

Response to Referee #1:

We thank referee #1 for their helpful comments. Our responses are given below in black with the referee's comments in blue. The new text in the modified manuscript is given in red (italicized).

Referee #1:

General comments

This paper describes the process to obtain the surface NO₂ mixing ratio using the NO₂ column from the ground-based Pandora observation. Some validations are performed by the comparison to the model outputs and in-situ measurement. The approaches look logical and the final products seem reasonable. Considering the rising importance of Pandora data for the air quality monitoring, publication of this work is useful to the research community and potential readers. But some more discussions should be included to improve this manuscript. Please refer to the comments below for the revision process.

Specific comments

P1, L27: Traffic is one of the important NO₂ sources. Please include.

Done.

It is primarily emitted from combustion processes such as fossil fuel combustion (e.g., traffic, electricity generation from power plants) and biomass burning, as well as from lightning. NO₂ is a nitrate aerosol precursor, and it also contributes to acid deposition and eutrophication (ECCC, 2016).

P2 L1-L12: Why is the total vertical column NO₂ significantly treated first instead of insitu measurement. For air quality monitoring, in-situ measurement has been performed basically. Before the importance of Pandora observation is addressed, the necessity of ground-based remote sensing of trace gases should be stated first, in spite of the well-structured ground-based in-situ networks.

A new paragraph has been included in the Introduction to describe the current in situ measurements and the monitoring network (NAPS).

As surface NO₂ concentrations are regulated by many environmental agencies (e.g., Environment and Climate Change Canada and U.S. Environmental Protection Agency), in situ NO₂ measurements are commonly carried out by many national monitoring networks, such as the National Air Pollution Surveillance (NAPS, <https://www.canada.ca/en/environment-climate-change/services/air-pollution/monitoring-networks-data/national-air-pollution-program.html>) network in Canada, which was established in 1969. The in situ methods used to measure surface NO₂ have evolved over the years; for example, luminol chemiluminescence (e.g., Kelly et al., 1990; Maeda et al., 1980; Wendel et al., 1983), long-path differential optical absorption spectroscopy (e.g., Platt, 1994), photolytic

conversion/chemiluminescence (e.g., Gao et al., 1994; Ryerson et al., 2000), and laser-induced fluorescence (e.g., Thornton et al., 2000) are all found to be reliable methods with an uncertainty within 10 % at the 1 ppbv and higher concentration levels (McClenny, 2000). Currently, the in situ approach used by NAPS for surface NO₂ air quality monitoring is the photolytic conversion/chemiluminescence technique, which converts NO₂ to NO and subsequently detects the NO by chemiluminescence reaction (McClenny, 2000; NRC, 1992). This in situ monitoring provides good measurements at ground level (0.4 ppbv accuracy), but NO₂ is not uniformly mixed through the atmosphere, and not even within the atmospheric boundary layer due to emission and removal processes taking place at the surface.

P2 L8-9: The reason to use zenith-sky observations under the cloudy condition is not clear. Why zenith sky mode has more reliable than direct-sun mode under the cloudy condition?

The direct-sun mode requires unobscured sun to make reliable measurements. We have modified the text as follows:

Zenith-sky observations have been widely used for stratospheric ozone and NO₂ observations, particularly under cloudy conditions when direct-sun measurements are unreliable (note that zenith-sky observations use scattered sunlight and are less sensitive to clouds, e.g., Zhao et al. (2019)).

P4 L20-21: Why do we need O₄ retrieval? Still, some potential readers do not have any idea about the meaning of O₄ retrieval.

Some background information about O₄ is included in this new paragraph.

The oxygen collision complex (O₂)₂ (referred here as O₄), which is created by the collision of two oxygen molecules, has broadband absorptions from UV to near IR spectral ranges (Greenblatt et al., 1990; Platt and Stutz, 2008; Thalman and Volkamer, 2013). O₄ is widely used as a reference gas by many DOAS applications to infer cloud and aerosol properties (e.g., Gielen et al., 2014; Wagner et al., 2004, 2014, 2016; Wang et al., 2015; Zhao et al., 2019).

P7 L1-2: In other Pandora paper (usually including Dr. Jay Herman in the author list), diurnal variation of stratospheric NO₂ is not much considered. It will be more interesting (and even useful) to add some arguments compared with their approaches.

In short, there are two major concepts we must consider here, 1) the NO₂ profile (weights between stratospheric and tropospheric NO₂), and 2) the observation geometry (direct-sun or zenith-sky).

To calculate proper AMFs, there are two things that must be considered: the day-night difference due to photochemistry and the morning and evening peaks of NO₂ due to local traffic. As discussed in the paper

(see Fig. A1b), the stratospheric diurnal variation contributes about 0.1 DU difference in total column NO_2 . Thus, this amount is large enough to matter for urban sites such as Toronto, and can be more significant for rural sites. In general, our conclusion is that an urban site with direct-sun observation should have smaller impact from the stratospheric diurnal variation. On the other hand, a rural site with zenith-sky observation should have a significant impact. This information has now been included in Appendix B as follows:

Please note that the strength of this bias is related to 1) the NO_2 profile (weights between stratospheric and tropospheric NO_2), and 2) the observation geometry (direct-sun or zenith-sky). In general, an urban site with direct-sun observation should have smaller impact from the stratospheric diurnal variation. On the other hand, a rural site with zenith-sky observation should have significant impact.

P8 L10-24: Notation is not clear. There are VCD_{DS} and VCE_{ZS} . Also, there are VCD_{EMP} and $\text{VCD}_{\text{NDACC}}$. VCD_{EMP} (and $\text{VCD}_{\text{NDACC}}$) is related to VCD_{DS} or VCD_{ZS} ? Better notation or more description of these notations look required.

VCD_{EMP} and $\text{VCD}_{\text{NDACC}}$ are both VCDs retrieved from zenith-sky observations. To make our notation clearer, we have changed VCD_{EMP} and $\text{VCD}_{\text{NDACC}}$ to $\text{VCD}_{\text{ZS-EMP}}$ and $\text{VCD}_{\text{ZS-NDACC}}$. The corresponding text has been changed as follows:

Total column NO_2 can then be retrieved using Eqn. (1) and these two sets of AMFs, where the one based on empirical AMFs is referred to as $\text{VCD}_{\text{ZS-EMP}}$ and the one based on NDACC AMFs is referred to as $\text{VCD}_{\text{ZS-NDACC}}$. The RCD value used in the retrievals is 0.39 ± 0.01 DU, which is retrieved along with $\text{AMF}_{\text{ZS-EMP}}$ (Appendix A).

In general, the $\text{VCD}_{\text{ZS-EMP}}$ and $\text{VCD}_{\text{ZS-NDACC}}$ performed as expected. Compared with VCD_{DS} , the $\text{VCD}_{\text{ZS-NDACC}}$ shows a -25% bias, while the $\text{VCD}_{\text{ZS-EMP}}$ only shows a -4 % bias (indicated by the red lines on each panel and their slopes). In addition, $\text{VCD}_{\text{ZS-EMP}}$ shows less SZA dependence than $\text{VCD}_{\text{ZS-NDACC}}$ (see the increased bias for measurements made in larger SZA conditions in Figure 2b).

P10 L10-12: V_{ftrop} is from the GEOS-chem simulations. How does GEOS-Chem consider the lightning NO_x in the free-troposphere? In other words, GEOS-Chem can catch the free-tropospheric lightening NO_x reasonably? It seems better to include some discussions about the estimation of lightning NO_x in the free troposphere and even stratosphere, which is related to the final quality of C_{pan} from the Pandora observation.

Lightning NO_x emissions are computed as a function of cloud-top height, and are scaled globally as described by Sauvage et al. (2007) to match Optical Transient Detector/Lightning Imaging Sensor (OTD/LIS) climatological observations of lightning flashes. The global source is imposed to be 6 Tg(N) yr^{-1}

(Martin et al., 2007). Higher NO_x yields per flashes are used at mid-latitudes than in the tropics (Hudman et al., 2007). Different versions of the GEOS-CHEM model have been used extensively in the retrieval of tropospheric VCDs, and have been shown capable of simulating the vertical distributions of NO_2 (e.g. Lamsal et al., 2008; Martin et al., 2002) and SO_2 (e.g. Lee et al., 2009). Some of this information about the estimation of lightning NO_x has now been included in Section 2.2.2.

The model has a detailed representation of tropospheric chemistry, including aerosols and their precursors (Park et al., 2004). In the simulation used in this study, a global lightning NO_x source of 6 Tg N yr^{-1} (Martin et al., 2002) was imposed. Lightning NO_x emissions are computed as a function of cloud-top height, and are scaled globally as described by Sauvage et al. (2007) to match Optical Transient Detector/Lightning Imaging Sensor (OTD/LIS) climatological observations of lightning flashes.

To provide some quality information about the final data product (C_{Pan}), a detailed uncertainty estimation has been provided in a new Appendix D. In general, the contribution from the free tropospheric NO_2 is limited compared to lower tropospheric and stratospheric NO_2 , as suggested by the following paragraph from Appendix D:

The estimated Pandora zenith-sky-based surface NO_2 data have uncertainties from 4.8 to 6.5 ppbv. In Eqn. 8, the contributions of the V_{Pan} , V_{Strat} , V_{ftrop} , and R_{CV} terms to the total uncertainty are 36%, 2%, 0.3%, and 62 %, respectively. This result indicates that the uncertainty in the Pandora zenith-sky-based surface NO_2 is dominated by the uncertainties of Pandora zenith-sky total column NO_2 and modelled column-to-surface conversion ratio (R_{CV}). However, please note that this uncertainty budget depends on the NO_2 vertical distributions, and may vary from site to site; e.g., in Toronto, the tropospheric NO_2 is typically 2-4 times higher than the stratospheric NO_2 , and thus, the contribution to uncertainty from V_{Pan} is much larger than the corresponding contributions from V_{Strat} and V_{ftrop} .

P10 L22-24: How about the year-to-year variation of conversion ratio? Is it larger or smaller than the diurnal variation of conversion ratio? If there is, it seems to include some discussion for how to treat or consider the year-to-year variation.

A year-to-year variation of conversion ratio is expected. However, the magnitude of this variation depends on the conditions of the site (e.g., local emission patterns, year round sunlight period). In other words, as the ratio in the look-up table is calculated using monthly mean (data period: April 2016 to December 2017), it is expected that for different period, this ratio will be different. For example, if the local emission patterns or meteorological conditions have significant changes for a site, the derived conversion ratio should reveal these changes. In general, the look-up table approach ($C_{\text{pan-LUT}}$) is aiming for a quick and near-real-time data delivery. Thus, for a given year, we recommend using a mean PSC-LUT that is calculated from model simulations of previous years. On the other hand, for the off-line data, the $C_{\text{pan-model}}$ is the final, high-quality, year-specific data product that will be delivered to users. This information has now been included in this section:

In general, the look-up table approach ($C_{pan-LUT}$) is aiming for a quick and near-real-time data delivery. Thus, to minimize year-to-year variation (e.g., from changing meteorological conditions or changing local emission patterns), for a given year we recommend using a mean PSC-LUT that is calculated from model simulations of previous years. On the other hand, the $C_{pan-model}$ is the off-line, high-quality, year-specific data product that will be delivered for air quality research and other applications.

P11 L21-28: How to connect the results and discussions here to those in Fig. 4? The lesson from Fig. 4 seems that the ZS Pandora NO₂ has better quality than the direct-sun Pandora NO₂, right? But here both ZS and direct-sun Pandora NO₂ shows consistent quality. Then what do we really learn from Fig. 4 in this work?

Please note the results in Fig. 4 are for total column NO₂, while the discussion in page 11 is about Pandora derived surface NO₂ concentrations. We also found slightly better agreement in Fig. 4 between Pandora zenith-sky NO₂ VCD with OMI. However, due to the limited number of coincident data, we think the zenith-sky and direct-sun Pandora NO₂ are of similar good quality. To further assess the quality, we need a longer period of observation and more coincident measurements. Some possible explanations are now included in Section 3.2.

The better correlation and lower bias for zenith-sky versus direct-sun might be a case of coincident errors, i.e., compare to Pandora direct-sun total column NO₂, both OMI and Pandora zenith-sky total column NO₂ underestimate the local NO₂ at Toronto (see Figure 2). When taking into account the standard error of the fitting and the confidence level of R , the difference between zenith-sky and direct-sun data is not significant (i.e., in Fig. 4 from panels a to d, the slopes with standard error are 0.64 ± 0.02 , 0.67 ± 0.02 , 0.70 ± 0.04 , and 0.71 ± 0.03 ; the 95% confidence intervals for R values are 0.45 to 0.63, 0.61 to 0.75, 0.43 to 0.77, and 0.60 to 0.86).

P12 L3-18 and Fig. 8: In this study, the validation is performed for the 2015-2017 period. Why only a month (APR 2017) is considered in this analysis? The better quality of ZS NO₂ than direct-sun NO₂ can be justified with this figure but whether the ZS NO₂ always provide the reliable NO₂ under the heavy clouds cannot be determined based on only 1-month situation. It seems necessary to deal with additional cloudy cases (at least in the supplements).

Fig. 8 is an example of the time series. It was chosen because of its good mixture of clear, moderate cloud, and heavy cloud conditions. The general performance of cloud filtering was included in Appendix C, which examines the quality of ZS NO₂ in different cloud conditions (categorized by the enhancement of O₄) for the period April 2016 to December 2017. We have included another month (October 2016) with good mixed conditions in Appendix C as another example.

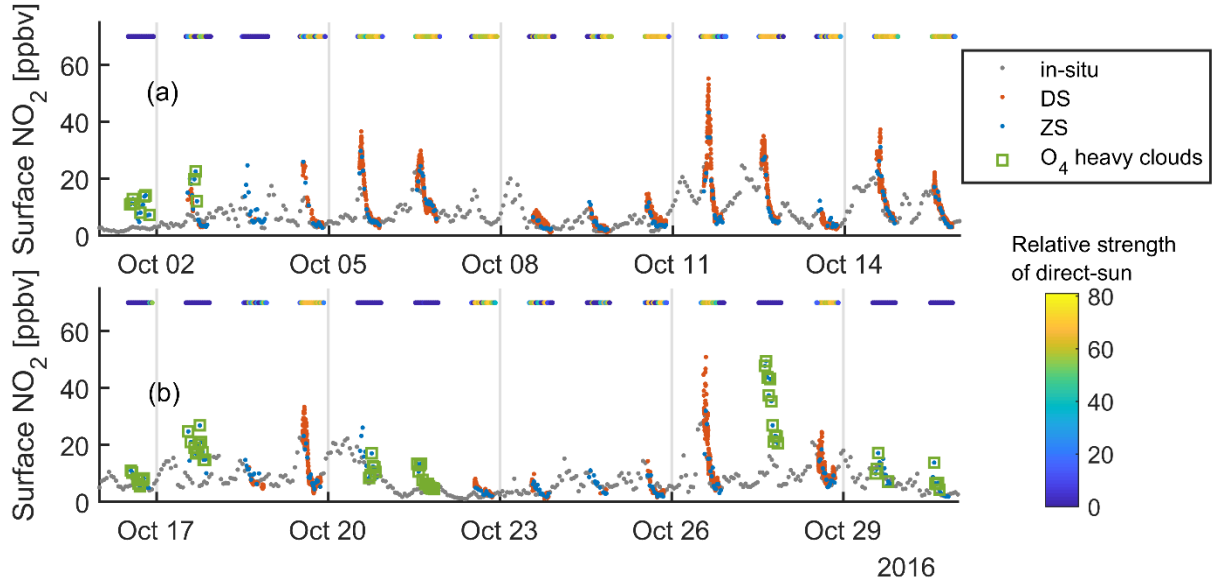
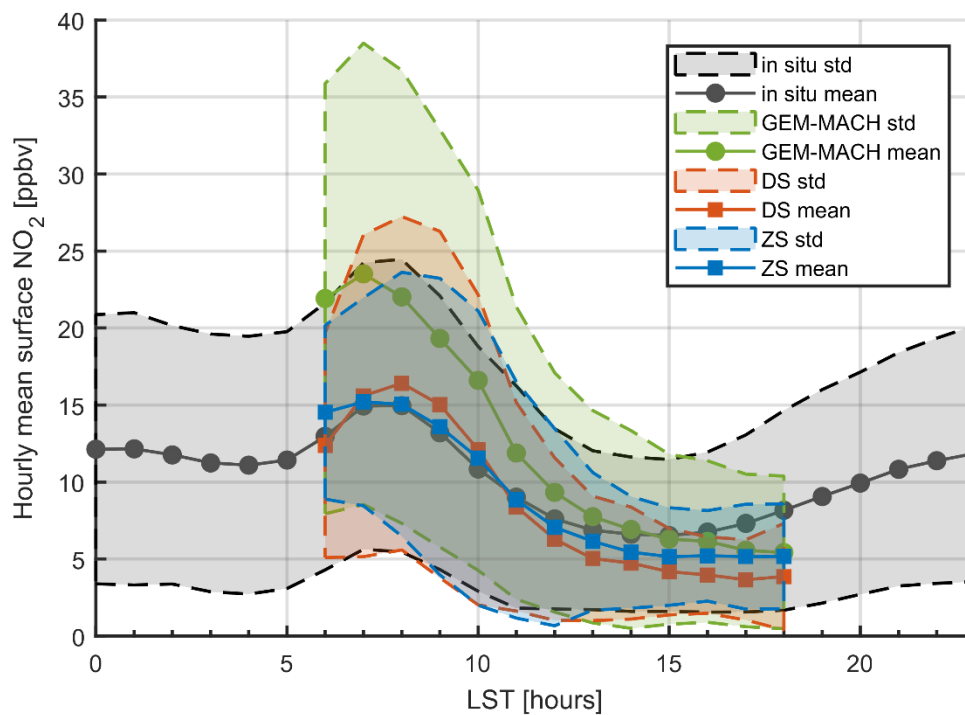


Figure A4. Example of surface NO₂ concentration time series in all conditions. The in situ, Pandora direct-sun (DS), and Pandora zenith-sky (ZS) surface NO₂ concentrations are shown by different coloured dots. The total sky imager relative strength of direct-sun data are plotted as a colour-coded horizontal dot-line on the top area of each panel. For Pandora zenith-sky data, the measurements with enhanced O₄ (heavy cloud indicator) are also labelled by green squares.

P13 L3-L20 and Figs. 9 and 10: In addition to the GEM-MACH model results, Pandora ZS and direct-sun NO₂ products also show the higher standard deviation after sunrise compared to the time close to sunset. Considering the radiative extinction is more disturbed during the time close to sunrise and sunset, the higher standard deviation in the morning is not well understood. Why NO₂ product in the morning has higher biases? This is also due to the propagation of GEM-MACH PBL uncertainty? (But authors mentioned that the surface NO₂ from ZS Pandora observation is less dependent on the PBL height in P13 L29-30).

The standard deviations of Pandora ZS and DS in the morning are comparable to those of the in situ data. In Fig. 9, at 7:00 LST., in situ NO₂ is 14.9 ± 9.3 ppbv, while GEM-MACH, Pandora DS, and Pandora ZS NO₂ are 23.5 ± 15.0 ppbv, 15.6 ± 10.5 ppbv, and 15.2 ± 6.8 ppbv, respectively. To make it more clear, the shaded areas, which represent the 1 σ envelope, have been modified to use dashed lines as boundaries.



For evening conditions, in Fig. 9, at 17:00 LST, in situ NO₂ is 7.3 ± 5.8 ppbv, while GEM-MACH, Pandora DS, and Pandora ZS NO₂ are 5.6 ± 5.0 ppbv, 3.6 ± 2.6 ppbv, and 5.2 ± 3.4 ppbv, respectively. The larger morning standard deviations in all observations and modelled data are related to different NO₂ emission patterns in the mornings (i.e., work-days and weekend). The lower PBL in the mornings also enhanced the impacts from emissions compared to afternoon conditions. This information has now been included in the paragraph:

For example, at 7:00 LST, in situ NO₂ is 14.9 ± 9.3 ppbv, while GEM-MACH, Pandora DS, and Pandora ZS NO₂ are 23.5 ± 15.0 ppbv, 15.6 ± 10.5 ppbv, and 15.2 ± 6.8 ppbv, respectively. At 17:00 LST, in situ NO₂ is 7.3 ± 5.8 ppbv, while GEM-MACH, Pandora DS, and Pandora ZS NO₂ are 5.6 ± 5.0 ppbv, 3.6 ± 2.6 ppbv, and 5.2 ± 3.4 ppbv, respectively. The larger standard deviations in the morning are due to the datasets not being divided into work-days and weekends.

Technical corrections

P6 L2: two times -> twice

Done.

P7 L30-32: The meaning of this statement is not clear to me. Please more clarify the explanation.

It has been modified to be clearer.

*Instead of finding the link between zenith-sky **spectral** intensity and total column **values** (i.e., following the Brewer and Dobson zenith-sky total ozone retrieval method), deriving empirical zenith-sky AMFs for Pandora zenith-sky measurements is more straightforward since Pandora zenith-sky spectra can be analyzed to produce NO₂ dSCDs.*

P8 L10-12: This statement looks redundant, just a repetition of statements above.

The statement in P8 L10-12 is not redundant. The notation of this part has been modified to make this clearer:

*If we make an assumption that the coincident direct-sun (DS) and zenith-sky (ZS) measurements sampled the same air mass, then the empirical zenith-sky AMFs (referred to here as **AMF_{ZS-Emp}**) can be calculated by assuming $VCD_{DS} = VCD_{ZS}$, which gives*

$$VCD_{DS}(SZA) = \frac{dSCD_{ZS}(SZA) + RCD_{ZS}}{AMF_{ZS-Emp}(SZA)}. \quad (2)$$

*Next, we can use nearly-coincident VCD_{DS} and $dSCD_{ZS}$ in a multi-non-linear regression to retrieve **AMF_{ZS-Emp}** and RCD_{ZS} together. To ensure the quality of the retrieved **AMF_{ZS-Emp}**, only high quality direct-sun total column NO₂ data are used with $SZA < 75^\circ$. Details about the empirical zenith-sky AMF calculation are shown in Appendix A.*

*Figure 1 shows a comparison of the empirical zenith-sky AMFs and NDACC AMFs (calculated for the Toronto measurements). Total column NO₂ can then be retrieved using Eqn. (1) and these two sets of AMFs, where the one based on empirical AMFs is referred to as **VCD_{ZS-Emp}** and the one based on NDACC AMFs is referred to as **VCD_{ZS-NDACC}**. The RCD value used in the retrievals is 0.39 ± 0.01 DU, which is retrieved along with **AMF_{ZS-Emp}** (Appendix A).*

P11 L10: This result (slightly lower correlation) -> This slightly lower correlation Figs. 4, 7, and A2: If a slope value is the only significant value, then 'slope = ***' looks better expression. If authors would show the information with the format of 'y = slope x', it seems better to add intercept values together, to have the original regression equation (the equation looks not completed without intercept values)

The fitting that we used here is not a simple linear fitting, but a non-linear fit with intercept set to zero (this information was provided in the captions of Figs. 4, 7, and A2). The reason to force the intercept to be zero is to provide a meaningful value to represent the bias between two datasets. The assumption we made here is that if atmospheric NO₂ is zero, then any instrument should measure zero NO₂. In other words, the assumption is that there should be no additive systematic bias between instruments (i.e., drifting from zero), but only multiplicative systematic bias exist. In reality, although this assumption might not be true, for any instrument that we know, if it has a drift from zero (an additive systematic

bias) we first correct this drifting before using its data. In other words, although the intercept value in simple linear fitting could be an indicator for multiplicative systematic bias between instruments, the value itself depends on which instrument we trust as truth (i.e., which dataset is used as y, and which dataset is used as x). Thus, the intercept is not really a meaningful indicator to show here.

In short, the fittings (intercept forced to be zero) are used to reveal the bias between two datasets, and the equations shown on those figures are complete.

Figure captions: It seems that all figure captions can be more clarified to better describe the figures. Please check and correct once more before the submission of revised manuscript (Now they are somewhat confused. Fig. 6 and 7 are examples. What really figure b and c imply? Direct description of figure a, b, c looks better to read).

The captions have been modified as requested. The implications of Panels (b) and (c) in each figure (Figures 6 and 7) are included at the end of Section 4.2.

Figure 6. Modelled and Pandora zenith-sky surface NO₂ vs. in situ NO₂. (a) shows the GEM-MACH modelled surface NO₂ data vs. in situ NO₂; (b) and (c) show the Pandora zenith-sky (ZS) surface NO₂ data vs. in situ NO₂. The Pandora ZS surface NO₂ data in (b) and (c) are derived using the hourly modelled conversion ratio and the monthly PSC-LUT, respectively. (d) to (f) are histograms corresponding to the data in (a) to (c). On each scatter plot, the red line is the linear fit with intercept set to 0 and the black line is the one-to-one line. The scatter plots are colour-coded by the normalized density of the points.

Figure 7. Modelled and Pandora direct-sun surface NO₂ vs. in situ NO₂. (a) shows the GEM-MACH modelled surface NO₂ data vs. in situ NO₂; (b) and (c) show the Pandora direct-sun (DS) surface NO₂ data vs. in situ NO₂. The Pandora DS surface NO₂ data in (b) and (c) are derived using the hourly modelled conversion ratio and the monthly PSC-LUT, respectively. (d) to (f) are histograms corresponding to the data in (a) to (c). On each scatter plot, the red line is the linear fit with intercept set to 0 and the black line is the one-to-one line. The scatter plots are colour-coded by the normalized density of the points.

