REPLY TO REFEREE #1

We thank the referee for the thorough review of our manuscript, as well as for the careful evaluation, constructive suggestions, and positive feedback. Referee comments are reproduced in black, and our replies are given in blue.

Review of the study „Validation of TROPOMI and WRF-Chem NO2 across seasons using SWING+ and surface observations over Bucharest“ by Antoine Pasternak et al.

The study focuses on validating a high-resolution chemistry-transport model and satellite retrievals of tropospheric nitrogen dioxide (NO_2) over the urban region of Bucharest. The authors run the WRF-Chem model with 1 km resolution over the Bucharest region for 17 two-day time series in 2021 and 2022 across the different seasons, and compare its output with ground-based in situ meteorological and chemical observations as well as airborne column measurements from the SWING+ instrument (17 flights between 2021–2022). They evaluate the satellite-based TROPOMI tropospheric NO_2 column product (v2.4.0) using the airborne SWING+ data as reference, and WRF-Chem as an intercomparison. The main findings are the different biases for different concentration ranges and the seasonality of the results. Results are compared to existing TROPOMI validation studies. They show that NOx emissions from the CAMS-REF inventory need to be scaled.

General comments:

The paper presents valuable and comprehensive validation work for both TROPOMI and WRF-Chem NO_2 data over the Bucharest region. It provides detailed comparisons combining satellite, model, airborne, and surface observations over an urban area with complex emission patterns. The analysis across different seasons and atmospheric conditions adds further strength.

I think it needs discussion, if not only a scaling of the CAMS-REG, but a seasonal dependent scaling is necessary. Can you also comment on how region-dependent your scaling factor is? I think the seasonal dependency in the validation results should, in general, be more highlighted in the abstract and conclusion.

We thank the referee for this invitation to further discussion. We acknowledge that a constant scaling factor is unlikely to capture the full complexity of emission variability. However, deriving credible seasonal scaling factors is beyond the scope of our analysis. This would require a robust mass-balance approach across the available flight dates to isolate seasonal emission effects, which is nontrivial given that some seasons are less represented and the influence of various factors besides emissions on NO_2 levels.

Moreover, our 1.5 factor is sufficient to close the gap with SWING+ magnitudes on average (see Table 6), which is adequate for our purpose of validating TROPOMI, since the remaining day-to-day discrepancies are handled through the bias-correction procedure described in Section 3.3.1.

We note that the preprint by Hohenberger et al. (2025) indicates that CAMS-REG underestimates road-transport urban NO_x emissions by about 35 % in European cities, based on a comparison with independent urban inventories. This is broadly consistent with our results, given the expected variability in uncertainties across cities, and helps explain why our factor applies to the Bucharest region.

Following the referee's advice, both this last comment and the seasonal dependency of TROPOMI bias have been highlighted in the conclusion. We also modified the abstract to add details on the seasonal analysis. These modifications are presented in response to the specific comments referring to lines L558 and L577/578 below.

I think section 4 would benefit from being less of a review and including more explicit comparisons of this study's results with previous validation studies.

Following the referee's advice, we have added details on our seasonal results in the context of the existing literature, where comparisons could be made. Please see our response to the specific comment referring to lines L577/578.

Overall, this is a well-structured, technically sound paper, and its results are relevant to the air quality satellite and modelling community. I recommend publication after minor adjustments.

Specific comments:

L23: Missing reference.

The reference of Seinfeld and Pandis (2016) has been added.

L53: Add a reference for the lifetime information.

The reference of Laughner et al. (2019) has been added.

L55: Add a reference for the expected seasonality.

The reference of Boersma et al. (2009) has been added.

L101: What is this highest resolution available?

The resolution varies between datasets; all are referenced on the website cited in the manuscript.

L103: Which ERA5 variables have been used?

We added a comprehensive list in Supplement 1.

L110: Instead of "each chemical species" maybe better "several chemical species". Which have you used?

The first paragraph of Sect. 2.1.2 has been updated to list the various chemical species included in the CAMS-REG inventory.

L113: How do the monthly factors vary by season for the main sectors contributing to the emissions over Bucharest? Do you know how these factors are determined?

