Response to AMT Reviewer #1

The authors present a novel study using Pandora MAX-DOAS measurements and the HeiPro retrieval algorithm to produce a three-year dataset of NO2 profiles and partial columns in Toronto, Canada. While the paper provides valuable insights into the spatial and temporal distribution of NO2. However, several areas require clarification, reorganization, and deeper analysis to strengthen the validity and comprehensibility of the study. My detailed comments are below.

We thank the reviewer for their helpful comments, which have helped to improve the manuscript. Our point-by-point responses are provided below in blue font, with new text added to the manuscript given in *red font and italics*. Line numbers refer to the new clean version of the manuscript.

General Comments

Stratospheric-Tropospheric Separation (STS) Method

- The authors use a complicated approach involving a box model and OMI observations to separate stratospheric and tropospheric columns for Pandora direct-sun (DS) measurements. The STS method is questionable due to multiple layers of assumptions and models, creating uncertainties.
    - Why not employ stratospheric-tropospheric column ratios from established models, such as CAMS, TM5, or GEM-MACH, to simplify and enhance the accuracy?
        - The STS method used in this work follows the approach of several previous studies (e.g., Zhao et al., 2019, Choi et al., 2020, Zhao et al., 2022) , while the OMI stratospheric NO_2 data product (v3) has been shown to have good agreement with other satellite and ground-based FTIR measurements (Krotkov et al., 2017). Additionally, TROPOMI vs. Pandora-DS tropospheric NO_2 show good agreement (discussed further in the next bullet point), providing an independent verification of the STS method used. Models such as CMAS and TM5 are not independent data sources, if we wish to compare the results with satellite instruments (such as TROPOMI). For example, TM5 has already been used in TROPOMI STS algorithm, CAMS assimilated TROPOMI data, etc. GEM-MACH is a good option, but the operational version involved in current work only has tropospheric column. However, we fully agree with the referee that we could explore some more advanced models for the STS in the future.
        - Additionally, the stratospheric portion removed from the Pandora-DS total columns is 34% $\pm$ 2.8%. For reference and comparison, we also calculated the stratospheric-to-total column portion for TROPOMI measurements and found this value to be 44% $\pm$ 9.7%. Please note that the former value represents data throughout the Pandora-DS measurement day while the latter represents data at a single measurement time per day. The following

text has been revised in the manuscript, to include the stratospheric percentage, on lines 215–216:

Due to the diurnal variation of NO_2, and the satellite's overpass time of 13:30 local time (LT), a photochemical box model (Pratmo, discussed further in Section 2.3.1) was used to calculate stratospheric NO_2 at various Pandora measurement times throughout the day. The stratospheric portion that was removed accounted for 34% $\pm$ 2.8% of the Pandora-DS NO_2 total columns.

- A direct comparison of Pandora-DS total columns with TROPOMI total columns would provide additional insights into discrepancies between ground-based and satellite observations.
  - A plot of TROPOMI vs. Pandora-DS NO_2 total columns has now been added in the manuscript in the Appendix (Figure A4) and is copied below for reference. For total column comparisons, the mean relative bias of TROPOMI to Pandora-DS is 1.53% $\pm$ 21.2% while the zero-intercept slope is 0.89 $\pm$ 0.02. For tropospheric column comparisons, the mean relative bias of TROPOMI to Pandora-DS is $-$0.85% $\pm$ 34% while the zero-intercept slope is 1.04 $\pm$ 0.03. At this measurement site, ground-based direct-Sun and satellite observations of NO_2 agree reasonably well. The agreement between the tropospheric columns provides an indication that the stratospheric-tropospheric separation method we are using is reasonable.
  - The following text has been added to the manuscript on lines 461–475:

    For reference, HeiPro (0–4 km) vs. Pandora-DS total columns are compared in panel (a) of Fig. A4, which shows that HeiPro partial columns exhibit a positive multiplicative bias of 16% $\pm$ 0.7% and a mean relative bias of 6.1% $\pm$ 4.8%. Not surprisingly, there is better agreement here as compared to Fig. 3a (i.e., HeiPro vs. Pandora-DS tropospheric NO_2) since the Pandora-DS total columns are larger. The TROPOMI vs. Pandora-DS NO_2 total and tropospheric column comparisons are shown in Fig. A4b–c, respectively. Pandora-DS and TROPOMI show good agreement with one another for both total column (multiplicative bias: $-$12% $\pm$ 1.9%; mean relative bias: 0.1% $\pm$ 21%) and tropospheric NO_2 (multiplicative bias: 4.4% $\pm$ 3.5%; mean relative bias: $-$0.9% $\pm$ 34%). Note that the large uncertainties are due to the relatively large TROPOMI total column and tropospheric NO_2 errors. Additionally, the tropospheric NO_2 agreement in panel (c) provides more confidence in the stratospheric-tropospheric separation method that was used in the study (i.e., Pratmo-OMI data).

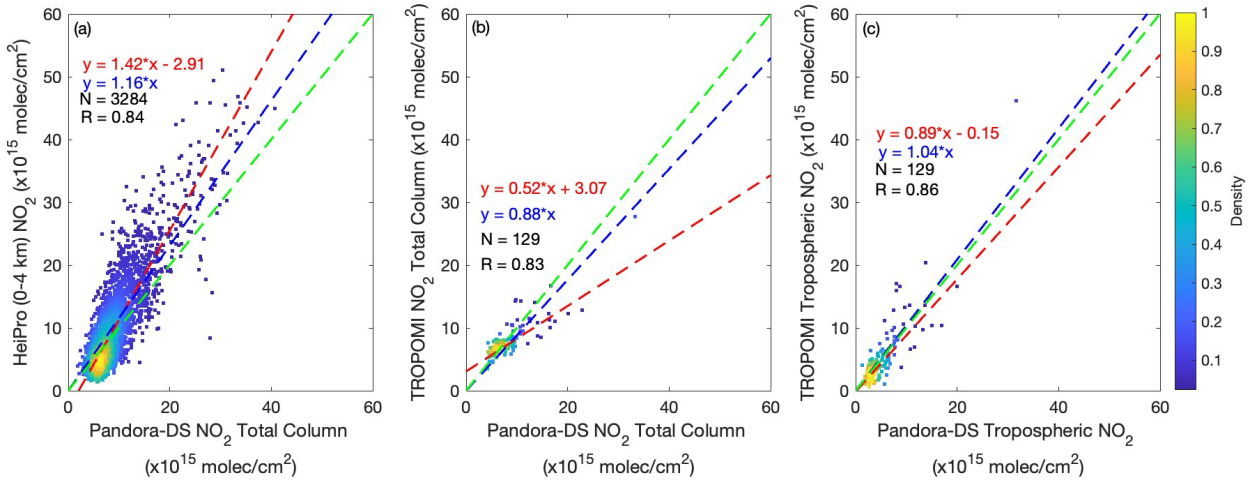


Figure A4. Comparisons (2018–2020) of (a) HeiPro (0–4 km) NO_2 partial columns vs. Pandora-DS NO_2 total columns, (b) TROPOMI vs. Pandora-DS NO_2 total columns, and (c) TROPOMI vs. Pandora-DS tropospheric NO_2 columns. The dashed lines and color bar are as indicated in Fig. 3. For panels (b) and (c), the coincidence criteria are 10 km and 10 minutes.

