<http://www.asdi.com/products-and-services/fieldspec-spectroradiometers/fieldspec-4-hi-res>, last access: 12-09-2016) and compared to the APEX surface albedo. For the wavelength 490nm, the average albedo error over all targets is 10%, which is assumed to be a realistic estimate of the uncertainty related to the a priori surface albedo. (3) Based on the sensitivity study performed in Sect. 4.3.2, the uncertainty related to the a priori NO₂ profile shape is lower than 8%. (4) According to the performed simulations, the uncertainty related to the assumption of a pure Rayleigh atmosphere is estimated to be less than 10%. (5) Both the viewing and sun geometry can be determined with high accuracy, thus the impact on the error in the AMF computation is expected to be small. Moreover, the performed sensitivity study, summarised in Table 5, has revealed that varying input for the viewing/sun geometry has a very low impact on the TAMF variability. Therefore it is assumed that the uncertainties related to RAA, VZA and SZA are less than 1%. Finally, all error sources contributing to the overall error σ_{AMF_i} are summed in quadrature and an estimate of approximately 15% is obtained.

The error budget analysis, based on 73000 observations of the Antwerp data set, is reported in Table 7 and Fig. 13. In Table 7, typical relative and absolute errors are given for different classes, depending on the NO₂ VCD amount. For each class, the typical error σ is provided for each individual error source in the retrieval approach, as well as for the overall error on the retrieved VCDs, σ_{VCD} .

An overall relative error, σ_{VCD} , of 21% on the retrieved NO₂ VCDs is reported in Table 7 for the Antwerp data set, or a mean VCD and absolute error of $17.4 \pm 3.7 \times 10^{15}$ molec cm⁻². For the Brussels data set the relative error is slightly higher (~28%), due to a lower mean VCD of $7.7 \pm 2.1 \times 10^{15}$ molec cm⁻². In general, the main contribution is coming from the error on the DOAS fit (1-sigma slant error), σ_{DSCD} , which can be quite substantial in case of small VCD retrievals (~29% for the Antwerp data set). For medium and high VCDs, the overall relative errors decrease to 21% and 18%, respectively. For the latter, the main contributions are from the error on the DOAS fit and the TAMF computation, σ_{TAMF} . Relative errors related to the DOAS fit and to the estimation of the residual amount, σ_{SCDref} , drop in case of larger VCDs, while errors originating from the TAMF computation are not affected. Of course the overall relative error is dependent on the NO₂ levels in the covered area. Low VCD retrievals will produce larger relative errors. The relatively low mean NO₂ VCD and related higher error for the Brussels data set is due to a combination of the seasonality of NO₂, the lack of significant industrial sources in the area and the fact that a fair amount of semi-rural area around the city was covered (see Sect. 5).

The overall absolute and relative errors are plotted in function of the retrieved NO₂ VCDs in Fig. 13, for the Antwerp data set. The largest absolute errors are obviously associated with the highest retrievals. The relative errors, on the other hand, which can be up to 100% in case of very low tropospheric contributions, show a steep and rapid drop in case of increasing VCDs. The relative error is almost constant (~21%) for NO₂ VCDs larger than 1.5×10^{16} molec cm⁻².

5 Discussion of the retrieval results

The generated NO₂ VCD distribution maps are shown in Figs. 15, 17 and 19 for respectively the Antwerp (day 105), Brussels (day 181) and Liège (day 105) data set. Meteorological conditions and the general flight pattern, both important to interpret the observed NO₂ field, are discussed in Sect. 2.2. In general, the obtained spectral and spatial resolution allows to map the fine scale NO₂ horizontal variability and spatial gradients, and to resolve individual emission sources. Patterns of enhanced NO₂ can be observed, largely consistent with and transported downwind from the emission inventory sources. The distribution maps show that the NO₂ concentrations can be highly variable in urban areas and can exhibit strong gradients. The maps are built from several adjacent flightlines with an approximated acquisition time of 8 to 15 min per flightline. As a dynamic NO₂ field is measured, minor biases can occur between adjacent flightlines.

The NO₂ VCD diurnal variation retrieved at the Uccle MAX-DOAS station on the campaign days, i.e. 15 April and 30 June 2015, are plotted in Fig. 14. The blue and red column indicate the flight time on April 15 over Antwerp (8:06 - 9:30 UTC) and on June 30 over Brussels (12:43 - 14:04 UTC), respectively. Further details of the station are provided in Sect. 6.2 and its location is indicated by a green dot in Fig. 17. Unfortunately there is no MAX-DOAS station so far in the Antwerp area. On both flight days, the Uccle station was upwind of the Brussels city center, thus in a semi-polluted area. During the April 15 flight, the NO₂ VCDs range between 4 and 21 x 10¹⁵ molec cm⁻², between 80° SZA sunrise and sunset, and VCDs increase approximately by 2.5 x 10¹⁵ molec cm⁻² during flight time. On June 30, NO₂ VCDs range between 4 and 7 x 10¹⁵ molec cm⁻². There is a slight decrease of 1.5 x 10¹⁵ molec cm⁻² during the APEX flight.

In the case of the Antwerp data set, the predominant anthropogenic emitters are mainly related to industrial activities in the harbor, but also to traffic in the southeast part. The port of Antwerp contains the biggest (petro)chemical cluster in Europe with branches of BASF, ExxonMobil, Solvay, Total, etc. The red dots in the NO₂ VCD distribution map (Fig. 15) represent the most significant stacks, emitting more than 25 kg of NO_x per hour, according to the emission inventory of the Belgian Interregional Environment Agency. The NO₂ field exhibits a strong gradient from west to east, consistent with the southwesterly wind direction. In the west, the NO₂ levels are low, around 3 to 7 x 10¹⁵ molec cm⁻² due to the lack of major contributing sources. Substantial uncertainties can occur in this area as the levels are close to the detection limit. Downwind of the sources, the transported NO₂ is building up and patterns of enhanced NO₂ can be observed with maxima up to 3.5 x 10¹⁶ molec cm⁻². The NO₂ VCDs are on average 1.7 ± 0.4 x 10¹⁶ molec cm⁻². The detected plumes are clearly related to and transported from the most contributing stacks in the area. The main central plume with large extent is a double plume with main emissions from two chimney stacks at site “A”, emitting respectively 30 and 95 kg of NO_x per hour at an altitude of 70 m AGL, and a third stack northeast of it at site “B”, emitting 145 kg of NO_x per hour at an altitude of 35 m AGL. The plume is approximately 12 km long and is unfortunately not fully covered by the flightplan. There is actually a cluster of 30 to 40 stacks at site “A” and “B” which are contributing as well to the central plume. However, according to the emission inventory their emissions are mostly lower than 10 kg of NO_x per hour. Smaller but clearly confined plumes origin from sites C and D, which are more isolated. Ship emissions are also expected to contribute to the observed NO₂ field, however these are hard to differentiate in this particular data set.

In the southeastern part of the NO₂ map (Fig. 15), NO₂ patterns can be observed, which are related to traffic emissions from the Antwerp city, the R1 ring road and the key highways E313 and E19. The two last flightlines are acquired around 11:00 - 11:30 AM LT and presumably the air masses containing the emissions from the rush hour are detected here, transported from the city and ring road R1 and building up, due to low wind speeds, northeast of the city. In the western part of the data set, an artefact is still present over the H-shaped docks due to too low retrieved VCDs over the water body and overestimation of the VCDs over the surrounding quays. Such an artefact is not present over other water bodies in the data set. The artefact could not be removed properly by selecting a larger amount of NO₂ in the reference spectrum. It is most probably caused by a combination of the overall low retrievals in this area, upwind of the sources, and computed AMFs that don't properly compensate for the lower backscatter over the water body (low albedo), and higher backscatter over the surrounding docks (high albedo). As discussed in Sect. 4.3.2, the albedo has a strong impact on the AMF computations. The albedo varies strongly from approximately 2% to 12% between the water body and the quay. A slight overestimation of the albedo, and subsequently the AMF, over the waterbody and underestimation of the albedo and AMF over the quay could explain this artefact.

On flightline eight of the Antwerp data set, an along-track profile is taken (see Fig. 15) for which the NO₂ DSCD and VCD time series are plotted from north to south in Fig. 16. Negative DSCDs point at a high amount of NO₂ in the reference/background, when compared to the analysed spectrum. The blue shaded error region on the DSCDs corresponds to the statistical error on the DOAS fit, being also a measure for the detection limit. In order for an absorber to be clearly identified, the retrieved column needs to be larger than this threshold. The 1 sigma detection limit of APEX retrievals has a typical value between 3.4 and 4.4×10^{15} molec cm⁻² on the DSCD, corresponding to 1.8 and 2.3×10^{15} molec cm⁻² on the VCD, assuming a typical AMF of 1.9 (See Sect. 4.3). The retrieved NO₂ signal is well above the detection limit. The 1-sigma slant error of APEX retrievals has a typical value between 3.4 and 4.4×10^{15} molec cm⁻² on the DSCD, corresponding to a detection limit of approximately 1.8 and 2.3×10^{15} molec cm⁻² on the VCD, assuming a typical AMF of 1.9 (See Sect. 4.3). In order for an absorber to be clearly identified, the retrieved column needs to be larger than this threshold.

The NO₂ VCD distribution map for the Brussels data set is shown in Fig. 16. The study area, consisting of the Brussels (sub)urban area and surrounding background, is covered by 95833 binned APEX pixels in approximately 80 minutes. In general, the NO₂ levels are almost 55 % lower than for the Antwerp data set, with minima and maxima of 1 and 20×10^{15} molec cm⁻², respectively. This is due to both the lack of significant industrial sources in the Brussels area, as well as the seasonality of NO₂ which tend to show maxima in winter and early spring. The NO₂ VCDs are on average $7.7 \pm 2.1 \times 10^{15}$ molec cm⁻². Again a strong gradient can be observed, consistent with the southeasterly wind direction, with low NO₂ VCDs above the Sonian forest in the southeast and increased levels downwind of Brussels city.

The NO₂ VCD distribution map for the Brussels data set is shown in Fig. 17. The study area, consisting of the Brussels (sub)urban area and surrounding background, is covered by 95833 binned APEX pixels in approximately 80 minutes. Again a strong gradient can be observed, consistent with the southeasterly wind direction, with low NO₂ VCDs above the Sonian

forest in the southeast and increased levels downwind of Brussels city and the ring road R0. The NO₂ VCDs in the Brussels data set are on average $7.7 \pm 2.1 \times 10^{15}$ molec cm⁻², with minima and maxima of 1 and 20×10^{15} molec cm⁻², respectively. In general, the NO₂ levels are almost 55 % lower than for the Antwerp data set. This is due to a combination of (1) the lack of significant industrial sources in the Brussels area, (2) the time of flight with respect to the traffic rush hours, and (3) the seasonality of NO₂ which tend to show maxima in winter and early spring. The NO₂ VCD diurnal variation retrieved at the Uccle MAX-DOAS station (Fig. 14) shows overall larger columns for the flight on 15 April, when compared to the flight on 30 June, with a mean NO₂ VCD of 10 and 5×10^{15} molec cm⁻², respectively, between 80° SZA sunrise and sunset. The seasonal NO₂ cycle, observed at the Uccle station, is discussed more in detail in Blechschmidt et al. (2017).

The NO_x sources in Brussels are predominantly related to traffic and concentrated along the R0 Brussels ring road and the junctions with the key highways E40 and E19. The R0 is one of the busiest highways in Belgium with traffic volumes of more than 70.000 cars per day. Patterns of enhanced NO₂ can be observed near the city center as well. In Fig. 18, a number of noticeable hotspots of interest in the Brussels data set, indicated by blue, dashed squares in Fig. 17, are highlighted and discussed here:

- Some of the highest emissions can be observed in the northeast where there is a major junction between the ring road R0 and the E19 (Fig. 18.a). Besides, there are two other main interchanges and some small industry in Vilvoorde. Strong NO_x emissions in this area can also be related to planes and airport traffic operations at the Brussels international airport. In June 2015, there were in total 22338 plane movements and on 30 June, approximately 50 planes took off in northeastern direction between 3 and 6 PM LT. However due to the proximity of the ring road and interchanges, it is nontrivial to differentiate the contributing sources.

- An increase of the NO₂ VCDs can be observed at “place Meiser” with values around 1.2×10^{16} molec cm⁻² (Fig. 18.b). It is a busy junction, close to the city center, where seven major roads are coming together.

- Fig. 18.c zooms in on the eastern part of the E40 highway, just past the junction with the ring road. The highway forms a clear segregation between the NO₂ levels in the northern and southern part. Due to the lack of considerable contributing sources, the air masses upwind of the highway contain little NO₂, while there is a significant increase noticeable downwind of it.