The figure of CAMS-REG NO_x sectoral emissions over Bucharest was replaced with a seasonal emission plot that retains sectoral information in Sect. 2.1.2. The text now highlights that the temporal factors from TNO rely primarily on energy consumption statistics and traffic counts (Denier van der Gon et al., 2011).

Table 2: I would suggest putting this table in the attachment. How is the NO/NO2 ratio determined for the mapping of CAMS-REG NOx into MOZART-4 NO and NO2 and how is this influencing your results?

We moved the Table to Supplement 2, next to the following comments.

"NO_x species are distributed as 90% NO and 10% NO_2 following the MOZCART users' guide (https://www.acom.ucar.edu/wrf-chem/MOZCART_UsersGuide.pdf, last access: 14 January 2026). Different values were reported in the literature, with NO_2/NO_x ratios ranging from 5.3% to 39% (Kuhn et al. (2024)). However, we expect this choice to have a negligible effect on the results, since NO has a very short lifetime during daytime (a few minutes; Seinfeld and Pandis (2016)) and photochemical equilibrium is rapidly established."

L160: "The first data point is recorded at 01:00 LT." Why is this different from the meteorology comparison?

We decided to keep 24 RNMCA data points per day, starting with the one recorded at 01:00 LT and effectively covering the first hour of each day to ensure daily representativeness. In comparison, MARS provides meteorological data at one-minute intervals, for which we retained 1440 measurements per day (60×24) to maintain daily representativeness.

Equation 1: How were the composition and ratios determined?

We added a note to emphasize that these ratios are prescribed by Lamsal et al (2008). They have also been used in previous modeling studies (e.g., Poraicu et al. (2023), Kuhn et al. (2024)).

L216: Just to clarify, with these errors, you mean the AMF, SCDre, and DSCDs errors?

We mention the contribution of AMF, SCD_{ref}, and DSCDs errors to the VCD error. This sentence has been revised to clarify our point:

"The combined contributions of the AMF, SCD_{ref}, and DSCD yield a total VCD uncertainty of 0.9-1.9 × 10^{15} molec. cm^{-2}."

L226: You are integrating over the troposphere, but SWING+ only measured below 3km, is this correct?

The light-path from the Sun to an airborne nadir instrument, such as SWING+, traverses the upper troposphere. As such, these measurements are sensitive to the entire troposphere, even though their sensitivity is highest below the aircraft altitude. Averaging kernels are defined over the full altitude range, in order to take these variations into account (Merlaud et al. (2018); Tack et al. (2021)).

L318-320: Do you mean, it is expected to see a morning, a late afternoon and an evening peak? Because you see two peaks (morning and evening), which I thought are both related to the rush hour. But if I understand your discussion correctly, you expect an late afternoon rush hour peak? What is causing the evening peak?

As explained by Poraicu et al. (2023), there is no late-afternoon peak due to the counterbalancing effects of chemistry and planetary boundary layer mixing. The evening peak occurs because emissions remain high (though much lower than in the late afternoon), while chemical loss and mixing are much weaker.

L331: This might be because the second day is always the flight day, so usually a clear-sky day.

This is correct. However, the original sentence was misleading, as we intended to emphasize that the second winter day shows better agreement compared to other seasons, rather than relative to the first day. We have revised the text accordingly: “The second day of the winter runs shows the best daytime agreement, relative to other seasons. However, because the winter analysis is based on only two time series, it is difficult to draw definitive conclusions regarding which season is best reproduced by the model.”

L373-376: Move “It also provides statistics per season and for the entire dataset. For two dates, reported in the table, we truncate data associated with the beginning of the flight for reasons explained in Sect. 3.2.2.” to L373 after “for each separate flight.”. Which is the other truncated day? Selecting 13:24 LT instead of what time?

The paragraph has been revised following the referee’ suggestion and adding some information on the truncated flight and start time.

L385: Be careful in the discussion, and remind the reader that some seasonal biases might compensate each other.

We adapted the paragraph accordingly:

“The comparable numbers of days with either positive (7) or negative (10) biases in Table 6 suggest a balanced model behavior on average. The small overall bias across all selected dates (MB of 0.5×10^{15} molec. cm^{-2} and RB of 13%), along with the underestimation in surface NO_2* found in Sect. 2.2.2 (MB of -8 μg m^{-3} and RB of -33%), provides a retrospective justification for increasing the CAMS-REG anthropogenic NO_x emissions by a factor of 1.5, as proposed in Sect. 2.1.2.