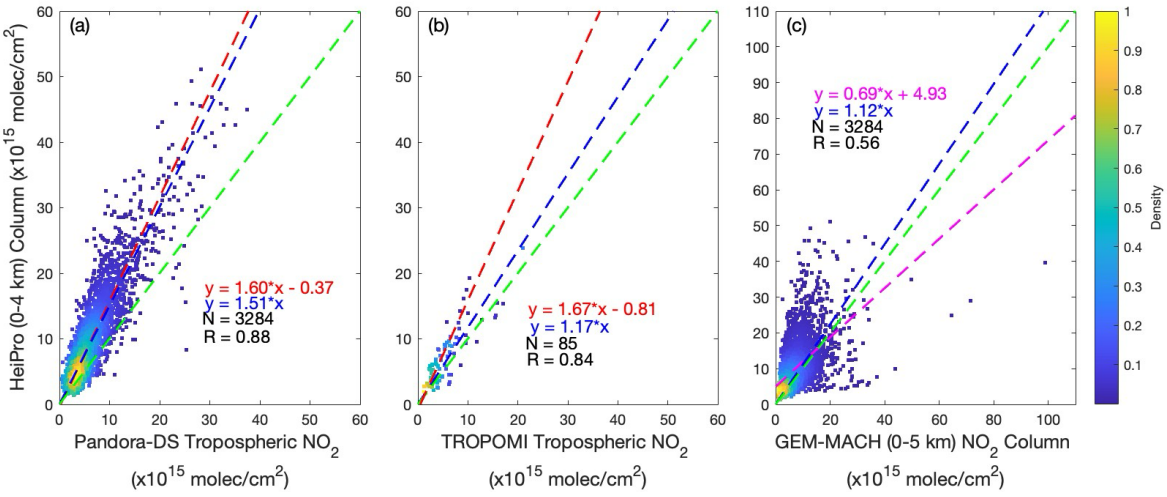
MAX-DOAS Retrieval and Atmospheric Profile

- The description of the MAX-DOAS profile retrieval lacks essential details:
  - Are the ERA5 atmospheric profiles daily averages or spatio-temporally interpolated to the measurement times?
    - The ERA5 atmospheric profiles are daily averages obtained from the ERA5 grid box nearest to the measurement site. The atmospheric profiles are calculated from the average of daytime (11:00, 14:00, 17:00, 20:00, 23:00 UTC) temperature and pressure profiles. This is discussed in Section 2.4.2 (lines 367–369), which describes the ERA5 data used in the study, but the following text has been added in Section 2.1.3 (lines 280–281) for clarification:
      Additional a priori inputs to HeiPro were daily pressure and temperature profiles from ERA5 reanalysis data at the grid box nearest to the measurement site (discussed further in Section 2.4.2).
  - What assumptions are made about NO2 above the retrieval height? Are these values based on standard atmospheric profiles or other sources?
    - The retrieval height is 0–4 km. For the HeiPro datasets presented here, NO_2 above the retrieval height is not considered. Therefore, the HeiPro profiles that are integrated to produce columns represent partial columns from 0–4 km. This has been stated in the manuscript text and all figures where necessary.

Data Consistency and Filtering

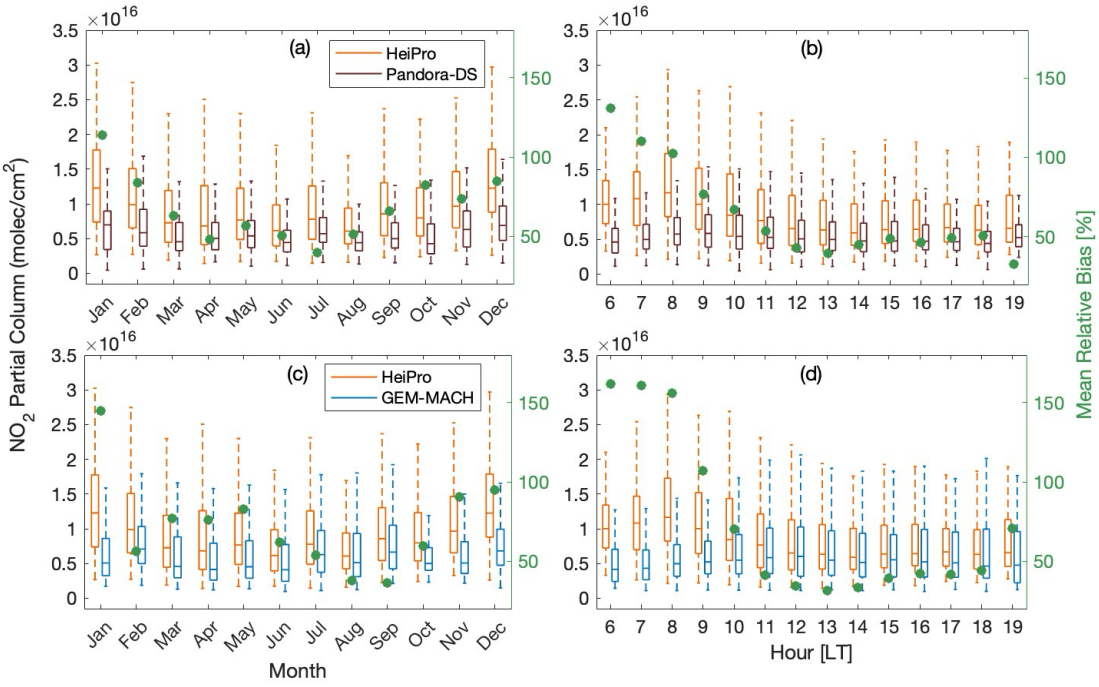
- The differences in coinciding data points in Figure 3(a) (direct sun) and Figure 3(c) (model comparisons) require clarification:
  - If Pandora DS and MAX-DOAS data originate from the same source, why is there a discrepancy in the data points?