- The gas turbine Drogenbos powerplant in the southwest is the only significant industrial source, within the covered area, emitting more than 25 kg of NO_x per hour. A confined NO₂ emission plume, transported downwind from the double stack close to the channel, can be clearly resolved with a typical VCD around 1.5×10^{16} molec cm⁻² (Fig. 18.d).

The NO₂ levels observed in Liège range between 1 and 32×10^{15} molec cm⁻², with a mean VCD of $13.3 \pm 3.1 \times 10^{15}$ molec cm⁻². The overall error on the retrieved NO₂ VCDs is on average 23%. In order to avoid repetition, the small Liège data set is not shown and further discussed here as it leads to similar findings and conclusions. In the Liège data set, the highest NO₂ emissions are observed in the northeast, in the industrialised area of Herstal. The NO₂ levels range between 1 and 32×10^{15}

molec cm⁻², with a mean VCD of $13.3 \pm 3.1 \times 10^{15}$ molec cm⁻². The overall error on the retrieved NO₂ VCDs is on average 23%.

6 Correlative data sets

6.1 mobile-DOAS observations

5 For Brussels and Antwerp, the APEX NO₂ retrievals have been compared with correlative data sets acquired by a car mobile-DOAS system. The routes strive to have a good spatial distribution by covering the city ring road, and transects to the city center and other main NO₂ sources. The BIRA-IASB mobile-DOAS setup follows the MAX-DOAS principle (Hönninger et al., 2004) and consists of a compact double Avantes spectrometer, recording scattered light spectra simultaneously in two channels, one at zenith and one at 30° off-axis. The spectral resolution is around 1.1 nm. The telescopes, yielding a 2.6° field of view, are assembled in an optical head and mounted on the roof of the car. They are connected to the spectrometers by two 400 μm optical fibers. A GPS is used to geolocalise the measurements and the whole set-up uses the standard 12v car battery. The integration time is fixed at 5 ms and in order to increase the SNR, spectra series are averaged to a final spectrum, each 10 s. Thus, the spatial resolution of the measurements is around 138 m in case of a car speed of 50 km h⁻¹. Further details of the mobile-DOAS setup and the applied retrieval approach are discussed in Merlaud (2013a).

6.2 Mini MAX-DOAS observations

During the Brussels flight on 30 June, there was an overpass at 15:21 LT over the Uccle NDACC (Network for the Detection of Atmospheric Composition Change) candidate station (50.78° N, 4.35° E, 100 m ASL), where a mini-MAX-DOAS instrument is deployed (Ma et al., 2013; Gielen et al., 2014). The station is located south of the Brussels city center. Its location is indicated by a green dot on Fig. 17. The operating instrument, a commercial system from Hoffmann Messtechnik GmbH which has been continuously running since 2011, is pointing towards the north (Brussels city center) with a field of view of 0.6° and has a spectral resolution of 0.6 nm FWHM. Each 15 minutes, the instrument performs a full MAX-DOAS scan, comprising 11 elevation angles, including a zenith measurement.

6.3 Correlative comparison of retrieved NO₂ VCDs

25 The APEX, mobile-DOAS, and mini-MAX-DOAS observations are targeting mainly NO₂ in the lower troposphere. To ensure comparability of the collocated data sets, the retrieval settings are harmonised as much as possible.

For the comparison with the mobile-DOAS data, a VCD is extracted from the generated APEX NO₂ distribution map for each co-located mobile measurement. Both the mobile and APEX NO₂ VCD time series are plotted in Figs. 20.a and 20.b, for Antwerp and Brussels respectively. Near-simultaneous observations with APEX and the mobile-DOAS system were

pursued. In most cases, however, a NO₂ column at a certain geolocation is not sampled at exactly the same time and variability in local emissions and meteorology can lead to differences. In order to facilitate the interpretation of the NO₂ VCD time series, the time offset is plotted as well. A positive bias implies that a certain air mass was sampled earlier by the mobile-DOAS system than by APEX:

$$t_{offset} = t_{APEX} - t_{Mobile} \quad (6)$$

For Antwerp and Brussels, the time series are in good agreement, both for low and high VCDs. In general, higher VCDs up to 3×10^{16} molec cm⁻² are measured in Antwerp, mainly originating from the industrial activities in the harbor area. The NO₂ VCDs measured in Antwerp by APEX and mobile-DOAS are respectively 1.8 and 1.6×10^{16} molec cm⁻² on average, and respectively 6.8 and 6.4×10^{15} molec cm⁻² for the Brussels data set. The mobile measurements are representative for the whole data set as the averages are close to the mean values for the full NO₂ VCD distribution maps, being 1.7×10^{16} and 7.7×10^{15} molec cm⁻² for Antwerp and Brussels, respectively. In general a positive bias of approximately 12 % and 6 % can be observed for the APEX retrievals, for Antwerp and Brussels, respectively. The larger bias for the Antwerp data set can be related to the larger time offset of up to 3 hours. The impact of variability in local emissions and meteorology is subsequently larger here. The parts without mobile observations in the time series are related to car stops or traffic jams. For the Brussels flight, additional efforts were done to minimise the time offset between the mobile and APEX observations. The commercial application flightradar24.com was used for this purpose, visualising the real-time position of the plane. The time offset is less than 1 hour and 2 overpasses can be identified which are synchronized both in space and time, one at 12:46 and one at 13:39 UTC. For both cases, the difference between the retrieved VCDs is 1.1×10^{15} molec cm⁻².

The overpass at 13:21 UTC over the Uccle mini-MAX-DOAS station, which is synchronised both in time and location, is also plotted on Fig. 20.b. As can be observed in Fig. 17, the station is located upwind of the city and other main sources, so an urban background was measured. Both measurements are in good agreement: A NO₂ VCD of 5.8 and 6.5×10^{15} molec cm⁻² was retrieved for the mini MAX-DOAS and APEX, respectively.

Scatter plots and linear regression analysis of the APEX and mobile-DOAS NO₂ VCDs (~~averaged by two subsequent retrievals~~) are given in Figs. 21.a and 21.b, for the Antwerp and Brussels data set, respectively. In order to reduce the impact of very local emissions and sampling of the same APEX pixel, the mobile data is averaged in bins of two consecutive measurements. In total 521 observations for the Antwerp data set and 342 for Brussels were compared. The correlation coefficients are 0.84 and 0.85 for Antwerp and Brussels, respectively. Slopes are within 7% of unity and a larger intercept can be observed for the Antwerp data set. For Brussels, we see a slight positive bias for the mobile-DOAS in case of small VCDs, while it gets negative at higher VCDs. Efforts were done to ensure the comparability of the correlative data sets, but nevertheless the scatter can be largely explained by observation time differences in combination with variability of the NO₂ signal, sampling of different air masses due to the viewing geometry, differences in the sensitivity to NO₂, and instrumental and algorithmic conceptual differences and related errors and uncertainties.

7 Summary and conclusions

A retrieval scheme is presented to successfully infer the NO₂ VCD field over cities, based on the DOAS analysis of APEX level0-DC spectra. This is currently one of the few studies reporting on NO₂ horizontal distribution mapping at this scale and level of detail. APEX flights were performed above **three** of the largest urban areas in Belgium, being the cities of Antwerp and Liège (15 April 2015), and Brussels (30 June 2015). This study demonstrates that (1) the urban atmospheric NO₂ field can be mapped at high spatial resolution in a relatively short time frame, based on a systematic flightplan, (2) contributing local emission sources can be resolved, and (3) fine-scale structures within the detected plumes can be exposed. For example, Antwerp city, the harbor and the surrounding semi-rural area, covering approximately 350 km², have been mapped in less than 90 minutes. A spatial resolution of approximately 60 x 80 m² is obtained, after increasing the SNR to 2500 by a 20 x 20 spatial binning, with a superior geolocation accuracy of less than 3 m. The mapped NO₂ field shows that hotspots of enhanced NO₂, related to heavy traffic, industrial facilities, etc. can be identified and large emission sources can be distinguished.

The main sources in the Antwerp area (15 April 2015) appear to be related to (petro)chemical industry. The NO₂ levels range between 3 x 10¹⁵ molec cm⁻² in the west and 35 x 10¹⁵ molec cm⁻² in the east, downwind of the sources, with a VCD of 17.4 ± 3.7 x 10¹⁵ molec cm⁻² on average. The NO₂ levels observed in Liège (15 April 2015) range between 1 and 32 x 10¹⁵ molec cm⁻², with a mean of 13.3 ± 3.1 x 10¹⁵ molec cm⁻². The NO₂ levels in the Brussels area (30 June 2015) are on average 55% lower than in Antwerp with minima and maxima of 1 and 20 x 10¹⁵ molec cm⁻², respectively, and a mean VCD of 7.7 ± 2.1 x 10¹⁵ molec cm⁻². The NO_x sources are mainly originating from traffic and are concentrated along the R0 ring road and junctions with the key highways. In order to quantitatively assess the APEX NO₂ retrievals, the Antwerp and Brussels data set are compared with correlative car mobile-DOAS measurements. Both data sets are in good agreement with correlation coefficients around 0.85 and slopes close to unity. However, APEX retrievals tend to be on average 12% and 6% higher for Antwerp and Brussels, respectively.

The APEX instrument was initially not designed for trace gas retrieval applications. Despite its outstanding spatial resolution and georeferencing when compared to other imaging systems, and the benefit of being radiometrically calibrated periodically (at least once a year), this study has revealed some limitations related to the spectral performance, i.e. spectral resolution, sampling rate and robustness of the slit function in operational conditions. The economised spectral information and discussed instabilities lead to (1) additional mandatory steps in the retrieval approach, e.g. accurate in-flight wavelength calibration (**Recently a new in-flight spectral calibration algorithm, based on a maximum a posteriori optimal estimation approach, is proposed by Kuhlmann et al. (2016), improving the quality of the fit.**), (2) the selection of different reference spectra, (3) a higher detection limit, and (4) higher uncertainties in the retrieval approach. The error budget assessment indicates that the overall error σ_{VCD} on the retrieved NO₂ VCDs is on average 21%, 23% and 28% for the Antwerp, Liège and Brussels data set, respectively. Low VCD retrievals are mainly limited by noise (error on the DOAS fit or 1-sigma slant

error), while large retrievals are mainly limited by both the slant error as systematic errors, e.g. related to the AMF computations.

5 High resolution quantitative information about the atmospheric NO₂ horizontal variability is currently rare, but can be very valuable for (air quality) studies in urbanised areas. Airborne observations of the NO₂ field complement and bridge the gap between local point observations of ground stations, global monitoring by spaceborne instruments, and model output. Recently, new APEX flights were performed over Antwerp, Belgium as well as over Berlin, Germany. Revisiting Antwerp in the framework of the BUMBA project (Belgian Urban NO₂ Monitoring Based on APEX remote sensing) will (1) further improve the characterisation of the (temporal) NO₂ distribution in the city and (2) further consolidate the developed retrieval algorithm. The APEX flights over Berlin were performed in the framework of the ESA-EUFAR supported AROMAPEX 10 campaign, clustered in the AROMAT project (Airborne ROmanian Measurements of Aerosols and Trace gases). Two additional imagers, AirMAP (Airborne imaging DOAS instrument for Measurements of Atmospheric Pollution; Schönhardt et al., 2015; Meier et al., 2016) and SWING (Small Whiskbroom Imager for atmospheric composition monitoriNG; Merlaud et al., 2013b), were simultaneously operated from a Cessna from the Free University of Berlin, providing a unique data set for intercomparison purposes. These new data sets are currently under analysis, taking into account the findings and lessons 15 learned from the research described in this paper.

Acknowledgements. The Belgian Federal Science Policy Office is gratefully appreciated for funding the presented work and the BUMBA (Belgian Urban NO₂ Monitoring Based on APEX remote sensing) project, through the STEREO III Research Programme for Earth Observation (SR/00/310). The authors also wish to express their gratitude to the other partners of the BUMBA project, i.e. the Flemish Institute for Technological Research (VITO) and the Belgian Interregional Environment 20 Agency (IRCEL-CELINE). The authors wish to thank Gerrit Kuhlman and Dominik Brunner from the Swiss Federal Laboratories for Materials Science and Technology (EMPA) and Steffen Dörner of the Max-Planck-Institute for Chemistry (MPIC) for fruitful discussions.

References

25 Anderson, G., Clough, S., Kneizys, F., Chetwynd, J., and Shettle, E.: AFGL atmospheric constituent profiles (0-120km), Hanscom AFB, MA: Optical Physics Division, Air Force Geophysics Laboratory,[1986]. AFGL-TR; 86-0110. US Air Force Geophysics Laboratory. Optical Physics Division., 1, 1986.

Baidar, S., Oetjen, H., Coburn, S., Dix, B., Ortega, I., Sinreich, R., and Volkamer, R.: The CU Airborne MAX-DOAS instrument: vertical profiling of aerosol extinction and trace gases, Atmos. Meas. Tech., 6, 719–739, 2013.