The small overall model bias against SWING+ reflects compensating seasonal biases of opposite sign, indicating that a temporally varying scaling factor for NO_x emissions may be more realistic. However, while finer, day-specific adjustments based on the column evaluations in Table 6 could be considered, they would likely introduce abrupt and potentially unrealistic temporal variations in emissions, e.g., in November 2021, when the mean model bias ranges from -5% to +125% across different days. This variability may reflect the fact that, in addition to emission uncertainties, the model daily performance (e.g., chemistry and transport) on a limited set of days can strongly influence seasonal statistics, particularly in winter and fall, whereas spring and summer appear more consistent."

Table 7: Do you have an idea why the agreement is so different between the days?

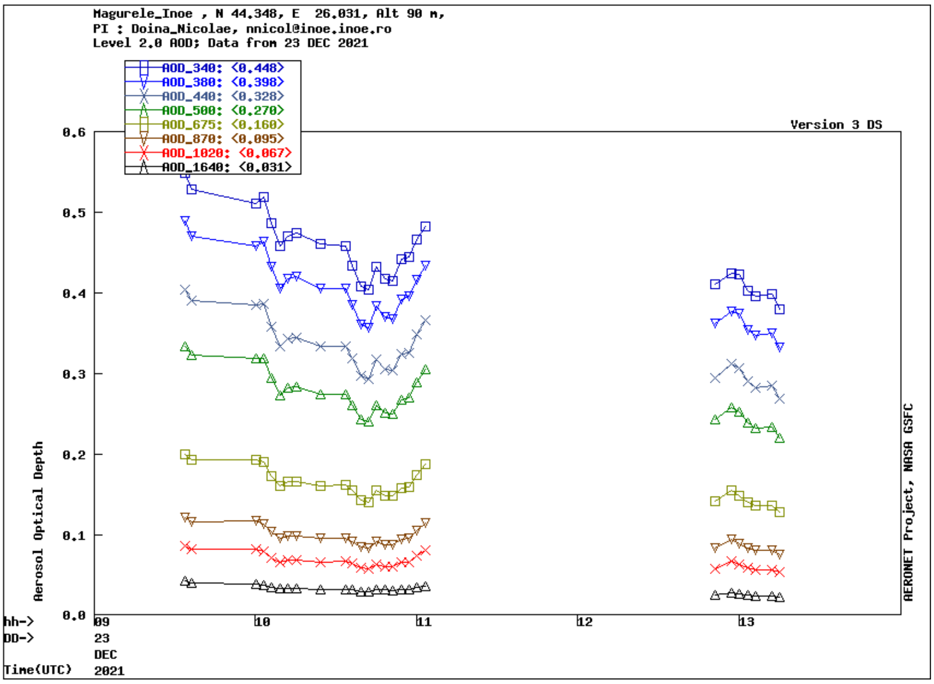
It is difficult to pinpoint the sources of model errors, as they can be numerous, including performance under specific meteorological conditions, boundary and initial conditions, or biases in emissions. The dramatic case of 22/11/2021 exemplifies a situation in which a wind direction change was poorly reproduced by the model, as discussed in the second paragraph of Section 3.2.3.