- The reason for the discrepancy in the number of coincident data points between Fig. 3a and Fig. 3c is due to data gaps in the OMI NO_2 stratospheric column data used in the calculation of the Pandora-DS tropospheric columns. While MAX-DOAS and direct-Sun have slightly different measurement schedules, they do have a similar measurement frequency as they are from the same instrument. However, the OMI stratospheric NO_2 data (which is used for the Pandora-DS tropospheric data) has gaps with missing days and hours of the day. Therefore, this shortens the number of coincident data points between Pandora-DS and the MAX-DOAS HeiPro data, while the MAX-DOAS HeiPro and GEM-MACH data do not have this temporal limitation. We have modified panels (a) and (c) in Fig. 3 to only include coincident data among HeiPro, Pandora-DS, and GEM-MACH, and so they have the same number of data points (N=3284). We have also adjusted any reference in the text to the biases of HeiPro to GEM-MACH partial columns (they changed very slightly due to the new number of data points). The modified Fig. 3 is copied below for reference:



- Did the authors apply data filtering, such as cloud filtering, before using Pandora data for MAX-DOAS retrievals?
  - Yes, Pandora O_4 and NO_2 dSCDs with fitting residual RMS values ≥ 0.003 for the dSCD retrieval were filtered out (only 5% of the data, mostly for SZA $> 80°$). This filter was selected to improve the quality of the fits and discard noisier fits from the analysis. This has now been clarified in Section 2.1 of the manuscript (lines 268–272):
    Following the QDOAS analysis, O_4 and NO_2 dSCDs with fitting residual root-mean-square (RMS) values $\geq 3\times10^{-3}$ were discarded (only 5% of the data, mostly at SZA $> 80°$). This filter was selected to improve the quality of the fits and discard noisier fits from the analysis.
  - Cloud filtering was not used.

- Similarly, in Figure 4, the differences between MAX-DOAS (HeiPro) results in panels (a) vs. (c) and (b) vs. (d) need to be explained.
  - The differences between HeiPro in panel 4a vs. HeiPro in panel 4c is due to the same reason listed under the bullet point above (i.e., that the OMI stratospheric NO_2 data limits the number of coinciding data points with MAX-DOAS and Pandora-DS). Figure 4 has been modified to plot data that is coincident among HeiPro, Pandora-DS, and GEM-MACH. Therefore, the HeiPro results in panels (a) vs. (c) and (b) vs. (d) are now identical. The modified Fig. 4 is copied below for reference.



Organization and Logical Flow

- Figure 1 is currently located in the introduction but would fit better in Section 2, "Instrument Description," to align with the discussion of the Pandora instrument and measurement conditions.
  - We agree with the suggestion that Figure 1 is better suited for Section 2. Figure 1 (now Figure 2) and corresponding text have been moved to Section 2.
- Section 2 should follow a logical sequence: instrument description, retrieval algorithm, and then the models/data used for comparison and validation. This reorganization would enhance the clarity of the methods section.
  - Section 2 has been rearranged so that Section 2.1 focuses solely on Pandora data. The Pandora is introduced and the three subsections that follow go into detail regarding the direct-Sun, dSCD, and profiling retrieval algorithms. Section 2.2 then discusses the TROPOMI instrument and measurements. Section 2.3 discusses the in situ instrument and measurements. Section 2.4 and the three subsections that follow discuss the model data (ERA5, PRATMO, and GEM-MACH). We believe that

this reorganization is clearer and more concise, with each dataset described along with, or shortly after, its respective instrument description.

Wavelength Range and Spectral Retrieval

- The wavelength ranges for direct-sun and MAX-DOAS spectral fits are unclear:
  - Did the authors use the same wavelength range for both retrievals? If not, provide justification for the differences.
    - The wavelength ranges for direct-Sun and MAX-DOAS retrievals are different. The wavelength range for Pandora direct-Sun is 400–440 nm, as per the PGN's direct-Sun NO_2 retrieval algorithm. This retrieval window has been verified, has a well-known performance, and is currently the official data product from PGN (e.g., Zhao et al., 2019). The wavelength range for MAX-DOAS is 338–370 nm, which is based on the MAX-DOAS UV retrieval guidelines for NO_2 outlined in Kreher et al. (2020) and is a recommended protocol by the Network for the Detection of Atmospheric Composition Change (NDACC) UV-visible working group. The wavelength ranges were previously stated in the manuscript in Section 2 where the retrieval algorithms are discussed and in the Conclusions section. Although the retrieval windows differ, it is worth noting that even when comparing the MAX-DOAS long vis HeiPro results to Pandora-DS, the results are very similar to the long UV HeiPro results (see table below), indicating that the differences introduced by the choice of fitting window is small. The following text has been added to the manuscript on lines 422–431:
      Although the direct-Sun and MAX-DOAS retrieval wavelengths are different due to the varying standard protocols for each, it is worthwhile to note that the HeiPro long vis versus long UV NO_2 partial column comparisons showed remarkable agreement with one another, with a zero-intercept slope of 0.97 $\pm$ 0.004 and mean relative bias of 0.7% $\pm$ 5.9%. We therefore do not expect the choice of retrieval window to significantly impact the HeiPro long UV partial column comparisons to Pandora-DS (see Table A1 for the HeiPro long vis NO_2 partial column comparisons).
  - Details of the spectral retrieval process should be included.
    - We have added more detail on the direct-Sun spectral retrieval process. The following text has been added to the manuscript on lines 201–208:
      The standard Pandora-DS total column NO_2 data product is obtained using Total Optical Absorption Spectroscopy (TOAS), as implemented by PGN's BlickP software (Cede, 2019). Direct-Sun spectra in the 400–440 nm range are fitted with cross-sections of NO_2 (at an effective temperature of 254.5 K, Vandaele et al., 1998), O_3 (at an effective temperature of 255 K, Brion et al., 1993, 1998; Daumont et al., 1992), and a fourth-order polynomial to produce SCDs of NO_2 with a clear-sky precision of 2.7×10^{14} molec cm^{-2} (Herman et al., 2009). A synthetic reference spectrum is used in the analysis and is obtained by taking an average of several measured spectra which are corrected for their total optical depth. Following this, NO_2 SCDs are

converted to vertical column densities (VCDs or total columns) using geometric AMFs. The Pandora-DS NO_2 VCD has an absolute accuracy of 1.3×10^{15} molec cm^{-2} (Herman et al., 2009).