30

Beirle, S., Kühl, S., Puķīte, J., and Wagner, T.: Retrieval of tropospheric column densities of NO₂ from combined SCIAMACHY nadir/limb measurements, *Atmos. Meas. Tech.*, 3, 283–299, doi:10.5194/amt-3-283-2010, 2010.

5 Beirle, S., Lampel, J., Lerot, C., Sihler, H., and Wagner, T.: Parameterizing the instrumental spectral response function and its changes by a super-Gaussian and its derivatives, *Atmos. Meas. Tech.*, 10, 581-598, doi:10.5194/amt-10-581-2017, 2017.

Berk, A., Andersonm, G.P., Bernstein, L., Acharya, P., Dothe, H., Matthew, M., Adler-Golden, S., Chetwynd Jr., J., Richtsmeier, S., Brian Pukall, B., Allred, C., Jeong, L. and Hoke, M.: MODTRAN4 radiative transfer modeling for atmospheric correction, in: Proceedings SPIE 3756, Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research III, Denver, CO, USA, 18 July, 1999.

10 Biesemans, J., Sterckx, S., Knaeps, E., Vreys, K., Adriaensen, S., Hooyberghs, J., Meuleman, K., Kempeneers, P., Deronde, B., Everaerts, J., Schlapfer, D., and Nieke, J.: Image processing workflows for airborne remote sensing, Proc. 5th EARSeL Workshop on Imaging Spectroscopy, 1-14, Bruges, Belgium, 2007.

15 Blechschmidt, A.-M., Arteta, J., Coman, A., Curier, L., Eskes, H., Foret, G., Gielen, C., Hendrick, F., Marécal, V., Meleux, F., Parmentier, J., Peters, E., Pinardi, G., Piders, A. J. M., Plu, M., Richter, A., Sofiev, M., Valdebenito, Á. M., Van Roozendael, M., Vira, J., Vlemmix, T., and Burrows, J. P.: Comparison of tropospheric NO₂ columns from MAX-DOAS retrievals and regional air quality model simulations, *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2016-1003, in review, 2017.

Boersma, K. F., Eskes, H. J., and Brinksma, E. J.: Error analysis for tropospheric NO₂ retrieval from space, *J. Geophys. Res.*, 109, D04311, doi:10.1029/2003JD003962, 2004.

20 Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D.: An improved tropospheric NO₂ column retrieval algorithm for the Ozone Monitoring Instrument, *Atmos. Meas. Tech.*, 4, 1905–1928, doi:10.5194/amt-4-1905-2011, 2011.

25 Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O. C., Vogel, A., Hartmann, M., Bovensmann, H., Frerik, J., and Burrows, J. P.: Measurements of molecular absorption spectra with the SCIAMACHY Pre-Flight Model: Instrument characterization and reference spectra for atmospheric remote sensing in the 230–2380 nm region, *J. Photochem. Photobiol. A*, 157, 167–184, 2003.

- Bruns, M., Buehler, S. A., Burrows, J. P., Richter, A., Rozanov, A., Wang, P., Heue, K. P., Platt, U., Pundt, I., and Wagner, T.: NO₂ Profile retrieval using airborne multi axis UV-visible skylight absorption measurements over central Europe, *Atmos. Chem. Phys.*, 6, 3049–3058, doi:10.5194/acp-6-3049-2006, 2006.
- 5 Bucsela, E. J., Krotkov, N. A., Celarier, E. A., Lamsal, L. N., Swartz, W. H., Bhartia, P. K., Boersma, K. F., Veeffkind, J. P., Gleason, J. F., and Pickering, K. E.: A new stratospheric and tropospheric NO₂ retrieval algorithm for nadir-viewing satellite instruments: applications to OMI, *Atmos. Meas. Tech. Discuss.*, 6, 1361–1407, doi:10.5194/amtd-6-1361-2013, 2013.
- Chan, A. W. H., Chan, M. N., Surratt, J. D., Chhabra, P. S., Loza, C. L., Crouse, J. D., Yee, L. D., Flagan, R. C.,
10 Wennberg, P. O., and Seinfeld, J. H.: Role of aldehyde chemistry and NO_x concentrations in secondary organic aerosol formation, *Atmos. Chem. Phys.*, 10, 7169–7188, doi:10.5194/acp-10-7169-2010, 2010.
- Chance, K. V. and Spurr, R. J. D.: Ring effect studies: Rayleigh scattering, including molecular parameters for rotational Raman scattering, and the Fraunhofer spectrum, *Appl. Optics*, 36, 5224–5230, 1997.
- 15 Chance, K., Kurosu, T. P., and Sioris, C. E.: Undersampling correction for array detector-based satellite spectrometers, *Appl. Opt.*, 44, 1296–1304, doi: 10.1364/AO.44.001296, 2005.
- Chance, K., and Kurucz, R. L.: An improved high-resolution solar reference spectrum for Earth's atmosphere measurements in the ultraviolet, visible, and near infrared, <http://www.cfa.harvard.edu/atmosphere> (last access: Sept 2013), 2010.
- 20 Constantin, D.-E., Merlaud, A., Van Roozendael, M., Voiculescu, M., Fayt, C., Hendrick, F., Pinardi, G., and Georgescu, L.: Measurements of tropospheric NO₂ in Romania using a zenith-sky mobile DOAS system and comparisons with satellite observations, *Sensors*, 13, 3922–3940, doi:10.3390/s130303922, 2013.
- Crutzen, P.: The influence of nitrogen oxides on the atmospheric ozone content, *Quart. J. Roy. Meteor. Soc.*, 96, 320–325,
25 1970.
- Danckaert, T., Fayt, C., and Van Roozendael, M.: QDOAS software user manual 2.108, IASB/BIRA, Uccle, Belgium, 2015, available at http://uv-vis.aeronomie.be/software/QDOAS/QDOAS_manual.pdf, last access: 12 December 2015.
- 30 Dix, B., Brenninkmeijer, C. A. M., Frieß, U., Wagner, T., and Platt, U.: Airborne multi-axis DOAS measurements of atmospheric trace gases on CARIBIC long-distance flights, *Atmos. Meas. Tech.*, 2, 639–652, doi: 10.5194/amt-2-639-2009, 2009.

D'Odorico, P., Guanter, L., Schaepman, M.E., Schläpfer, D.: Performance assessment of onboard and scene-based methods for Airborne Prism Experiment spectral characterization, *Appl. Opt.*, 50, 4755–4764, doi: <https://doi.org/10.1364/AO.50.004755>, 2011.

5 D'Odorico, P.: Monitoring the spectral performance of the APEX imaging spectrometer for inter-calibration of satellite missions, Remote Sensing Laboratories, Department of Geography, University of Zurich, 2012.

General, S., Pöhler, D., Sihler, H., Bobrowski, N., Frieß, U., Zielcke, J., Horbanski, M., Shepson, P. B., Stirm, B. H., Simpson, W. R., Weber, K., Fischer, C., and Platt, U.: The Heidelberg Airborne Imaging DOAS Instrument (HAIDI) - a
10 novel imaging DOAS device for 2-D and 3-D imaging of trace gases and aerosols, *Atmos. Meas. Tech.*, 7, 3459–3485, doi:10.5194/amt-7-3459-2014, 2014.

Gielen, C., Van Roozendaal, M., Hendrick, F., Pinardi, G., Vlemmix, T., De Bock, V., De Backer, H., Fayt, C., Hermans, C., Gillotay, D., and Wang, P.: A simple and versatile cloud-screening method for MAX-DOAS retrievals, *Atmos. Meas. Tech.*,
15 7, 3509-3527, doi:10.5194/amt-7-3509-2014, 2014.

Grainger, J. F. and Ring, J.: Anomalous Fraunhofer line profiles, *Nature*, 193, p. 762, 1962.

Harder, J. W. and Brault, J. W.: Atmospheric measurements of water vapor in the 442-nm region, *J. Geophys. Res.*, 102,
20 6245–6252, doi: 10.1029/96JD01730, 1997.

Hermans, C., Vandaele, A. C., Fally, S., Carleer, M., Colin, R., Coquart, B., Jenouvrier, A., and Mérienne, M.-F.: Absorption cross-section of the collision-induced bands of oxygen from the UV to the NIR, in: Proceedings of the NATO Advanced Research Workshop, Weakly Interacting Molecular Pairs: Unconventional Absorbers of Radiation in the
25 Atmosphere, Fontevraud, France, 24 April – 2 May 2002, edited by: Camy-Peyret, C. and Vigasin, A. A., NATO Science Series IV Earth and Environmental Sciences, vol. 27, Kluwer Academic Publishers, Boston, 193–202, 2003.

Heue, K.-P., Wagner, T., Broccardo, S. P., Walter, D., Piketh, S. J., Ross, K. E., Beirle, S., and Platt, U.: Direct observation of two dimensional trace gas distributions with an airborne Imaging DOAS instrument, *Atmos. Chem. Phys.*, 8, 6707-6717,
30 2008.

Hilboll, A., Richter, A., Rozanov, A., Hodnebrog, Ø., Heckel, A., Solberg, S., Stordal, F., and Burrows, J. P.: Improvements to the retrieval of tropospheric NO₂ from satellite – stratospheric correction using SCIAMACHY limb/nadir matching and comparison to Oslo CTM2 simulations, *Atmos. Meas. Tech.*, 6, 565–584, doi:10.5194/amt-6-565-2013, 2013.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET – A federated instrument network and data archive for aerosol characterization, *Remote Sens. Environ.*, 66, 1–16, 1998.

5

Hönninger, G., von Friedeburg, C., and Platt, U.: Multi axis differential optical absorption spectroscopy (MAX-DOAS), *Atmos. Chem. Phys.*, 4, 231–254, doi:10.5194/acp-4-231-2004, 2004.

Hueni, A., Biesemans, J., Meuleman, K., Dell’Endice, F., Schlapfer, D., Odermatt, D., Kneubuehler, M., Adriaensen, S.,
10 Kempenaers, S., Nieke, J., and Itten, K.: Structure, components, and interfaces of the Airborne Prism Experiment (APEX) Processing and Archiving Facility, *IEEE Transactions on Geoscience and Remote Sensing*, 47(1), 29-43, 2008.

Itten, K. I., Dell’Endice, F., Hueni, A., Kneubühler, M., Schläpfer, D., Odermatt, D., Seidel, F., Huber, S., Schopfer, J.,
15 Kellenberger, T., Bühler, Y., D’Odorico, P., Nieke, J., Alberti, E., and Meuleman, K.: APEX – the Hyperspectral ESA Airborne Prism Experiment, *Sensors*, 8, 6235–6259, doi:10.3390/s8106235, 2008.

Jehle, M., Hueni, A., Damm, A., D’Odorico, P., Kneubühler, M., Schläpfer, D., and Schaepman, M. E.: APEX – current status, performance and product generation, in: *Proc. IEEE Sensors 2010 Conference*, Waikoloa, Hawaii, USA, 1 - 4 November, 2010, doi:10.1109/ICSENS.2010.5690122, 2010.

20

Kuhlmann, G., Hueni, A., Damm, A. and Brunner, D.: An Algorithm for In-Flight Spectral Calibration of Imaging Spectrometers, *Remote Sensing*, 8(12), 1017, doi:10.3390/rs8121017, 2016.

Lauwaet, D., Viaene, P., Brisson, E., van Lipzig, N. P. M., van Noije, T., Strunk, A., Van Looy, S., Veldeman, N., Blyth, L.,
25 De Ridder, K., and Janssen, S.: The effect of climate change and emission scenarios on ozone concentrations over Belgium: a high-resolution model study for policy support, *Atmos. Chem. Phys.*, 14, 5893-5904, doi:10.5194/acp-14-5893-2014, 2014.

Lawrence, J. P., Anand, J. S., Vande Hey, J. D., White, J., Leigh, R. R., Monks, P. S., and Leigh, R. J.: High-resolution measurements from the airborne Atmospheric Nitrogen Dioxide Imager (ANDI), *Atmos. Meas. Tech.*, 8, 4735-4754, doi:10.5194/amt-8-4735-2015, 2015.

30 Lefebvre, W., Van Poppel, M., Maiheu, B., Janssen, S. and Dons, E.: Evaluation of the RIO-IFDM-street canyon model chain, *Atmospheric Environment* 77, 325-337, 2013.