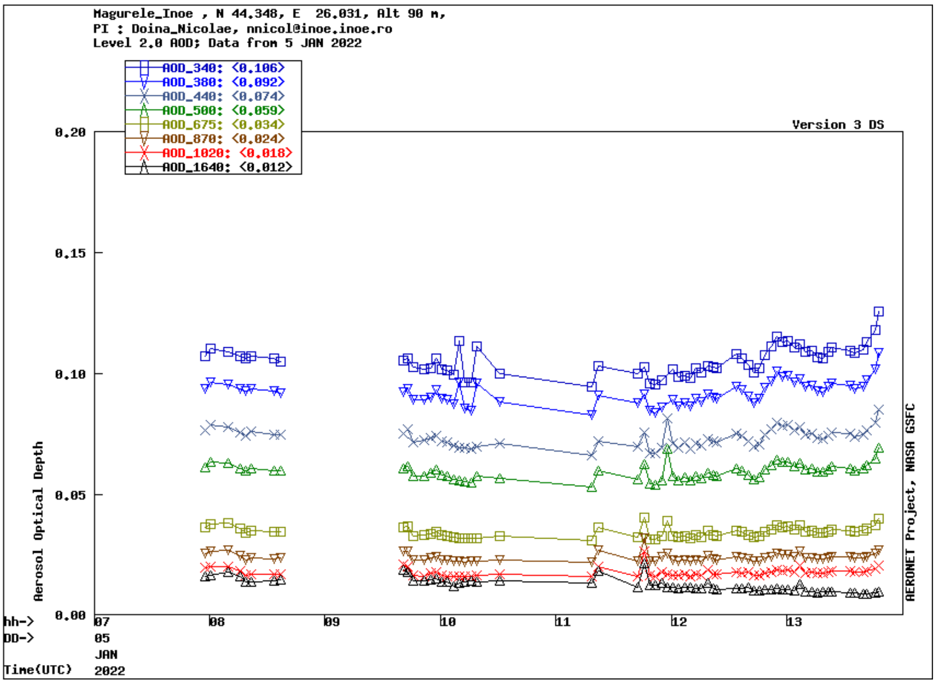
Line 471: Do you have any ideas why this day you have excluded is not working well?

We added a note in Sect. 3.3.2:

"The flight day of 23/12/2021 shows less convincing results (Table 6) and is characterized by consistently high modeled background values (see Supplement 4), which may be due to inaccurate initial or boundary conditions for NO_x species, oxidant concentrations, and/or heterogeneous chemistry on aerosols."

This last possibility is hinted at by a significant difference in aerosol optical depth (AOD) values measured at the INOE station in Măgurele on 23/12/2021 compared with 05/01/2022, as shown in the plots below (top: 23/12/2021, bottom: 05/01/2022).





Line 481: ...span from 2019 to 2025, cover several TROPOMI product versions and focus...

Done.

L487: and the upper bound from 10^16 to 1.5x10^16.

The upper bound was already placed at 1.5 × 10^{16} molec. cm^{-2} in Sect. 3.3.2. We adapted the sentence to refer to the upper bound, for clarity.

L558: See general comments. Please comment if seasonal scaling would be better and how region-dependent your factor might be.

We have adapted the line referring to our scaling of CAMS-REG into a paragraph in the conclusion, following the referee's advice.

"The underestimation of WRF-Chem NO and NO_2 daytime surface levels, along with the small positive bias for NO_2 modeled column magnitudes across different flight dates, supports an empirical upscaling of CAMS-REG v7.0 anthropogenic NO_x emissions over Bucharest. It is also consistent with the documented low bias in CAMS-REG road-traffic NO_x emissions in European cities with respect to independent urban inventories, estimated at approximately -35% (Hohenberger et al., 2025). The factor of 1.5 was sufficient for our purpose of validating TROPOMI. However, for a more in-depth assessment of the CAMS-REG inventory, different temporal profiles could be tested (e.g., Guevara et al. (2021)), and the overall magnitude could be adjusted seasonally using mass-balance inversion techniques (e.g., (Cooper et al., 2017; Poraicu et al., 2023))."

L577/578 Connect this to what you have found. How much seasonality was studied in these evaluations of the TROPOMI product? Might this be something that should be investigated further?

We added some details in Sect. 4:

"Finally, we assess the seasonal dependence of the TROPOMI bias. In our study, low-column biases range from -17% to 19% across seasons, while high-column biases range from -1% to -18%. Our summer results (-3% and -1% for low and high columns, respectively) agree well with the aircraft-based analysis for the PAL v2.3 product of Poraicu et al. (2023), with differences of 8% or less. Our fall results (19% and -4%) are consistent with those of Dimitropoulou et al. (2020) using recalculated AMF, with differences within 6%. They are also in line with Lange et al. (2023), showing positive biases for low columns. However, for high columns in fall, both our study and Lange et al. (2023) report negative biases, a feature captured by Dimitropoulou et al. (2020) only when using the original AMF. In contrast, winter and spring results show weaker consistency with Dimitropoulou et al. (2020). However, differences in methodology (notably the use of dual-scan MAX-DOAS observations) and in the TROPOMI product version limit direct comparability. This underscores the need for further validation studies, particularly in winter and spring, where comparable aircraft campaigns are lacking."