- We also now include more detail on the MAX-DOAS spectral retrieval process. The following text has been added to the manuscript on lines 260–268:

  O_4 and NO_2 dSCDs were retrieved in both the ultraviolet (UV, 338–370 nm, only measurements with the UV band pass filters) and visible (vis, 425–490 nm) windows. Differential cross-sections of NO_2 at 294 and 220 K for both windows (Vandaele et al., 1998), O_4 at 293 K for both windows (Thalman and Volkamer, 2013), O_3 at 223 and 243 K for UV and 223 K for vis (Serdyuchenko et al., 2014), BrO at 223 K for UV only (Fleischmann et al., 2004), HCHO at 297 K for UV only (Meller and Moortgat, 2000), and H_2O for vis only (Rothman et al., 2010) were convolved using the instrument slit function and the nominal wavelength calibration file from PGN. A fifth-degree polynomial, linear offset, and first-order shift and stretch were used in both windows (Kreher et al., 2020).

- Line 217-218 mentions NO2 retrieval in both UV and VIS bands—did the authors compare results from these bands? Specify which band was used for comparisons with direct sun, satellite, and model data.
  - The manuscript had previously stated that only the UV results were used for all comparisons in the manuscript (line 308). We did not initially include comparisons of HeiPro long UV vs. long vis in the manuscript, but the results are shown in (i) the scatter plot below which shows a zero-intercept slope of 0.97 ± 0.004 and mean relative bias of 0.67% ± 5.9%, and (ii) the table below which shows that the HeiPro long UV and long vis biases towards the partial columns are similar (mostly the same within uncertainties). The table has been added to the manuscript in the Appendix (Table A1).

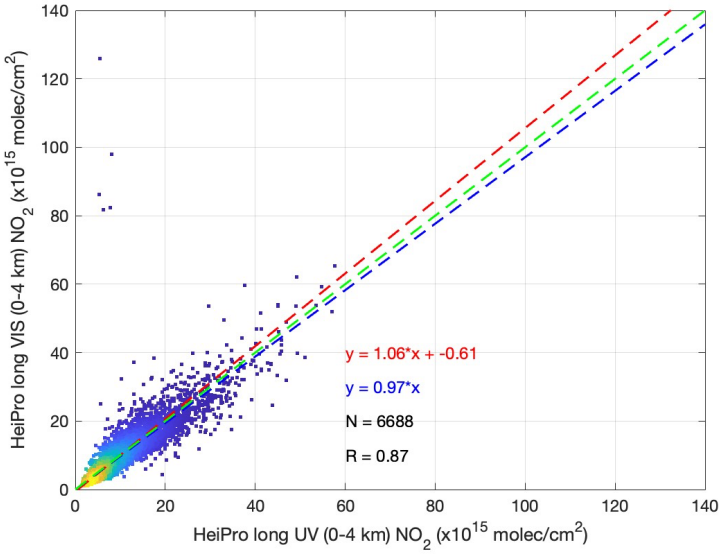
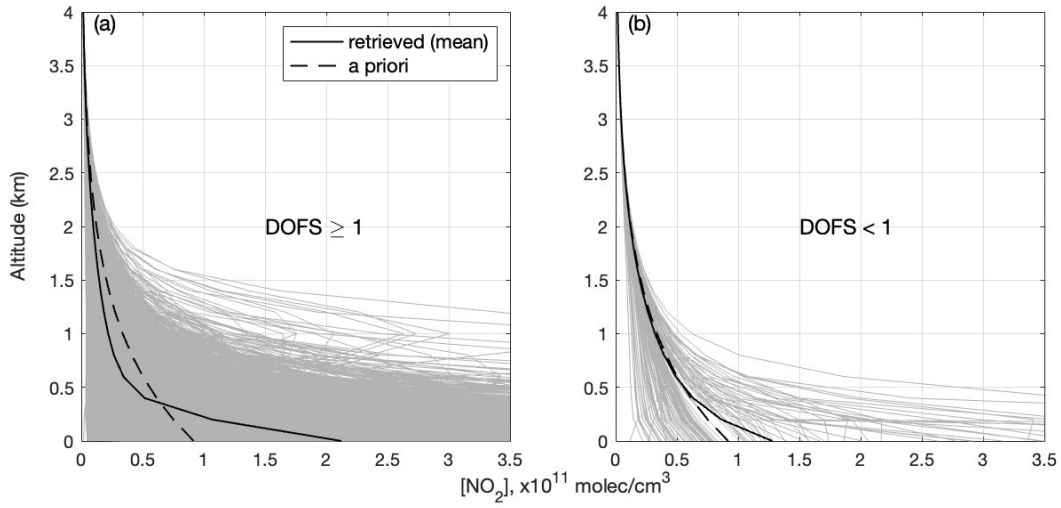


Table A1. Multiplicative biases and mean relative biases (± uncertainties) of HeiPro towards NO_2 partial columns from Pandora-DS, TROPOMI, and GEM-MACH, for both the HeiPro long UV and long vis results.

NO_2 Partial Column Comparison		HeiPro Scan Type	
		long UV	long vis
HeiPro vs. Pandora-DS partial columns	Multiplicative Bias	51% ± 0.8%	49% ± 0.9%
	Mean Relative Bias	61% ± 9.7%	61% ± 6.8%
HeiPro vs. TROPOMI	Multiplicative Bias	17% ± 4.0%	13% ± 4.6%
	Mean Relative Bias	37% ± 51%	40% ± 45%
HeiPro vs. GEM-MACH	Multiplicative Bias	12% ± 1.2%	13% ± 1.3%
	Mean Relative Bias	67% ± 7.1%	64% ± 2.4%

Quality Filtering and Averaging Kernels

- Section 2.2.2 discusses filtering criteria (e.g., DOFS < 1), but the rationale behind these criteria is not well explained:
    - Provide theoretical or empirical justification for the chosen threshold.
        - Thank-you for flagging this. A DOFS threshold of 1 is a conservative value and indicates that there is one independent piece of information from the retrieval. It is used, for example, in a study by Vlemmix et al. (2015) for quality control of aerosol, formaldehyde, and NO_2 profile retrievals. The plot below shows in panel (a) the profiles that meet the DOFS threshold and in panel (b) the profiles that do not. As shown in panel (b), the mean of the NO_2 profiles with DOFS < 1 (solid black line) more closely resembles the a priori (dashed black line) when compared to the mean of the NO_2 profiles that meet the threshold in panel (a). Because profiles with DOFS $<$ 1 largely resemble the a priori profile, they are discarded as they contain limited information from the measurements. We modified the manuscript on lines 327–329 to now say:
        Lastly, retrievals for which both the NO_2 profiles and aerosol extinction profiles had DOFS < 1 were excluded from the analysis for quality control purposes (e.g., Vlemmix et al., 2015) and represented 19% of the dataset. Such retrievals contain limited information from the measurements and are more influenced by the a priori profile.