- Leitão, J., Richter, A., Vrekoussis, M., Kokhanovsky, A., Zhang, Q. J., Beekmann, M., and Burrows, J. P.: On the improvement of NO₂satellite retrievals – aerosol impact on the airmass factors, *Atmos. Meas. Tech.*, 3, 475-493, doi:10.5194/amt-3-475-2010, 2010.
- 5
- Liu, C., Liu, X., Kowalewski, M. G., Janz, S. J., González Abad, G., Pickering, K. E., Chance, K., and Lamsal, L. N.: Characterization and verification of ACAM slit functions for trace-gas retrievals during the 2011 DISCOVER-AQ flight campaign, *Atmos. Meas. Tech.*, 8, 751-759, doi:10.5194/amt-8-751-2015, 2015.
- 10
- Ma, J. Z., Beirle, S., Jin, J. L., Shaiganfar, R., Yan, P., and Wagner, T.: Tropospheric NO₂ vertical column densities over Beijing: results of the first three years of ground-based MAX-DOAS measurements (2008–2011) and satellite validation, *Atmos. Chem. Phys.*, 13, 1547–1567, doi:10.5194/acp-13-1547-2013, 2013.
- Mayer, B. and Kylling, A.: Technical note: The libRadtran software package for radiative transfer calculations - description and examples of use, *Atmos. Chem. Phys.*, 5, 1855–1877, doi:10.5194/acp-5-1855-2005, 2005.
- 15
- Meier, A. C., Schönhardt, A., Bösch, T., Richter, A., Seyler, A., Ruhtz, T., Constantin, D.-E., Shaiganfar, R., Wagner, T., Merlaud, A., Van Roozendaal, M., Belegante, L., Nicolae, D., Georgescu, L., and Burrows, J. P.: High-resolution airborne imaging DOAS-measurements of NO₂ above Bucharest during AROMAT, *Atmos. Meas. Tech. Discuss.*, doi:10.5194/amt-2016-329, in review, 2016.
- 20
- Melamed, M. L., Solomon, S., Daniel, J. S., Langford, A. O., Portmann, R. W., Ryerson, T. B., Jr., D. K. N., and McKeen, S. A.: Measuring reactive nitrogen emissions from point sources using visible spectroscopy from aircraft, *Journal of Environmental Monitoring*, 5, 29–34, 2003.
- 25
- Merlaud, A., Van Roozendaal, M., Theys, N., Fayt, C., Hermans, C., Quennehen, B., Schwarzenboeck, A., Ancellet, G., Pommier, M., Pelon, J., Burkhardt, J., Stohl, A., and De Mazière, M.: Airborne DOAS measurements in Arctic: vertical distributions of aerosol extinction coefficient and NO₂ concentration, *Atmos. Chem. Phys.*, 11, 9219-9236, doi:10.5194/acp-11-9219-2011, 2011.
- 30
- Merlaud, A.: Development and use of compact instruments for tropospheric investigations based on optical spectroscopy from mobile platforms, Presses Universitaires de Louvain, 2013a.

Merlaud, A., Constantin, D., Mingireanu, F., Mocanu, I., Maes, J., Fayt, C., Voiculescu, M., Murariu, G., Georgescu, L., and Van Roozendaal, M. : Small whiskbroom imager for atmospheric composition monitoring (SWING) from an unmanned aerial vehicle (UAV), in: Proceedings of the 21st ESA Symposium on European Rocket & Balloon Programmes and related Research, Thun, Switzerland, 9 - 13 June, 2013b.

5

Mostafa, M.M.R.: Boresight calibration of integrated inertial/camera systems, in: Proceedings of the International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation, Banff, Canada, 5-8 June 2001, 440-445, 2001.

Mostafa, M.M.R. and Hutton, J.: Direct positioning and orientation systems: how do they work? What is the attainable accuracy?, in: Proceedings American Society of Photogrammetry and Remote Sensing (ASPRS) Annual Meeting, St. Louis, MO, USA, 22-27 April, 2001.

10

Nowlan, C. R., Liu, X., Leitch, J. W., Chance, K., González Abad, G., Liu, C., Zoogman, P., Cole, J., Delker, T., Good, W., Murcray, F., Ruppert, L., Soo, D., Follette-Cook, M. B., Janz, S. J., Kowalewski, M. G., Loughner, C. P., Pickering, K. E., Herman, J. R., Beaver, M. R., Long, R. W., Szykman, J. J., Judd, L. M., Kelley, P., Luke, W. T., Ren, X., and Al-Saadi, J. A.: Nitrogen dioxide observations from the Geostationary Trace gas and Aerosol Sensor Optimization (GeoTASO) airborne instrument: Retrieval algorithm and measurements during DISCOVER-AQ Texas 2013, *Atmos. Meas. Tech.*, 9, 2647-2668, doi:10.5194/amt-9-2647-2016, 2016.

15

Petritoli, A., Ravegnani, F., Giovanelli, G., Bortoli, D., Bonaf, U., Kostadinov, I., and Oulanovsky, A.: Off-axis measurements of atmospheric trace gases by use of an airborne ultraviolet-visible spectrometer, *Applied Optics*, 41, 5593–5599, 2002.

20

Platt, U. and Stutz, J.: *Differential Optical Absorption Spectroscopy: Principles and Applications*, Springer-Verlag, Berlin, Germany, 2008.

Popp, C., Brunner, D., Damm, A., Van Roozendaal, M., Fayt, C., and Buchmann, B.: High-resolution NO₂ remote sensing from the Airborne Prism EXperiment (APEX) imaging spectrometer, *Atmos. Meas. Tech.*, 5, 2211–2225, doi:10.5194/amt-5-2211-2012, 2012.

25

QGIS development team, 2009. QGIS Geographic Information System, Open Source Geospatial Foundation, <http://qgis.osgeo.org>.

30

- Richter, A. and Burrows, J.P.: Retrieval of Tropospheric NO₂ from GOME Measurements, *Adv. Space Res.*, 29, 1673–1683, 2002.
- 5 Richter, R., Schläpfer, D., and Müller, A.: Operational Atmospheric Correction for Imaging Spectrometers Accounting for the Smile Effect, *IEEE Trans. on Geoscience and Remote Sensing*, 49, 1772-1780, 2011.
- Savitzky, A., and Golay, M. J. E.: Smoothing and differentiation of data by simplified least squares procedures, *Anal. Chem. J.*, 36, 1627-1639, 1964.
- 10 Schaepman, M. E., Jehle, M., Hueni, A., D'Odorico, P., Damm, A., Weyermann, J., Schneider, F. D., Laurent, V., Popp, C., Seidel, F.C., Lenhard, K., Gege, P., Küchler, C., Brazile, J., Kohler, P., De Vos, L., Meuleman, K., Meynart, R., Schläpfer, D., Kneubühler, M., Itten, K. I.: Advanced radiometry measurements and Earth science applications with the Airborne Prism Experiment (APEX), *Remote Sensing of Environment*, 158, 207-219, 2015.
- 15 Schafer, R. W.: What is a Savitzky-Golay Filter?, *IEEE Signal Process. Mag.*, 28, 111-117, 2011.
- Schönhardt, A., Altube, P., Gerilowski, K., Krautwurst, S., Hartmann, J., Meier, A. C., Richter, A., and Burrows, J. P.: A wide field-of-view imaging DOAS instrument for two-dimensional trace gas mapping from aircraft, *Atmos. Meas. Tech.*, 8, 5113-5131, doi:10.5194/amt-8-5113-2015, 2015.
- 20 Solomon, S., Schmeltekopf, A. L., and Sanders, R. W.: On the interpretation of zenith sky measurements, *J. Geophys. Res.*, 92, 8311–8319, doi:10.1029/JD092iD07p08311, 1987.
- Solomon, S., Portmann, R. W., Sanders, R. W., Daniel, J. S., Madsen, W., Bartram, B., and Dutton, E. G.: On the role of nitrogen dioxide in the absorption of solar radiation, *J. Geophys. Res.*, 104, 12047–12058, doi:10.1029/1999JD900035, 25 1999.
- Spurr, R., *LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems*, *Light Scattering Reviews*, Volume 3, ed. A. Kokhanovsky, Springer, 2008.
- 30 Sterckx, S., Vreys, K., Biesemans, J., Iordache, M.-D., Bertels, L., and Meuleman, K.: Atmospheric correction of APEX hyperspectral data, *Miscellanea Geographica – Regional studies on development*, 20(1), 16-20, 2016.

- Tack, F., Hendrick, F., Goutail, F., Fayt, C., Merlaud, A., Pinardi, G., Hermans, C., Pommereau, J.-P., and Van Roozendael, M.: Tropospheric nitrogen dioxide column retrieval from ground-based zenith-sky DOAS observations, *Atmos. Meas. Tech.*, 8, 2417-2435, doi:10.5194/amt-8-2417-2015, 2015.
- 5 Valks, P. J. M., Pinardi, G., Richter, A., Lambert, J., Hao, N., Loyola, D., Van Roozendael, M., and Emmadi, S.: Operational total and tropospheric NO₂ column retrieval for GOME-2, *Atmos. Meas. Tech.*, 4, 1491–1514, doi:10.5194/amt-4-1491-2011, 2011.
- Vandaele, A.-C., Hermans, C., Simon, P.C., Carleer, M., Colin, R., Fally, S., Mérienne, M.-F., Jenouvrier, A. and Coquart, B.: Measurements of the NO₂ absorption cross-section from 42000 cm⁻¹ to 10000 cm⁻¹ (238-1000 nm) at 220 K and 294 K, *J. Quant. Spectrosc. Radiat. Transfer*, 59, 171–184, 1998.
- 10 Vreys, K., Iordache, M.-D., Bomans, B., and Meuleman, K.: Data acquisition with the APEX hyperspectral sensor, *Miscellanea Geographica – Regional studies on development*, 20(1), pp. 5-10. Retrieved 23 Nov. 2016, from doi:10.1515/mgrsd-2016-0001, 2016a.
- 15 Vreys, K., Iordache, M.-D., Biesemans, J., Meuleman, K. : Geometric correction of APEX hyperspectral data, *Miscellanea Geographica – Regional studies on development*, 20, doi:10.1515/mgrsd-2016-0006, 2016b.
- 20 Wagner, T., Burrows, J. P., Deutschmann, T., Dix, B., von Friedeburg, C., Frieß, U., Hendrick, F., Heue, K., Irie, H., Iwabuchi, H., Kanaya, Y., Keller, J., McLinden, C. A., Oetjen, H., Palazzi, E., Petritoli, A., Platt, U., Postlyakov, O., Pukite, J., Richter, A., Van Roozendael, M., Rozanov, A., Rozanov, V., Sinreich, R., Sanghavi, S., and Wittrock, F.: Comparison of box-air-mass-factors and radiances for Multiple-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) geometries calculated from different UV/visible radiative transfer models, *Atmos. Chem. Phys.*, 7, 1809-1833, 2007.
- 25 Wang, P., Richter, A., Bruns, M., Rozanov, V. V., Burrows, J. P., Heue, K.-P., Wagner, T., Pundt, I., and Platt, U.: Measurements of tropospheric NO₂ with an airborne multi-axis DOAS instrument, *Atmos. Chem. Phys.*, 5, 337–343, doi:10.5194/acp-5-337-2005, 2005.
- 30

Table 1. APEX spatial performance and instrumental specifications.

Spatial performance (at 6100 m AGL)	
Spatial CCD	1000 pixels
FOV (across-track)	28°
Swath width	3000 m
IFOV (across-track)	0.028°
Ground speed	72 m·s ⁻¹
Integration time	58 ms
Spatial resolution (across-track)	3 m
Spatial resolution (along-track)	4 m

5 **Table 2.** Acquired data sets and flight characteristics. Wind and temperature data are collected from weather stations of the Royal Meteorological Institute (RMI) and averaged over the time of flight. Population data from 1 January 2015 is retrieved from Statistics Belgium: <http://statbel.fgov.be/>, last access: 15 January 2016. For each flight, the day number of the year 2015 is mentioned between brackets and will be used further in the manuscript to refer to the different flights.

	Brussels	Antwerp	Liège
Date (day of year)	30-06-2015 (181)	15-04-2015 (105)	15-04-2015 (105)
Flight time LT (UTC + 2)	14:43 - 16:04	10:06 - 11:30	11:55 - 12:18
# flightlines	8	9	3
Flight pattern (Heading °)	0, 180	0, 180	40, 220
SZA (°)	29.7 - 38.6	60.4 - 49.6	46.0 - 44.1
Wind direction (°)	125	235	240
Wind speed (Bft)	2	3	3
Temperature (°C)	27.2	18.7	20.8
PBL height (m)	1200	500	700
Lat (°N) / Long (°E)	50.8 / 4.4	51.2 / 4.4	50.6 / 5.6
Terrain altitude (m ASL)	76	10	66
Total population	1.175.173	513.570	195.968
Population density (#/km²)	6751	2496	2828

10

15

Table 3. APEX spectral performance for the NO₂ calibration window. Both nominal (laboratory performance) and effective (in-flight performance) parameters are provided. Due to the wavelength- and pixel location-dependency, FWHM and shift are provided for 490 nm, i.e. the middle of the analysis window, and for the nadir looking detector **pixel** of the pushbroom sensor.