The good consistency between $C_{pan-model}$ and $C_{pan-LUT}$ implies that two versions of Pandora surface NO₂ data can be delivered in the future, i.e., an off-line version that relies on hourly inputs from the model, and a near-real-time version that only needs a pre-calculated LUT.

Time period: Please clearly suggest the time period of each analysis (at least in each figure caption). I think each analysis has a different time period of analysis, which is a little complex to me. While I can accept this difference, the clear time information should be provided to help the readers' understanding.

Done. Captions of Figures 2, 4, and 6 to 10 are modified to include the period of each analysis.

Reference

ECCC: Canadian Environmental Sustainability Indicators: Air Quality. Environment and Climate Change Canada, ISBN: 978-0-660-06016-3, 2016.

Gao, R. S., Keim, E. R., Woodbridge, E. L., Ciciora, S. J., Proffitt, M. H., Thompson, T. L., McLaughlin, R. J. and Fahey, D. W.: New photolysis system for NO₂ measurements in the lower stratosphere, *J Geophys Res-Atmos*, 99(D10), 20673–20681, doi:10.1029/94JD01521, 1994.

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.*, 7, 3509–3527, doi:10.5194/amt-7-3509-2014, 2014.

Greenblatt, G. D., Orlando, J. J., Burkholder, J. B. and Ravishankara, A. R.: Absorption measurements of oxygen between 330 and 1140 nm, *J. Geophys. Res.*, 95, 18577–18582, doi:10.1029/JD095iD11p18577, 1990.

Hudman, R. C., Jacob, D. J., Turquety, S., Leibensperger, E. M., Murray, L. T., Wu, S., Gilliland, A. B., Avery, M., Bertram, T. H., Brune, W., Cohen, R. C., Dibb, J. E., Flocke, F. M., Fried, A., Holloway, J., Neuman, J. A., Orville, R., Perring, A., Ren, X., Sachse, G. W., Singh, H. B., Swanson, A. and Wooldridge, P. J.: Surface and lightning sources of nitrogen oxides over the United States: Magnitudes, chemical evolution, and outflow, *J Geophys Res-Atmos*, 112(D12), doi:10.1029/2006JD007912, 2007.

Kelly, T. J., Spicer, C. W. and Ward, G. F.: An assessment of the luminol chemiluminescence technique for measurement of NO₂ in ambient air, *Atmos. Environ.*, 24(9), 2397–2403, doi:10.1016/0960-1686(90)90332-H, 1990.

Lamsal, L. N., Martin, R. V., Donkelaar, A. van, Steinbacher, M., Celarier, E. A., Bucsela, E., Dunlea, E. J. and Pinto, J. P.: Ground-level nitrogen dioxide concentrations inferred from the satellite-borne Ozone Monitoring Instrument, *J. Geophys. Res.*, 113(D16), doi:10.1029/2007JD009235, 2008.

Lee, C., Martin, R. V., Donkelaar, A. van, O’Byrne, G., Krotkov, N., Richter, A., Huey, L. G. and Holloway, J. S.: Retrieval of vertical columns of sulfur dioxide from SCIAMACHY and OMI: Air mass factor algorithm development, validation, and error analysis, *J Geophys Res-Atmos*, 114(D22), doi:10.1029/2009JD012123, 2009.

Maeda, Yasuaki, Aoki, Kenji. and Munemori, Makoto.: Chemiluminescence method for the determination of nitrogen dioxide, *Anal. Chem.*, 52(2), 307–311, doi:10.1021/ac50052a022, 1980.

Martin, R. V., Chance, K., Jacob, D. J., Kurosu, T. P., Spurr, R. J. D., Bucsela, E., Gleason, J. F., Palmer, P. I., Bey, I., Fiore, A. M., Li, Q., Yantosca, R. M. and Koelemeijer, R. B. A.: An improved retrieval of tropospheric nitrogen dioxide from GOME, *J. Geophys. Res.*, 107(D20), ACH 9-1-ACH 9-21, doi:10.1029/2001JD001027, 2002.

- Martin, R. V., Sauvage, B., Folkins, I., Sioris, C. E., Boone, C., Bernath, P. and Ziemke, J.: Space-based constraints on the production of nitric oxide by lightning, *J Geophys Res-Atmos*, 112(D9), doi:10.1029/2006JD007831, 2007.
- McClenny, W. A.: Recommended methods for ambient air monitoring of NO, NO₂, NO_y and individual NO_z species, Tech. rep., US Environmental Protection Agency, Triangle Park, NC 27711., 2000.
- NRC: Rethinking the Ozone Problem in Urban and Regional Air Pollution, National Research Council, National Academy Press, Washington, D.C., 1992.
- Park, R. J., Jacob, D. J., Field, B. D., Yantosca, R. M. and Chin, M.: Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy, *J. Geophys. Res.*, 109(D15), doi:10.1029/2003JD004473, 2004.
- Platt, U.: Differential optical absorption spectroscopy (DOAS), in *Air Monitoring by Spectroscopic Techniques*, pp. 27–84, John Wiley, New York., 1994.
- Platt, U. and Stutz, J.: *Differential Optical Absorption Spectroscopy: Principles and Applications*, Springer, Germany, Berlin., 2008.
- Ryerson, T. B., Williams, E. J. and Fehsenfeld, F. C.: An efficient photolysis system for fast-response NO₂ measurements, *J. Geophys. Res.*, 105(D21), 26447–26461, doi:10.1029/2000JD900389, 2000.
- Sauvage, B., Martin, R. V., Donkelaar, A. van, Liu, X., Chance, K., Jaeglé, L., Palmer, P. I., Wu, S. and Fu, T.-M.: Remote sensed and in situ constraints on processes affecting tropical tropospheric ozone, *Atmos. Chem. Phys.*, 7(3), 815–838, doi:10.5194/acp-7-815-2007, 2007.
- Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O₂-O₂ collision pairs between 340 and 630 nm and at atmospherically relevant pressure, *Phys. Chem. Chem. Phys.*, 15, 15371–15381, 2013.
- Thornton, J. A., Wooldridge, P. J. and Cohen, R. C.: Atmospheric NO₂: In Situ Laser-Induced Fluorescence Detection at Parts per Trillion Mixing Ratios, *Anal. Chem.*, 72(3), 528–539, doi:10.1021/ac9908905, 2000.
- Wagner, T., Dix, B., Friedeburg, C. v, Frieß, U., Sanghavi, S., Sinreich, R. and Platt, U.: MAX-DOAS O₄ measurements: A new technique to derive information on atmospheric aerosols: Principles and information content, *J. Geophys. Res.*, 109, D22205, doi:10.1029/2004jd004904, 2004.
- Wagner, T., Apituley, A., Beirle, S., Dörner, S., Friess, U., Remmers, J. and Shaiganfar, R.: Cloud detection and classification based on MAX-DOAS observations, *Atmos. Meas. Tech.*, 7, 1289–1320, doi:10.5194/amt-7-1289-2014, 2014.
- Wagner, T., Beirle, S., Remmers, J., Shaiganfar, R. and Wang, Y.: Absolute calibration of the colour index and O₄ absorption derived from Multi AXis (MAX-)DOAS measurements and their application to a standardised cloud classification algorithm, *Atmos. Meas. Tech.*, 9, 4803–4823, doi:10.5194/amt-9-4803-2016, 2016.

Wang, Y., Penning de Vries, M., Xie, P. H., Beirle, S., Dörner, S., Remmers, J., Li, A. and Wagner, T.: Cloud and aerosol classification for 2.5 years of MAX-DOAS observations in Wuxi (China) and comparison to independent data sets, *Atmos. Meas. Tech.*, 8, 5133–5156, doi:10.5194/amt-8-5133-2015, 2015.

Wendel, G. J., Stedman, D. H., Cantrell, C. A. and Damrauer, Lenore.: Luminol-based nitrogen dioxide detector, *Anal. Chem.*, 55(6), 937–940, doi:10.1021/ac00257a027, 1983.

Zhao, X., Bognar, K., Fioletov, V., Pazmino, A., Goutail, F., Millán, L., Manney, G., Adams, C. and Strong, K.: Assessing the impact of clouds on ground-based UV–visible total column ozone measurements in the high Arctic, *Atmos. Meas. Tech.*, 12(4), 2463–2483, doi:10.5194/amt-12-2463-2019, 2019.

Response to Referee #2:

We thank referee #2 for their helpful comments. Our responses are given below in black with the referee's comments in blue. The new text in the modified manuscript is given in red (italicized).

Referee #2:

General comments

This study presents a new approach to extend direct-sun NO₂ measurements from Pandora instruments with a zenith-sky mode also applicable under cloudy conditions. In addition, attempts are made to also derive surface concentration estimates from total column observations. The general methodology is strongly inspired from the empirical zenith-sky mode developed for total ozone measurements by Dobson and Brewer spectrophotometers. Although the adopted approach implies many approximations not always well described or even identified (see detailed comments below), results are surprisingly good and certainly of interest for the ACP readership. I found the manuscript well written, concise and easy to read; also figures are of good quality and adequate in number and the appendices provide useful additional information. With one exception, credit to existing literature is appropriate. I therefore recommend publication in ACP, after careful consideration for the comments and suggestions below.

Specific comments

Section 2.1.2, L. 20: what was the temperature used for the NO₂ cross-section in the zenith-sky QDOAS retrievals? Is it consistent with the effective temperature assumed for the direct-sun retrieval (254.5 K)?

The effective temperatures for NO₂ and ozone have now been included in the manuscript. The temperature used for the ZS NO₂ was 254.5 K, which is consistent with the effective temperature used for the direct-sun retrieval.

Cross sections of NO₂ at an effective temperature of 254.5 K (Vandaele et al., 1998), ozone at an effective temperature of 223 K (Bogumil et al., 2003), H₂O (Rothman et al., 2005), O₄ (Hermans et al., 2003), and Ring (Chance and Spurr, 1997) are all fitted; a fifth-order polynomial and a first-order linear offset are also included in the DOAS analysis.

Section 2.1.2, L. 25: how was the NO₂ residual amount in the reference spectrum determined here? Generally speaking, this paper lacks a proper analysis of the uncertainties. It would be useful to add a section describing the estimated uncertainties for the zenith-sky column and surface concentrations (which are new data products introduced in this study).

The residual amount in the reference spectrum (RCD) was determined by using the proposed multi-non-linear regressions (see Appendix A). In short, the RCD was retrieved along with the empirical AMFs.

From the multi-non-linear model, the standard fitting error of RCD is 0.01 DU. This information was included in Section 3.1.

The RCD value used in the retrievals is 0.39 ± 0.01 DU, which is retrieved along with AMF_{ZS-Emp} (Appendix A).

In addition, as suggested by the referee the following new section (Appendix D) has been included that describes the estimated uncertainties for the zenith-sky column and surface concentrations:

D. Uncertainty estimation

The uncertainties of retrieved Pandora zenith-sky NO_2 data products (total column and surface concentration) are estimated and discussed here to assess the quality of the data products. The uncertainties of total column and surface concentrations are estimated first using the uncertainty propagation method (referred to here as the UP method) based on Eqns. 2 and 3. The combined uncertainties of total column can be calculated using:

$$\sigma_{VCD_{ZS}} = \sqrt{\left(\frac{\sigma_{dSCD}}{AMF}\right)^2 + \left(\frac{\sigma_{RCD}}{AMF}\right)^2 + \left(\frac{\sigma_{AMF \times SCD}}{AMF^2}\right)^2}, \quad (7)$$

where σ_{dSCD} is the statistical uncertainty on the DOAS fit (output of QDOAS) and σ_{RCD} and σ_{AMF} are the estimated statistical uncertainties using standard errors of the RCD and the zenith-sky empirical AMF regression, respectively (Eqn. 4). To estimate the upper limit of the nominal uncertainty, AMF and SCD are used as median and maximum values in the dataset, respectively.

The combined uncertainties of the surface concentration can be calculated using:

$$\sigma_{C_{Pan}} = \sqrt{\left(R_{CV}\sigma_{V_{Pan}}\right)^2 + \left(R_{CV}\sigma_{V_{strat}}\right)^2 + \left(R_{CV}\sigma_{V_{ftrop}}\right)^2 + \left(V_{Pan} - V_{strat} - V_{ftrop}\right)^2 \sigma_R^2}, \quad (8)$$

where $\sigma_{V_{Pan}}$ is the uncertainty of Pandora zenith-sky total column NO_2 , (here we use the derived σ_{VCD} in Eqn. 7), $\sigma_{V_{strat}}$ is the uncertainty of the stratospheric NO_2 column (estimated using the 1-sigma standard deviation of the V_{strat}), $\sigma_{V_{ftrop}}$ is the uncertainty of the free troposphere NO_2 column (estimated using the 1-sigma standard deviation of the V_{ftrop}). R_{CV} is the GEM-MACH calculated surface VMR to PBL column ratio, and σ_R is the uncertainty of that ratio (estimated using the 1-sigma standard deviation of the R_{CV}). The means of R_{CV} , V_{Pan} , V_{strat} , and V_{ftrop} are used in the uncertainty estimation.

Besides the UP method, another simple approach to estimate uncertainty is to compare the data product with another high-quality (lower uncertainty) coincident data. For example, if we assume that the Pandora direct-sun total column NO_2 data can represent the true value, we can estimate the uncertainty of Pandora zenith-sky total column NO_2 by calculating the 1-sigma standard deviation of their difference (referred to here as the SDD method):

$$\sigma_{VCD_{ZS}} = \sigma(VCD_{DS} - VCD_{ZS}). \quad (9)$$

Similarly, if we assume that the in situ surface NO_2 VMR can represent the true value, the uncertainty of Pandora zenith-sky-based surface NO_2 VMR can be given by:

$$\sigma_{C_{Pan}} = \sigma(C_{insitu} - C_{pan}). \quad (10)$$

Also, if there is systematic bias between the two datasets, it can be removed and the random uncertainty can be calculated by:

$$\sigma_{VCD_{ZS}} = \sigma(VCD_{DS} - k_1 VCD_{ZS}), \quad (11)$$

$$\sigma_{C_{Pan}} = \sigma(C_{insitu} - k_2 C_{pan}), \quad (12)$$

where k_1 and k_2 are the slopes in the linear fits with intercept set to zero (e.g., slopes in Figs. 2 and 6). This method is referred to here as the unbiased SDD. These three uncertainty estimation methods (UP, SDD, and unbiased SDD) were all implemented, and the results are summarized in Table A1. The results show that Pandora zenith-sky total column NO_2 data have a 0.09-0.12 DU uncertainty that is about twice to the Pandora direct-sun total column nominal accuracy (0.05 DU, at 1-sigma level). When using the UP method, for the worst-case scenario, the Pandora zenith-sky total column NO_2 have a 0.17 DU uncertainty (i.e. using minimum of AMFs to estimate the upper limit of uncertainty). The estimated Pandora zenith-sky-based surface NO_2 VMR data have uncertainties from 4.8 to 6.5 ppbv. In Eqn. 8, the contributions of the V_{Pan} , V_{Strat} , V_{ftrop} , and R_{CV} terms to the total uncertainty are 36%, 2%, 0.3%, and 62 %, respectively. This result indicates that the uncertainty in the Pandora zenith-sky-based surface NO_2 VMR is dominated by the uncertainties of Pandora zenith-sky total column NO_2 and the modelled column-to-surface conversion ratio (R_{CV}). However, note that this uncertainty budget depends on the NO_2 vertical distributions, and hence may vary from site to site; e.g., in Toronto, tropospheric column NO_2 is typically 2-4 times higher than stratospheric column NO_2 , and thus, the contribution to uncertainty from V_{Pan} is much larger than the corresponding contributions from V_{Strat} and V_{ftrop} . In addition, the uncertainty of Pandora direct-sun surface NO_2 VMR is also estimated and provided in Table 1. It shows slightly better results than for zenith-sky-based surface NO_2 VMR.

Table 1. Estimated uncertainties for Pandora zenith-sky total column and surface NO_2 .

Estimation method	$\sigma_{VCD_{ZS}}$ (DU)	$\sigma_{C_{Pan-ZS}}$ (ppbv)	$\sigma_{C_{Pan-DS}}$ (ppbv)
UP	0.12	6.5	5.4
SDD	0.09	5.1	5.0
unbiased SDD	0.09	4.8	4.8

Section 3.1: The approach introduced for the zenith-sky AMF calculation is fully empirical and strongly inspired from the zenith-sky measurement mode used with Dobson and Brewer total ozone spectrophotometers. Basically the idea is to use simultaneous direct-sun and zenith-sky measurements to infer effective AMFs for the zenith-sky geometry. It is then assumed that the established relationship remains valid under moderately cloudy conditions (O4 is used to exclude thick diffusing clouds). I first note that the authors do not refer to the AMT publication by Tack et al. (2015) (<https://www.atmosmeas-tech.net/8/2417/2015/>) where a more physical approach to derive total and

tropospheric NO₂ columns from zenith-sky measurements is described. Second, I see a major drawback in the empirical approach used here, which is that the total AMF for zenith-sky measurements is expected to be a strong function of not only the solar zenith angle but also the tropospheric column itself. In first approximation, one can assume that the stratospheric AMF will mostly follow the solar geometry (geometrical AMF) while the PBL AMF is approximately constant and close to one at any solar position. In consequence, for intermediate and low sun conditions, the stratospheric and PBL AMFs differ quite strongly and the total AMF depends on the relative amount of NO₂ present in the PBL and in the stratosphere. I think that the classification could be improved substantially by taking this dependence into account (probably within an iterative scheme). The dependence on the season accounts somehow for this effect (since it implicitly accounts for the seasonality of the stratospheric NO₂ column), but only in a very crude way.

The information from Tack et al. (2015) has now been included in Section 3.1 and this publication has been included in the References section:

In Tack et al. (2015), a more sophisticated four-step approach to derive total and tropospheric NO₂ columns from zenith-sky measurements was proposed, which involves using a RTM to calculate appropriate tropospheric AMFs. However, due to benefits from using the high-quality Pandora direct-sun total column NO₂ measurements, this work took a different but simple and robust approach to derive zenith-sky total column NO₂.

We thank the referee for their insightful suggestions about improving the empirical AMF calculations. The current empirical AMFs are limited to high and intermediate sun conditions (i.e., SZA < 75°). In the future, when deriving low-sun empirical ZS AMFs, we will try the suggested iterative scheme to account for the stronger influence of PBL NO₂. We think this suggestion is valuable, and we have now included this information in the manuscript (Appendix A) as follows:

In addition, the current empirical AMFs are limited to high and intermediate sun conditions (i.e., SZA < 75°). For low-sun conditions, the total AMF for zenith-sky measurements is expected to be a strong function of not only the SZA, but also the tropospheric column itself. Thus, for future work to derive low-sun empirical zenith-sky AMFs, the stronger influence of PBL NO₂ has to be accounted for (i.e., the geometry form AMFs are not enough).

Section 3.1, L. 19: “ : : VCD_Emp shows less SZA dependence than VCD_DS : : “ -> I think that VCD_NDACC is meant here instead of VCD_DS

Corrected.

In addition, VCD_{Emp} shows less SZA dependence than VCD_{NDACC} (see the increased bias for measurements made in larger SZA conditions in Figure 2b).

Section 3.1, last paragraph: note that the zenith-sky AMF could also be affected by aerosols present in PBL (together with NO₂), due to their impact on the light path. Although the impact of aerosols is likely to be moderate, it certainly contribute to the uncertainty of the measurements and this should be mentioned.

Thank you for this comment. The following new text has been included in Section 3.1, last paragraph:

The derived zenith-sky total column NO₂ values are affected by both clouds and aerosols due to their impact on the light path. The presence of clouds and aerosols contributes to the uncertainty of the measurements. However, the impact of aerosols is expected to be moderate in most cases compared to that of clouds (e.g., Hendrick et al., 2011; Tack et al., 2015). Thus, this work has focused on evaluating the impact from clouds.

Section 3.2, L. 5-10: the finding that zenith-sky columns despite their larger uncertainties (in comparison to direct-sun data) show a better agreement with satellite measurements is quite surprising and interesting. I am not convinced by the argument stating that the air mass sampled by zenith-sky measurements is more representative of the air mass sampled by the satellite than the direct-sun. Considering the size of typical OMI pixels (approx. 20x20 km²), one can argue that both direct-sun and zenith-sky measurements are more local in nature, and therefore maybe another explanation can be found. Could it be that direct-sun and zenith-sky measurements sample different meteorological conditions (e.g. different wind patterns), or maybe that zenith-sky are generally more homogeneously distributed around the overpass time of the satellite so that the average value becomes more representative? Please comment on these issues.

We were not able to find a solid reason for this finding (zenith-sky measurements have better agreement with OMI). The measurement frequency was given in Section 2.1.1 and the coincidence criteria were provided in Section 3.2. We agree with the referee that both direct-sun and zenith-sky measurements are more local when compared with OMI measurements. However, with the coincidence criteria used (± 30 min), we are not sure that meteorological conditions are the key factor. Also, we did not average the ground-based data. The nearest (in time) measurement that was within ± 30 min of OMI overpass time was used (see Section 3.2). Therefore, it is not an averaging issue. In addition, when taking the standard error of the fitting and the confidence level of R into account, the difference between zenith-sky and direct-sun data is not significant (i.e., in Fig. 4 from panels a to d, the slopes with standard error are 0.64 ± 0.02 , 0.67 ± 0.02 , 0.70 ± 0.04 , and 0.71 ± 0.03 ; the 95% confidence interval for R values are 0.45 to 0.63, 0.61 to 0.75, 0.43 to 0.77, and 0.60 to 0.86). In general, we think it is probably a case of coincident error, i.e., compared to Pandora DS, both OMI and Pandora ZS underestimate the local NO₂ at Toronto (see Section 3.2 and Figure 2). More discussion of this issue has now been included in Section 3.2:

The better correlation and lower bias for zenith-sky versus direct-sun might be a case of coincident error, i.e., compared to Pandora direct-sun, both OMI and Pandora zenith-sky total column NO₂ underestimate

the local NO₂ at Toronto (see Figure 2). When taking into account the standard error of the fitting and the confidence level of R, the difference between zenith-sky and direct-sun data is not significant (i.e., in Fig. 4 from panesl a to d, the slopes with standard error are 0.64 ± 0.02 , 0.67 ± 0.02 , 0.70 ± 0.04 , and 0.71 ± 0.03 ; the 95% confidence intervals for R values are 0.45 to 0.63, 0.61 to 0.75, 0.43 to 0.77, and 0.60 to 0.86).

Section 4.1: the method used to convert NO₂ column measurements into surface concentrations implies a lot of approximations/assumptions. The uncertainties associated to these assumptions should be better described. Equation 3 starts from the total NO₂. For the zenith-sky case, this column already contains quite a large uncertainty (cf. previous point). Then a correction for the stratospheric column and the free tropospheric column is made. The stratospheric column is taken from photochemically-corrected stratospheric NO₂ OMI measurements without any further verification. Can we exclude any possible systematic bias between OMI and ground-based measurements? Was there any attempt to verify that OMI and Pandora measurements do agree well under clean conditions? Also monthly mean OMI data are used. Can we safely neglect the day-to-day variability in the stratospheric NO₂ content (e.g. in Spring it is known that transport patterns can produce short term variations of the stratospheric NO₂)?

There could be a systematic bias between OMI and ground-based data but most likely the bias found at Toronto is related to tropospheric NO₂. Thus, we decided to assess the OMI stratospheric NO₂ data and used it in the algorithm. In this work, we tested the algorithm using both OMI daily and monthly mean stratospheric NO₂, and we did not find any significant differences in the derived surface NO₂. The reason is that for Toronto, tropospheric NO₂ accounts for 73 % of the total column amounts on average (see Section 3.1) around local noon. During rush hours, the percentage of tropospheric NO₂ should be even higher. Also, when using OMI daily stratospheric NO₂, we need to interpolate the data for cloud days, which will also increase uncertainty. One of the reasons that we decided to use monthly mean OMI stratospheric NO₂ was also to reveal the robustness of the method (i.e., not highly dependent on satellite measurements). Some of this information has now been included in Appendix B:

Note that the strength of this bias is related to 1) the NO₂ profile (weights between stratospheric and tropospheric NO₂), and 2) the observation geometry (direct-sun or zenith-sky). In general, an urban site with direct-sun observation should have less impact from the stratospheric diurnal variation. On the other hand, a rural site with zenith-sky observation should have significant impact.

As regards the free-tropospheric NO₂ content, it is taken directly from model data. How large and uncertain is this contribution? Finally the last step is based on the assumption that the modelled column to surface concentration ratio is representative of the actual ratio. Is this assumption expected to be correct in all cases? What about the impact of the limited horizontal resolution of the model? In any case, I think that all these uncertainties are responsible for the large scatter of the correlations shown in Figure 6. As such they should be discussed with a little bit more of attention. This being said, I agree that

the average behaviors found (and illustrated in Figs. 9-11) are quite convincing and demonstrate well the potential of the data set.