- What is the typical DOFS value in the retrieval? Including an averaging kernel plot would help visualize the retrieval sensitivity.
  - For the NO_2 UV retrievals, the mean DOFS value is 2.35 and the median DOFS value is 2.48. For the AOD UV retrieval, the mean DOFS value is 1.61 and the median DOFS value is 1.78. Histograms of the DOFS for the various NO_2 retrievals are already shown in Figure A1(a), but the mean values of the long UV retrievals of AOD and NO2 before and after filtering are now stated in the manuscript (lines 329–331):
    The mean DOFS values before and after filtering are 1.61 $\pm$ 0.68 and 1.88 $\pm$ 0.42 for aerosol extinction, respectively, and 2.35 $\pm$ 0.49 and 2.39 $\pm$ 0.40 for NO_2, respectively.
  - Thanks for this suggestion. Averaging kernel plots for O_4 and NO_2, together with timeseries of O_4 and NO_2 dSCDs, have been added to the Appendix (Figure A3). It is also displayed below for reference:

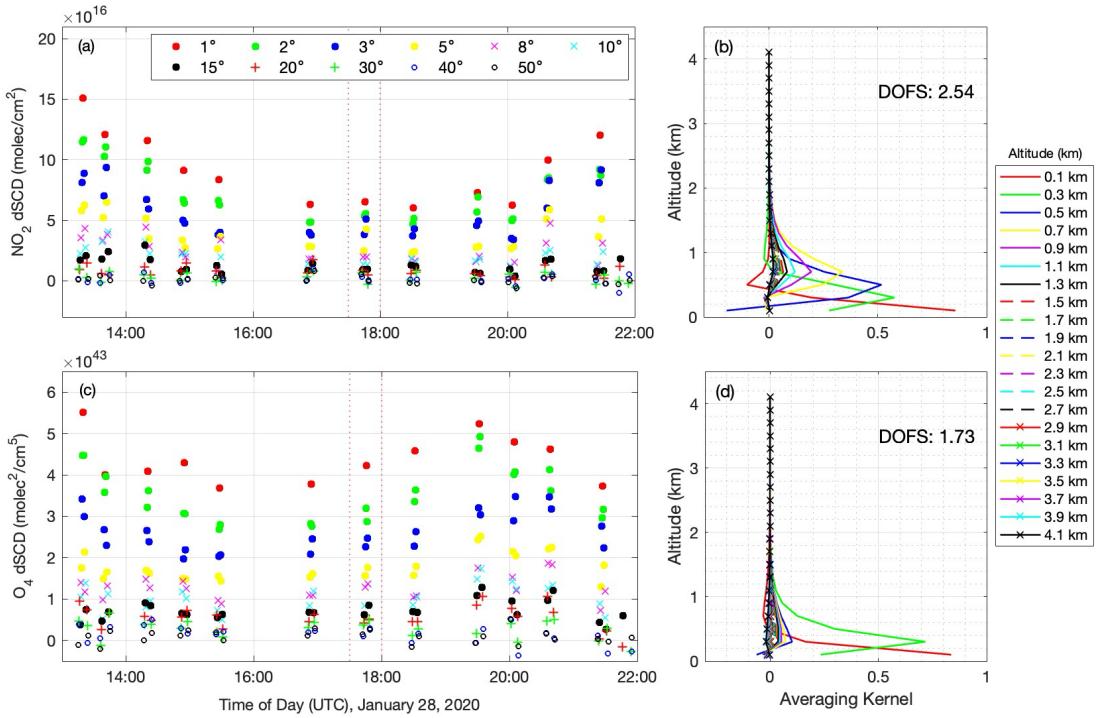


Figure A3: MAX-DOAS dSCDs (UV, 338–370 nm) of (a) NO_2 and (c) O_4 retrieved on January 28, 2020. The right panels show the averaging kernels and respective DOFS for a single HeiPro profile retrieval of (b) NO_2 and (d) aerosol extinction. The dSCDs used in the NO2 and aerosol extinction profile retrievals are indicated by the red dashed lines in panels (a) and (c), respectively.

- Demonstrating the impact of data filtering on the results would also improve transparency.

  The profiles that were filtered out (with AOD and NO_2 DOFS < 1) represent 19% of the dataset, with the majority of this 19% removed due to the AOD DOFS being below 1. This mostly does not have a large impact on the data (for example, partial column comparisons of HeiPro to Pandora-DS and TROPOMI remain essentially unchanged before and after filtering, but the comparisons to GEM-MACH do change). We have now included a statement that the DOFS filter excludes 19% of the profile retrievals as noted above (lines 327–329).

Comparison of Data Sources

- In the results section, the biases between HeiPro, Pandora-DS, TROPOMI, and GEM-MACH are analyzed, but the spatial, temporal, and observational characteristics of these datasets are not sufficiently discussed:
  - Summarize the resolutions, error characteristics, and limitations of each dataset in a table for clarity.
    - We have added Table 1 in the manuscript, which includes the temporal, spatial resolutions, and errors/uncertainties of each dataset and is copied below for reference. We also add the following on lines 178–179:

Table 1 provides a summary of the datasets used in the study as well as some characteristics of the datasets such as resolutions and uncertainties.

Table 1. Overview of the key attributes of the NO_2 datasets used in this study.