Spectral performance for NO₂ calibration window	
Spectral interval	370 - 600 nm
Spectral pixels	249 (unbinned mode)
Nominal FWHM	1.5 nm
In-flight FWHM	> 2.8 and < 3.3 nm
Nominal spectral shift from CW	< 0.2 nm (single flight)
In-flight spectral shift from CW	> 0.05 and < 0.8 nm
Spectral sampling interval (SSI)	0.9 nm
Sampling rate	3.1 to 3.6 pixels per FWHM

5

Table 4. Main DOAS spectral fitting analysis parameters for NO₂ slant column retrieval.

Parameter	Settings
Wavelength calibration method	Reference solar atlas (Chance and Kurucz, 2010) - Gaussian
Calibration interval	370 – 600 nm (5 sub-windows)
Fitting interval	470 – 510 nm
Cross-sections	
NO₂	Vandaele et al. (1998), 298 K
O₄	Hermans et al. (2003)
Ring effect correction method	Chance and Spurr (1997)
Resol	Small diff. in spectral resolution
Polynomial term	Polynomial order 5
Intensity offset correction	Polynomial order 1

10

15

20

Table 5. Sensitivity study with varying input parameters in the radiative transfer model based on 73000 and 96000 observations for the Antwerp and Brussels data set, respectively. For the albedo, sun and viewing geometry, each time two scenarios are provided based on the $\mu \pm 1\sigma$ level and corresponding TAMFs are derived by BAMF integration along a mean AURORA NO₂ profile.

RTM parameter	Parameter $\mu - 1\sigma$	Parameter $\mu + 1\sigma$	TAMF $\mu - 1\sigma$	TAMF $\mu + 1\sigma$	TAMF variability
Antwerp					
Albedo	2%	8%	1.3	2.2	65%
RAA	37.1°	151.1°	1.9	1.9	4%
VZA	4.4°	11.5°	1.9	1.9	1%
SZA	51.2°	58°	1.9	2.0	6%
Brussels					
Albedo	2%	8%	1.3	2.2	66%
RAA	48°	152.6°	1.9	1.8	3%
VZA	5.1°	11.7°	1.8	1.8	1%
SZA	31.9°	37.4°	1.6	1.7	3%

Table 6. A priori NO₂ profile shape sensitivity study with four scenarios: integration along (1) local NO₂ vertical profiles A_{interp} interpolated on the AURORA model grid; (2) a fixed AURORA NO₂ profile A_{harbor} over a polluted area; (3) a well-mixed NO₂ box profile of 0.5 km and (4) 1 km height in the lowest layer. TAMFs for the Antwerp data set are calculated based on the four scenarios, as well as the variability between the first scenario and the other three.

NO ₂ profile	TAMF	TAMF variability
A_{interp}	1.9	-
A_{harbor}	1.8	8%
Box _{0.5km}	1.8	7%
Box _{1km}	1.9	1%

Table 7. Error budget analysis based on 73000 retrieved NO₂ VCDs of the Antwerp data set. Typical relative errors (percent) and absolute errors ($\times 10^{15}$ molec cm⁻² for σ_{DSCD} , σ_{SCDref} and σ_{VCD}) are provided in column two to four, for small (< 33th percentile or < 1.4×10^{16} molec cm⁻²), medium (33th to 66th percentile or 1.4 to 2.0×10^{16} molec cm⁻²) and high (> 66th percentile or > 2.0×10^{16} molec cm⁻²) NO₂ VCD retrievals, respectively. Column five shows the mean error for all retrieved VCDs.

Error source	Small VCDs	Medium VCDs	High VCDs	All VCDs
σ_{DSCD}	40% (3.8)	23% (3.9)	15% (3.9)	22% (3.9)
σ_{SCDref}	19% (1.8)	11% (1.8)	7% (1.8)	10% (1.8)
σ_{TAMF}	15% (0.3)	15% (0.3)	15% (0.3)	15% (0.3)
σ_{VCD}	29% (2.8)	21% (3.5)	18% (4.7)	21% (3.7)

5

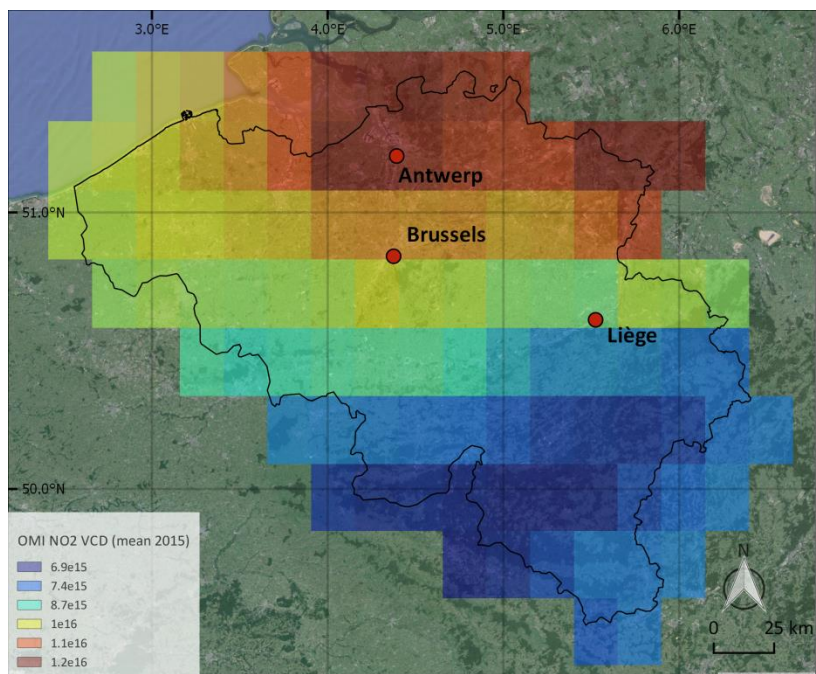


Figure 1. OMI annual NO₂ VCD map for Belgium, 2015 and overview of the **three** Belgian cities where APEX flights were performed (Google, TerraMetrics, Giovanni, NASA GES DISC, OMNO2d, units in molec cm⁻²).

10

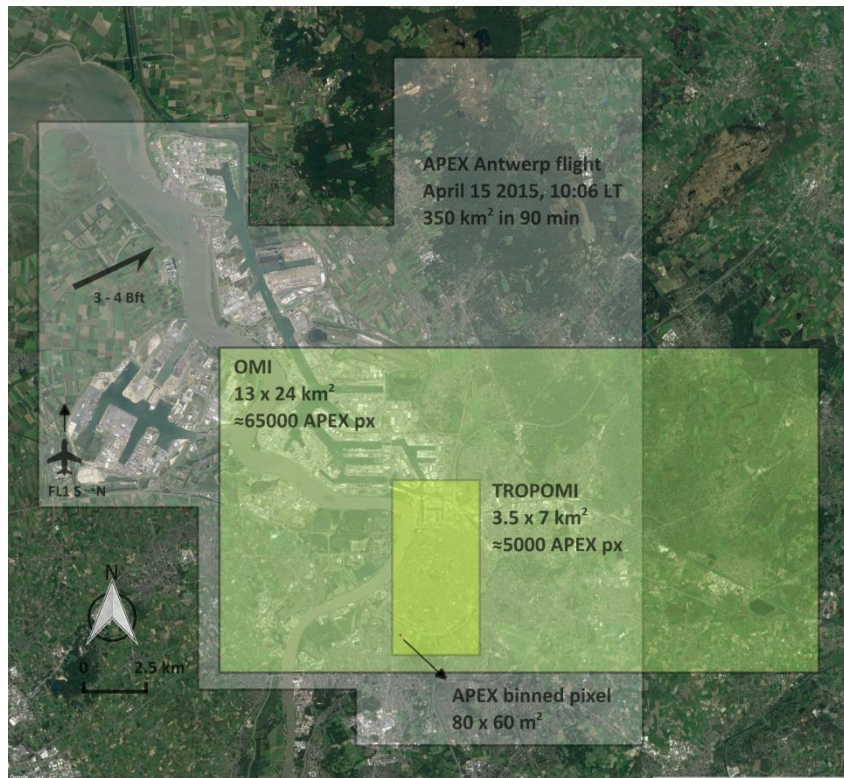


Figure 2. Details of the APEX flightplan over Antwerp on 15 April 2015 and comparison with a spaceborne nadir OMI and TROPOMI pixel (Google, DigitalGlobe).

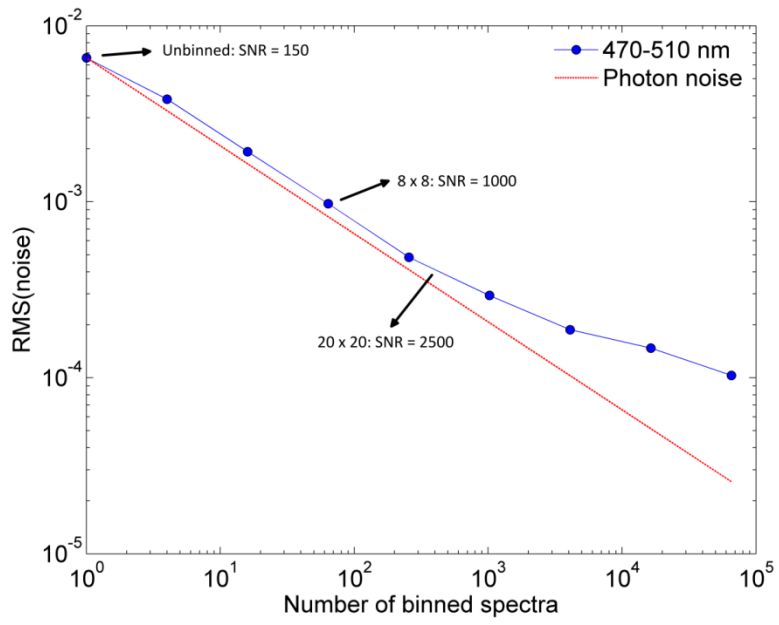


Figure 3. Allan plot illustrating the impact of spatial binning of the raw spectra on the RMS of the noise, plotted on a double logarithmic scale.

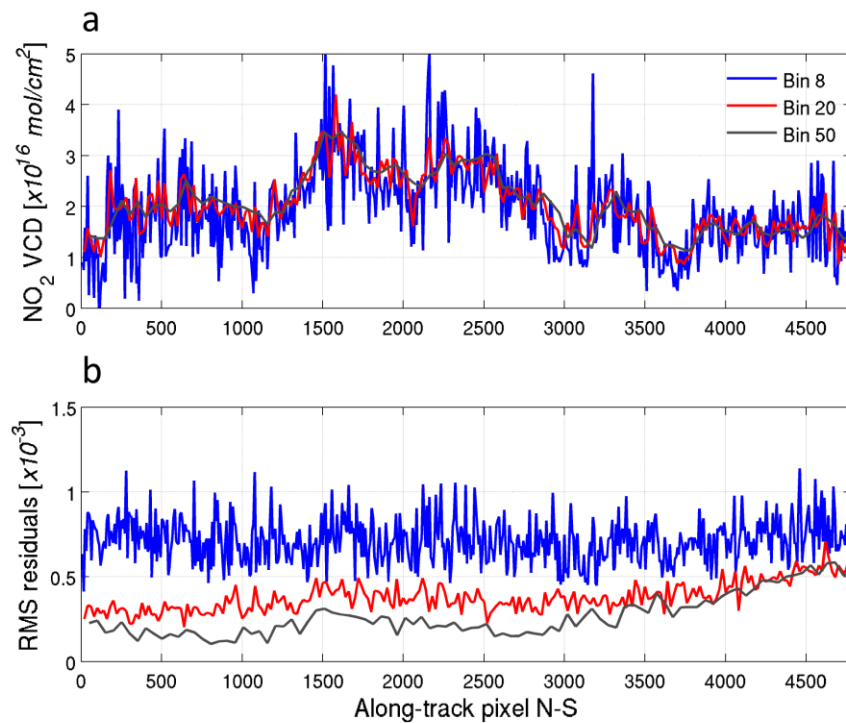


Figure 4. Along-track profile of a) NO₂ VCDs and b) RMS error of DOAS fits for 3 different levels of aggregated spectra, i.e. 8 by 8, 20 by 20, and 50 by 50 pixels. The VCDs are retrieved from an overlapping column on flightline eight of the Antwerp data set and are plotted from north to south.