This is linked to an addition in the conclusion:

"Good agreement is found with seasonal studies comparing TROPOMI with aircraft (summer) and MAX-DOAS (fall) measurements, with differences relative to our results below 10%. The scarcity of seasonal studies and the differences in methodology, however, limit the comparability and highlight the need for more dedicated validation campaigns, particularly in winter and spring."

We also added a line to the abstract addressing the seasonality of our results:

"Seasonal diagnostics indicate variability in the bias for low columns, showing a marked positive bias in fall and negative biases in other seasons, whereas the negative bias at higher columns remains stable."

Technical corrections:

L6/7: Split the sentence into two: ...are underestimated. Satisfactory agreement with observations is achieved ...

Done.

L29: like the Global Ozone ...

Done.

L120: A preliminary evaluation ...

Done.

Fig.3: stationary instead of staionnary for sector C

This figure has been adapted into a table in Supplement 2.

L142: ...for the first 15 of the 17 SWING+ dates...

Done.

L143/144: Move the sentence "When available, ..." to the end of this subsection.

This sentence is specific to the MARS measurements and so we left it untouched.

Table 3: I would replace the 0 with "-", since the meaning is not that there are 0 overpasses but just no O3 measurements at these stations

Done.

L156: including NO and NO2, and for some of them also O3.

Done.

L201: Typical and typically so close to each other doesn't read very well.

Done.

L204: remove with

Done.

L237: DOAS was already introduced.

Done.

L245: and the offline product instead of and offline products

Done.

L252: To avoid confusions: Does "for its" refer to SWING+ or WRF-Chem?

We replaced "its" with "WRF-Chem".

L326: Add …modeled outputs, separated by seasons, defined as summer (June, July, August), …

Done.

L367: Thereafter, the modeled columns correlate…

Done.

L433: precision of the regression method

Done.

L476&478: For better readability you can try to add the seasons directly after each bias: -6+/-25% (summer), …

Done.

L504/505: Hard to read, check rewriting.

We replaced "Several studies recalculated the air mass factor (AMF) for versions v1.1 to v1.3 using alternative a priori profiles in place of the a priori profiles of the TROPOMI data based on the TM5-MP model." with "Several studies recalculated the air mass factor (AMF) for versions v1.1 to v1.4 using alternative a priori profiles in place of those from the TM5-MP model…"

L537: We assess the WRF-Chem performance…

Done.

References

- Boersma, K. F., Jacob, D. J., Trainic, M., Rudich, Y., De Smedt, I., Dirksen, R., and Eskes, H. J.: Validation of urban NO2 concentrations and their diurnal and seasonal variations observed from the SCIAMACHY and OMI sensors using in situ surface measurements in Israeli cities, Atmospheric Chemistry and Physics, 9, 3867–3879, https://doi.org/10.5194/acp-9-3867-2009, 2009.
- Cooper, M., Martin, R. V., Padmanabhan, A., and Henze, D. K.: Comparing mass balance and adjoint methods for inverse modelling of nitrogen dioxide columns for global nitrogen oxide emissions, Journal of Geophysical Research:

Atmospheres, 122, 4718–4734, https://doi.org/https://doi.org/10.1002/2016JD025985, 2017.