Dataset	Temporal resolution	Horizontal resolution	Errors/uncertainties
Pandora-DS total column NO_2	90 s	< 4 km (for SZA < 50°) 4–17 km (for SZA 50°–80°) (Herman et al., 2009)	1.3×10^{15} molec/cm^2 (Herman et al., 2009, 2018)
OMI stratospheric NO_2	1 day	13 km × 24 km (Krotkov et al., 2017)	2×10^{14} molec/cm^2 (Krotkov et al., 2017)
HeiPro	22 minutes	5–10 km (effective pathlength, Ortega et al., 2015)	4.4×10^{14} molec/cm^2 (this work)
TROPOMI tropospheric NO_2	1 day	7 km × 3.5 km (5.5 km × 3.5 km since Aug. 2019)	8.5×10^{14} molec/cm^2 (Eskes and Eichmann, 2019)
GEM-MACH	1 hour	10 km × 10 km	N/A
In situ NO_2	1 minute	Point measurement	0.4 ppbv (Thermo Scientific, 2015)

- Quantify the contributions of individual factors (e.g., PBL height, SAA, and seasonality) to observed biases rather than relying solely on trend descriptions.
  - Contributions of individual factors to the bias of MAX-DOAS (HeiPro) to Pandora-DS have been quantified. Since the contributions of these individual factors varied throughout the day, we report ranges and maximum values rather than a single contribution percentage. For example, we stated the following in the Conclusions section of the manuscript on lines 900–902: "The PBL height, combined with the missing 0–15 m partial column in the Pandora-DS measurements, contributed a maximum of 8.4% of the multiplicative bias and 18% of the mean relative bias in the morning hours, with these values declining to < 5% in the evening hours". This was similarly done for the solar azimuth angle contribution and seasonal contribution. In summary, the PBL, solar azimuth angle, and seasonality contributed up to 8%, 27%, and 39% of the multiplicative bias, respectively, and 18%, 52%, and 85% of the mean relative bias, respectively.
  - For HeiPro comparisons to TROPOMI and GEM-MACH, quantifying the contributions of individual factors was not feasible:
    - Since the SAA does not affect the GEM-MACH data, it was not analyzed as a factor contributing to the bias. Additionally, for comparisons of HeiPro and GEM-MACH partial columns, effects of PBL height were not analyzed (whereas for HeiPro vs. Pandora-DS

comparisons, they were due to the missing 15 m in the Pandora-DS measurements).
        - Factors such as PBL height and SAA could not be quantified for HeiPro vs. TROPOMI due to the low temporal frequency of measurements per day, which limited the PBL heights and SAA ranges sampled. Similarly, seasonality could not be assessed for HeiPro vs. TROPOMI due to the low number of coincident measurements when separated by season.

Seasonal and Diurnal Trends

- Provide more detailed explanations for observed trends:
    - How do lower PBL heights in the early morning contribute to NO2 accumulation? Why are concentrations higher in winter? Correlate these trends with emission sources and meteorological conditions at the observation site.
        - The following text has been added to the manuscript on lines 481–484:
        This difference in detection may further be amplified by shallower PBL heights during winter months and morning hours when the PBL height is smaller due to lower surface temperatures and less boundary layer dynamics. This leads to less vertical mixing of pollutants, with NO_2 accumulating near the surface (e.g., Lin and McElroy, 2010; Chan et al., 2018; Schreier et al., 2019) where it is not captured by the Pandora-DS measurements in the first 15 m.
        - We now mention in the manuscript that NO_2 concentrations are higher in the winter due to (i) reduced sunlight and increased lifetime of NO_2, and (ii) increased heat-associated emissions during the winter months (lines 486–489):
        Lastly, while NO_2 increases during the wintertime due to greater anthropogenic emissions from heat sources (e.g., Meng et al., 2018) and increased lifetimes due to decreased solar radiation, it is possible that increased emissions can contribute to the bias if there are more NO_2 emissions coming from the multi-axis azimuth viewing direction of 255° compared to the various direct-Sun viewing angles.
            - We do not have NO_2 emissions data for the measurement site and surrounding region, but potential sources can be the local residential homes. Therefore, we only speculate that wintertime emissions may be higher in the direction of the multi-axis scans (255°).

Extrapolation and Vertical Profiles

- Discuss the limitations of linear extrapolation methods used in vertical NO2 profile retrievals. Explore non-linear methods for more accurate surface concentration estimates.
    - We agree that non-linear extrapolation methods may provide more accurate surface concentration estimates, and we will explore this for future studies when

performing such extrapolations. The linear extrapolation we used may be underestimating the NO_2 gradient from 100 m to 0 m. On average, the linear extrapolation used in this work increases the NO_2 by 1 ppbv from 100 m to 0 m. Without NO_2 profile information at the measurement site, it is difficult to know what a realistic increase would be from 100 m to 0 m, but we will explore this in the future. The following text has been added to the manuscript on lines 295–296 to address the limitations of the linear extrapolation:

The linear extrapolation method used in this study produces NO_2 surface values that are, on average, 1 ppbv larger than the NO_2 value at 100 m. Therefore, this extrapolation method may underestimate the HeiPro surface NO_2 values that we report.

- Analyze the changes in GEM-MACH profiles before and after smoothing across different height ranges, identifying layers with the most significant changes.
    - In the manuscript, we now state the following on lines 837–846:

      Across all seasons, the HeiPro NO_2 median profiles from 0–200 m underestimate the unsmoothed GEM-MACH median values, while from 1.5–4 km, the HeiPro median profiles then overestimate the unsmoothed GEM-MACH median values. For the 0–200 m layer, the mean relative bias of HeiPro towards GEM-MACH decreases from −37% (unsmoothed) to −6.1% (smoothed). Note that these biases are representative of the integrated 0–200 m layer and may differ slightly from the surface values reported in Table 2. The most significant changes occur in the layer from 1.5–4 km, where the HeiPro bias towards GEM-MACH decreases from > 1000% (unsmoothed) to 2.6% (smoothed). The HeiPro surface underestimation and free tropospheric overestimation of the unsmoothed GEM-MACH profiles can probably be explained by the NO_2 inventories used in the GEM-MACH model, which, respectively, (i) utilize older inventories that do not account for reduced emissions over the years, and (ii) do not account for free tropospheric NO_2 sources while the a priori NO_2 profile contains free tropospheric NO_2.

Minor Comments

Technical Issues

1. **Line 225**: The term "DOF" is inconsistently introduced. Ensure its full form and abbreviation are aligned throughout the text.
    - The full form (degrees of freedom for signal) and abbreviation (DOFS) are defined on lines 314–315. We have checked that DOFS is subsequently used consistently within the text.
2. Standardize terminology for "Pandora direct-Sun" and "Pandora-DS" to avoid confusion.
    - The full form (Pandora direct-Sun) and abbreviation (Pandora-DS) are defined on line 120. We have checked that Pandora-DS is subsequently used consistently within the text.

Figures and Tables

- Figures 3 and 4: Clarify the differences in datasets and ensure consistent labeling.

- Differences in: (i) the number of points in Figs. 3a versus 3c, and (ii) the box-and-whisker plots in Fig. 4 are as discussed above. The y-axis label in Fig. 3a has been changed to "HeiPro (0–4 km) NO_2 Column ($\times 10^{15}$ molec/cm^2)" for consistency.