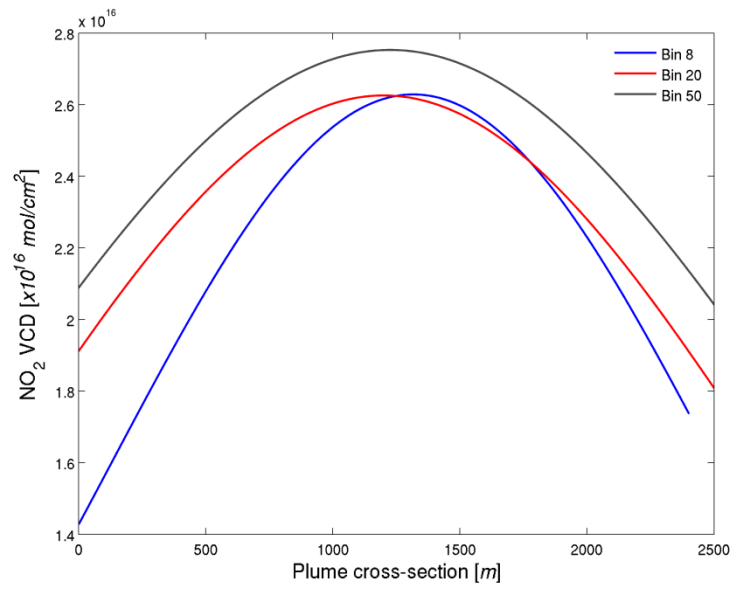


Figure 5. Gaussian model fit to the NO₂ VCDs of a 2500 m horizontal profile, perpendicular to a major industrial NO₂ plume, for three different binning levels, i.e. 8², 20² and 50² pixels, respectively.

5

10

15

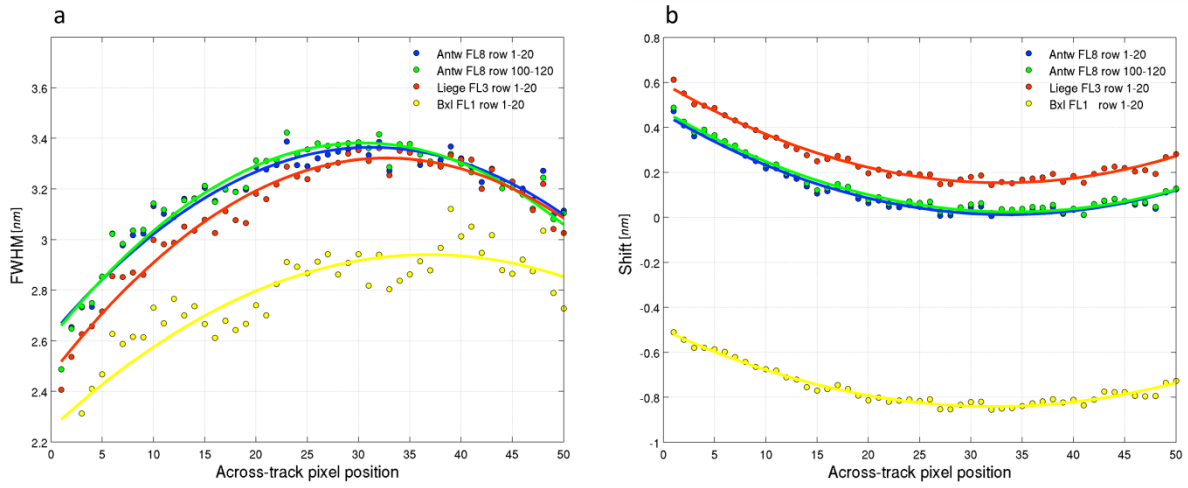


Figure 6. In-flight spectral calibration: **a)** the spectral resolution (FWHM) and **b)** the spectral shift and its dependency on the across-track scanline pixel position (spectral smile), plotted for 490 nm, i.e. the middle of the analysis window, for different flightlines. A second order polynomial has been fitted to each calibration set.

5

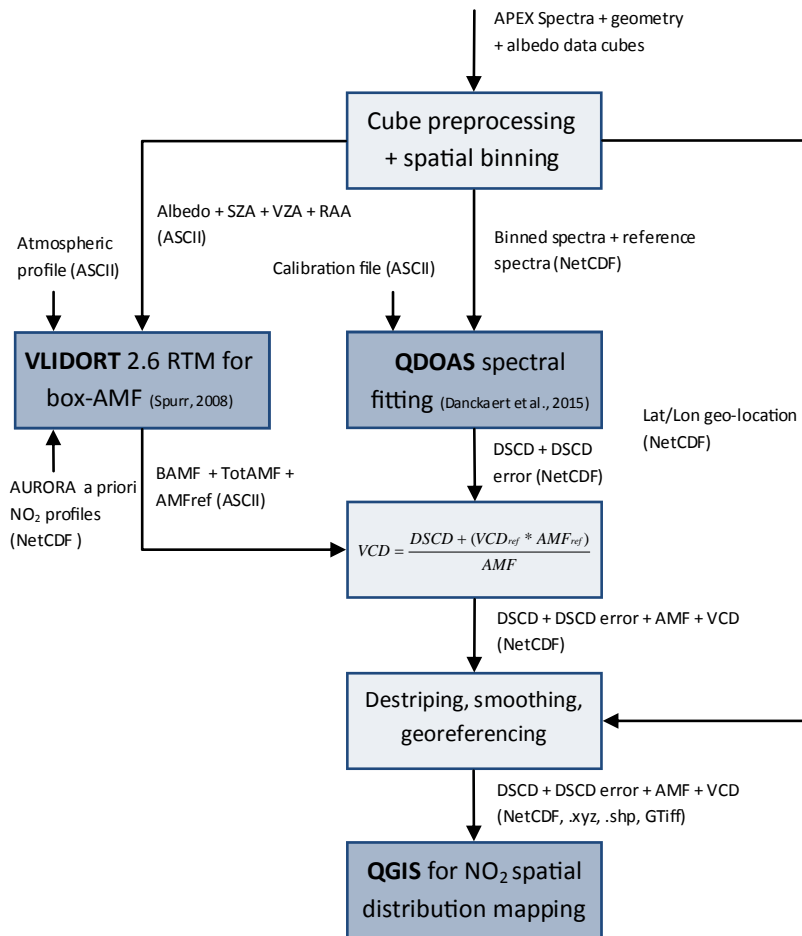


Figure 7. Flowchart of the APEX NO₂ VCD retrieval algorithm.

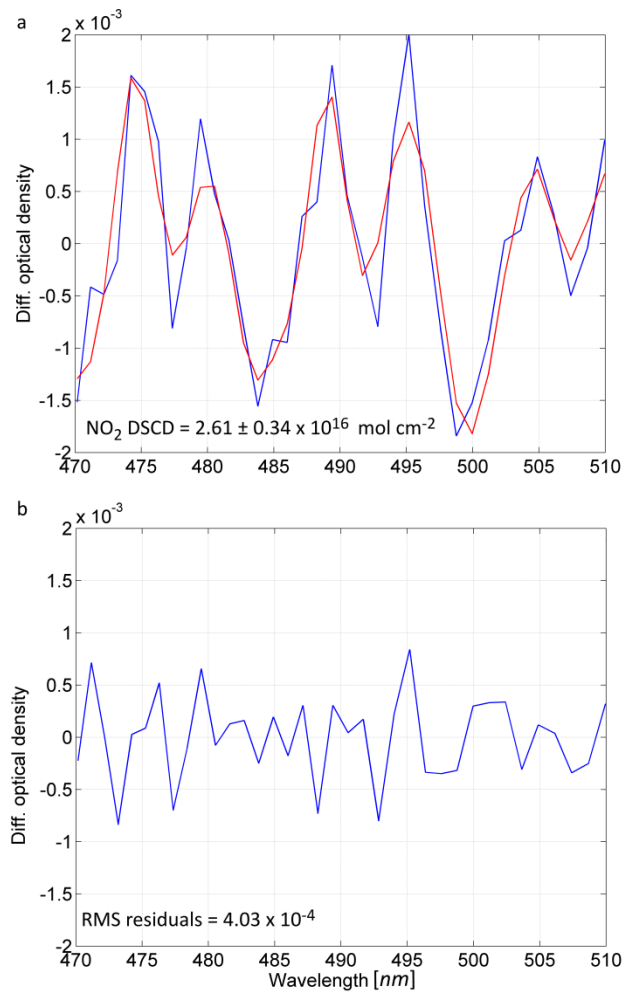


Figure 8. Typical DOAS fit with **a)** red line, corresponding to the NO_2 molecular cross-section, convolved with the instrument slit function and scaled to the detected absorption in the measured spectrum (blue line) and **b)** the remaining residuals of the spectral fit.

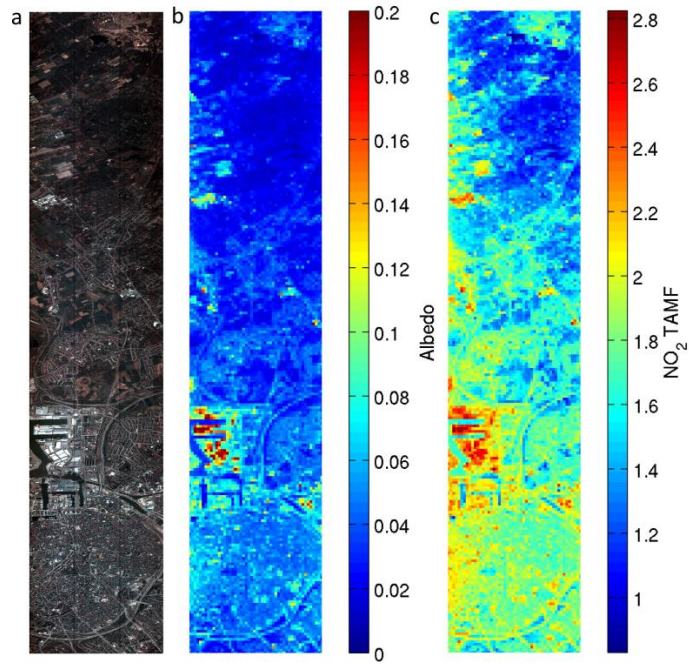


Figure 9. **a)** APEX true color composite, **b)** APEX albedo level 2 product and **c)** computed TAMFs, for flightline **eight** of the Antwerp data set (15 April 2015). The strong dependency of the AMF to the albedo can be observed.

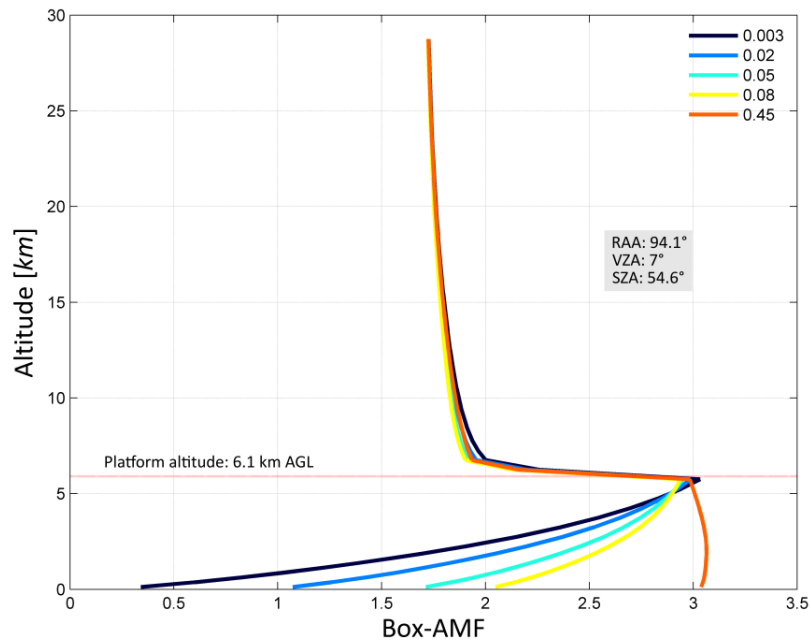


Figure 10. BAMF profiles illustrating the vertical sensitivity of the APEX instrument to NO_2 . The high impact of the surface albedo, mainly on the lowest atmospheric layers, is shown based on **five** different scenarios, ranging from low to high albedo. Scenarios are based on the minimum, $\mu - 1\sigma$, mean, $\mu + 1\sigma$ and maximum albedo in the Antwerp data set.