The current method is sophisticated in that outputs from several different chemical transport models (with different strengths) were included. Currently, we are working to provide estimate of uncertainties of the final data products (Pandora surface NO₂). Based on the Toronto datasets, the mean of free tropospheric NO₂ is 0.03 ± 0.01 DU (mean $\pm 1\sigma$) (GEOS-Chem), while the boundary layer NO₂ is 0.37 ± 0.29 DU (GEM-MACH), and stratospheric NO₂ is 0.10 ± 0.02 DU (OMI monthly mean stratospheric NO₂, with Pratmo to account for diurnal variation). Thus, the uncertainty contributions from GEOS-Chem, OMI, and Pratmo are much less than the uncertainty contribution from GEM-MACH. An uncertainty estimation section is now included as Appendix D (see the reply to previous comments), which addresses the uncertainties associated with the model inputs.

The assumption that the modelled column to surface concentration ratio is representative of the actual ratio has challenges, especially for shallow boundary layer conditions (e.g., Figure 11). However, we think the model can better represent the real conditions for most cases (taking emissions, atmospheric dynamics, and chemistry into account), when comparing to other existing methods (which simply assume that NO₂ is uniformly mixed through the entire PBL). The current GEM-MACH operational forecast model has spatial resolution of 10 km \times 10 km (information was provided in Section 2.2.1). Based on our current results, the model shows some bias compared to in situ data, which could have contributions from the spatial resolution.

Finally this study makes use of Pandora direct-sun completed by zenith-sky measurements. At no point in the paper, the potential of extending the data set with multi-axis measurements providing more information on the tropospheric NO₂ is mentioned, although this would be a logical evolution for a follow up activity. Please consider adding this in the perspectives.

Thank you for this comment. Text about the potential of Pandora multi-axis measurements has been added to the Conclusions section.

Currently, the standard Pandora observation schedule includes direct-sun, zenith-sky, and multi-axis scanning measurements (i.e., measuring at multiple viewing angles). At present, multi-axis measurement algorithms are still under development, but in the future, by using the multi-axis measurements and optimal estimation techniques (e.g., Rodgers, 2000) or the five angles O2O2-ratio algorithm (Cede, 2019), it may be possible for Pandora measurements to be used to derive NO₂ tropospheric profiles and columns.

Reference

Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O. C., Vogel, A., Hartmann, M., Kromminga, H., Bovensmann, H., Frerick, J. and Burrows, J. P.: Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: instrument characterization and reference data for atmospheric remote-sensing in the 230-2380 nm region, *J. Photochem. Photobiol. A*, 157, 167–184, doi:10.1016/s1010-6030(03)00062-5, 2003.

Cede, A.: Manual for Blick Software Suite 1.6, 2019.

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. Opt.*, 36, 5224–5230, doi:10.1364/AO.36.005224, 1997.

Hendrick, F., Pommereau, J. P., Goutail, F., Evans, R. D., Ionov, D., Pazmino, A., Kyrö, E., Held, G., Eriksen, P., Dorokhov, V., Gil, M. and Van Roozendaal, M.: NDACC/SAOZ UV-visible total ozone measurements: improved retrieval and comparison with correlative ground-based and satellite observations, *Atmos. Chem. Phys.*, 11, 5975–5995, doi:10.5194/acp-11-5975-2011, 2011.

Hermans, C., Vandaele, A. C., Fally, S., Carleer, M., Colin, R., Coquart, B., Jenouvrier, A. and Merienne, M.-F.: Absorption cross-section of the collision-induced bands of oxygen from the UV to the NIR, in *Weakly Interacting Molecular Pairs: Unconventional Absorbers of Radiation in the Atmosphere*, edited by C. Camy-Peyret and A. A. Vigasin, pp. 193–202, Springer, Germany., 2003.

Rodgers, C. D.: *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific, Singapore., 2000.

Rothman, L. S., Jacquemart, D., Barbe, A., Chris Benner, D., Birk, M., Brown, L. R., Carleer, M. R., Chackerian, J. C., Chance, K., Coudert, L. H., Dana, V., Devi, V. M., Flaud, J. M., Gamache, R. R., Goldman, A., Hartmann, J. M., Jucks, K. W., Maki, A. G., Mandin, J. Y., Massie, S. T., Orphal, J., Perrin, A., Rinsland, C. P., Smith, M. A. H., Tennyson, J., Tolchenov, R. N., Toth, R. A., Vander Auwera, J., Varanasi, P. and Wagner, G.: The HITRAN 2004 molecular spectroscopic database, *J. Quant. Spectrosc. Radiat. Transfer*, 96, 139–204, doi:10.1016/j.jqsrt.2004.10.008, 2005.

Tack, F., Hendrick, F., Goutail, F., Fayt, C., Merlaud, A., Pinardi, G., Hermans, C., Pommereau, J.-P. and Van Roozendaal, M.: Tropospheric nitrogen dioxide column retrieval from ground-based zenith–sky DOAS observations, *Atmos. Meas. Tech.*, 8(6), 2417–2435, doi:10.5194/amt-8-2417-2015, 2015.

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 42 000 cm⁻¹ to 10 000 cm⁻¹ (238 -1000 nm) at 220 K and 294 K, *J. Quant. Spectrosc. Radiat. Transfer*, 59, 171–184, doi:10.1016/s0022-4073(97)00168-4, 1998.

Retrieval of total column and surface NO₂ from Pandora zenith-sky measurements

Xiaoyi Zhao¹, Debora Griffin¹, Vitali Fioletov¹, Chris McLinden¹, Jonathan Davies¹, Akira Ogyu¹, Sum Chi Lee¹, Alexandru Lupu¹, Michael D. Moran¹, Alexander Cede^{2,3}, Martin Tiefengraber^{3,4}, Moritz Müller^{3,4}

¹Air Quality Research Division, Environment and Climate Change Canada, Toronto, M3H 5T4, Canada.

²NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.

³LuftBlick, Kreith 39A, 6162 Mutter, Austria.

⁴Department of Atmospheric and Cryospheric Sciences, University of Innsbruck, Innsbruck, Austria.

Correspondence to: Xiaoyi Zhao (xiaoyi.zhao@canada.ca)

Abstract. Pandora spectrometers can retrieve nitrogen dioxide (NO₂) vertical column densities (VCDs) via two viewing geometries: direct-sun and zenith-sky. The direct-sun NO₂ VCD measurements have high quality (0.1 DU accuracy in clear-sky conditions) and do not rely on any radiative transfer model to calculate air mass factors (AMFs); however, they are not available when the sun is obscured by clouds. To perform NO₂ measurements in cloudy conditions, a simple but robust NO₂ retrieval algorithm is developed for Pandora zenith-sky measurements. This algorithm derives empirical zenith-sky NO₂ AMFs from coincident high-quality direct-sun NO₂ observations. Moreover, the retrieved Pandora zenith-sky NO₂ VCD data are converted to surface NO₂ concentrations with a scaling algorithm that uses chemical-transport-model predictions and satellite measurements as inputs. NO₂ VCDs and surface concentrations are retrieved from Pandora zenith-sky measurements made in Toronto, Canada, from 2015 to 2017. The retrieved Pandora zenith-sky NO₂ data (VCD and surface concentration) show good agreement with both satellite and in situ measurements. The diurnal and seasonal variations of derived Pandora zenith-sky surface NO₂ data also agree well with in situ measurements (diurnal difference within ± 2 ppbv). Overall, this work shows that the new Pandora zenith-sky NO₂ products have the potential to be used in various applications such as future satellite validation in moderate cloudy scenes and air quality monitoring.

1 Introduction

Nitrogen dioxide (NO₂) is an important air pollutant and plays a critical role in tropospheric photochemistry (e.g., ECCC, 2016; EPA, 2014). It is primarily emitted from combustion processes such as fossil fuel combustion (e.g., traffic, electricity generation from power plants) and biomass burning, as well as from lightning. NO₂ is a nitrate aerosol precursor, and it also contributes to acid deposition and eutrophication (ECCC, 2016). Exposure to NO₂ can lead to adverse health effects, such as irritation of the lungs, a decrease in lung function, and an increase in susceptibility to allergens for people with asthma (EEA, 2017; WHO, 2017).

As surface NO₂ concentrations are regulated by many environmental agencies (e.g., Environment and Climate Change Canada and U.S. Environment Protection Agency), in situ NO₂ measurements are commonly carried out by many national monitoring networks, such as the National Air Pollution Surveillance (NAPS, <https://www.canada.ca/en/environment-climate-change/services/air-pollution/monitoring-networks-data/national-air-pollution-program.html>) network in Canada, which was established in 1969. The in situ methods used to measure surface NO₂ have evolved over the years; for example, luminol chemiluminescence (e.g., Kelly et al., 1990; Maeda et al., 1980; Wendel et al., 1983), long-path differential optical absorption spectroscopy (e.g., Platt, 1994), photolytic conversion/chemiluminescence (e.g., Gao et al., 1994; Ryerson et al., 2000), and laser-induced fluorescence (e.g., Thornton et al., 2000) are all found to be reliable methods with an uncertainty within 10 % at the 1 ppbv and higher concentration levels (McClenny, 2000). Currently, the in situ approach used by NAPS for surface NO₂ air quality monitoring is the photolytic conversion/chemiluminescence technique, which converts NO₂ to NO and subsequently detects the NO by chemiluminescence reaction (McClenny, 2000; NRC, 1992). This in situ monitoring measurements provides good measurements at ground level (0.4 ppbv accuracy), but NO₂ is not uniformly mixed through the atmosphere, and not even within the atmospheric boundary layer due to emission and removal processes taking place at the surface.

Total vertical column NO₂ can be measured by many ground-based UV-visible remote sensing instruments using direct-sun, zenith-sky, or off-axis spectroscopy techniques (Cede et al., 2006; Drosoglou et al., 2017; Herman et al., 2009; Lee et al., 1994; Noxon, 1975; Piters et al., 2012; Roscoe et al., 2010; Tack et al., 2015; Vaughan et al., 1997). These measurements are of high quality and good precision, and have been widely used for atmospheric chemistry studies (e.g. Adams et al., 2012; Hendrick et al., 2014) and satellite validations (e.g., Celarier et al., 2008; Drosoglou et al., 2018; Irie et al., 2008; Wenig et al., 2008). Among all these different viewing geometries, direct-sun measurements are of high accuracy, and are not dependent on radiative transfer models (RTMs) to calculate air mass factors (AMFs) (Herman et al., 2009) or on knowledge of other atmospheric constituents. Zenith-sky observations have been widely used for stratospheric ozone and NO₂ observations, particularly under cloudy conditions when direct-sun measurements are unreliable (note that zenith-sky observations use scattered sunlight and are less sensitive to clouds, e.g., Zhao et al. (2019)). Off-axis measurements have good sensitivity in the boundary layer and could provide tropospheric trace gas profiles and surface concentrations (Frieß et al., 2011; Hendrick et al., 2014; Kramer et al., 2008; Wagner et al., 2011), but they are more sensitive to cloud cover than zenith-sky measurements.

The Pandora sun spectrometer is a new instrument developed to measure vertical column densities (total columns) of trace gases in the atmosphere using sun and sky radiation in the UV-visible part of the spectrum (Herman et al., 2009). One of its primary data products is NO₂ total vertical column density (VCD) from the direct-sun viewing mode, where VCD represents the vertically integrated number of molecules per unit area and is reported in units of molec cm⁻² or Dobson Unit (1 DU = 2.6870×10¹⁶ molec cm⁻²). The Pandora direct-sun NO₂ VCD products have been validated through many field campaigns (Flynn et al., 2014; Lamsal et al., 2017; Martins et al., 2016; Piters et al., 2012; Reed et al., 2015), ground-based comparisons (Herman et al., 2009; Wang et al., 2010), and satellite validations (Ialongo et al., 2016; Lamsal et al., 2014).

Since their introduction in 2006, Pandora spectrometers have been deployed at more than 50 sites globally. The Pandora no. 103 instrument used in this study has been deployed in Toronto, Canada since 2013 to perform direct-sun measurements (Zhao

et al., 2016). Since 2015, the observation schedule of Pandora no. 103 has been modified to perform alternating direct-sun and zenith-sky measurements. Knepp et al. (2017) assessed Pandora's capability to derive stratospheric NO₂ using zenith-sky viewing geometry (in twilight periods), but their study was limited to slant column densities (SCDs). At this time, there are no standard Pandora zenith-sky NO₂ VCD data products available. As one goal of this work, we have focused on developing a new NO₂ retrieval algorithm for zenith-sky measurements to expand Pandora NO₂ measurements into cloudy scenes.

In addition to retrieval of zenith-sky total column NO₂, another goal of this work is to derive surface NO₂ concentration from total column measurements. Surface NO₂ has been a focus of scientific studies due to its strong correlation with air quality (AQ) and health issues (ECCC, 2016), with NO₂ as one of the three components (along with ozone and PM_{2.5}) used to compute the Air Quality Health Index (AQHI: Stieb et al., 2008) in Canada's AQ public awareness programs. Efforts to link total column NO₂ with its surface concentrations have been made by many researchers (Flynn et al., 2014; Knepp et al., 2015; Kollonige et al., 2017; Lamsal et al., 2008, 2014; McLinden et al., 2014). For example, Knepp et al. (2015) proposed a method to estimate NO₂ surface mixing ratios from Pandora direct-sun total column NO₂ via application of a planetary boundary-layer (PBL) height correction factor. Kollonige et al. (2017) adapted this method and compared Pandora direct-sun surface NO₂ and OMI surface NO₂. They concluded that the two main sources of error for the conversion of the total column NO₂ to surface NO₂ are (1) poor weather conditions (e.g., cloud cover and precipitation) and (2) PBL height estimation, both of which affect the NO₂ column-surface relationship and instrument sensitivities to boundary layer NO₂. Thus, in this work we present a simple but robust algorithm for deriving surface NO₂ concentration from Pandora zenith-sky measurements, which has several advantages such as the ability (1) to extend Pandora NO₂ measurements to cloudy conditions and (2) to provide more accurate surface NO₂ concentration estimates that are less sensitive to PBL height. This work also provides reliable total column NO₂ measurements in cloudy conditions and could be used in satellite validations in partially cloudy scenes.

This paper is organized as follows. Section 2 describes the measured and modelled NO₂ data used in this study. In Section 3, the empirical AMFs for Pandora zenith-sky NO₂ measurements are derived using high-quality Pandora direct-sun total column NO₂ data. These empirical AMFs and the Network for the Detection of Atmospheric Composition Change (NDACC) AMFs (Hendrick et al., 2011; Sarkissian et al., 1995; Van Roozendael et al., 1998; Van Roozendael and Hendrick, 2009; Vaughan et al., 1997) are both applied to Pandora zenith-sky total column NO₂ retrievals to help evaluate the performance of the empirical AMFs. Also, the retrieved Pandora zenith-sky total column NO₂ data are evaluated by comparison with satellite measurements. In Section 4, the zenith-sky total column NO₂ data are converted to surface concentration by using a scaling algorithm. The zenith-sky surface NO₂ concentration data are assessed by comparison with in situ measurements. Lastly, in Section 5, several aspects of this zenith-sky surface NO₂ dataset are discussed, which include: diurnal and seasonal variation, and PBL effect, followed by conclusions in Section 6.

2 Datasets and models

2.1 Measurements

2.1.1 Pandora direct-sun total column NO₂

The Pandora instrument records spectra between 280 and 530 nm with resolution of 0.6 nm (Herman et al., 2009, 2015; Tzortziou et al., 2012). It uses a temperature-stabilized Czerny-Turner spectrometer, with a 50 μm entrance slit, 1200 groove mm^{-1} grating, and a 2048×64 back-thinned Hamamatsu charge-coupled device (CCD) detector. The spectra are analysed using a total optical absorption spectroscopy (TOAS) technique (Cede, 2019), in which absorption cross sections for multiple atmospheric absorbers, such as ozone, NO₂, and sulphur dioxide (SO₂), are fitted to the spectra.

The Pandora direct-sun total column NO₂ data are produced using Pandora's standard NO₂ algorithm implemented in the BlinkP software (Cede, 2019). The measured direct-sun spectra from 400 to 440 nm are used in the TOAS analysis. A synthetic reference spectrum is produced by averaging multiple measured spectra and corrected for the estimated total optical depth included in it. Cross sections of NO₂ at an effective temperature of 254.5 K (Vandaele et al., 1998), ozone at an effective temperature of 225 K (Brion et al., 1993, 1998; Daumont et al., 1992), and a fourth-order polynomial are all fitted. The resulting NO₂ SCDs are then converted to total column VCDs by using direct-sun geometry AMFs. Herman et al. (2009) show that Pandora direct-sun total column NO₂ has a clear-sky precision of 0.01 DU (in slant column) and a nominal accuracy of 0.1 DU (in vertical column, 2-sigma level). Additional information on Pandora calibrations, operation, and retrieval algorithms can be found in Herman et al. (2009) and Cede (2019).

The Pandora no. 103 instrument has been deployed in Toronto since September 2013 to perform direct-sun observations (Zhao et al., 2016). The instrument is installed on the roof of the Environment and Climate Change Canada (ECCC) Downsview building (43.7810°N, -79.4680°W) in Toronto. The building is located in a suburban area with multiple roads nearby. Since 2015, the instrument has employed an alternating direct-sun and zenith-sky observation schedule, which consists of direct-sun measurements every 90 seconds and zenith-sky measurements every 30 minutes during the sunlit period. About two-and-a-half years (February 2015 to September 2017) of continuous alternating measurements are used in this study.

2.1.2 Pandora zenith-sky total column NO₂

Retrieval of trace gases from Pandora's zenith-sky measurements is not included in the standard BlinkP processing software (Cede, 2019). The Pandora zenith-sky spectra for this study are processed using the differential optical absorption spectroscopy (DOAS) technique (Noxon, 1975; Platt, 1994; Platt and Stutz, 2008; Solomon et al., 1987) with the QDOAS software (Danckaert et al., 2015). A single reference spectrum is used, which was obtained from a zenith-sky measurement at local noon from a day that had low total column NO₂. Following the NDACC recommendations (Van Roozendael and Hendrick, 2012), NO₂ differential slant column densities (dSCDs) are retrieved in the 425-490 nm window (to retrieve oxygen collision complex simultaneously). The oxygen collision complex (O₂)₂ (referred here as O₄), which is created by the collision of two oxygen molecules, has broadband absorptions from UV to near IR spectral ranges (Greenblatt et al., 1990; Platt and Stutz,

2008; Thalman and Volkamer, 2013). O₄ is widely used as a reference gas by many DOAS applications to infer cloud and aerosol properties (e.g., Gielen et al., 2014; Wagner et al., 2004, 2014, 2016; Wang et al., 2015; Zhao et al., 2019). Cross sections of NO₂ at an effective temperature of 254.5 K (Vandaele et al., 1998), ozone at an effective temperature of 223 K (Bogumil et al., 2003), H₂O (Rothman et al., 2005), O₄ (Hermans et al., 2003), and Ring (Chance and Spurr, 1997) are all fitted; a fifth-order polynomial and a first-order linear offset are also included in the DOAS analysis.

The output of QDOAS is NO₂ dSCDs, which can be converted to total column NO₂ via the Langley plot method with the use of the NDACC NO₂ AMF look-up table (LUT) (Van Roozendaal and Hendrick, 2012). The NDACC AMF LUT is used here only as a reference since it was primarily developed for retrieval of stratospheric NO₂. Other empirical zenith-sky NO₂ AMFs have been developed and are used to convert NO₂ dSCDs to total columns. Details about these two different AMFs are given in Section 3.1.

2.1.3 OMI SPv3 data

The Ozone Monitoring Instrument (OMI) is a Dutch-Finnish nadir-viewing UV-visible spectrometer aboard the National Aeronautics and Space Administration (NASA)'s Earth Observing System (EOS) Aura satellite that was launched in July 2004. The OMI instrument measures the solar radiation backscattered by the Earth's atmosphere and surface between 270 and 500 nm with resolution of 0.5 nm (Levelt et al., 2006, 2018). OMI has a 780×576 CCD detector that measures at 60 across-track positions simultaneously, and thus, does not require across-track scanning. Due to this approach, the spatial resolution of the CCD pixels varies significantly along the across-track direction: those pixels near the track centre have ground footprint of $13 \text{ km} \times 24 \text{ km}$ (along-track \times across-track), whereas those close to the track edge (e.g. view zenith angle = 56°) have a ground footprint roughly of $23 \text{ km} \times 126 \text{ km}$ (de Graaf et al., 2016). Note that from 2012 onwards the smallest pixels (across-track positions) can no longer be used and are excluded from the analysis (known as the "row anomaly", i.e. Levelt et al., 2018). This means the "smallest" pixels available for an OMI comparison are larger than $13 \text{ km} \times 24 \text{ km}$.

The OMI NO₂ data used in this work are the NASA standard product (SP) (Bucsela et al., 2013; Wenig et al., 2008) version 3.0 Level 2 (SPv3.0) (Krotkov et al., 2017). The NO₂ SCDs are derived using the DOAS technique in the 405-465 nm window (Marchenko et al., 2015). The AMFs used in SPv3.0 are calculated by using $1^\circ \times 1.25^\circ$ (latitude \times longitude) resolution *a priori* NO₂ and temperature profiles from the Global Modeling Initiative (GMI) chemistry-transport model with yearly varying emissions (Krotkov et al., 2017).

2.1.4 In situ measurements

The National Air Pollution Surveillance (NAPS) network was established in 1969 to monitor and assess the quality of ambient (outdoor) air in the populated regions of Canada (<https://www.canada.ca/en/environment-climate-change/services/air-pollution/monitoring-networks-data/national-air-pollution-program.html>, accessed 23 November 2018). NAPS provides accurate long-term air quality data (ozone, NO₂, SO₂, carbon monoxide (CO), fine particulate matter, etc.) of a uniform standard across Canada (e.g., Dabek-Zlotorzynska et al., 2011; Reid and Aherne, 2016).

The in situ NO₂ data used in this study were collected at the NAPS Toronto North station (located 100 m away from the Pandora instrument). The site is 186 m above sea level, and the height of the air intake is 4 m above the ground.

The in situ NO₂ concentration is measured using a photolytic NO₂ instrument (Thermo 42i) that is also sensitive to other gaseous inorganic nitrogen compounds (e.g., nitric acid (HNO₃) and peroxyacetyl nitrate (PAN)) (McLinden et al., 2014).

- 5 Thus in areas where direct NO_x (nitrogen oxides) emission sources are limited and other nitrogen compounds are present, NO₂ may be overestimated (e.g., in rural areas). For the current site, however, this positive bias has been found to be only about 5%, except for very low NO₂ concentrations (<5 ppbv) (Yushan Su, Ontario Ministry of the Environment, Conservation and Parks, personal communication, October 2018).

2.2 Numerical models

- 10 Predicted NO₂ fields from three atmospheric chemistry models are used in the algorithm described in Section 4.1 to derive surface NO₂ concentration from Pandora zenith-sky total column NO₂ data. Following McLinden et al. (2014), this work uses the Global Environmental Multi-scale Modelling Air quality and CHemistry (GEM-MACH) regional chemical transport model (CTM) and the GEOS-Chem global CTM to simulate total columns and vertical profiles of tropospheric NO₂ and surface NO₂ concentration. The stratospheric NO₂ partial columns are estimated using OMI satellite data and the Pratmo box-model.

15 2.2.1 GEM-MACH

- GEM-MACH is ECCC's regional air quality forecast model. It is run operationally **twice** times per day to predict hourly surface pollutant concentrations over North America for the next 48 hours (Moran et al., 2009; Pavlovic et al., 2016; Pendlebury et al., 2018). The model consists of an online tropospheric chemistry module (Akingunola et al., 2018; Pavlovic et al., 2016) embedded within the ECCC Global Environmental Multi-scale (GEM) numerical weather prediction model (Côté et al., 1998). Physical and chemical processes represented in GEM-MACH include emissions, dispersion, gas- and aqueous-phase chemistry, inorganic heterogeneous chemistry, aerosol dynamics, and wet and dry removal. The model uses gridded hourly emissions fields based on U.S. and Mexican national inventories from the U.S. Environmental Protection Agency (EPA) Air Emissions Modeling Platform and on Canada's national Air Pollutant Emission Inventory (APEI, <https://pollution-waste.canada.ca/air-emission-inventory>, accessed 23 November 2018) (Zhang et al., 2018). Currently, only NO_x emissions in
25 the PBL are included in the operational model; free-tropospheric NO_x emissions from lightning and in-flight aircraft are not considered. In this work, the GEM-MACH hourly NO₂ vertical profiles from 0 to 1.5 km and surface concentrations are retrieved from archived operational forecasts on the native model grid covering North America at 10 km × 10 km horizontal resolution for the period April 2016 to December 2017. The corresponding grid-box closest to the Pandora location was used in this study.

2.2.2 GEOS-Chem

The GEOS-Chem chemical transport model (Bey et al., 2001) has been used extensively in the retrieval of tropospheric columns, and has been shown to be capable of reasonably simulating the vertical distributions of NO₂ (Lamsal et al., 2008; Martin et al., 2002; McLinden et al., 2014). The model has a detailed representation of tropospheric chemistry, including aerosols and their precursors (Park et al., 2004). In the simulation used in this study, a global lightning NO_x source of 6 Tg N yr⁻¹ (Martin et al., 2002) was imposed. **Lightning NO_x emissions are computed as a function of cloud-top height, and are scaled globally as described by Sauvage et al. (2007) to match Optical Transient Detector/Lightning Imaging Sensor (OTD/LIS) climatological observations of lightning flashes.** The model was run on a 1/2° × 2/3° (latitude × longitude) grid in nested mode over North America, and was driven by assimilated meteorology from the Goddard Earth Observing System (GEOS-5). The modelled NO₂ profiles were used to calculate monthly mean NO₂ partial columns in the free troposphere (1.5 to 12 km), as the GEM-MACH model does not include free-tropospheric NO₂ sources (lightning, in-flight aircraft emissions).