- Add a table summarizing key attributes of the datasets (e.g., resolution, uncertainties).
  - Table 1 (shown above) summarizing the key attributes of all datasets has now been added to the manuscript.

Section 3.1.1

- The calculation of the 0–15 meter column concentration relies on surface NO2 measurements. Discuss the reliability of this assumption and its impact on comparisons. Quantify the average 0–15 meter column in absolute and relative terms.
  - The mean 0–15 meter column for time periods coincident with the Pandora-DS data is 2.84×10^{14} molec/cm^2. It represents 4.6% $\pm$ 3.1% of the modified Pandora-DS tropospheric NO_2 column (mean and standard deviation).
  - We do not have information on how the NO_2 varies from 0 m to 15 m, and so it is difficult to discuss the reliability of this assumption and its impact on comparisons. This vertical resolution of 15 m is finer than the grids of the HeiPro and GEM-MACH datasets we are using. In the best-case scenario, the 0 m and 15 m NO_2 values are the same. In this scenario, we can use the NAPS in situ error estimate of 0.4 ppbv, converted to a column value from 0–15 m, to obtain an error estimate for the 0–15 meter column for each data point (mean error estimate of 1.53×10^{13} molec/cm^2). On average, the errors represent 9.8% of the 0–15 meter column and 0.34% of the modified Pandora-DS tropospheric NO_2 column. In the worst-case scenario, the NO_2 value at 0 m drops to 0 ppbv at 15 m. In this extreme scenario, the mean 0–15 meter column for time periods coincident with the Pandora-DS data is 1.42×10^{14} molec/cm^2. It represents 2.4% $\pm$ 1.7% of the modified Pandora-DS tropospheric NO_2 column (mean and standard deviation). On average, the error estimates (which comes from the 0.4 ppbv uncertainty in the in situ measurements) now represent 0.35% of the modified Pandora-DS tropospheric columns. Therefore, in the worst-case scenario, the 0–15 m column represents ~2% less of the modified Pandora-DS tropospheric columns currently used in the study. The effect that this has on Fig. 5, where the PBL contribution + missing 15 m contribution in the Pandora-DS measurements are quantified, is that the PBL contribution is effectively reduced by about half (e.g., from 19% mean relative bias to 8.5% at the SAA range of 66°–85°). Although not an insignificant reduction, this reduction will only occur with the unrealistic scenario that the surface NO_2 at some VMR drastically reduces to 0 ppbv at 15 m. Therefore, we now state in the manuscript the following on lines 554–557:
    Because the 0–15 m column derivation assumes a constant NO_2 VMR from 0 to 15 m, the estimations provided here of the PBL contribution to the bias represent upper limits (aside from cases with lofted plumes below 15 m). In cases where the NO_2 VMR at 15 m $<$ NO_2 VMR at 0 m, the contribution of the 0–15 m column would be lower, and therefore, the PBL contribution to the bias would be lower.

References

Brion, J., Chakir, A., Daumont, D., Malicet, J., and Parisse, C.: High-resolution laboratory absorption cross section of O_3. Temperature effect, Chem. Phys. Lett., 213, 610–612, https://doi.org/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, https://doi.org/10.1023/a:1006036924364, 1998.

Cede, A.: Manual for Blick Software Suite 1.8, available at: https://www.pandonia-global-network.org/wpcontent/uploads/2021/09/BlickSoftwareSuite_Manual_v1-8-4.pdf, last accessed: 16 April 2024.

Chan, K. L., Wiegner, M., Wenig, M., and Pöhler, D.: Observations of tropospheric aerosols and NO_2 in Hong Kong over 5 years using ground based MAX-DOAS, Science of The Total Environment, 619–620, 1545–1556, https://doi.org/10.1016/j.scitotenv.2017.10.153, 2018.

Choi, S., Lamsal, L. N., Follette-Cook, M., Joiner, J., Krotkov, N. A., Swartz, W. H., Pickering, K. E., Loughner, C. P., Appel, W., Pfister, G., Saide, P. E., Cohen, R. C., Weinheimer, A. J., and Herman, J. R.: Assessment of NO_2 observations during DISCOVER-AQ and KORUS-AQ field campaigns, Atmos. Meas. Tech., 13, 2523–2546, https://doi.org/10.5194/amt-13-2523-2020, 2020.

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, https://doi.org/10.1007/bf00053756, 1992.

Eskes, H. J. and Eichmann, K.-U.: S5P Mission Performance Centre Nitrogen Dioxide [L2 NO_2] Readme, 2019.

Fleischmann, O. C., Hartmann, M., Burrows, J. P., and Orphal, J.: New ultraviolet absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy, J. Photoch. Photobio. A, 168, 117–132, https://doi.org/10.1016/j.jphotochem.2004.03.026, 2004.

Herman, J., Cede, A., Spinei, E., Mount, G., Tzortziou, M., and Abuhassan, N.: NO_2 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, D13307, https://doi.org/10.1029/2009JD011848, 2009.

Herman, J., Spinei, E., Fried, A., Kim, J., Kim, J., Kim, W., Cede, A., Abuhassan, N., and Segal-Rozenhaimer, M.: NO_2 and HCHO measurements in Korea from 2012 to 2016 from Pandora spectrometer instruments compared with OMI retrievals and with aircraft measurements during

the KORUS-AQ campaign, Atmos. Meas. Tech., 11, 4583–4603, https://doi.org/10.5194/amt-11-4583-2018, 2018.

Kreher, K., Van Roozendael, M., Hendrick, F., Apituley, A., Dimitropoulou, E., Frieß, U., Richter, A., Wagner, T., Lampel, J., Abuhassan, N., Ang, L., Anguas, M., Bais, A., Benavent, N., Bösch, T., Bognar, K., Borovski, A., Bruchkouski, I., Cede, A., Chan, K. L., Donner, S., Drosoglou, T., Fayt, C., Finkenzeller, H., Garcia-Nieto, D., Gielen, C., Gómez-Martín, L., Hao, N., Henzing, B., Herman, J. R., Hermans, C., Hoque, S., Irie, H., Jin, J., Johnston, P., Khayyam Butt, J., Khokhar, F., Koenig, T. K., Kuhn, J., Kumar, V., Liu, C., Ma, J., Merlaud, A., Mishra, A. K., Müller, M., Navarro-Comas, M., Ostendorf, M., Pazmino, A., Peters, E., Pinardi, G., Pinharanda, M., Piters, A., Platt, U., Postylyakov, O., Prados-Roman, C., Puentedura, O., Querel, R., Saiz-Lopez, A., Schönhardt, A., Schreier, S. F., Seyler, A., Sinha, V., Spinei, E., Strong, K., Tack, F., Tian, X., Tiefengraber, M., Tirpitz, J.-L., van Gent, J., Volkamer, R., Vrekoussis, M., Wang, S., Wang, Z., Wenig, M., Wittrock, F., Xie, P. H., Xu, J., Yela, M., Zhang, C., and Zhao, X.: Intercomparison of NO_2, O_4, O_3 and HCHO slant column measurements by MAX-DOAS and zenith-sky UV–visible spectrometers during CINDI-2, Atmos. Meas. Tech., 13, 2169–2208, https://doi.org/10.5194/amt-13-2169-2020, 2020.