5

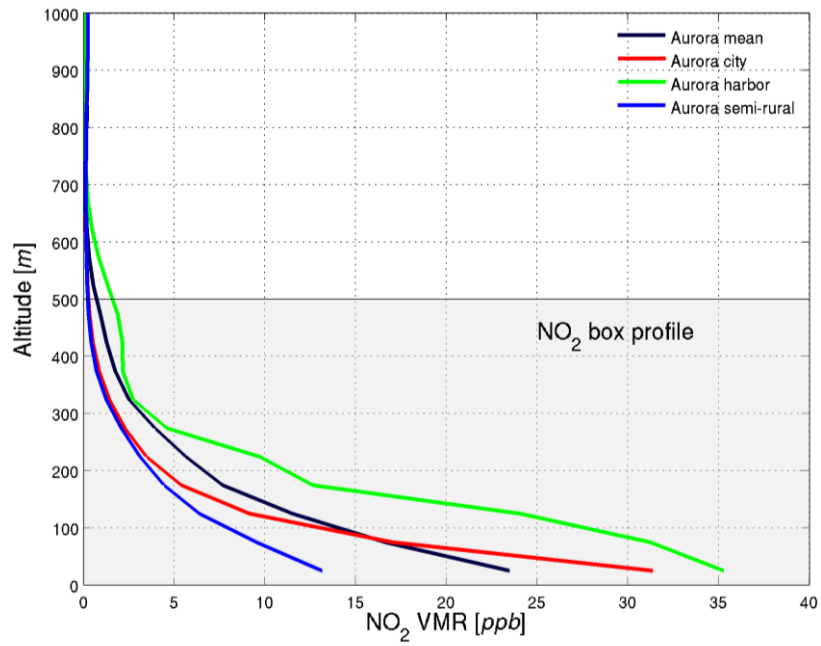


Figure 11. Representative AURORA a priori NO₂ profiles, used for the Antwerp data set RTM calculations. A simple NO₂ box profile of 500 m height which is well-mixed in the boundary layer is used as well in the sensitivity study.

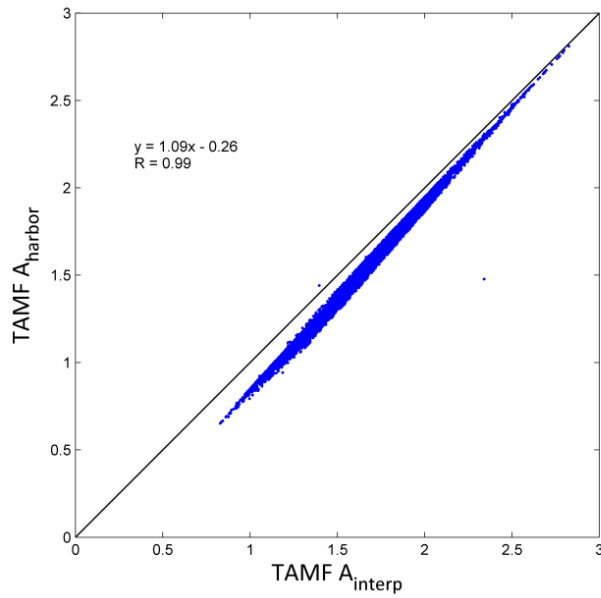


Figure 12. Scatter plot and linear regression analysis for the TAMF computation, based on **two** scenarios: integration of the BAMFs along (1) local NO_2 vertical profiles A_{interp} , interpolated on the AURORA model grid and (2) a fixed AURORA NO_2 profile A_{harbor} , with high NO_2 concentrations related to industrial sources.

5

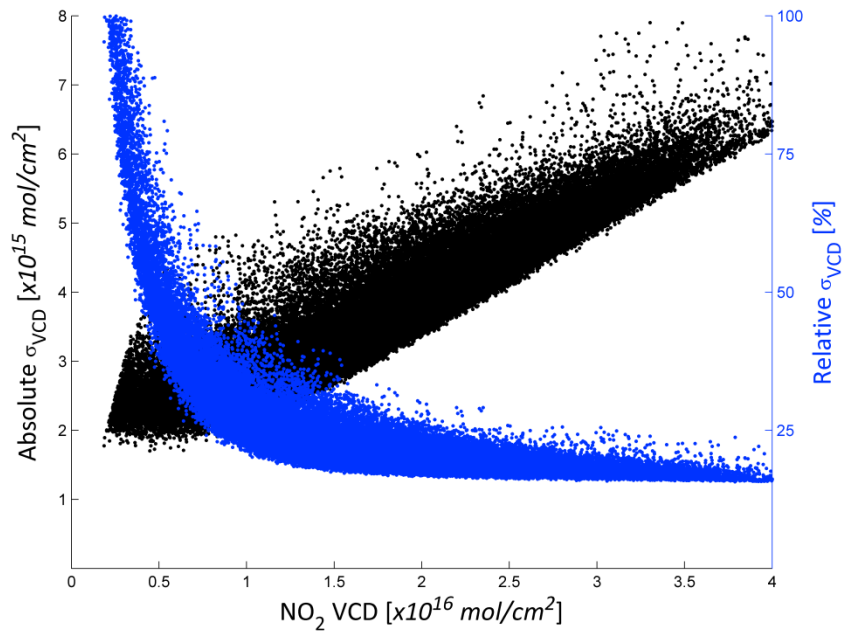


Figure 13. Overall absolute (black dots) and relative errors (blue dots), σ_{VCD} , on the retrieved NO₂ VCDs, based on the Antwerp data set.

5

10

15

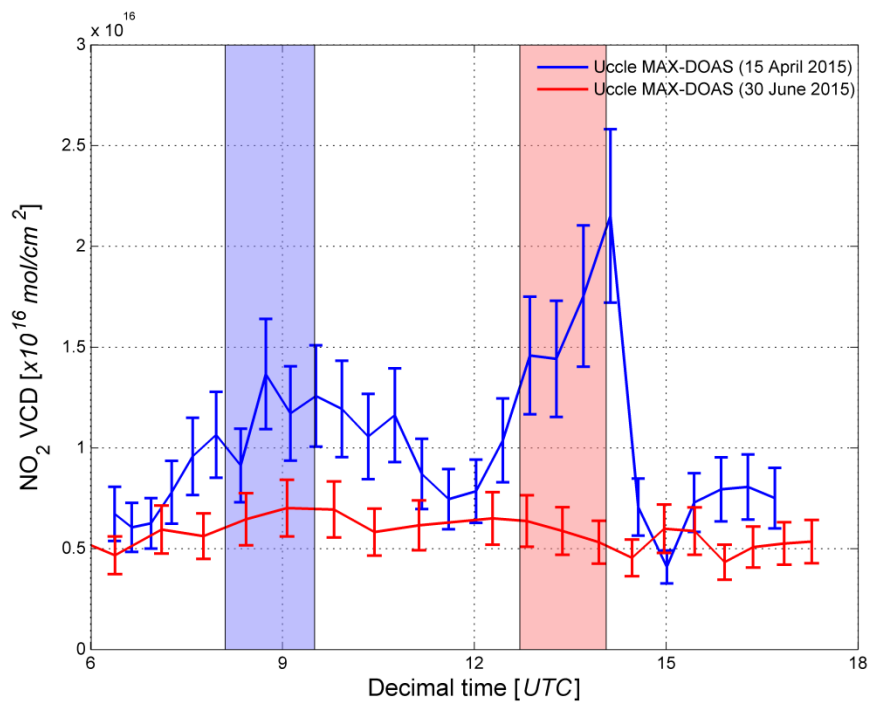


Figure 14. NO₂ VCD diurnal variation retrieved from the Uccle MAX-DOAS station on 15 April and 30 June 2015. The blue and red column indicate the flight time on April 15 over Antwerp and on June 30 over Brussels, respectively.

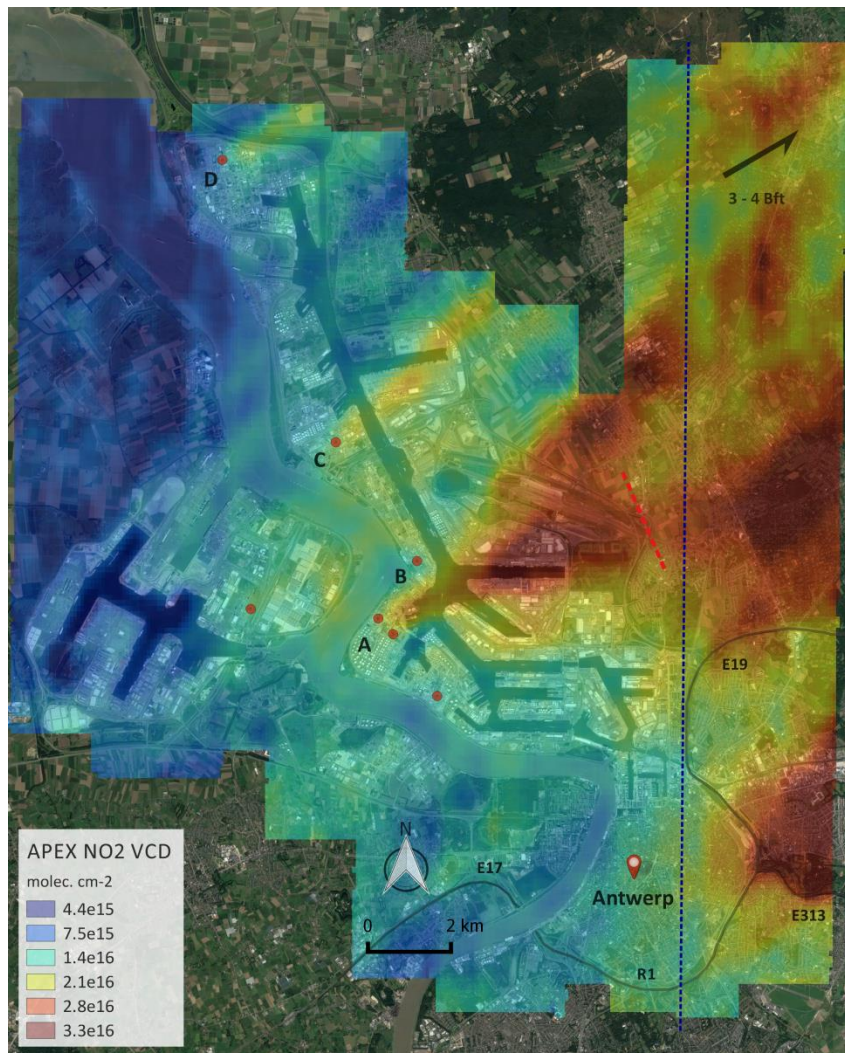


Figure 15. Retrieved NO₂ VCD field for Antwerp (15 April 2015) (Google, TerraMetrics). Red dots indicate the chimney stacks, emitting more than 25 kg of NO_x per hour, according to the emission inventory of the Belgian Interregional Environment Agency. **Four industrial sites**, which are further discussed in Sect. 5, are indicated by the labels A-D. The red dotted line indicates the plume cross-section, plotted in Fig. 5. The blue vertical line indicates an along-track profile, for which the NO₂ DSCD and VCD time series are plotted in Fig. 16.

5

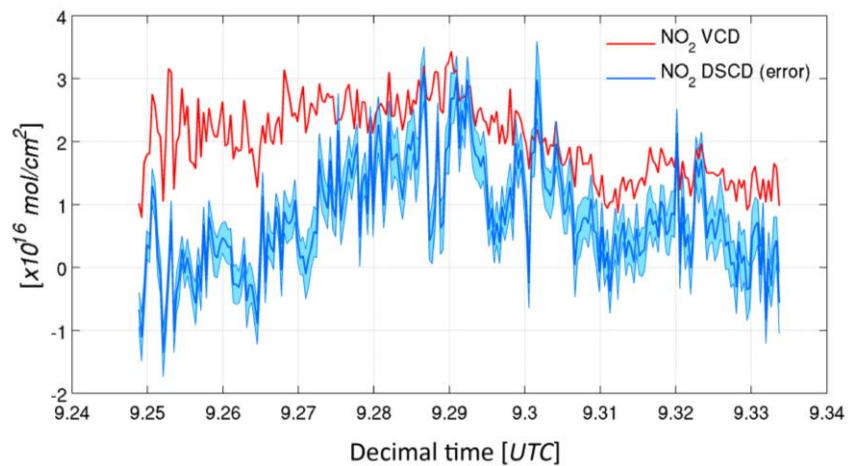


Figure 16. NO₂ DSCD and VCD time series for an along-track profile from north to south taken on flightline eight of the Antwerp data set (15 April 2015). The retrieved NO₂ signal VCDs is are well above the detection limit of approximately 1.8 to 2.3 x 10¹⁵ molec cm⁻². The profile, crossing the main plume from the harbor and the city center, is indicated by a dashed (blue) vertical line in Fig. 15.