2.2.3 Pratmo box-model

Pratmo is a stratospheric photochemical box-model (Brohede et al., 2008; Lindenmaier et al., 2011; McLinden et al., 2000). The model has detailed stratospheric chemistry that includes long-lived species (nitrous oxide (N₂O), methane (CH₄), and water vapor (H₂O)) and halogen families (NO_y, Cl_y, and Br_y) that are based on a combination of three-dimensional model output and tracer correlations (Adams et al., 2017). Heterogeneous chemistry of background stratospheric sulfate aerosols is also included. The model is constrained with climatological profiles of ozone and temperature. Stratospheric NO₂ has a strong diurnal variation; therefore, diurnal corrections must be applied when OMI stratospheric NO₂ measurements (around local noon) are interpolated to Pandora measurement times. Ratios of modelled stratospheric NO₂ columns are calculated at OMI overpass time and Pandora measurement time. These ratios are multiplied by the OMI measured stratospheric NO₂ to produce stratospheric NO₂ columns corresponding to the time of Pandora measurements. Details about the use of the Pratmo box-model and the calculation of stratospheric NO₂ partial columns are provided in Section 4.1.

3 Total column NO₂ retrieval

3.1 Zenith-sky air mass factor

The NDACC UV-visible network uses zenith-sky AMFs in its total column NO₂ retrievals. To improve the overall homogeneity of the UV-visible NO₂ column measurements, NDACC recommended using the NO₂ AMF LUT (Van Roozendaal and Hendrick, 2012). This LUT is based on climatological NO₂ profiles that are composed of (1) 20-60 km NO₂ profiles developed by Lambert et al. (1999, 2000) and (2) 12-20 km NO₂ profiles derived from SAOZ balloon observations (Van Roozendaal and Hendrick, 2012). The NO₂ concentration is set to zero below 12 km altitude. The NO₂ AMFs have been calculated using the UVSPEC/DISORT RTM (Hendrick et al., 2006; Wagner et al., 2007). The parameters used in

building the LUT are: wavelength, ground albedo, altitude of the station, and solar zenith angle (SZA). Aerosol extinction, ozone, and temperature profiles come from an aerosol model (Shettle, 1989), the U.S. Standard Atmosphere, and the TOMS V8 Climatology, respectively.

The NDACC LUT is designed for stratospheric NO₂ retrievals. Note that the absence of tropospheric NO₂ in the NDACC LUT construction will lead to an underestimation of the total column NO₂ in urban areas. For example, from 2015 to 2017, tropospheric NO₂ accounted for $73 \pm 11 \%$ (1σ) of the total column amounts in Toronto (OMI SPv3.0 data). To account for this significant tropospheric NO₂ in urban areas, new empirical AMFs were developed in this study and the NDACC AMF LUT is used for comparison purposes only. In Tack et al. (2015), a more sophisticated four-step approach to derive total and tropospheric NO₂ columns from zenith-sky measurements was proposed, which involved with using a RTM to calculate appropriate tropospheric AMFs. However, due to benefits from using the high-quality Pandora direct-sun total column NO₂ measurements, this work took a different but simple and robust approach to derive zenith-sky total column NO₂.

Empirical AMFs are calculated for Pandora zenith-sky NO₂ measurements in such a way that they can be used to retrieve zenith-sky total column NO₂ values that match the high-quality Pandora direct-sun total column NO₂ values. Inferring total columns from zenith-sky observations through comparisons with accurate direct-sun observations is a common approach for Brewer and Dobson zenith-sky total ozone measurements (Kerr et al., 1988). For example, in the Brewer instrument zenith-sky ozone algorithm, weighted zenith-sky light intensities measured at four wavelengths (F) are expressed as a function of the slant path (μ) and total column ozone (Kerr et al., 1981). The nine semi-empirical coefficients used to derive total column ozone from measured F in the equation are estimated from a set of direct-sun and zenith-sky observations made nearly simultaneously (Fioletov et al., 2011). Instead of finding the link between zenith-sky spectral intensity and total column values (i.e., following the Brewer and Dobson zenith-sky total ozone retrieval method), deriving empirical zenith-sky AMFs for Pandora zenith-sky measurements is more straightforward since Pandora zenith-sky spectra can be analyzed to produce NO₂ dSCDs.

The relation between VCD and dSCD can be expressed as:

$$VCD = \frac{dSCD + RCD}{AMF}, \quad (1)$$

where, RCD is the reference column density that shows the slant column amount of the trace gas in the reference spectrum (Section 2.1.2). If we make an assumption that the coincident direct-sun (DS) and zenith-sky (ZS) measurements sampled the same air mass, then the empirical zenith-sky AMFs (referred to here as AMF_{ZS-Emp}) can be calculated by assuming $VCD_{DS} = VCD_{ZS}$, which gives

$$VCD_{DS}(SZA) = \frac{dSCD_{ZS}(SZA) + RCD_{ZS}}{AMF_{ZS-Emp}(SZA)}. \quad (2)$$

Next, we can use nearly-coincident VCD_{DS} and $dSCD_{ZS}$ in a multi-non-linear regression to retrieve AMF_{ZS-Emp} and RCD_{ZS} together. To ensure the quality of the retrieved AMF_{ZS-Emp} , only high quality direct-sun total column NO₂ data are used with $SZA < 75^\circ$. Details about the empirical zenith-sky AMF calculation are shown in Appendix A.

Figure 1 shows a comparison of the empirical zenith-sky AMFs and NDACC AMFs (calculated for the Toronto measurements). Total column NO₂ can then be retrieved using Eqn. (1) and these two sets of AMFs, where the one based on empirical AMFs is referred to as VCD_{ZS-Emp} and the one based on NDACC AMFs is referred to as VCD_{ZS-NDACC}. The RCD value used in the retrievals is 0.39 ± 0.01 DU, which is retrieved along with AMF_{ZS-Emp} (Appendix A). Figure 2 shows the comparisons of the NO₂ columns measured by zenith-sky and direct-sun methods. The regression analyses were performed by using the following coincidence criteria: (1) nearest Pandora direct-sun measurement that was within ± 5 min of Pandora zenith-sky measurement, (2) SZA < 75°, and (3) Pandora direct-sun total column NO₂ data have assured high quality (BlickP L2 data quality flag for nitrogen dioxide = 0). In general, the VCD_{ZS-Emp} and VCD_{ZS-NDACC} performed as expected. Compared with VCD_{DS}, the VCD_{ZS-NDACC} shows a -25% bias, while the VCD_{ZS-Emp} only shows a -4 % bias (indicated by the red lines on each panel and their slopes). In addition, VCD_{ZS-Emp} shows less SZA dependence than VCD_{ZS-NDACC} (see the increased bias for measurements made in larger SZA conditions in Figure 2b). These results confirm that, for urban sites, the tropospheric NO₂ profile should be included when calculating empirical zenith-sky AMFs. In the rest of the paper, only the zenith-sky NO₂ retrieved using empirical AMFs will be discussed. The derived zenith-sky total column NO₂ values are affected by both clouds and aerosols due to their impact on the light path. The presence of clouds and aerosols contributes to the uncertainty of the measurements. However, the impact of aerosols is expected to be moderate in most cases compare to that of clouds (e.g., Hendrick et al., 2011; Tack et al., 2015). Thus, this work has focused on evaluating the impact from clouds. Note that the Pandora zenith-sky total column NO₂ data discussed in Sections 3 are a “clear-sky subset” of Pandora zenith-sky measurements. The assessment of Pandora zenith-sky NO₂ measurements in cloudy conditions are provided in Section 4.

3.2 Comparison with satellite measurements

To illustrate the NO₂ variability over Toronto, Figure 3 shows the time series (2015-2017) from Pandora direct-sun, zenith-sky, and OMI SPv3.0 total column NO₂. In general, the NO₂ datasets from the ground-based Pandora instrument and the satellite follow the same pattern. However, the satellite data are likely to miss the peak NO₂ values in the morning since OMI only passes over Toronto once per day around 1:30 p.m. (local time).

We also performed regression analyses by using the following coincidence criteria: (1) nearest (in time) measurement that was within ± 30 min of OMI overpass time, (2) closest OMI ground pixel (having a distance from the ground pixel centre to the location of the Pandora instrument less than 20 km), and (3) cloud fraction ≤ 0.3 (the effective geometric cloud fraction, as determined by the OMCLDO2 algorithm; Celarier et al., 2016). In this comparison, only high-quality OMI data are used (VcdQualityFlags = 0) (Celarier et al., 2016). Figures 4a and 4b show the scatter plots of OMI vs. Pandora direct-sun and OMI vs. Pandora zenith-sky total column NO₂, respectively. Figures 4c and 4d show similar comparisons but only use OMI NO₂ measured by “small pixels” (i.e., having viewing zenith angle of less than 35°). The better correlation and lower bias for zenith-sky versus direct-sun might be a case of coincident errors, i.e., compare to Pandora direct-sun total column NO₂, both OMI and Pandora zenith-sky total column NO₂ underestimate the local NO₂ at Toronto (see Figure 2). When taking into account the standard error of the fitting and the confidence level of R, the difference between zenith-sky and direct-sun data is not

significant (i.e, in Fig. 4 from panels a to d, the slopes with standard error are 0.64 ± 0.02 , 0.67 ± 0.02 , 0.70 ± 0.04 , and 0.71 ± 0.03 ; the 95% confidence intervals for R values are 0.45 to 0.63, 0.61 to 0.75, 0.43 to 0.77, and 0.60 to 0.86). The comparison results indicate that, at the Toronto site, OMI underestimates the total column by about 30 %. This underestimation is qualitatively consistent with the fact that the Pandora location is near the northern edge of peak Toronto NO₂, and the relatively large OMI pixels are also generally sampling areas of less NO₂ in the vicinity. The use of the relatively coarse (1°) GMI model for profiles shapes (Section 2.1.3) will also lead to a low bias considering the peak NO_x emissions span roughly $0.5^\circ \times 0.5^\circ$. Similar results have been found elsewhere.

Ialongo et al. (2016) reported a similar negative bias using OMI SPv3.0 and Pandora direct-sun total column NO₂ in Helsinki (-32 % bias and R = 0.51), and they suggested this was due to the difference between the OMI pixel and the relatively small Pandora field-of-view. In Reed et al. (2015), Pandora measurements at 11 sites were evaluated; the authors found that the best correlation between OMI SPv3.0 and Pandora direct-sun total column NO₂ data is for rural sites. They concluded this could be due to smaller atmospheric variability in the rural region. Other studies such as Goldberg et al. (2017) found an even worse OMI-Pandora comparison between these two data products with striking negative bias at high values and poor correlation (R = 0.3). The authors attributed the poor agreement to the coarse resolution of OMI and its AMFs computed with GMI *a priori* NO₂ profiles. In general, our comparison results show that: (1) the Pandora direct-sun total column NO₂ data measured in Toronto have a reasonable agreement with OMI, and (2) the Pandora zenith-sky total column NO₂ data show results similar to those for direct-sun total column when compared with OMI SPv3.0.

4 Surface NO₂ concentration retrieval

The performance of the clear-sky Pandora zenith-sky total column NO₂ data has been assessed by using OMI and Pandora direct-sun data as described in Section 3.2. However, the validation of cloudy-scene Pandora zenith-sky total column data is not simple, since near-simultaneous good quality direct-sun or satellite measurements in most cloudy conditions are not available. This cloudy-scene validation can be done by comparison with in situ NO₂ measurements that are not affected by weather. In general, the comparison between total columns and surface concentrations can be done by two approaches: (1) convert Pandora zenith-sky total columns to surface concentrations; and (2) convert in situ surface concentrations to total column values. For example, Spinei et al. (2018) calculated “ground-up” VCDs from in situ surface concentrations by using additional measurements of PBL height or assuming trace gas profiles. In this work, the first approach is employed since the surface NO₂ data products from Pandora remote-sensing measurements have direct applications in areas such as air quality monitoring.

4.1 Column-to-surface conversion algorithm

A simple but robust scaling method is adapted to derive surface NO₂ concentration from Pandora zenith-sky total column NO₂ measurements. Following Lamsal et al. (2008) and McLinden et al. (2014), the surface NO₂ concentration is estimated using the modelled profile and surface concentration,

$$C_{pan} = (V_{pan} - V_{strat} - V_{ftrop}) \times \left(\frac{C}{V_{PBL}} \right)_{G-M}, \quad (3)$$

where C_{pan} is the surface NO₂ volume mixing ratio (VMR) to be estimated, C is the surface NO₂ VMR from GEM-MACH (or G-M), V_{pan} is the total column NO₂ measured by Pandora, V_{strat} is the stratospheric NO₂ partial column, V_{ftrop} is the NO₂ partial column in the free troposphere, and V_{PBL} is the NO₂ partial column in the PBL. This equation assumes the chemical transport models can effectively capture the spatial and temporal behaviour of the concentration-to-partial-column ratio.

- 10 In this work, V_{PBL} (0-1.5 km) is integrated from the GEM-MACH NO₂ profile and V_{ftrop} (1.5-12 km) is integrated from the GEOS-Chem NO₂ profile. Both GEM-MACH and GEOS-Chem have an hourly temporal resolution. Thus, the integrated V_{PBL} and V_{ftrop} can account for NO₂ diurnal variation. However, V_{strat} is from OMI monthly mean stratospheric NO₂, which does not have diurnal variation. Thus, the Pratmo box-model is used to calculate stratospheric NO₂ diurnal ratios. The OMI stratospheric NO₂ columns are interpolated to morning and evening hours by multiplying by the box-model diurnal ratios. Details about the
- 15 calculation of V_{strat} as well as references are provided in Appendix B.

The $(C/V_{PBL})_{G-M}$ ratio in Eqn. 3 is provided by GEM-MACH, and has hourly temporal resolution. This modelled $(C/V_{PBL})_{G-M}$ ratio is referred to here as a conversion ratio R_{cv} . Besides the hourly modelled conversion ratio, a simple monthly look-up table is built using an average of the one-and-a-half years of GEM-MACH model outputs (April 2016 to December 2017) that were available. The look-up table (referred to here as the Pandora surface-concentration look-up table, or PSC-LUT) is

20 composed of monthly conversion ratios with hourly resolution as shown in Figure 5. For example, assuming that a Pandora NO₂ total column measurement is made on a day in December at 15:00 LST, then the corresponding conversion ratio from the PSC-LUT is 28 ppbv DU⁻¹ (see the black arrow). Our results in Figure 5 show that the conversion ratio changes throughout the day as well as with season: 0.1 DU (partial column NO₂ in the PBL) corresponds to 5-8 pptv of surface NO₂ in the morning (8:00 LST), 2-3 pptv around local noon (13:00 LST), and 2-4 pptv in the evening (18:00 LST). In general, the variation of

25 conversion ratios demonstrates that the surface NO₂ concentration is controlled not only by PBL height, but also by both boundary-layer dynamics and photochemistry. The surface NO₂ derived using the hourly modelled R_{cv} ratio is referred to here as $C_{pan-model}$, while the surface NO₂ derived using the monthly mean PSC-LUT is referred to here as $C_{pan-LUT}$. In general, $C_{pan-model}$ is a data product that depends on daily model outputs, but $C_{pan-LUT}$ only needs the pre-calculated PSC-LUT and is thus less dependent on the model. In general, the look-up table approach ($C_{pan-LUT}$) is aiming for a quick and near-real-time data delivery.

- 30 Thus, to minimize year-to-year variation (e.g., from changing meteorological conditions or changing local emission patterns), for a given year we recommend using a mean PSC-LUT that is calculated from model simulations of previous years. On the other hand, the $C_{pan-model}$ is the off-line, high-quality, year-specific data product that will be delivered for air quality research and other applications. Details of these two different surface NO₂ data products are discussed in the next section.

4.2 Comparison with measurements and model

Figure 6 shows the evaluation of modelled and Pandora zenith-sky surface NO₂ concentrations, both using in situ NO₂ measurements as the reference. The Pandora data have been filtered for heavy clouds (details are given in Section 4.3). The GEM-MACH modelled surface concentrations in Toronto reproduce the in situ measurements very well with the comparison showing high correlation ($R = 0.78$) and moderate positive bias (37 %, Figure 6a). The Pandora zenith-sky surface NO₂ data, $C_{pan-model}$, shows almost the same correlation ($R = 0.77$), with only -7 % bias (Figure 6b). The better performance of $C_{pan-model}$ is expected since the conversion method for Pandora zenith-sky measurements relies on the GEM-MACH modelled NO₂ profile (see Eqn. 3); in other words, the Pandora zenith-sky surface NO₂ has at least one more piece of information (i.e., NO₂ total column) than GEM-MACH surface NO₂ concentrations. The $C_{pan-LUT}$ shows a similar correlation coefficient ($R = 0.73$) and has improved bias (-3 %, Figure 6c). This result (slightly lower correlation) is also reasonable and acceptable since $C_{pan-LUT}$ is derived with the monthly PSC-LUT, which has less accurate information than the hourly modelled data.

Besides the improved bias, Pandora zenith-sky surface NO₂ concentrations, $C_{pan-model}$ and $C_{pan-LUT}$ (Figures 6e and 6f) also have better frequency distributions than the GEM-MACH (Figure 6d). Figure 6d shows that the NO₂ surface concentrations peaks (ambient background concentrations) from model and in situ data are misaligned. This indicates that the GEM-MACH NO₂ background surface concentrations have a 1ppbv low bias at this site. In contrast, the zenith-sky surface NO₂ at peak-frequency matches the in situ data (Figures 6e and 6f), indicating that the low bias of the background surface NO₂ value has been corrected with this additional information from Pandora zenith-sky total column measurements. In addition, in high NO₂ concentration conditions (> 20 ppbv), the zenith-sky surface NO₂ also shows better agreement with the in situ NO₂ than do the modelled data. The mean of the top 10 % of the in situ data is 26 ± 1 ppbv (uncertainty of the mean), whereas the corresponding values for GEM-MACH, $C_{pan-model}$, and $C_{pan-LUT}$ are 39 ± 1 ppbv, 26 ± 1 ppbv, and 27 ± 1 ppbv, respectively.

The total column-to-surface concentration conversion algorithm has also been applied to the Pandora direct-sun total column NO₂ (see Figure 7). Figure 7b shows that the direct-sun surface NO₂ data have a similar agreement with the in situ data (-8 % bias and $R = 0.80$) as the zenith-sky surface NO₂. In high NO₂ concentration conditions, direct-sun data have a similarly good agreement with the in situ measurements. For this direct-sun based dataset, the mean of the top 10 % of the in situ data is 27 ± 1 ppbv, whereas the corresponding values for GEM-MACH, $C_{pan-model}$, and $C_{pan-LUT}$ are 40 ± 1 ppbv, 27 ± 1 ppbv, and 27 ± 1 ppbv, respectively.

Thus, in general, both Pandora zenith-sky and direct-sun surface NO₂ datasets can be used reliably to obtain surface concentrations. The good consistency between $C_{pan-model}$ and $C_{pan-LUT}$ implies that two versions of Pandora surface NO₂ data can be delivered in the future, i.e., an off-line version that relies on the inputs from hourly model, and a near-real-time version that only needs a pre-calculated LUT.

4.3 Measurements in different sky conditions

Although zenith-sky observations are less sensitive to cloud conditions than direct-sun observations, we still need to be cautious about the derived zenith-sky surface NO₂ in heavy cloud conditions. Due to enhanced scattering, heavy clouds could lead to a significant overestimation of surface NO₂ derived from zenith-sky measurements. A cloud filtering method based on retrieved O₄ dSCDs is used to identify these conditions. High retrieved O₄ values correspond to long optical path lengths and therefore it is expected that corresponding NO₂ values are overestimated as discussed in Appendix C.

The effectiveness of the zenith-sky NO₂ in cloudy scenes is demonstrated by the time series plots (Figure 8) of in situ and Pandora direct-sun and zenith-sky data (in their original temporal resolutions). Under clear-sky conditions (for example, April 8-14), both Pandora direct-sun and zenith-sky-based surface concentrations correlate well with the in situ measurements. Under moderately cloudy conditions, when Pandora direct-sun observations cannot provide high-quality data, Pandora zenith-sky observation still can yield good measurements that compare well with in situ data (for example, April 26-29). Under heavy cloud conditions, however, which are identified by enhanced O₄ (Appendix C), Pandora zenith-sky-derived surface NO₂ yielded higher than in situ measurements (for example, April 4 and 6, see the green squares). This feature is due to the enhanced multi-scattering in heavy cloud conditions, which leads to enhanced NO₂ absorption in the measured spectra.

Sensitivity tests (Appendix C) show that only 10 % of all zenith-sky measurements are strongly affected by this enhanced absorption, indicating the zenith-sky NO₂ algorithm is applicable to most measurements made in thin and moderate cloud conditions (Toronto has about 44 % of daylight hours with clear-sky conditions per year). The relative strength of direct-sun measured by a collocated Total Sky Imager (model TSI-880) is plotted on top of each panel in Figure 8 as an additional indicator of sky conditions. The relative strength of direct-sun is from the integration of blocking-strip luminance. In general, when the relative strength of direct-sun is high (> 60), good quality direct-sun and zenith-sky NO₂ data can both be produced. However, when sun strength is moderate (30-60), only zenith-sky NO₂ data are reliable. When sun strength is low (< 30), zenith-sky NO₂ has increased bias and needs to be filtered out.

5 Discussion

This study evaluated the performance of Pandora zenith-sky measurements with Pandora direct-sun measurements, satellite measurements, and in situ measurements. In general, the quality of zenith-sky data is affected by three main factors: (1) quality of empirical zenith-sky AMFs; (2) cloud conditions (heavy clouds or moderate/thin clouds); and (3) quality of modelled NO₂ profile (this factor only applies to Pandora surface NO₂ data). The quality of empirical zenith-sky AMFs and the cloud effect have been addressed in Appendices A and C, respectively. The third factor is discussed in Sections 5.1 and 5.2. **The uncertainty estimations for Pandora zenith-sky and direct-sun data products are provided in Appendix D.**

5.1 Diurnal and seasonal variation

From the Pandora zenith-sky and direct-sun measurements, and modelled NO₂ profiles, surface NO₂ concentrations were obtained that agree well with in situ measurements collected at the same location. The Pandora surface NO₂ data were also analyzed in more detail with a focus on temporal variations. Figure 9 shows the averaged surface NO₂ diurnal variations of four different datasets. The in situ instrument produces continuous measurements 24 hours per day, whereas Pandora only has measurements when sunlight is available. The diurnal variation of surface NO₂ concentration is controlled by dynamics (e.g., vertical mixing, wind direction), photochemistry, and local emissions. Thus, the diurnal variations are calculated using only the hours when in situ, direct-sun, and zenith-sky data are all available.

Figure 9 shows that all four datasets/curves captured the enhanced morning surface NO₂ and the decreasing trend afterwards.

However, the model has a positive offset (6-9 ppbv) in the morning (due in part to the use of older emissions inventories: Moran et al., 2018) and a negative offset (1-3 ppbv) in the evening relative to the in situ data. For example, at 7:00 LST, in situ NO₂ is 14.9 ± 9.3 ppbv, while GEM-MACH, Pandora DS, and Pandora ZS NO₂ are 23.5 ± 15.0 ppbv, 15.6 ± 10.5 ppbv, and 15.2 ± 6.8 ppbv, respectively. At 17:00 LST, in situ NO₂ is 7.3 ± 5.8 ppbv, while GEM-MACH, Pandora DS, and Pandora ZS NO₂ are 5.6 ± 5.0 ppbv, 3.6 ± 2.6 ppbv, and 5.2 ± 3.4 ppbv, respectively. The larger standard deviations in the morning are due to the datasets not being divided into work-days and weekends. Compared to the modelled data, the Pandora direct-sun and zenith-sky data show improvements in the morning, but almost no changes for the evening. This feature is investigated and found to be correlated with the GEM-MACH modelled PBL height (details in Section 5.2).

The diurnal variation is also examined by grouping the data by seasons. Figure 10 shows that the surface NO₂ concentrations in winter (December, January, and February) are higher than the corresponding values in summer (June, July, and August).

This difference is mainly due to short sunlit periods and less solar radiation (e.g., increased lifetime of NO₂ and decreased PBL height) in winter. The model has better agreement with the in situ data in summer than in the colder seasons. The best performance of the model is found around local noon, and this feature is not dependent on seasons. Figure 10 also shows that the quality of Pandora zenith-sky and direct-sun surface NO₂ estimates is affected by the quality of GEM-MACH modelled data. For example, Figure 10c shows that in autumn (September, October, and November), GEM-MACH has the largest offset in the morning. This error is thus propagated to the Pandora direct-sun surface data, and leads to a larger offset in the morning (than any other season). On the other hand, when GEM-MACH shows a better agreement with in situ measurements (e.g., in spring and summer), Pandora zenith-sky and direct-sun estimates also show better agreement with in situ observations. In general, both Pandora direct-sun and zenith-sky surface NO₂ data show good agreement with in situ measurements in all seasons; the hourly mean values of Pandora surface NO₂ are all well within the 1σ envelope of the in situ measurements.