Krotkov, N. A., Lamsal, L. N., Celarier, E. A., Swartz, W. H., Marchenko, S. V., Bucsela, E. J., Chan, K. L., Wenig, M., and Zara, M.: The version 3 OMI NO_2 standard product, Atmos. Meas. Tech., 10, 3133–3149, https://doi.org/10.5194/amt-10-3133-2017, 2017.

Lin, J.-T. and McElroy, M. B.: Impacts of boundary layer mixing on pollutant vertical profiles in the lower troposphere: Implications to satellite remote sensing, Atmospheric Environment, 44, 1726–1739, https://doi.org/10.1016/j.atmosenv.2010.02.009, 2010.

Meller, R. and Moortgat, G. K.: Temperature dependence of the absorption cross sections of formaldehyde between 223 and 323 K in the wavelength range 225–375 nm, J. Geophys. Res., 105, 7089–7101, https://doi.org/10.1029/1999JD901074, 2000.

Meng, K., Xu, X., Cheng, X., Xu, X., Qu, X., Zhu, W., Ma, C., Yang, Y., and Zhao, Y.: Spatio-temporal variations in SO_2 and NO_2 emissions caused by heating over the Beijing-Tianjin-Hebei Region constrained by an adaptive nudging method with OMI data, Science of The Total Environment, 642, 543–552, https://doi.org/10.1016/j.scitotenv.2018.06.021, 2018.

Ortega, I., Koenig, T., Sinreich, R., Thomson, D., and Volkamer, R.: The CU 2-D-MAX-DOAS instrument – Part 1: Retrieval of 3-D distributions of NO_2 and azimuth-dependent OVOC ratios, Atmos. Meas. Tech., 8, 2371–2395, https://doi.org/10.5194/amt-8-2371-2015, 2015.

Rothman, L., Gordon, I., Barber, R., Dothe, H., Gamache, R., Goldman, A., Perevalov, V., Tashkun, S., and Tennyson, J.: HITEMP, the high-temperature molecular spectroscopic database, J. Quant. Spectrosc. Ra., 111, 2139–2150, https://doi.org/10.1016/j.jqsrt.2010.05.001, 2010.

Schreier, S. F., Richter, A., and Burrows, J. P.: Near-surface and path-averaged mixing ratios of NO_2 derived from car DOAS zenith-sky and tower DOAS off-axis measurements in Vienna: a case study, Atmos. Chem. Phys., 19, 5853–5879, https://doi.org/10.5194/acp-19-5853-2019, 2019.

Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W., and Burrows, J. P.: High spectral resolution ozone absorption crosssections – Part 2: Temperature dependence, Atmos. Meas. Tech., 7, 625–636, https://doi.org/10.5194/amt-7-625-2014, 2014.

Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O_2–O_2 collision pairs between 340 and 630 nm and at atmospherically relevant pressure, Phys. Chem. Chem. Phys., 15, 15371–15381, https://doi.org/10.1039/C3CP50968K, 2013.

Thermo Scientific. (2015). Model 42i instruction manual: Chemiluminescence NO-NO_2-NO_x analyzer (Part No. 101350-00, 25JUL2015). Thermo Fisher Scientific.

Vandaele, A., Hermans, C., Simon, P., Carleer, M., Colin, R., Fally, S., Mérienne, M., Jenouvrier, A., and Coquart, B.: Measurements of the NO_2 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. Ra., 59, 171–184, https://doi.org/10.1016/S0022- 4073(97)00168-4, 1998.

Vlemmix, T., Hendrick, F., Pinardi, G., De Smedt, I., Fayt, C., Hermans, C., Piters, A., Wang, P., Levelt, P., and Van Roozendael, M.: MAX-DOAS observations of aerosols, formaldehyde and nitrogen dioxide in the Beijing area: comparison of two profile retrieval approaches, Atmos. Meas. Tech., 8, 941–963, https://doi.org/10.5194/amt-8-941-2015, 2015.

York, D., Evensen, N. M., Martínez, M. L., and De Basabe Delgado, J.: Unified equations for the slope, intercept, and standard errors of the best straight line, American Journal of Physics, 72, 367–375, https://doi.org/10.1119/1.1632486, 2004.

Zhao, X., Griffin, D., Fioletov, V., McLinden, C., Davies, J., Ogyu, A., Lee, S. C., Lupu, A., Moran, M. D., Cede, A., Tiefengraber, M., and Müller, M.: Retrieval of total column and surface NO_2 from Pandora zenith-sky measurements, Atmos. Chem. Phys., 19, 10619–10642, https://doi.org/10.5194/acp-19-10619-2019, 2019.

Zhao, X., Griffin, D., Fioletov, V., McLinden, C., Cede, A., Tiefengraber, M., Müller, M., Bognar, K., Strong, K., Boersma, F., Eskes, H., Davies, J., Ogyu, A., and Lee, S. C.: Assessment of the quality of TROPOMI high-spatial-resolution NO_2 data products in the Greater Toronto Area, Atmos. Meas. Tech., 13, 2131–2159, https://doi.org/10.5194/amt-13-2131-2020, 2020.

Zhao, X., Fioletov, V., Alwarda, R., Su, Y., Griffin, D., Weaver, D., Strong, K., Cede, A., Hanisco, T., Tiefengraber, M., McLinden, C., Eskes, H., Davies, J., Ogyu, A., Sit, R., Abboud, I., and Lee, S. C.: Tropospheric and Surface Nitrogen Dioxide Changes in the Greater Toronto Area during the First Two Years of the COVID-19 Pandemic, Remote Sensing, 14, 1625, https://doi.org/10.3390/rs14071625, 2022.