- Denier van der Gon, H., Hendriks, C., Kuenen, J., Segers, A., and Visschedijk, A.: Description of current temporal emission patterns and sensitivity of predicted AQ for temporal emission patterns, Tech. rep., TNO, Princetonlaan 6, 3584 CB Utrecht, The Netherlands, https://atmosphere.copernicus.eu/sites/default/files/2019-07/MACC_TNO_del_1_3_v2.pdf, 2011.
- Dimitropoulou, E., Hendrick, F., Pinardi, G., Friedrich, M. M., Merlaud, A., Tack, F., De Longueville, H., Fayt, C., Hermans, C., Laffineur, Q., Fierens, F., and Van Roozendael, M.: Validation of TROPOMI tropospheric NO2 columns using dual-scan multi-axis differential optical absorption spectroscopy (MAX-DOAS) measurements in Uccle, Brussels, Atmospheric Measurement Techniques, 13, 5165–5191, https://doi.org/10.5194/amt-13-5165-2020, 2020.
- Guevara, M., Jorba, O., Tena, C., Denier van der Gon, H., Kuenen, J., Elguindi, N., Darras, S., Granier, C., and Perez Garcia-Pando, C.: Copernicus Atmosphere Monitoring Service TEMPOral profiles (CAMS-TEMPO): global and European emission temporal profile maps for atmospheric chemistry modelling, Earth System Science Data, 13, 367–404, https://doi.org/10.5194/essd-13-367-2021, 2021.
- Hohenberger, T. L., Malki, M. E., Visschedijk, A., Guevara, M., Ramacher, P., Marongiu, A., Lanzani, G. G., Fossati, G., Kousa, A., Athanasopoulou, E., Kakouri, A., and Kuenen, J.: Link-based European road transport emissions for CAMS-REG v8.1 and a comparison to city inventories, Earth System Science Data Discussions, 2025, 1–28, https://doi.org/10.5194/essd-2025-428, 2025.
- Kuhn, L., Beirle, S., Kumar, V., Osipov, S., Pozzer, A., Bosch, T., Kumar, R., and Wagner, T.: On the influence of vertical mixing, boundary layer schemes, and temporal emission profiles on tropospheric NO2 in WRF-Chem – comparisons to in situ, satellite, and MAX-DOAS observations, Atmospheric Chemistry and Physics, 24, 185–217, https://doi.org/10.5194/acp-24-185-2024, 2024.
- Lamsal, L. N., Martin, R. V., van Donkelaar, A., 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, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2007JD009235, 2008.
- Lange, K., Richter, A., Schonhardt, A., Meier, A. C., Bosch, T., Seyler, A., Krause, K., Behrens, L. K., Wittrock, F., Merlaud, A., Tack, F., Fayt, C., Friedrich, M. M.,

Dimitropoulou, E., Van Roozendael, M., Kumar, V., Donner, S., Dorner, S., Lauster, B., Razi, M., Borger, C., Uhlmannsiek, K., Wagner, T., Ruhtz, T., Eskes, H., Bohn, B., Santana Diaz, D., Abuhassan, N., Schuttemeyer, D., and Burrows, J. P.: Validation of Sentinel-5P TROPOMI tropospheric NO2 products by comparison with NO2 measurements from airborne imaging DOAS, ground-based stationary DOAS, and mobile car DOAS measurements during the S5P-VAL-DE-Ruhr campaign, Atmospheric Measurement Techniques, 16, 1357–1389, https://doi.org/10.5194/amt-16-1357-2023, 2023.

- Laughner, J. L. and Cohen, R. C.: Direct observation of changing NOx lifetime in North American cities, Science, 366, 723–727, https://doi.org/10.1126/science.aax6832, 2019.
- Merlaud, A., Tack, F., Constantin, D., Georgescu, L., Maes, J., Fayt, C., Mingireanu, F., Schuettemeyer, D., Meier, A. C., Schonardt, A., Ruhtz, T., Bellegante, L., Nicolae, D., Den Hoed, M., Allaart, M., and Van Roozendael, M.: The Small Whiskbroom Imager for atmospheric compositioN monitorinG (SWING) and its operations from an unmanned aerial vehicle (UAV) during the AROMAT campaign, Atmospheric Measurement Techniques, 11, 551–567, https://doi.org/10.5194/amt-11-551-2018, 2018.
- Poraicu, C., Muller, J.-F., Stavrakou, T., Fonteyn, D., Tack, F., Deutsch, F., Laffineur, Q., Van Malderen, R., and Veldeman, N.: Cross-evaluating WRF-Chem v4.1.2, TROPOMI, APEX, and in situ NO2 measurements over Antwerp, Belgium, Geoscientific Model Development, 16, 479–508, https://doi.org/10.5194/gmd-16-479-2023, 2023.
- Seinfeld, J. H. and Pandis, S. N.: Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, 3rd Edition, John Wiley and Sons, New York, ISBN 978-1-118-94740-1, 2016.
- Tack, F., Merlaud, A., Iordache, M.-D., Pinardi, G., Dimitropoulou, E., Eskes, H., Bomans, B., Veefkind, P., and Van Roozendael, M.: Assessment of the TROPOMI tropospheric NO2 product based on airborne APEX observations, Atmospheric Measurement Techniques, 14, 615–646, https://doi.org/10.5194/amt-14-615-2021, 2021.