5.2 Planetary boundary-layer effect

The larger morning offset in modelled surface NO₂ may indicate that the GEM-MACH modelled PBL heights are biased in the morning when the boundary layer is shallow. Figure 11 (left column) shows the modelled PBL height plotted as a function

of the difference between modelled and in situ surface NO₂. Figure 11a shows that, in general, the difference between modelled and in situ NO₂ decreases with an increase of PBL height. When the modelled PBL height is less than 100 m, the mean difference is 18 ± 12 ppbv (1σ), while when the modelled PBL height is 1 km, the mean difference is only 2.9 ± 6.4 ppbv.

Even though the modelled surface concentrations are significantly impacted by the PBL, the modelled conversion ratio (from column to surface concentrations) seems unaffected since the surface NO₂ concentrations derived from Pandora zenith-sky data ($C_{pan-model}$) show much less dependence on the PBL height (Figure 11b). When the modelled PBL height is less than 100 m, the mean difference is 0.9 ± 8.9 ppbv. When the modelled PBL height is 1 km, the mean difference is slightly improved to 0.1 ± 4.4 ppbv. Figures 11c and 11h show similar plots as Figure 11a and 11b, but the dataset has been divided into three time-bins (before 9:00, 11:00 to 13:59, and after 15:00). Figures 11c, 11e, and 11f confirm that whenever the modelled PBL height is low, the relative difference between the model and in situ data is high. However, in general, most of these shallow PBL height conditions occur in the morning, and thus the modelled surface NO₂ has larger bias compared to in situ data in the morning. Figures 11d, 11f, and 11h show that Pandora zenith-sky surface NO₂ data have similar performance for all these three time-bins, which indicates that the data have less PBL height dependency than the modelled data. In other words, the model is able to capture the ratio between the boundary layer partial column and surface NO₂, although the PBL height may not be correct in the model. When this ratio is applied to both Pandora direct-sun and zenith-sky data, the estimated surface concentrations agree better with the in situ measurements.

6 Conclusions

The Pandora spectrometer was originally designed to retrieve total columns of trace gases such as ozone and NO₂ from direct-sun spectral measurements in the UV-visible spectrum. In this work, a new zenith-sky total column NO₂ retrieval algorithm has been developed. The algorithm is based on empirical AMFs derived from nearly simultaneous direct-sun and zenith-sky measurements. It is demonstrated that this algorithm can retrieve total columns in thin and moderate cloud conditions when direct-sun measurements are not available: only 10 % of the measurements affected by heavy cloud have to be filtered out due to large systematic biases (68 %). The new Pandora zenith-sky total column NO₂ data shows only -4% bias compared to the standard Pandora direct-sun data product. In addition, OMI NO₂ SPv3.0 data demonstrate similar biases (-30 % and -29 %, respectively) when compared to direct-sun and zenith-sky Pandora total column NO₂ data.

Surface NO₂ concentrations were calculated from Pandora direct-sun and zenith-sky total column NO₂ using column-to-surface ratios derived from GEM-MACH regional chemical transport model. The bias between Pandora-based direct-sun and zenith-sky NO₂ surface concentration estimates and in situ measurements is only -8 % and -7 % (with correlation coefficients 0.80 and 0.77), respectively, while the bias between the modelled concentrations and in situ measurements is up to 37 %. The Pandora-based surface NO₂ concentrations also show good diurnal and seasonal variation when compared to the in situ data. High surface NO₂ concentrations in the morning (from 6:00 to 9:00, local standard time) are present in all measured and modelled datasets, while, on average, the model overestimates surface NO₂ in the morning by 8.6 ppbv (at 7:00 LST). It

appears that the bias in modelled surface NO₂ is related at least in part to an incorrectly diagnosed PBL height. In contrast, the difference between Pandora-based and in situ NO₂ does not show any significant dependence on the PBL height. Thus, to enable a fast and practical Pandora surface NO₂ data production, the use of a pre-calculated conversion ratio PSC-LUT is recommended.

5 The new retrieval algorithm for Pandora zenith-sky NO₂ measurements can provide high-quality NO₂ data (both total column and surface concentration) not only in clear-sky conditions, but also in thin and moderate cloud conditions, when direct-sun observations are not available. Long-term Pandora zenith-sky NO₂ data could be used in future satellite validation for the medium cloudy scenes. Moreover, a column-to-surface conversion look-up table was produced for the Pandora instruments deployed in Toronto; therefore, quick and practical Pandora-based surface NO₂ concentration data can be obtained for air
10 quality monitoring purposes. The variation of conversion ratios in the PSC-LUT demonstrates that the surface NO₂ concentration is controlled not only by the PBL height, but also by both boundary-layer dynamics and photochemistry. This conversion approach can also be used to derive surface concentrations from satellite VCD measurements and thus can be particularly useful for the new generation of geostationary satellite instruments for air quality monitoring such as the Tropospheric Emissions: Monitoring of Pollution (TEMPO, Zoogman et al., 2014). **Currently, the standard Pandora
15 observation schedule includes direct-sun, zenith-sky, and multi-axis scanning measurements (i.e., measuring at multiple viewing angles). At present, multi-axis measurement algorithms are still under development, but in the future, by using the multi-axis measurements and optimal estimation techniques (e.g., Rodgers, 2000) or the five angles O2O2-ratio algorithm (Cede, 2019), it may be possible for Pandora measurements to be used to derive NO₂ tropospheric profiles and columns.**

20

Data availability. Pandora data are available from the Pandonia network (<http://pandonia.net/data/>). In situ surface NO₂ data are available from the National Air Pollution Surveillance (NAPS) program (<http://maps-cartes.ec.gc.ca/rnspa-naps/data.aspx>). OMI NO₂ SPv3.0 data are available from <https://disc.gsfc.nasa.gov/>. Any additional data may be obtained from Xiaoyi Zhao (xiaoyi.zhao@canada.ca).

25

Author contributions. XZ analyzed the data and prepared the manuscript, with significant conceptual input from DG, VF, and CM, and critical feedbacks from all co-authors. JD, AO, VF, XZ, and SCL operated and managed the Canadian Pandora network. AL, MDM, and DG performed and analyzed the GEM-MACH simulations. AC, MT, and MM operated the Pandonia network and provided critical technical support to the Canadian Pandora measurements and subsequent data analysis.

30

Acknowledgements. Xiaoyi Zhao was supported by the NSERC Visiting Fellowships in Canadian Government Laboratories program. We thank Ihab Abboud and Reno Sit for their technical support of Pandora measurements. We thank NAPS for providing surface NO₂ data. We acknowledge the NASA Earth Science Division for providing OMI NO₂ SPv3.0 data. We also thank Thomas Danckaert, Caroline Fayt, Michel Van Roozendaal, and others from IASB-BIRA for providing the QDOAS

software, the NDACC UV-visible working group for providing NDACC UV-visible NO₂ AMF LUT, and Yushan Su from the Ontario Ministry of the Environment, Conservation and Parks for providing NAPS Toronto North station in situ NO₂ information. **We thank two anonymous referees for their helpful and insightful comments, which improved the overall quality of this work.**

5 References

- Adams, C., Strong, K., Zhao, X., Bassford, M. R., Chipperfield, M. P., Daffer, W., Drummond, J. R., Farahani, E. E., Feng, W., Fraser, A., Goutail, F., Manney, G., McLinden, C. A., Pazmino, A., Rex, M. and Walker, K. A.: Severe 2011 ozone depletion assessed with 11 years of ozone, NO₂, and OClO measurements at 80°N, *Geophys. Res. Lett.*, 39, L05806, doi:10.1029/2011gl050478, 2012.
- 10 Adams, C., Normand, E. N., McLinden, C. A., Bourassa, A. E., Lloyd, N. D., Degenstein, D. A., Krotkov, N. A., Belmonte Rivas, M., Boersma, K. F. and Eskes, H.: Limb–nadir matching using non-coincident NO₂ observations: proof of concept and the OMI-minus-OSIRIS prototype product, *Atmos. Meas. Tech.*, 9(8), 4103–4122, doi:10.5194/amt-9-4103-2016, 2016.
- Adams, C., Bourassa, A. E., McLinden, C. A., Sioris, C. E., Clarmann, T. von, Funke, B., Rieger, L. A. and Degenstein, D. A.: Effect of volcanic aerosol on stratospheric NO₂ and N₂O₅ from 2002–2014 as measured by Odin-OSIRIS and Envisat-MIPAS, *Atmos. Chem. Phys.*, 17(13), 8063–8080, doi:10.5194/acp-17-8063-2017, 2017.
- 15 Akingunola, A., Makar, P. A., Zhang, J., Darlington, A., Li, S.-M., Gordon, M., Moran, M. D. and Zheng, Q.: A chemical transport model study of plume-rise and particle size distribution for the Athabasca oil sands, *Atmos. Chem. Phys.*, 18(12), 8667–8688, doi:10.5194/acp-18-8667-2018, 2018.
- Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J. and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, *J. Geophys. Res.*, 106(D19), 23073–23095, doi:10.1029/2001JD000807, 2001.
- 20 Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O. C., Vogel, A., Hartmann, M., Kromminga, H., Bovensmann, H., Frerick, J. and Burrows, J. P.: Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: instrument characterization and reference data for atmospheric remote-sensing in the 230–2380 nm region, *J. Photochem. Photobiol. A*, 157, 167–184, doi:10.1016/s1010-6030(03)00062-5, 2003.
- 25 Brion, J., Chakir, A., Daumont, D., Malicet, J. and Parisse, C.: High-resolution laboratory absorption cross section of O₃. Temperature effect, *Chem. Phys. Lett.*, 213, 610–612, doi:10.1016/0009-2614(93)89169-I, 1993.
- Brion, J., Chakir, A., Charbonnier, J., Daumont, D., Parisse, C. and Malicet, J.: Absorption Spectra Measurements for the Ozone Molecule in the 350–830 nm Region, *J. Atmos. Chem.*, 30, 291–299, doi:10.1023/a:1006036924364, 1998.
- 30 Brohede, S., McLinden, C. A., Urban, J., Haley, C. S., Jonsson, A. I. and Murtagh, D.: Odin stratospheric proxy NO_y measurements and climatology, *Atmos. Chem. Phys.*, 8(19), 5731–5754, doi:10.5194/acp-8-5731-2008, 2008.
- Bucsela, E. J., Krotkov, N. A., Celarier, E. A., Lamsal, L. N., Swartz, W. H., Bhartia, P. K., Boersma, K. F., Veefkind, 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.*, 6(10), 2607–2626, doi:10.5194/amt-6-2607-2013, 2013.
- 35 Cede, A.: Manual for Blick Software Suite **1.6**, 2019.

- Cede, A., Herman, J., Richter, A., Krotkov, N. and Burrows, J.: Measurements of nitrogen dioxide total column amounts using a Brewer double spectrophotometer in direct Sun mode, *J. Geophys. Res.*, 111(D5), doi:10.1029/2005JD006585, 2006.
- Celarier, E. A., Brinksma, E. J., Gleason, J. F., Veefkind, J. P., Cede, A., Herman, J. R., Ionov, D., Goutail, F., Pommereau, J.-P., Lambert, J.-C., Roozendael, M. van, Pinardi, G., Wittrock, F., Schönhardt, A., Richter, A., Ibrahim, O. W., Wagner, T.,
5 Bojkov, B., Mount, G., Spinei, E., Chen, C. M., Pongetti, T. J., Sander, S. P., Bucsela, E. J., Wenig, M. O., Swart, D. P. J., Volten, H., Kroon, M. and Levelt, P. F.: Validation of Ozone Monitoring Instrument nitrogen dioxide columns, *J. Geophys. Res.*, doi:10.1029/2007JD008908, 2008.
- Celarier, E. A., Lamsal, L. N., Bucsela, E. J., Marchenko, S. V. and Krotkov, N. A.: OMNO2 Version 3.0 Level 2 File Description (Document version 3.1), NASA Goddard Space Flight Center., 2016.
- 10 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. Opt.*, 36, 5224–5230, doi:10.1364/AO.36.005224, 1997.
- Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M. and Staniforth, A.: The Operational CMC–MRB Global Environmental Multiscale (GEM) Model. Part I: Design Considerations and Formulation, *Mon. Wea. Rev.*, 126(6), 1373–1395, doi:10.1175/1520-0493(1998)126<1373:TOCMGE>2.0.CO;2, 1998.
- 15 Dabek-Zlotorzynska, E., Dann, T. F., Kalyani Martinelango, P., Celo, V., Brook, J. R., Mathieu, D., Ding, L. and Austin, C. C.: Canadian National Air Pollution Surveillance (NAPS) PM_{2.5} speciation program: Methodology and PM_{2.5} chemical composition for the years 2003–2008, *Atmos. Environ.*, 45(3), 673–686, doi:10.1016/j.atmosenv.2010.10.024, 2011.
- Danckaert, T., Fayt, C., Van Roozendael, M., de Smedt, I., Letocart, V., Merlaud, A. and Pinardi, G.: QDOAS Software user manual version 2.109, Software user manual, Belgian Institute for Space Aeronomy, Brussels., 2015.
- 20 Daumont, D., Brion, J., Charbonnier, J. and Malicet, J.: Ozone UV spectroscopy I: Absorption cross-sections at room temperature, *J. Atmos. Chem.*, 15, 145–155, doi:10.1007/bf00053756, 1992.
- Drosoglou, T., Bais, A. F., Zyrichidou, I., Kouremeti, N., Poupkou, A., Liora, N., Giannaros, C., Koukouli, M. E., Balis, D. and Melas, D.: Comparisons of ground-based tropospheric NO₂ MAX-DOAS measurements to satellite observations with the aid of an air quality model over the Thessaloniki area, Greece, *Atmos. Chem. Phys.*, 17(9), 5829–5849, doi:10.5194/acp-17-
25 5829-2017, 2017.
- Drosoglou, T., Koukouli, M. E., Kouremeti, N., Bais, A. F., Zyrichidou, I., Balis, D., van der A, R. J., Xu, J. and Li, A.: MAX-DOAS NO₂ observations over Guangzhou, China; ground-based and satellite comparisons, *Atmos. Meas. Tech.*, 11(4), 2239–2255, doi:10.5194/amt-11-2239-2018, 2018.
- ECCC: Canadian Environmental Sustainability Indicators: Air Quality. Environment and Climate Change Canada, ISBN: 978-
30 0-660-06016-3, 2016.
- EEA: Air quality in Europe — 2017 report, European Environment Agency, EEA Report No. 13/2017, 2017.
- EPA: Air Quality Index-A Guide to Air Quality and Your Health, U.S. Environmental Protection Agency, EPA-454/R-00-005, 2014.
- Fioletov, V. E., McLinden, C. A., McElroy, C. T. and Savastiouk, V.: New method for deriving total ozone from Brewer zenith
35 sky observations, *J. Geophys. Res.*, 116, D08301, doi:10.1029/2010JD015399, 2011.

- Flynn, C. M., Pickering, K. E., Crawford, J. H., Lamsal, L., Krotkov, N., Herman, J., Weinheimer, A., Chen, G., Liu, X., Szykman, J., Tsay, S.-C., Loughner, C., Hains, J., Lee, P., Dickerson, R. R., Stehr, J. W. and Brent, L.: Relationship between column-density and surface mixing ratio: Statistical analysis of O₃ and NO₂ data from the July 2011 Maryland DISCOVER-AQ mission, *Atmos. Environ.*, 92, 429–441, doi:10.1016/j.atmosenv.2014.04.041, 2014.
- 5 Frieß, U., Sihler, H., Sander, R., Pöhler, D., Yilmaz, S. and Platt, U.: The vertical distribution of BrO and aerosols in the Arctic: Measurements by active and passive differential optical absorption spectroscopy, *J. Geophys. Res.*, 116, D00R04, doi:10.1029/2011jd015938, 2011.
- Gao, R. S., Keim, E. R., Woodbridge, E. L., Ciciora, S. J., Proffitt, M. H., Thompson, T. L., McLaughlin, R. J. and Fahey, D. W.: New photolysis system for NO₂ measurements in the lower stratosphere, *J Geophys Res-Atmos*, 99(D10), 20673–20681, doi:10.1029/94JD01521, 1994.
- 10 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.*, 7, 3509–3527, doi:10.5194/amt-7-3509-2014, 2014.
- Goldberg, D. L., Lamsal, L. N., Loughner, C. P., Swartz, W. H., Lu, Z. and Streets, D. G.: A high-resolution and observationally constrained OMI NO₂ satellite retrieval, *Atmos. Chem. Phys.*, 17(18), 11403–11421, doi:10.5194/acp-17-11403-2017, 2017.
- 15 de Graaf, M., Sihler, H., Tilstra, L. G. and Stammes, P.: How big is an OMI pixel?, *Atmos. Meas. Tech.*, 9(8), 3607–3618, doi:10.5194/amt-9-3607-2016, 2016.
- Greenblatt, G. D., Orlando, J. J., Burkholder, J. B. and Ravishankara, A. R.: Absorption measurements of oxygen between 330 and 1140 nm, *J. Geophys. Res.*, 95, 18577–18582, doi:10.1029/JD095iD11p18577, 1990.
- 20 Hendrick, F., Van Roozendaal, M., Kylling, A., Petritoli, A., Rozanov, A., Sanghavi, S., Schofield, R., von Friedeburg, C., Wagner, T., Wittrock, F., Fonteyn, D. and De Mazière, M.: Intercomparison exercise between different radiative transfer models used for the interpretation of ground-based zenith-sky and multi-axis DOAS observations, *Atmos. Chem. Phys.*, 6(1), 93–108, doi:10.5194/acp-6-93-2006, 2006.
- 25 Hendrick, F., Pommereau, J. P., Goutail, F., Evans, R. D., Ionov, D., Pazmino, A., Kyrö, E., Held, G., Eriksen, P., Dorokhov, V., Gil, M. and Van Roozendaal, M.: NDACC/SAOZ UV-visible total ozone measurements: improved retrieval and comparison with correlative ground-based and satellite observations, *Atmos. Chem. Phys.*, 11, 5975–5995, doi:10.5194/acp-11-5975-2011, 2011.
- Hendrick, F., Müller, J.-F., Clémer, K., Wang, P., Mazière, M. D., Fayt, C., Gielen, C., Hermans, C., Ma, J. Z., Pinardi, G., Stavrou, T., Vlemmix, T. and Roozendaal, M. V.: Four years of ground-based MAX-DOAS observations of HONO and NO₂ in the Beijing area, *Atmos. Chem. Phys.*, 14(2), 765–781, doi:10.5194/acp-14-765-2014, 2014.
- 30 Herman, J., Cede, A., Spinei, E., Mount, G., Tzortziou, M. and Abuhassan, N.: NO₂ column amounts from ground-based Pandora and MFDOAS spectrometers using the direct-sun DOAS technique: Intercomparisons and application to OMI validation, *J. Geophys. Res.*, 114(D13), D13307, doi:10.1029/2009JD011848, 2009.
- 35 Herman, J., Evans, R., Cede, A., Abuhassan, N., Petropavlovskikh, I. and McConville, G.: Comparison of ozone retrievals from the Pandora spectrometer system and Dobson spectrophotometer in Boulder, Colorado, *Atmos. Meas. Tech.*, 8, 3407–3418, doi:10.5194/amt-8-3407-2015, 2015.

- Hermans, C., Vandaele, A. C., Fally, S., Carleer, M., Colin, R., Coquart, B., Jenouvrier, A. and Merienne, M.-F.: Absorption cross-section of the collision-induced bands of oxygen from the UV to the NIR, in *Weakly Interacting Molecular Pairs: Unconventional Absorbers of Radiation in the Atmosphere*, edited by C. Camy-Peyret and A. A. Vigasin, pp. 193–202, Springer, Germany., 2003.
- 5 Ialongo, I., Herman, J., Krotkov, N., Lamsal, L., Boersma, K. F., Hovila, J. and Tamminen, J.: Comparison of OMI NO₂ observations and their seasonal and weekly cycles with ground-based measurements in Helsinki, *Atmos. Meas. Tech.*, 9(10), 5203–5212, doi:10.5194/amt-9-5203-2016, 2016.
- Irie, H., Kanaya, Y., Akimoto, H., Tanimoto, H., Wang, Z., Gleason, J. F. and Bucsel, E. J.: Validation of OMI tropospheric NO₂ column data using MAX-DOAS measurements deep inside the North China Plain in June 2006: Mount Tai Experiment
 10 2006, *Atmos. Chem. Phys.*, 8(22), 6577–6586, doi:10.5194/acp-8-6577-2008, 2008.
- Kelly, T. J., Spicer, C. W. and Ward, G. F.: An assessment of the luminol chemiluminescence technique for measurement of NO₂ in ambient air, *Atmos. Environ.*, 24(9), 2397–2403, doi:10.1016/0960-1686(90)90332-H, 1990.
- Kerr, J. B., McElroy, C. T. and Olafson, R. A.: Measurements of ozone with the Brewer ozone spectrophotometer, in *Proceedings of the Quadrennial Ozone Symposium*, pp. 74–79, Boulder, U.S.A., 1981.
- 15 Kerr, J. B., Asbridge, I. A. and Evans, W. F. J.: Intercomparison of total ozone measured by the Brewer and Dobson spectrophotometers at Toronto, *J. Geophys. Res.*, 93, 11129–11140, doi:10.1029/JD093iD09p11129, 1988.
- Knepp, T., Pippin, M., Crawford, J., Chen, G., Szykman, J., Long, R., Cowen, L., Cede, A., Abuhassan, N., Herman, J., Delgado, R., Compton, J., Berkoff, T., Fishman, J., Martins, D., Stauffer, R., Thompson, A. M., Weinheimer, A., Knapp, D., Montzka, D., Lenschow, D. and Neil, D.: Estimating surface NO₂ and SO₂ mixing ratios from fast-response total column
 20 observations and potential application to geostationary missions, *J. Atmos. Chem.*, 72(3–4), 261–286, doi:10.1007/s10874-013-9257-6, 2015.
- Knepp, T. N., Querel, R., Johnston, P., Thomason, L., Flittner, D. and Zawodny, J. M.: Intercomparison of Pandora stratospheric NO₂ slant column product with the NDACC-certified M07 spectrometer in Lauder, New Zealand, *Atmos. Meas. Tech.*, 10(11), 4363–4372, doi:10.5194/amt-10-4363-2017, 2017.
- 25 Koenker, R. and Hallock, K. F.: Quantile Regression, *J. Econo. Perspec.*, 15(4), 143–156, doi:10.1257/jep.15.4.143, 2001.
- Kollonige, D. E., Thompson, A. M., Josipovic, M., Tzortziou, M., Beukes, J. P., Burger, R., Martins, D. K., Zyl, P. G. van, Vakkari, V. and Laakso, L.: OMI Satellite and Ground-Based Pandora Observations and Their Application to Surface NO₂ Estimations at Terrestrial and Marine Sites, *J. Geophys. Res.*, 123(2), 1441–1459, doi:10.1002/2017JD026518, 2017.
- Kramer, L. J., Leigh, R. J., Remedios, J. J. and Monks, P. S.: Comparison of OMI and ground-based in situ and MAX-DOAS measurements of tropospheric nitrogen dioxide in an urban area, *J. Geophys. Res.*, 113(D16), doi:10.1029/2007JD009168, 2008.
 30
- Krotkov, N. A., Lamsal, L. N., Celarier, E. A., Swartz, W. H., Marchenko, S. V., Bucsel, E. J., Chan, K. L., Wenig, M. and Zara, M.: The version 3 OMI NO₂ standard product, *Atmos. Meas. Tech.*, 10(9), 3133–3149, doi:10.5194/amt-10-3133-2017, 2017.
- 35 Lambert, J.-C., Granville, J., Van Roozendaal, M., Sarkissian, A., Goutail, F., Müller, J.-F., Pommereau, J.-P. and Russell III, J. M.: A climatology of NO₂ profile for improved Air Mass Factors for ground-based vertical column measurements, in “Stratospheric Ozone 1999”, Air Pollution Research Report 73 (CEC DG XII)., 1999.