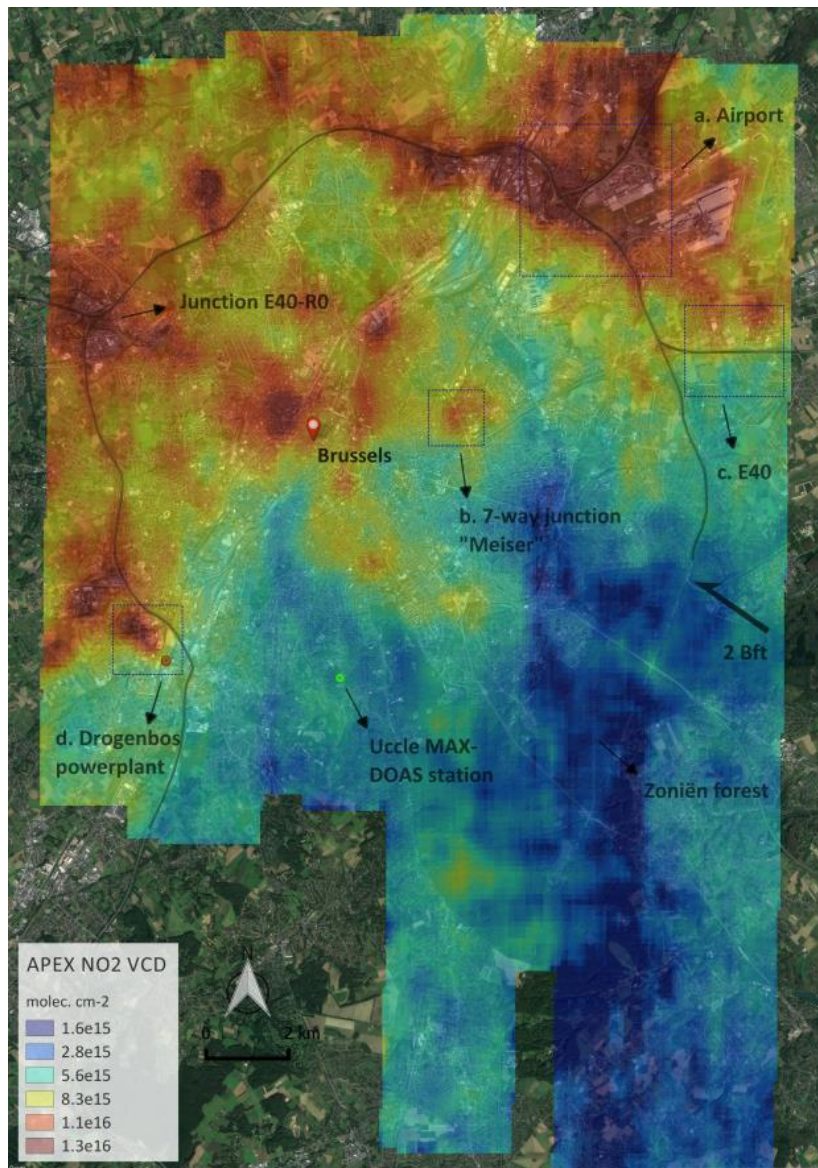


Figure 17. Retrieved NO_2 VCD field for Brussels (30 June 2015) (Google, TerraMetrics). Blue squares indicate four NO_2 VCD hotspots, highlighted in Fig. 18. The green dot is the location of the Uccle MAX-DOAS station.

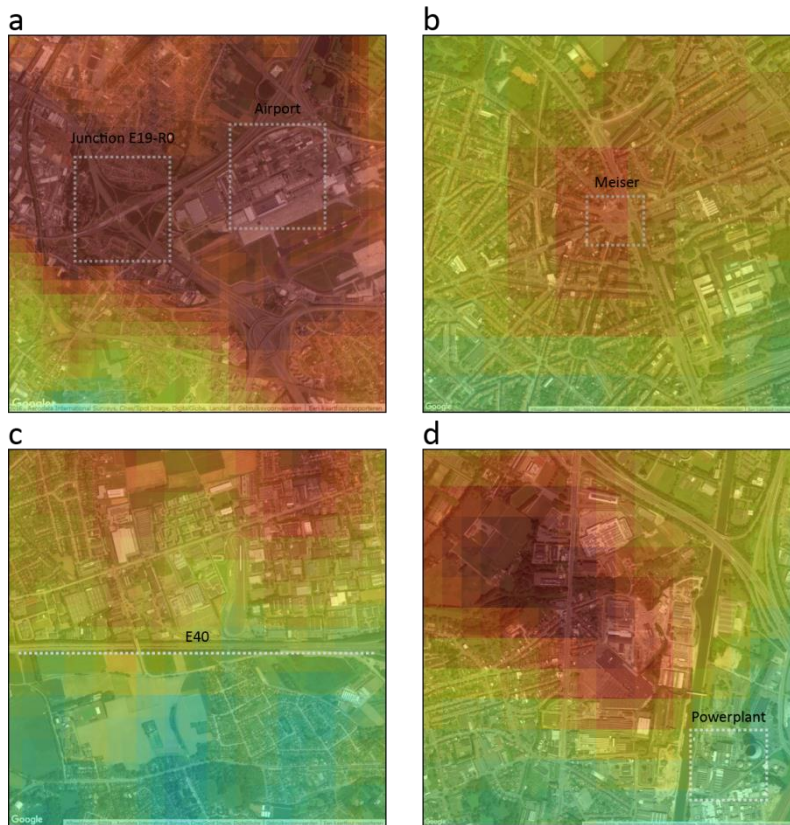


Figure 18. Enlargement of **four** peculiar NO₂ VCD hotspots in the Brussels data set (30 June 2015): **a**) major junction between the ring road R0 and the E19, and Brussels international airport; **b**) junction “place Meiser”, close to the city center; **c**) eastern part of the E40 highway; **d**) gas turbine Drogenbos powerplant. The locations of the **four** zooms are indicated in Fig. 17 by blue squares (Google, TerraMetrics).

5

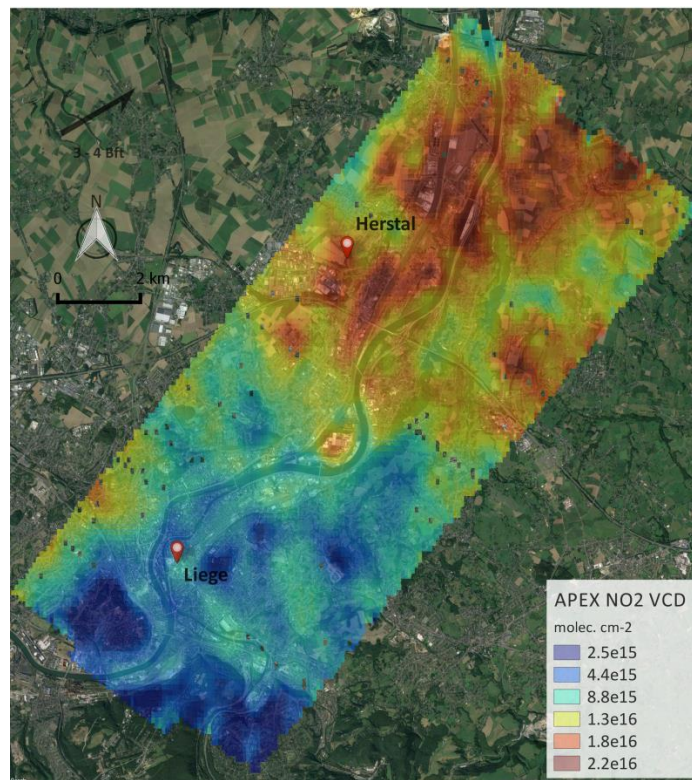


Figure 19. Retrieved NO₂ VCD field for Liège (15 April 2015) (Google, TerraMetrics).

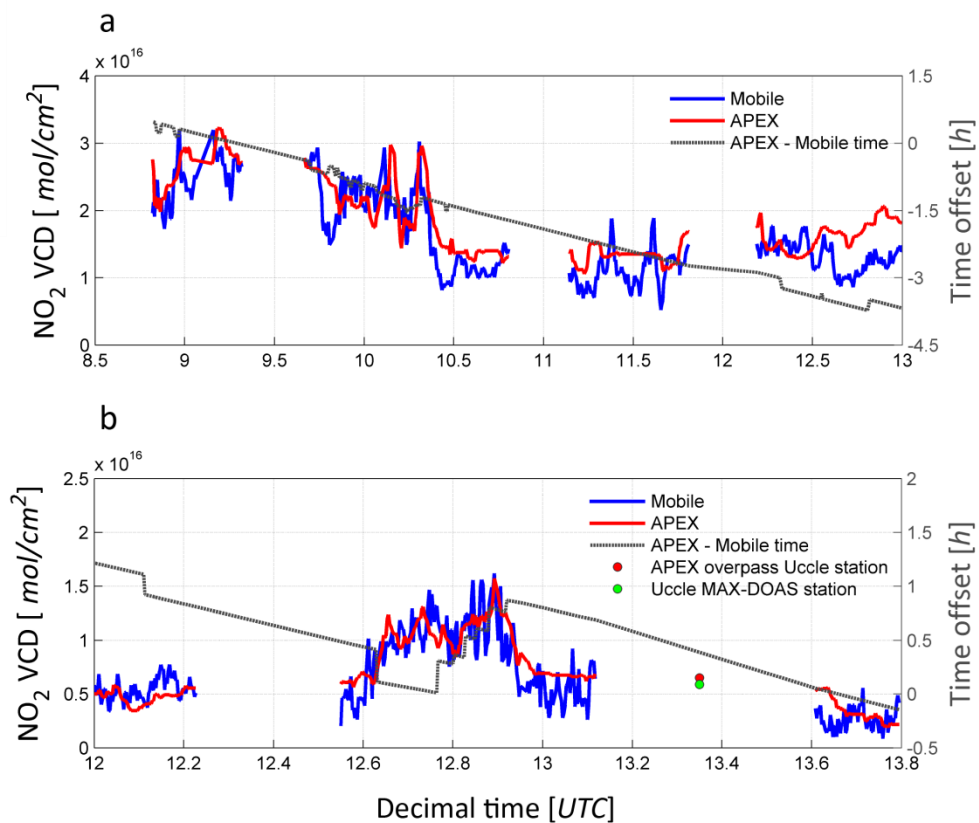
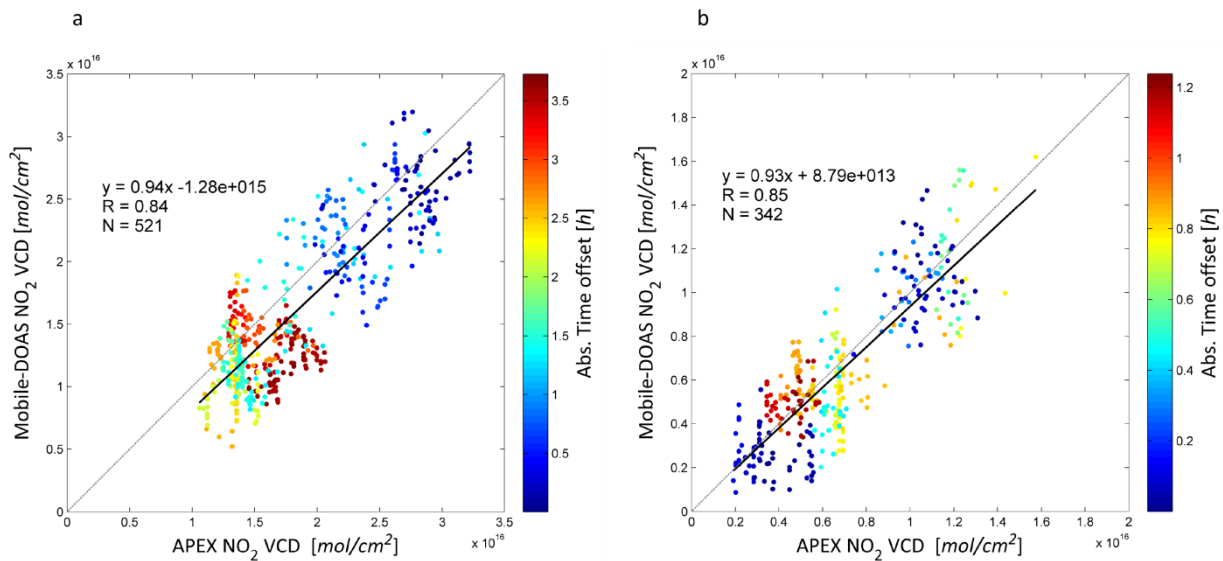


Figure 20. APEX and Mobile-DOAS NO₂ VCD time series for **a**) the Antwerp flight (day 105) and **b**) the Brussels flight (day 181), respectively. The time offset between APEX and mobile-DOAS observations is plotted in dark grey. The NO₂ VCDs, measured by APEX (red dot) and the mini-MAX-DOAS (green dot) at the overpass over the Uccle station (13:21 UTC) are indicated on Fig. 20.b.

5

10



5 **Figure 21.** Scatter plot and linear regression analysis of the position-synchronised NO₂ VCDs, retrieved from APEX and mobile-DOAS for **a)** the Antwerp flight (day 105) and **b)** the Brussels flight (day 181), respectively. The points are color-coded based on the absolute time offset between APEX and mobile-DOAS observations.