- Lambert, J.-C., Granville, J., Van Roozendaal, M., Müller, J.-F., Goutail, F., Pommereau, J.-P., Sarkissian, A., Johnston, P. V. and Russell III, J. M.: in *Global Behaviour of Atmospheric NO₂ as Derived from the Integrated Use of Satellite, Ground-based Network and Balloon Observations*, pp. 201–202, Sapporo, Japan., 2000.
- 5 Lamsal, L. N., Martin, R. V., Donkelaar, A. van, Steinbacher, M., Celarier, E. A., Bucsela, E., Dunlea, E. J. and Pinto, J. P.: Ground-level nitrogen dioxide concentrations inferred from the satellite-borne Ozone Monitoring Instrument, *J. Geophys. Res.*, 113(D16), doi:10.1029/2007JD009235, 2008.
- 10 Lamsal, L. N., Krotkov, N. A., Celarier, E. A., Swartz, W. H., Pickering, K. E., Bucsela, E. J., Gleason, J. F., Martin, R. V., Philip, S., Irie, H., Cede, A., Herman, J., Weinheimer, A., Szykman, J. J. and Knepp, T. N.: Evaluation of OMI operational standard NO₂ column retrievals using in situ and surface-based NO₂ observations, *Atmos. Chem. Phys.*, 14(21), 11587–11609, doi:10.5194/acp-14-11587-2014, 2014.
- Lamsal, L. N., Janz, S. J., Krotkov, N. A., Pickering, K. E., Spurr, R. J. D., Kowalewski, M. G., Loughner, C. P., Crawford, J. H., Swartz, W. H. and Herman, J. R.: High-resolution NO₂ observations from the Airborne Compact Atmospheric Mapper: Retrieval and validation, *J. Geophys. Res.*, 122(3), 1953–1970, doi:10.1002/2016JD025483, 2017.
- 15 Lee, A. M., Roscoe, H. K., Oldham, D. J., Squires, J. A. C., Sarkissian, A., Pommereau, J.-P. and Gardiner, B. G.: Improvements to the accuracy of measurements of NO₂ by zenith-sky visible spectrometers, *J. Quant. Spectrosc. Radiat. Transfer*, 52(5), 649–657, doi:10.1016/0022-4073(94)90031-0, 1994.
- Levelt, P. F., Hilsenrath, E., Leppelmeier, G. W., Van den Oord, G. H., Bhartia, P. K., Tamminen, J., De Haan, J. F. and Veefkind, J. P.: Science objectives of the ozone monitoring instrument, *IEEE Trans. Geosci. Remote Sens.*, 44, 1199–1208, doi:10.1109/TGRS.2006.872336, 2006.
- 20 Levelt, P. F., Joiner, J., Tamminen, J., Veefkind, J. P., Bhartia, P. K., Stein Zweers, D. C., Duncan, B. N., Streets, D. G., Eskes, H., A. R. van der, McLinden, C., Fioletov, V., Carn, S., Laatz, J. de, DeLand, M., Marchenko, S., McPeters, R., Ziemke, J., Fu, D., Liu, X., Pickering, K., Apituley, A., González Abad, G., Arola, A., Boersma, F., Chan Miller, C., Chance, K., Graaf, M. de, Hakkarainen, J., Hassinen, S., Ialongo, I., Kleipool, Q., Krotkov, N., Li, C., Lamsal, L., Newman, P., Nowlan, C., Suleiman, R., Tilstra, L. G., Torres, O., Wang, H. and Wargan, K.: The Ozone Monitoring Instrument: overview of 14 years in space, 25 *Atmos. Chem. Phys.*, 18(8), 5699–5745, doi:10.5194/acp-18-5699-2018, 2018.
- Lindenmaier, R., Strong, K., Batchelor, R. L., Bernath, P. F., Chabrilat, S., Chipperfield, M. P., Daffer, W. H., Drummond, J. R., Feng, W., Jonsson, A. I., Kolonjari, F., Manney, G. L., McLinden, C., Ménard, R. and Walker, K. A.: A study of the Arctic NO_y budget above Eureka, Canada, *J. Geophys. Res.*, 116(D23), doi:10.1029/2011JD016207, 2011.
- 30 Maeda, Yasuaki., Aoki, Kenji. and Munemori, Makoto.: Chemiluminescence method for the determination of nitrogen dioxide, *Anal. Chem.*, 52(2), 307–311, doi:10.1021/ac50052a022, 1980.
- Marchenko, S., Krotkov, N. A., Lamsal, L. N., Celarier, E. A., Swartz, W. H. and Bucsela, E. J.: Revising the slant column density retrieval of nitrogen dioxide observed by the Ozone Monitoring Instrument, *J. Geophys. Res.*, 120(11), 5670–5692, doi:10.1002/2014JD022913, 2015.
- 35 Martin, R. V., Chance, K., Jacob, D. J., Kurosu, T. P., Spurr, R. J. D., Bucsela, E., Gleason, J. F., Palmer, P. I., Bey, I., Fiore, A. M., Li, Q., Yantosca, R. M. and Koelmeijer, R. B. A.: An improved retrieval of tropospheric nitrogen dioxide from GOME, *J. Geophys. Res.*, 107(D20), ACH 9-1-ACH 9-21, doi:10.1029/2001JD001027, 2002.
- Martins, D. K., Najjar, R. G., Tzortziou, M., Abuhassan, N., Thompson, A. M. and Kollonige, D. E.: Spatial and temporal variability of ground and satellite column measurements of NO₂ and O₃ over the Atlantic Ocean during the Deposition of

- Atmospheric Nitrogen to Coastal Ecosystems Experiment, J. Geophys. Res., 121(23), 14,175–14,187, doi:10.1002/2016JD024998, 2016.
- McClenny, W. A.: Recommended methods for ambient air monitoring of NO, NO₂, NO_y and individual NO₂ species, Tech. rep., US Environmental Protection Agency, Triangle Park, NC 27711., 2000.
- 5 McLinden, C. A., Olsen, S. C., Hannegan, B., Wild, O., Prather, M. J. and Sundet, J.: Stratospheric ozone in 3-D models: A simple chemistry and the cross-tropopause flux, J. Geophys. Res., 105, 14653–14665, doi:10.1029/2000jd900124, 2000.
- McLinden, C. A., Fioletov, V., Boersma, K. F., Kharol, S. K., Krotkov, N., Lamsal, L., Makar, P. A., Martin, R. V., Veefkind, J. P. and Yang, K.: Improved satellite retrievals of NO₂ and SO₂ over the Canadian oil sands and comparisons with surface measurements, Atmos. Chem. Phys., 14(7), 3637–3656, doi:10.5194/acp-14-3637-2014, 2014.
- 10 Moran, M. D., Menard, S., Talbot, D., Huang, P., Makar, P. A., Gong, W., Landry, H., Gong, S., Gravel, S., Crevier, L.-P. and Kallaur, A.: Particulate-matter forecasting with GEM-MACH15, a new Canadian operational air quality forecast model, in 30th NATO/SPS ITM on Air Pollution Modelling and Its Application, pp. 289–293, Springer, San Francisco., 2009.
- Moran, M. D., Pavlovic, R. and Anselmo, D.: Regional Air Quality Deterministic Prediction System (RAQDPS): Update from version 019 to version 020. Technical note, Sept., Canadian Centre for Meteorological and Environmental Prediction,
- 15 Montreal, 43 pp., [online] Available from: http://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/tech_notes/technote_raqdps-v20_20180918_e.pdf, 2018.
- Noxon, J. F.: Nitrogen dioxide in the stratosphere and troposphere measured by ground-based absorption spectroscopy, Science, 189, 547–549, 1975.
- NRC: Rethinking the Ozone Problem in Urban and Regional Air Pollution, National Research Council, National Academy Press, Washington, D.C., 1992.
- 20 Park, R. J., Jacob, D. J., Field, B. D., Yantosca, R. M. and Chin, M.: Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy, J. Geophys. Res., 109(D15), doi:10.1029/2003JD004473, 2004.
- Pavlovic, R., Chen, J., Anderson, K., Moran, M. D., Beaulieu, P.-A., Davignon, D. and Cousineau, S.: The FireWork air quality forecast system with near-real-time biomass burning emissions: Recent developments and evaluation of performance for the 2015 North American wildfire season, J. Air Waste Manag. Assoc., 66(9), 819–841, doi:10.1080/10962247.2016.1158214, 2016.
- 25 Pendlebury, D., Gravel, S., Moran, M. D. and Lupu, A.: Impact of chemical lateral boundary conditions in a regional air quality forecast model on surface ozone predictions during stratospheric intrusions, Atmos. Environ., 174, 148–170, doi:10.1016/j.atmosenv.2017.10.052, 2018.
- 30 Piters, A. J. M., Boersma, K. F., Kroon, M., Hains, J. C., Van Roozendael, M., Wittrock, F., Abuhassan, N., Adams, C., Akrami, M., Allaart, M. A. F., Apituley, A., Beirle, S., Bergwerff, J. B., Berkhout, A. J. C., Brunner, D., Cede, A., Chong, J., Clémer, K., Fayt, C., Frieß, U., Gast, L. F. L., Gil-Ojeda, M., Goutail, F., Graves, R., Griesfeller, A., Großmann, K., Hemerijckx, G., Hendrick, F., Henzing, B., Herman, J., Hermans, C., Hoexum, M., van der Hoff, G. R., Irie, H., Johnston, P. V., Kanaya, Y., Kim, Y. J., Klein Baltink, H., Kreher, K., de Leeuw, G., Leigh, R., Merlaud, A., Moerman, M. M., Monks, P. S., Mount, G. H., Navarro-Comas, M., Oetjen, H., Pazmino, A., Perez-Camacho, M., Peters, E., du Piesanie, A., Pinardi, G., Puentedura, O., Richter, A., Roscoe, H. K., Schönhardt, A., Schwarzenbach, B., Shaiganfar, R., Sluis, W., Spinei, E., Stolk, A. P., Strong, K., Swart, D. P. J., Takashima, H., Vlemmix, T., Vrekoussis, M., Wagner, T., Whyte, C., Wilson, K. M., Yela,

- M., Yilmaz, S., Zieger, P. and Zhou, Y.: The Cabauw Intercomparison campaign for Nitrogen Dioxide measuring Instruments (CINDI): design, execution, and early results, *Atmos. Meas. Tech.*, 5, 457–485, doi:10.5194/amt-5-457-2012, 2012.
- Platt, U.: Differential optical absorption spectroscopy (DOAS), in *Air Monitoring by Spectroscopic Techniques*, pp. 27–84, John Wiley, New York., 1994.
- 5 Platt, U. and Stutz, J.: *Differential Optical Absorption Spectroscopy: Principles and Applications*, Springer, Germany, Berlin., 2008.
- Reed, A. J., Thompson, A. M., Kollonige, D. E., Martins, D. K., Tzortziou, M. A., Herman, J. R., Berkoff, T. A., Abuhassan, N. K. and Cede, A.: Effects of local meteorology and aerosols on ozone and nitrogen dioxide retrievals from OMI and pandora spectrometers in Maryland, USA during DISCOVER-AQ 2011, *J. Atmos. Chem.*, 72(3–4), 455–482, doi:10.1007/s10874-013-9254-9, 2015.
- 10 Reid, H. and Aherne, J.: Staggering reductions in atmospheric nitrogen dioxide across Canada in response to legislated transportation emissions reductions, *Atmos. Environ.*, 146, 252–260, doi:10.1016/j.atmosenv.2016.09.032, 2016.
- Rodgers, C. D.: *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific, Singapore., 2000.**
- Roscoe, H. K., Van Roozendaal, M., Fayt, C., du Piesanie, A., Abuhassan, N., Adams, C., Akrami, M., Cede, A., Chong, J., Clemer, K., Friess, U., Ojeda, M. G., Goutail, F., Graves, R., Griesfeller, A., Grossmann, K., Hemerijckx, G., Hendrick, F., Herman, J., Hermans, C., Irie, H., Johnston, P. V., Kanaya, Y., Kreher, K., Leigh, R., Merlaud, A., Mount, G. H., Navarro, M., Oetjen, H., Pazmino, A., Perez-Camacho, M., Peters, E., Pinardi, G., Puentedura, O., Richter, A., Schonhardt, A., Shaiganfar, R., Spinei, E., Strong, K., Takashima, H., Vlemmix, T., Vrekoussis, M., Wagner, T., Wittrock, F., Yela, M., Yilmaz, S., Boersma, F., Hains, J., Kroon, M., Peters, A. and Kim, Y. J.: Intercomparison of slant column measurements of NO₂ and O₄ by MAX-DOAS and zenith-sky UV and visible spectrometers, *Atmos. Meas. Tech.*, 3, 1629–1646, doi:10.5194/amt-3-1629-2010, 2010.
- 15 20 Rothman, L. S., Jacquemart, D., Barbe, A., Chris Benner, D., Birk, M., Brown, L. R., Carleer, M. R., Chackerian, J. C., Chance, K., Coudert, L. H., Dana, V., Devi, V. M., Flaud, J. M., Gamache, R. R., Goldman, A., Hartmann, J. M., Jucks, K. W., Maki, A. G., Mandin, J. Y., Massie, S. T., Orphal, J., Perrin, A., Rinsland, C. P., Smith, M. A. H., Tennyson, J., Tolchenov, R. N., Toth, R. A., Vander Auwera, J., Varanasi, P. and Wagner, G.: The HITRAN 2004 molecular spectroscopic database, *J. Quant. Spectrosc. Radiat. Transfer*, 96, 139–204, doi:10.1016/j.jqsrt.2004.10.008, 2005.
- 25 Ryerson, T. B., Williams, E. J. and Fehsenfeld, F. C.: An efficient photolysis system for fast-response NO₂ measurements, *J. Geophys. Res.*, 105(D21), 26447–26461, doi:10.1029/2000JD900389, 2000.
- Sarkissian, A., Roscoe, H. K., Fish, D., Van Roozendaal, M., Gil, M., Chen, H. B., Wang, P., Pommereau, J. P. and Lenoble, J.: Ozone and NO₂ Air-Mass Factors for Zenith-Sky Spectrometers - Intercomparison of Calculations with Different Radiative-Transfer Models, *Geophys. Res. Lett.*, 22, 1113–1116, 1995.
- 30 Sauvage, B., Martin, R. V., Donkelaar, A. van, Liu, X., Chance, K., Jaeglé, L., Palmer, P. I., Wu, S. and Fu, T.-M.: Remote sensed and in situ constraints on processes affecting tropical tropospheric ozone, *Atmos. Chem. Phys.*, 7(3), 815–838, doi:10.5194/acp-7-815-2007, 2007.
- 35 Seber, G. A. F. and Wild, C. J.: *Nonlinear Regression*, Wiley-Interscience, Hoboken, NJ., 2003.
- Shettle, E. P.: Models of aerosols, clouds, and precipitation for atmospheric propagation studies, in *AGARD Conference Proceedings No. 454: Atmospheric Propagation in the UV, Visible, IR, and MM-Wave Region and Related Systems Aspects*, Neuilly sur Seine, France., 1989.

- Solomon, S., Schmeltekopf, A. and Sanders, R.: On the interpretation of zenith sky absorption measurements, *J. Geophys. Res.*, 2, 8311–8319, 1987.
- Spinei, E., Whitehill, A., Fried, A., Tiefengraber, M., Knepp, T. N., Herndon, S., Herman, J. R., Müller, M., Abuhassan, N., Cede, A., Richter, D., Walega, J., Crawford, J., Szykman, J., Valin, L., Williams, D. J., Long, R., Swap, R. J., Lee, Y., Nowak, N. and Poche, B.: The first evaluation of formaldehyde column observations by improved Pandora spectrometers during the KORUS-AQ field study, *Atmos. Meas. Tech.*, 11(9), 4943–4961, doi:10.5194/amt-11-4943-2018, 2018.
- Stieb, D. M., Burnett, R. T., Smith-Doiron, M., Brion, O., Shin, H. H. and Economou, V.: A New Multipollutant, No-Threshold Air Quality Health Index Based on Short-Term Associations Observed in Daily Time-Series Analyses, *J. Air Waste Manag. Assoc.*, 58(3), 435–450, doi:10.3155/1047-3289.58.3.435, 2008.
- 10 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(6), 2417–2435, doi:10.5194/amt-8-2417-2015, 2015.
- Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O₂-O₂ collision pairs between 340 and 630 nm and at atmospherically relevant pressure, *Phys. Chem. Chem. Phys.*, 15, 15371–15381, 2013.
- 15 Thornton, J. A., Wooldridge, P. J. and Cohen, R. C.: Atmospheric NO₂: In Situ Laser-Induced Fluorescence Detection at Parts per Trillion Mixing Ratios, *Anal. Chem.*, 72(3), 528–539, doi:10.1021/ac9908905, 2000.
- Tzortziou, M., Herman, J. R., Cede, A. and Abuhassan, N.: High precision, absolute total column ozone measurements from the Pandora spectrometer system: Comparisons with data from a Brewer double monochromator and Aura OMI, *J. Geophys. Res.*, 117, doi:10.1029/2012JD017814, 2012.
- 20 Van Roozendael, M. and Hendrick, F.: Recommendations for total ozone retrieval from NDACC zenith-sky UV-VIS spectrometers, Belgian Institute for Space Aeronomy, Brussels., 2009.
- Van Roozendael, M. and Hendrick, F.: Recommendations for NO₂ column retrieval from NDACC zenith-sky UV-VIS spectrometers, Belgian Institute for Space Aeronomy, Brussels., 2012.
- 25 Van Roozendael, M., Peeters, P., Roscoe, H. K., De Backer, H., Jones, A. E., Bartlett, L., Vaughan, G., Goutail, F., Pommereau, J.-P. and Kyro, E.: Validation of ground-based visible measurements of total ozone by comparison with Dobson and Brewer spectrophotometers, *J. Atmos. Chem.*, 29, 55–83, doi:10.1023/A:1005815902581, 1998.
- 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 42 000 cm⁻¹ to 10 000 cm⁻¹ (238 -1000 nm) at 220 K and 294 K, *J. Quant. Spectrosc. Radiat. Transfer*, 59, 171–184, doi:10.1016/s0022-4073(97)00168-4, 1998.
- 30 Vaughan, G., Roscoe, H. K., Bartlett, L. M., OConnor, F. M., Sarkissian, A., Van Roozendael, M., Lambert, J. C., Simon, P. C., Karlsen, K., Hoiskar, B. A. K., Fish, D. J., Jones, R. L., Freshwater, R. A., Pommereau, J. P., Goutail, F., Andersen, S. B., Drew, D. G., Hughes, P. A., Moore, D., Mellqvist, J., Hegels, E., Klupfel, T., Erle, F., Pfeilsticker, K. and Platt, U.: An intercomparison of ground-based UV-visible sensors of ozone and NO₂, *J. Geophys. Res.*, 102, 1411–1422, doi:10.1029/96JD00515, 1997.
- 35 Wagner, T., Dix, B., Friedeburg, C. v., Frieß, U., Sanghavi, S., Sinreich, R. and Platt, U.: MAX-DOAS O₄ measurements: A new technique to derive information on atmospheric aerosols: Principles and information content, *J. Geophys. Res.*, 109, D22205, doi:10.1029/2004jd004904, 2004.

- Wagner, T., Burrows, J. P., Deutschmann, T., Dix, B., von Friedeburg, C., Frieß, U., Hendrick, F., Heue, K. P., 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 Roozendaal, 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, doi:10.5194/acp-7-1809-2007, 2007.
- Wagner, T., Beirle, S., Brauers, T., Deutschmann, T., Frieß, U., Hak, C., Halla, J. D., Heue, K. P., Junkermann, W., Li, X., Platt, U. and Pundt-Gruber, I.: Inversion of tropospheric profiles of aerosol extinction and HCHO and NO₂ mixing ratios from MAX-DOAS observations in Milano during the summer of 2003 and comparison with independent data sets, *Atmos. Meas. Tech.*, 4, 2685–2715, doi:10.5194/amt-4-2685-2011, 2011.
- Wagner, T., Apituley, A., Beirle, S., Dörner, S., Friess, U., Remmers, J. and Shaiganfar, R.: Cloud detection and classification based on MAX-DOAS observations, *Atmos. Meas. Tech.*, 7, 1289–1320, doi:10.5194/amt-7-1289-2014, 2014.
- Wagner, T., Beirle, S., Remmers, J., Shaiganfar, R. and Wang, Y.: Absolute calibration of the colour index and O₄ absorption derived from Multi AXis (MAX-)DOAS measurements and their application to a standardised cloud classification algorithm, *Atmos. Meas. Tech.*, 9, 4803–4823, doi:10.5194/amt-9-4803-2016, 2016.
- Wagner, T., Beirle, S., Benavent, N., Bösch, T., Chan, K. L., Donner, S., Dörner, S., Fayt, C., Frieß, U., García-Nieto, D., Gielen, C., González-Bartolome, D., Gomez, L., Hendrick, F., Henzing, B., Jin, J. L., Lampel, J., Ma, J., Mies, K., Navarro, M., Peters, E., Pinardi, G., Puertedura, O., Puķīte, J., Remmers, J., Richter, A., Saiz-Lopez, A., Shaiganfar, R., Sihler, H., Roozendaal, M. V., Wang, Y. and Yela, M.: Is a scaling factor required to obtain closure between measured and modelled atmospheric O₄ absorptions? An assessment of uncertainties of measurements and radiative transfer simulations for 2 selected days during the MAD-CAT campaign, *Atmospheric Measurement Techniques*, 12(5), 2745–2817, doi:https://doi.org/10.5194/amt-12-2745-2019, 2019.
- Wang, S., Pongetti, T. J., Sander, S. P., Spinei, E., Mount, G. H., Cede, A. and Herman, J.: Direct Sun measurements of NO₂ column abundances from Table Mountain, California: Intercomparison of low- and high-resolution spectrometers, *J. Geophys. Res.*, 115(D13), doi:10.1029/2009JD013503, 2010.
- Wang, Y., Penning de Vries, M., Xie, P. H., Beirle, S., Dörner, S., Remmers, J., Li, A. and Wagner, T.: Cloud and aerosol classification for 2.5 years of MAX-DOAS observations in Wuxi (China) and comparison to independent data sets, *Atmos. Meas. Tech.*, 8, 5133–5156, doi:10.5194/amt-8-5133-2015, 2015.
- Wendel, G. J., Stedman, D. H., Cantrell, C. A. and Damrauer, Lenore.: Luminol-based nitrogen dioxide detector, *Anal. Chem.*, 55(6), 937–940, doi:10.1021/ac00257a027, 1983.
- Wenig, M. O., Cede, A. M., Bucsela, E. J., Celarier, E. A., Boersma, K. F., Veefkind, J. P., Brinksma, E. J., Gleason, J. F. and Herman, J. R.: Validation of OMI tropospheric NO₂ column densities using direct-Sun mode Brewer measurements at NASA Goddard Space Flight Center, *J. Geophys. Res.*, 113(D16), doi:10.1029/2007JD008988, 2008.
- WHO: Evolution of WHO air quality guidelines: past, present and future, WHO Regional Office for Europe, Copenhagen, 2017.
- Zhang, J., Moran, M. D., Zheng, Q., Makar, P. A., Baratzadeh, P., Marson, G., Liu, P. and Li, S.-M.: Emissions preparation and analysis for multiscale air quality modeling over the Athabasca Oil Sands Region of Alberta, Canada, *Atmos. Chem. Phys.*, 18(14), 10459–10481, doi:doi.org/10.5194/acp-18-10459-2018, 2018.

Zhao, X., Fioletov, V., Cede, A., Davies, J. and Strong, K.: Accuracy, precision, and temperature dependence of Pandora total ozone measurements estimated from a comparison with the Brewer triad in Toronto, *Atmos. Meas. Tech.*, 9, 5747–5761, doi:10.5194/amt-9-5747-2016, 2016.

- 5 Zhao, X., Bogner, K., Fioletov, V., Pazmino, A., Goutail, F., Millán, L., Manney, G., Adams, C. and Strong, K.: Assessing the impact of clouds on ground-based UV–visible total column ozone measurements in the high Arctic, *Atmos. Meas. Tech.*, 12(4), 2463–2483, doi:10.5194/amt-12-2463-2019, 2019.

Zoogman, P., Jacob, D. J., Chance, K., Liu, X., Lin, M., Fiore, A. and Travis, K.: Monitoring high-ozone events in the US Intermountain West using TEMPO geostationary satellite observations, *Atmos. Chem. Phys.*, 14(12), 6261–6271, doi:10.5194/acp-14-6261-2014, 2014.

10

Appendices

A. Empirical zenith-sky AMF

Before calculating the empirical zenith-sky AMF, the VCD_{DS} and $dSCD_{ZS}$ have both been strictly filtered to ensure any measurements used in this calculation have the highest quality. For VCD_{DS} , data are filtered following Cede (2019) with
 15 several factors being considered, such as wavelength shift and residual in spectra fitting, direct-sun AMF, and estimated uncertainties for the vertical column. For $dSCD_{ZS}$, data are filtered using similar criteria as for VCD_{DS} , with adjustments for zenith-sky observations.

The VCD_{DS} and $dSCD_{ZS}$ data are merged and divided into several SZA bins. Each bin covers 5° . A multi-non-linear regression is performed by using the following equation:

$$20 \begin{bmatrix} VCD_1 \\ \vdots \\ VCD_n \end{bmatrix} = \begin{bmatrix} dSCD_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & dSCD_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} + RCD \begin{bmatrix} I_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & I_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \quad (4)$$

where, VCD_n is not a single direct-sun VCD data point, but is an $m \times 1$ matrix (m is the total number of measurements in SZA bin number n); the VCD_n represents all direct-sun VCDs in a 5° SZA bin, and each element of the $m \times 1$ matrix is a single VCD in that SZA bin. Similarly, $dSCD_n$ is also an $m \times 1$ matrix, with each element representing a single coincident zenith-sky dSCD in SZA bin number n . I_n is an $m \times 1$ indicator function, where the elements of I_n are set to 1. The RCD and b_1 to b_n are
 25 the parameters to be retrieved. In short, the design of this regression is based on Eqn. 2 (Section 3.1). The idea is to retrieve zenith-sky AMFs in several SZA bins, and, at the same time, all these regressions in different SZA bins are constrained to share a common predictor (RCD). The regression model can be solved by using an iterative procedure (Seber and Wild, 2003) to yield the estimated coefficients, b_1 to b_n and RCD . The b_n is the reciprocal of zenith-sky AMF in SZA bin n .

This regression model has been evaluated by using different sizes for the SZA bins. A 5° SZA bin is selected because the SZA
 30 bin must be small enough to capture the SZA dependency on zenith-sky AMFs, and, at the same time, it must also be large enough to ensure a sufficient number of measurements in each SZA bin (to perform reliable regressions). In order to deal with

the diurnal variation of NO₂ concentration and changing of profile shape (e.g., due to changing of boundary layer heights), the dataset has been divided into morning and evening sets, and discrete AMFs are retrieved for a.m. and p.m. separately (see the blue and red squares with error bars in Figure 1).

Next, these discrete AMF values are used to fit an empirical zenith-sky NO₂ AMF function, which has the expression:

$$5 \quad AMF = a_1 + (1.02 - a_1)/\cos(SZA) \quad (5).$$

The fitted empirical zenith-sky AMFs are shown in Figure 1 as blue and red lines (data regression period from February 2015 to September 2017). Several sensitivity tests have been performed to assess the quality of the empirical zenith-sky AMFs, including fitting the AMFs with/without a diurnal difference, fitting the AMFs with different empirical functions (e.g., exponential and simple geometry approximation) and fitting the AMFs by seasons. All these different choices of empirical

10 AMFs fitting functions or methods only introduce less than 5 % difference in the retrieved empirical AMFs. Thus, to make the empirical AMFs simple and robust, we selected to fit with a diurnal difference (Eqn. 5). **In addition, the current empirical AMFs are limited to high and intermediate sun conditions (i.e., $SZA < 75^\circ$). For low-sun conditions, the total AMF for zenith-sky measurements is expected to be a strong function of not only the SZA, but also the tropospheric column itself. Thus, for future work to derive low-sun empirical zenith-sky AMFs, the stronger influence of PBL NO₂ has to be accounted (i.e., the**
 15 **geometry form AMFs are not enough).**

B. Stratospheric NO₂ column

Several stratospheric NO₂ column values were tested and used in the surface NO₂ concentration algorithm (Eqn. 3). Figure A1a shows the OMI monthly mean (referred to as OMI) and Prato box-model stratospheric column NO₂ (Adams et al., 2016; McLinden et al., 2000) (referred to as box). Since the satellite only samples Toronto once per day, the OMI stratospheric NO₂
 20 lacks diurnal variation. To account for the diurnal variation, diurnal ratios of NO₂ VCD have been calculated and applied to OMI monthly mean data. The stratospheric NO₂ columns are calculated using

$$V_{OMI}(t) = \frac{V_{box}(t)}{V_{box}(t_0)} \times V_{OMI}(t_0), \quad (6)$$

where, $V_{OMI}(t_0)$ is the OMI measured stratospheric column, t_0 is OMI overpass time, $V_{box}(t_0)$ is the modelled stratospheric column at OMI overpass time, $V_{box}(t)$ is the modelled stratospheric column at time t , and $V_{OMI}(t)$ is the interpolated stratospheric
 25 column at time t . The interpolated OMI stratospheric columns are referred to as OMI-box. The grey dots on Figure A1b are OMI-box stratospheric NO₂ columns. The monthly mean of the box model (blue line) and OMI-box (black line) show that the amplitude of OMI-box is larger than the amplitude of the box model.

To justify why this diurnal variation has to be included, Figure A1c shows the total column NO₂ time series. The diurnal stratospheric NO₂ variation is about 0.1 DU in the summer (see grey dots in Figure A1b) when Pandora measured monthly
 30 mean total column is about 0.5 DU (Figure A1c). Thus, neglecting this diurnal variation will lead to diurnal biases in the derived surface NO₂ data (e.g., in the morning, this will lead to the overestimation of the stratospheric NO₂ and thus the underestimation of surface NO₂). **Please note that the strength of this bias is related to 1) the NO₂ profile (weights between**

stratospheric and tropospheric NO₂), and 2) the observation geometry (direct-sun or zenith-sky). In general, an urban site with direct-sun observation should have smaller impact from the stratospheric diurnal variation. On the other hand, a rural site with zenith-sky observation should have a significant impact.

C. Cloud effect and heavy cloud filtration

5 Direct-sun measurements need an unobscured sun. Even thin clouds could decrease the quality of retrieved NO₂ total columns, especially for low altitude clouds. Unlike direct-sun measurements, zenith-sky observations are made with scattered sunlight and have limited sensitivity to cloud cover. For example, Hendrick et al. (2011) calculated that, for NDACC UV-visible zenith-sky ozone measurements, clouds only contribute 3.3 % to the total random error. This is because a trace gas that is mostly distributed in the stratosphere has the mean scattering layer located at a higher altitude than the cloud layer. However, this
10 assumption may not be valid for NO₂. Depending on the properties of the clouds and the NO₂ profile, the clouds could have non-negligible impacts on zenith-sky NO₂ observations.

A typical method of removing zenith-sky measurements affected by heavy clouds is to eliminate measurements with large enhancements of O₄ and/or H₂O (Van Roozendaal and Hendrick, 2012). In the Pandora zenith-sky NO₂ retrieval, we use the O₄ dSCDs. Since the measured O₄ dSCDs has SZA dependency, all measured O₄ dSCDs are plotted against SZA and a second
15 order quantile regression (Koenker and Hallock, 2001) is applied to select the top few percentile of the measured O₄ dSCDs. Figure A2 shows examples of selected Pandora zenith-sky NO₂ data and their corresponding O₄ dSCDs values. For example, Figure A2a shows the O₄ dSCDs versus SZA and the top 10 percentile of the data with enhanced O₄ are marked in grey. The corresponding Pandora zenith-sky data are plotted against in situ data in Figure A2b, which shows low correlation ($R = 0.34$) and high bias (68 %). This result indicates that the enhanced scattering due to heavy cloud caused a positive bias in the Pandora
20 zenith-sky NO₂ retrieval. Figures A2c and A2d are similar to Figures A2a and A2b, but for selected Pandora zenith-sky NO₂ data that have O₄ values within the 40th to 50th percentile range. Figure A2d shows that when O₄ is not enhanced, the derived zenith-sky NO₂ has good agreement with in situ data ($R = 0.8$ and bias = -5 %). To summarize how the retrieved O₄ dSCDs can indicate the quality of the Pandora zenith-sky NO₂, the results from the other percentile bins are shown in Figure A3. In general, besides the top 10 percentile of data, the results from all the other bins show good correlation (above 0.6) and low
25 bias. Thus, in this study, the Pandora zenith-sky NO₂ data that have O₄ values within only the top 10 percentile are considered to be affected by heavy clouds and are removed. Some examples of this heavy cloud effect are shown in Figure A4 and Figure 8 in Section 4.3.

D. Uncertainty estimation

The uncertainties of retrieved Pandora zenith-sky NO₂ data products (total column and surface concentration) are estimated
30 and discussed here to assess the quality of the data products. The uncertainties of total column and surface concentrations are estimated first using the uncertainty propagation method (referred to here as the UP method) based on Eqns. 2 and 3. The combined uncertainties of total column can be calculated using:

$$\sigma_{VCD_{ZS}} = \sqrt{\left(\frac{\sigma_{dSCD}}{AMF}\right)^2 + \left(\frac{\sigma_{RCD}}{AMF}\right)^2 + \left(\frac{\sigma_{AMF \times SCD}}{AMF^2}\right)^2}, \quad (7)$$

where σ_{dSCD} is the statistical uncertainty on the DOAS fit (output of QDOAS) and σ_{RCD} and σ_{AMF} are the estimated statistical uncertainties using standard errors of the RCD and the zenith-sky empirical AMF regression, respectively (Eqn. 4). To estimate the upper limit of the nominal uncertainty, AMF and SCD are used as median and maximum values in the dataset, respectively.

5 The combined uncertainties of the surface concentration can be calculated using:

$$\sigma_{C_{Pan}} = \sqrt{\left(R_{CV}\sigma_{V_{Pan}}\right)^2 + \left(R_{CV}\sigma_{V_{strat}}\right)^2 + \left(R_{CV}\sigma_{V_{ftrop}}\right)^2 + \left(V_{Pan} - V_{strat} - V_{ftrop}\right)^2 \sigma_R^2}, \quad (8)$$

where $\sigma_{V_{pan}}$ is the uncertainty of Pandora zenith-sky total column NO_2 , (here we use the derived σ_{VCD} in Eqn. 7), $\sigma_{V_{strat}}$ is the uncertainty of the stratospheric NO_2 column (estimated using the 1-sigma standard deviation of the V_{strat}), $\sigma_{V_{ftrop}}$ is the uncertainty of the free troposphere NO_2 column (estimated using the 1-sigma standard deviation of the V_{ftrop}). R_{CV} is the GEM-
10 MACH calculated surface VMR to PBL column ratio, and σ_R is the uncertainty of that ratio (estimated using the 1-sigma standard deviation of the R_{CV}). The means of R_{CV} , V_{Pan} , V_{strat} , and V_{ftrop} are used in the uncertainty estimation.

Besides the UP method, another simple approach to estimate uncertainty is to compare the data product with another high-quality (lower uncertainty) coincident data. For example, if we assume that the Pandora direct-sun total column NO_2 data can represent the true value, we can estimate the uncertainty of Pandora zenith-sky total column NO_2 by calculating the 1-sigma
15 standard deviation of their difference (referred to here as the SDD method):

$$\sigma_{VCD_{ZS}} = \sigma(VCD_{DS} - VCD_{ZS}). \quad (9)$$

Similarly, if we assume that the in situ surface NO_2 VMR can represent the true value, the uncertainty of Pandora zenith-sky-based surface NO_2 VMR can be given by:

$$\sigma_{C_{Pan}} = \sigma(C_{insitu} - C_{pan}). \quad (10)$$

20 Also, if there is systematic bias between the two datasets, it can be removed and the random uncertainty can be calculated by:

$$\sigma_{VCD_{ZS}} = \sigma(VCD_{DS} - k_1 VCD_{ZS}), \quad (11)$$

$$\sigma_{C_{Pan}} = \sigma(C_{insitu} - k_2 C_{pan}), \quad (12)$$

where k_1 and k_2 are the slopes in the linear fits with intercept set to zero (e.g., slopes in Figs. 2 and 6). This method is referred to here as the unbiased SDD. These three uncertainty estimation methods (UP, SDD, and unbiased SDD) were all implemented,
25 and the results are summarized in Table A1. The results show that Pandora zenith-sky total column NO_2 data have a 0.09-0.12 DU uncertainty that is about twice to the Pandora direct-sun total column nominal accuracy (0.05 DU, at 1-sigma level). When using the UP method, for the worst-case scenario, the Pandora zenith-sky total column NO_2 have a 0.17 DU uncertainty (i.e. using minimum of AMFs to estimate the upper limit of uncertainty). The estimated Pandora zenith-sky-based surface NO_2 VMR data have uncertainties from 4.8 to 6.5 ppbv. In Eqn. 8, the contributions of the V_{Pan} , V_{strat} , V_{ftrop} , and R_{CV} terms to the
30 total uncertainty are 36%, 2%, 0.3%, and 62 %, respectively. This result indicates that the uncertainty in the Pandora zenith-sky-based surface NO_2 VMR is dominated by the uncertainties of Pandora zenith-sky total column NO_2 and the modelled column-to-surface conversion ratio (R_{CV}). However, note that this uncertainty budget depends on the NO_2 vertical distributions,

and hence may vary from site to site; e.g., in Toronto, tropospheric column NO₂ is typically 2-4 times higher than stratospheric column NO₂, and thus, the contribution to uncertainty from V_{Pan} is much larger than the corresponding contributions from V_{Strat} and V_{trop} . In addition, the uncertainty of Pandora direct-sun surface NO₂ VMR is also estimated and provided in Table 1. It shows slightly better results than for zenith-sky-based surface NO₂ VMR.

5

Table 1. Estimated uncertainties for Pandora zenith-sky total column and surface NO₂.

Estimation method	$\sigma_{VCD_{ZS}}$ (DU)	$\sigma_{C_{Pan-ZS}}$ (ppbv)	$\sigma_{C_{Pan-DS}}$ (ppbv)
UP	0.12	6.5	5.4
SDD	0.09	5.1	5.0
unbiased SDD	0.09	4.8	4.8

10

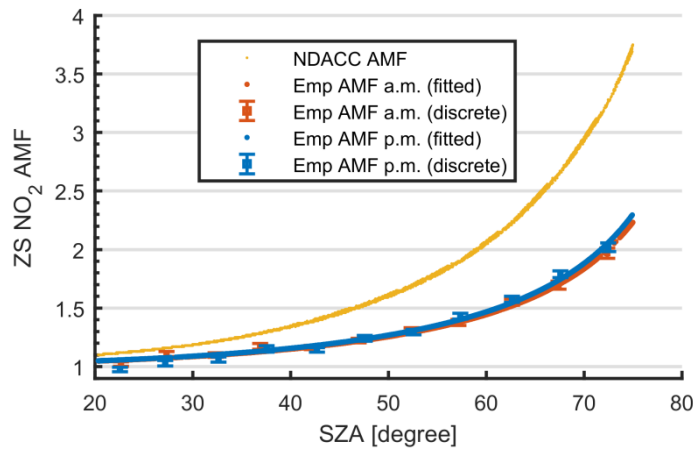


Figure 1. Comparison of zenith-sky NO₂ air mass factors. Blue and red squares with error bars (standard error) represent the empirical discrete zenith-sky NO₂ AMFs in each SZA bin for Toronto for the period Feb. 2015 to Sept. 2017. Blue and red lines show the fitted empirical zenith-sky NO₂ AMFs. NDACC AMFs calculated using the NDACC look-up table and assuming no NO₂ in the troposphere are shown in yellow.

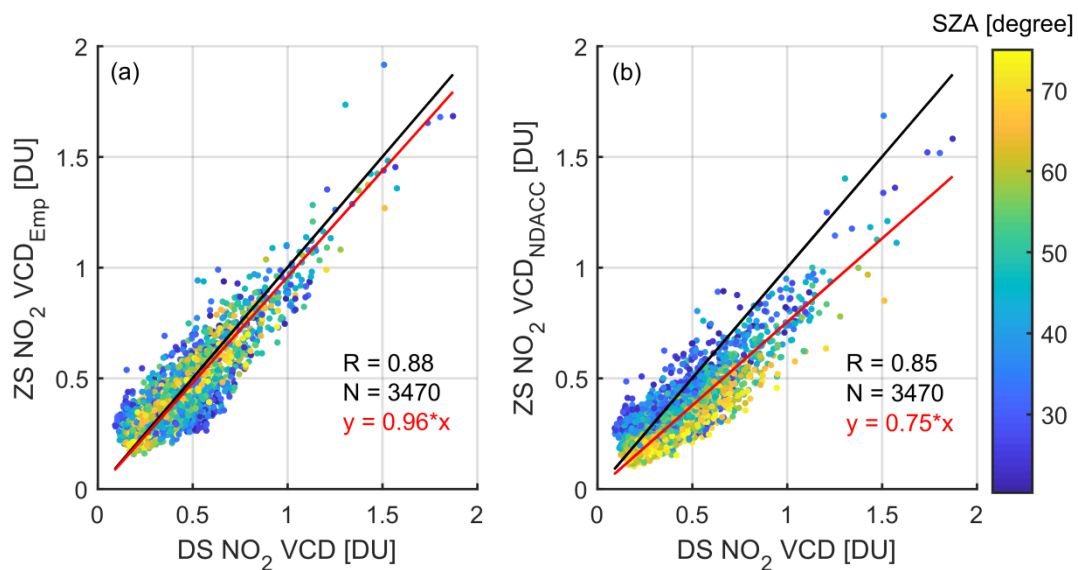


Figure 2. Comparisons of NO₂ total columns (2015-2017): (a) zenith-sky total column NO₂ retrieved using empirical AMFs vs. direct-sun total column NO₂, (b) zenith-sky total column NO₂ retrieved using NDACC AMFs vs. direct-sun total column NO₂. On each scatter plot, the red line is the linear fit with intercept set to 0, and the black line is the one-to-one line. The scatter plot is colour-coded by solar zenith angle (SA).

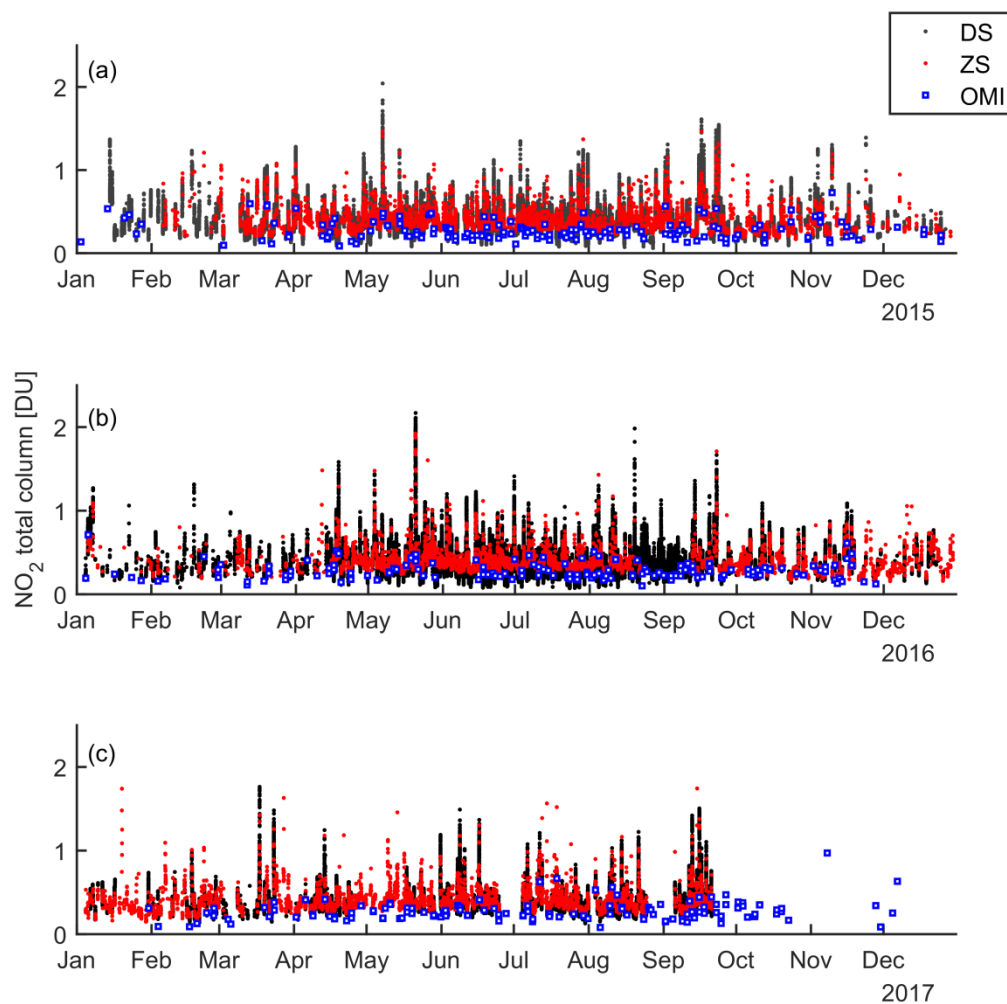


Figure 3. Annual time series of Pandora direct-sun (DS), Pandora zenith-sky (ZS), and OMI SPv3 total column NO_2 in Toronto from 2015-2017.

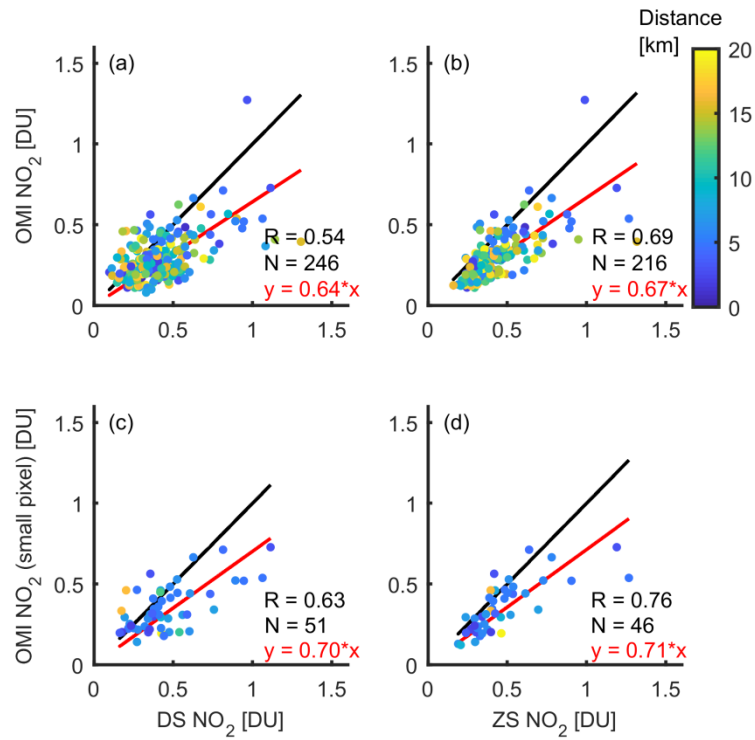


Figure 4. OMI vs. Pandora total column NO₂ (2015-2017). (a) and (c) show OMI vs. Pandora direct-sun NO₂, and (b) and (d) show OMI vs. Pandora zenith-sky NO₂. (a) and (b) show all available OMI measurements, while (c) and (d) show OMI data from small pixels only. On each scatter plot, the red line is the linear fit with intercept set to 0 and the black line is the one-to-one line. All scatter plots are colour-coded by the distance from the centre of an OMI ground-pixel to the location of Pandora.

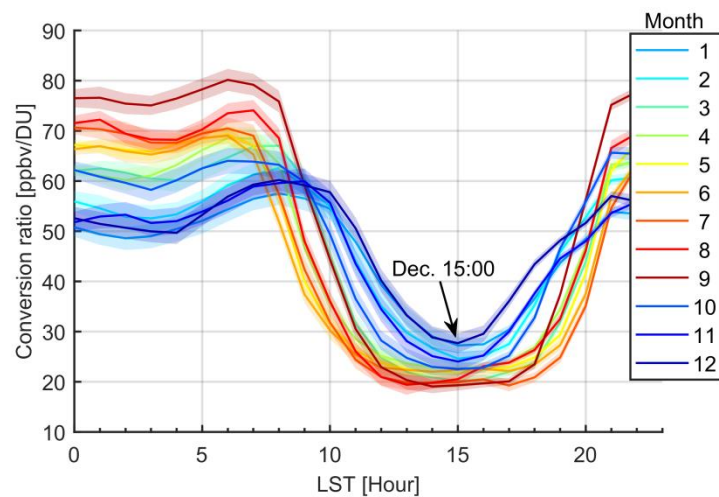


Figure 5. Dependence of the Pandora surface NO₂ concentration look-up table (PSC-LUT) on month of year and hour of day. The PSC-LUT is constructed using the GEM-MACH modelled NO₂ conversion ratios. Solid lines are monthly mean conversion ratios colour-coded by month. The shaded envelopes are the standard error of the mean.

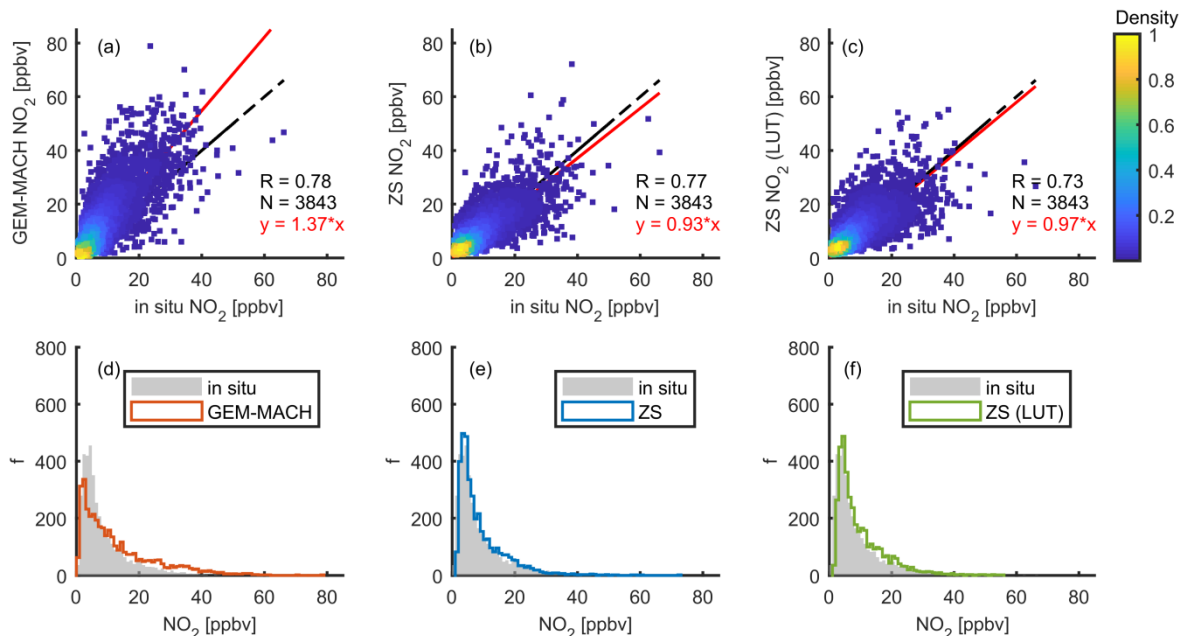


Figure 6. Modelled and Pandora zenith-sky surface NO₂ vs. in situ NO₂ (2016-2017). (a) shows the GEM-MACH modelled surface NO₂ data vs. in situ NO₂; (b) and (c) show the Pandora zenith-sky (ZS) surface NO₂ data vs. in situ NO₂. The Pandora ZS surface NO₂ data in (b) and (c) are derived using the hourly modelled conversion ratio and the monthly PSC-LUT, respectively. (d) to (f) are histograms corresponding to the data in (a) to (c). On each scatter plot, the red line is the linear fit with intercept set to 0 and the black line is the one-to-one line. The scatter plots are colour-coded by the normalized density of the points.

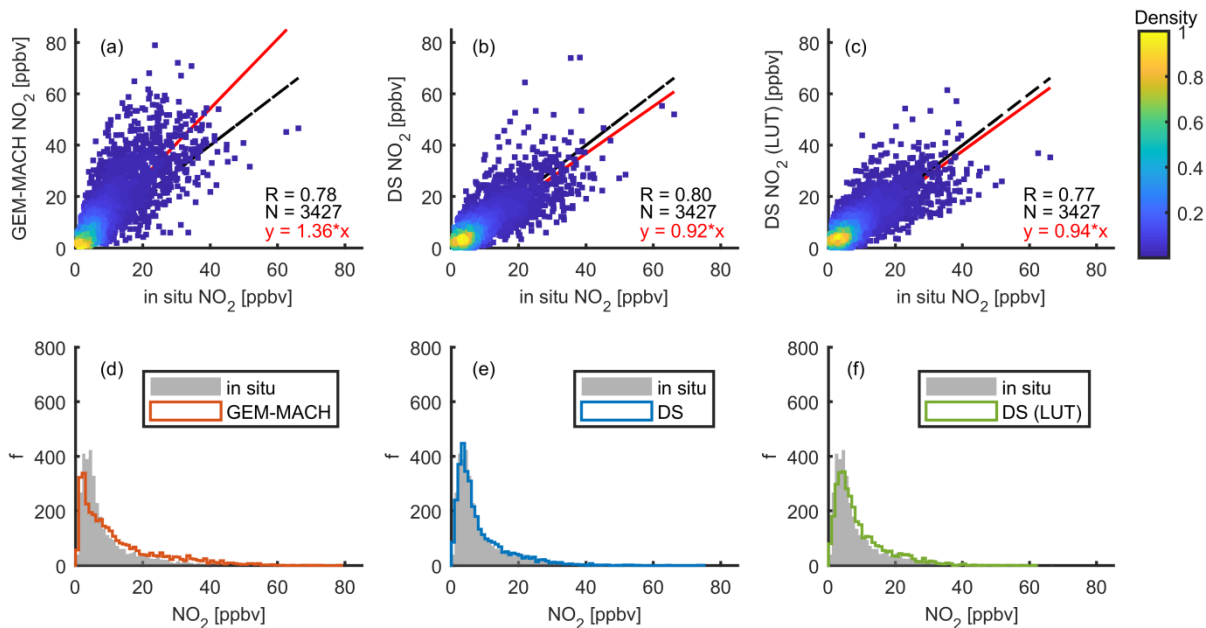


Figure 7. Modelled and Pandora direct-sun surface NO₂ vs. in situ NO₂ (2016-2017). (a) shows the GEM-MACH modelled surface NO₂ data vs. in situ NO₂; (b) and (c) show the Pandora direct-sun (DS) surface NO₂ data vs. in situ NO₂. The Pandora DS surface NO₂ data in (b) and (c) are derived using the hourly modelled conversion ratio and the monthly PSC-LUT, respectively. (d) to (f) are histograms corresponding to the data in (a) to (c). On each scatter plot, the red line is the linear fit with intercept set to 0 and the black line is the one-to-one line. The scatter plots are colour-coded by the normalized density of the points.

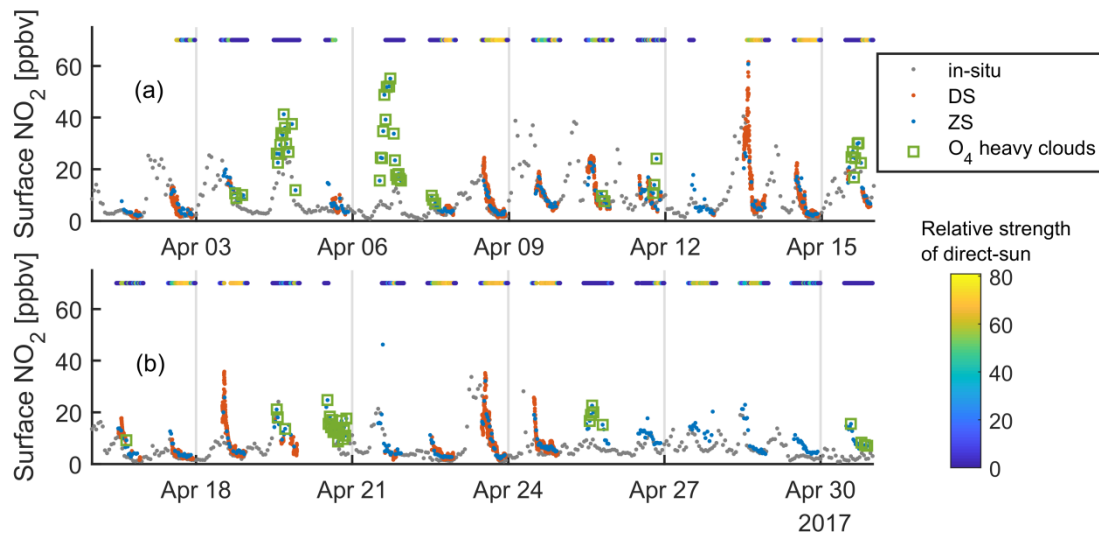


Figure 8. Example of surface NO₂ concentration time series in all conditions (April 2017). The in situ, Pandora direct-sun (DS), and Pandora zenith-sky (ZS) surface NO₂ concentrations are shown by different coloured dots. The total sky imager relative strength of direct-sun data are plotted as a colour-coded horizontal dot-line on the top area of each panel. For Pandora zenith-sky data, the measurements with enhanced O₄ (heavy cloud indicator) are also labelled by green squares.

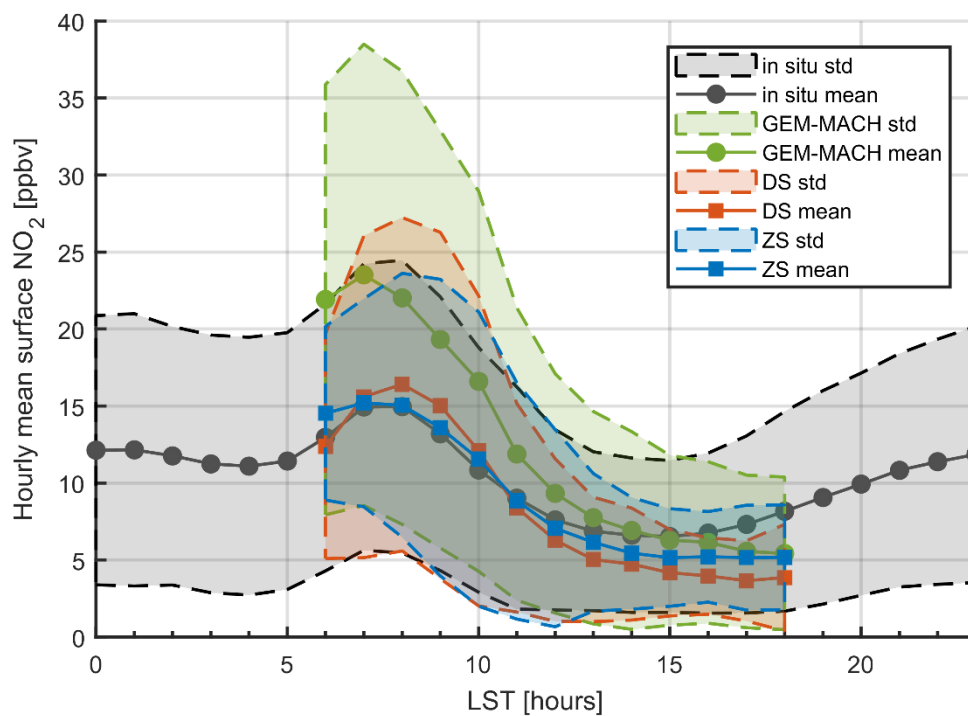


Figure 9. Diurnal variation of surface NO_2 concentration (2016-2017). The x-axis is the local standard time (LST). Lines with dot/square symbols represent the hourly mean of corresponding data indicated by the legend. The shaded area represents the 1σ envelope.

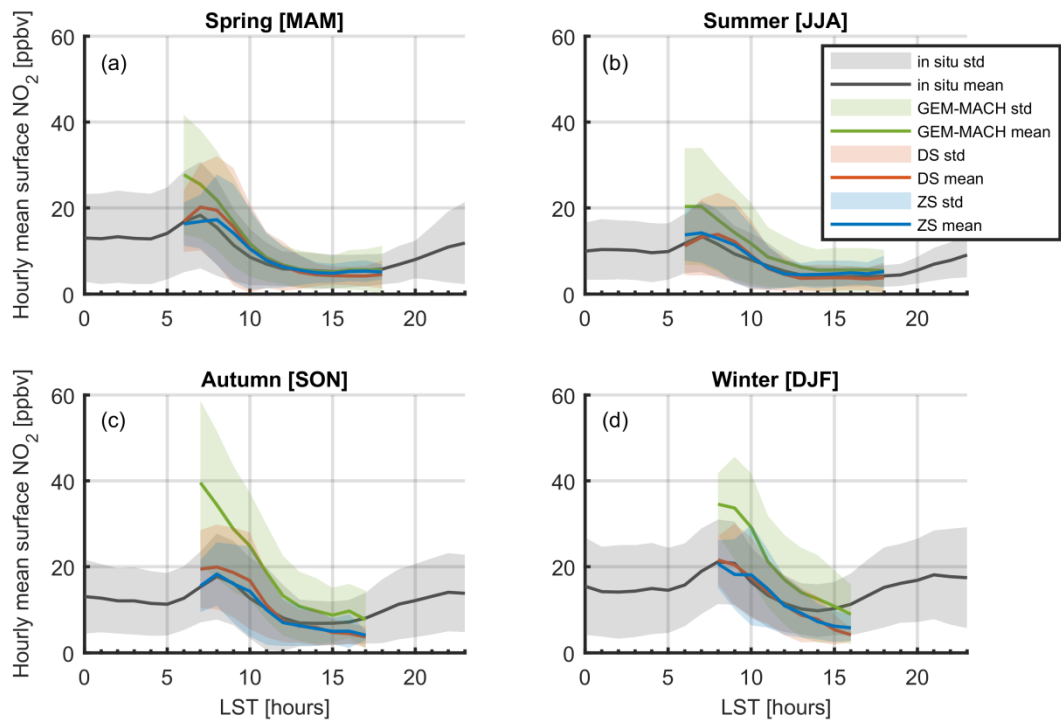


Figure 10. Diurnal variation of surface NO₂ concentration by season (2016-2017). The x-axis is the local standard time (LST). Each panel represents data collected in one season (spring, summer, autumn or winter). Solid lines represent mean of corresponding data indicated by the legend. The shaded area represents the 1 σ envelope.

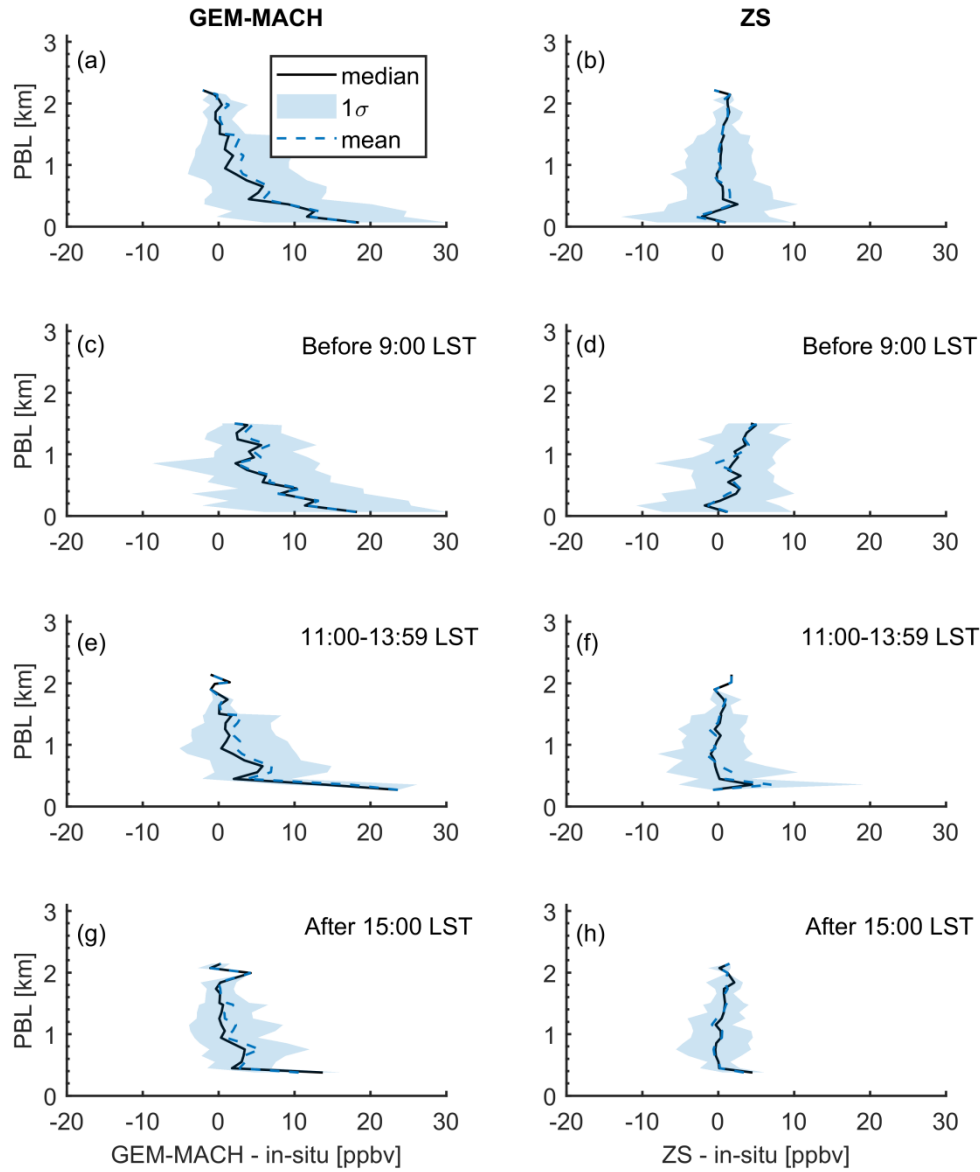


Figure 11. Illustration of planetary boundary-layer (PBL) effect (2016-2017). The y-axis is planetary boundary-layer height in km. The x-axes for the left column are the difference between GEM-MACH and in situ surface NO₂ concentrations; the x-axes for the right column are the difference between Pandora zenith-sky ($C_{pan-model}$) and in situ surface NO₂ concentration. Panels a and b show all available data, panels c and d show the morning data (before 9:00, local standard time), panels e and f show the noon data (from 11:00 to 13:59), and panels g and h show the evening data (after 15:00).

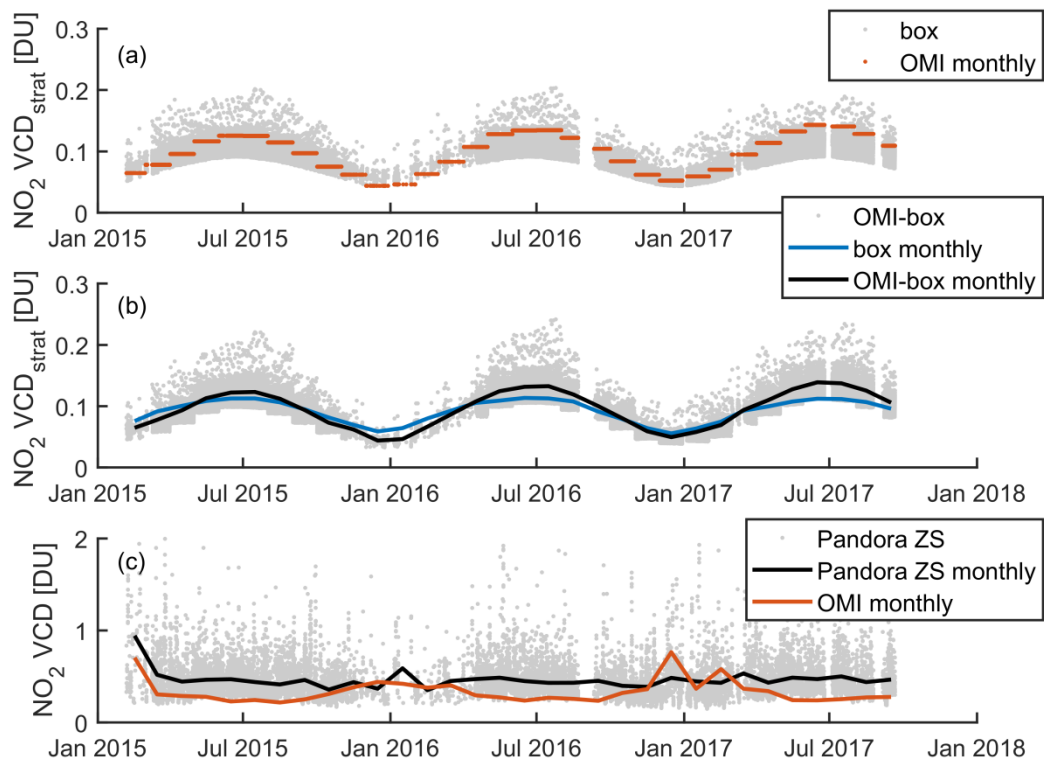


Figure A1. Time series of measured and modelled NO₂ columns: (a) stratospheric columns from the box model (hourly) and OMI (monthly), (b) stratospheric columns from OMI-box (hourly), box (monthly) and OMI-box (monthly), and (c) total columns from Pandora zenith-sky and OMI.

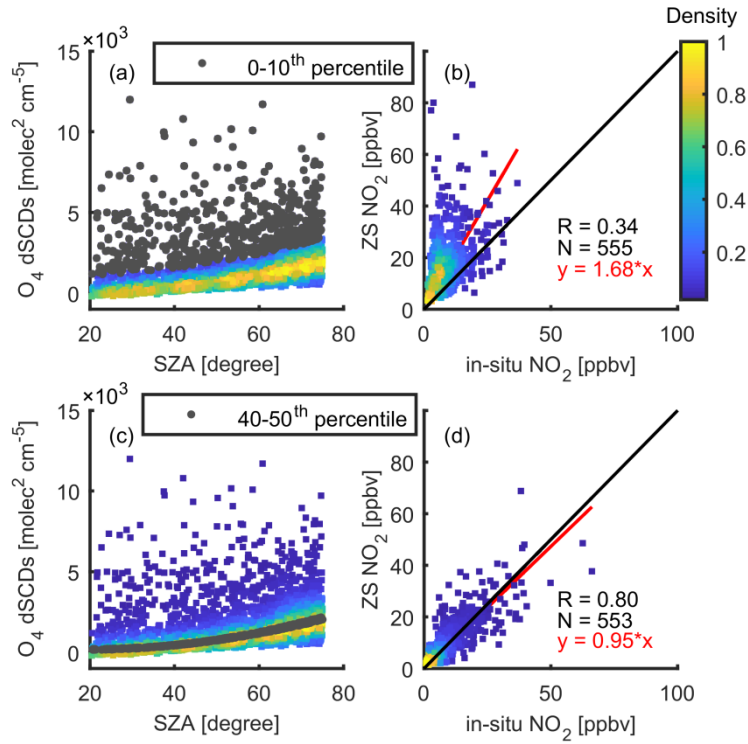


Figure A2. Illustration of cloud effect and heavy-cloud data filtration: (a) shows measured O₄ differential slant column densities versus solar zenith angle, the gray dots represent the top 0-10th percentile range of O₄, (b) shows the scatter plot of zenith-sky versus in situ surface NO₂ use data that has O₄ value within 0-10th percentile range (as identified in (a)), (c) is similar to (a) but the gray dots represent the 40-50th percentile range of O₄, (d) is similar to (b) but use the data that has O₄ value within the 40-50th percentile range. On scatter plot (b) and (c), the blue line is the linear fit with intercept set to 0, the red line is a simple linear fit, and the black line is the one-to-one line. All plots are colour-coded by the normalized density of the points.

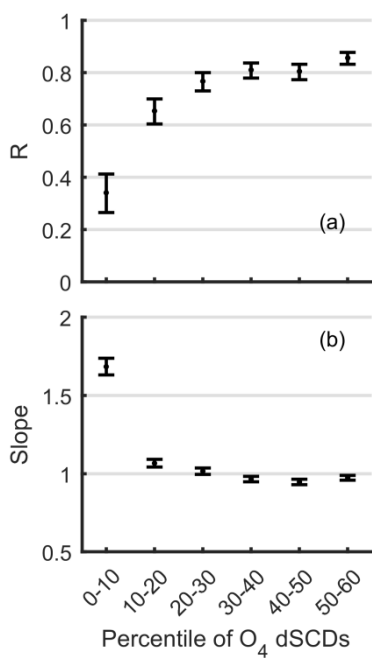


Figure A3. Correlation coefficient and bias (slope) between zenith-sky and in situ surface NO₂ data in different O₄ dSCDs percentile bins. (a) shows the correlation coefficients, (b) shows the slopes of linear fit with intercept set to 0.

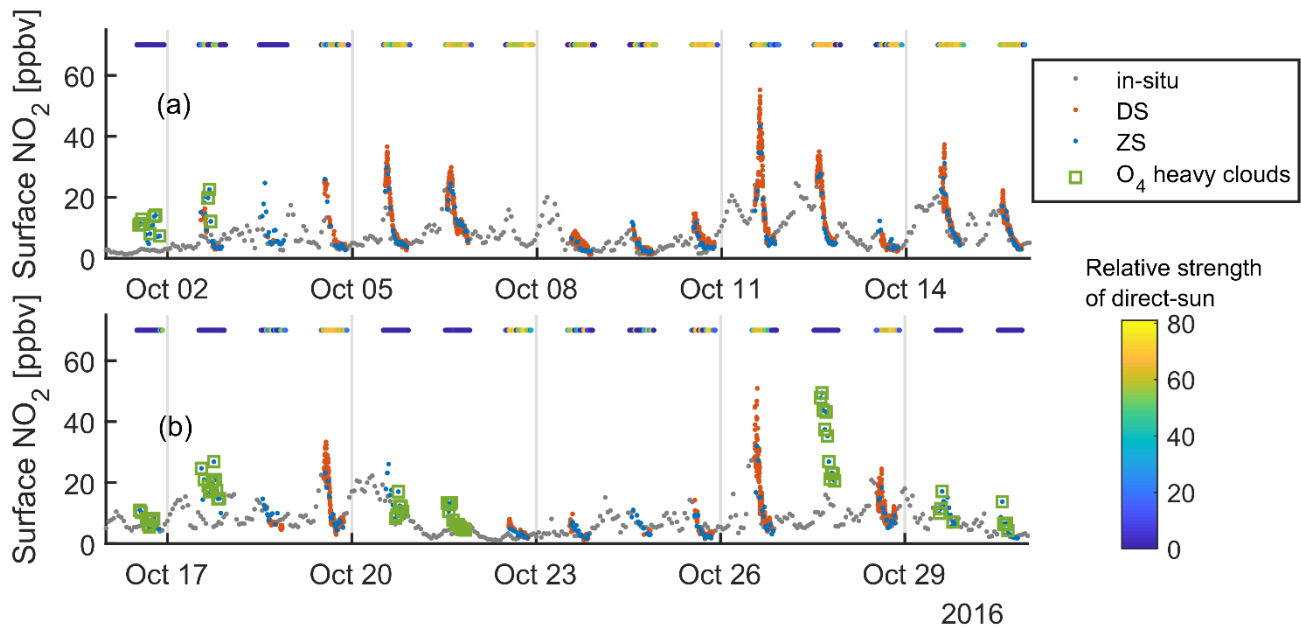


Figure A4. Example of surface NO_2 concentration time series in all conditions. The in situ, Pandora direct-sun (DS), and Pandora zenith-sky (ZS) surface NO_2 concentrations are shown by different coloured dots. The total sky imager relative strength of direct-sun data are plotted as a colour-coded horizontal dot-line on the top area of each panel. For Pandora zenith-sky data, the measurements with enhanced O_4 (heavy cloud indicator) are also labelled by green squares.