Authors' response to Referee Comments #2 regarding

'To new heights by flying low: Comparison of aircraft vertical NO2 profiles to model simulations and implications for TROPOMI NO2 retrievals'

by
T. Christoph V. W. Riess et al.

The original comments are in black, the authors' replies in blue.

Reviewer #2

This paper by Riess et al. compares aircraft-based vertical NO2 profiles to model simulations from TM5, CAMS-Europe, and LOTOS-EUROS. It investigates the influence of the a priori vertical NO2 profile on the TROPOMI tropospheric NO2 column retrieval. The comparison between actual aircraft measurements of NO2 vertical profiles and model simulations can provide valuable insights into the accuracy and reliability of model outputs. The analysis focuses on the polluted North Sea and the influence of shipping emissions. It is known that the a priori vertical NO2 profile is a dominating factor in the tropospheric NO2 column retrieval and that good knowledge of the vertical distribution is necessary. Since the TROPOMI NO2 product has started to be extensively used for the analysis of NO2 shipping signals, shipping emissions, and even single ships, analyzing the influence of the NO2 profile, especially over the North Sea, can provide valuable insights and improvements on the NO2 columns product. The study contributes to the field of atmospheric measurements by analyzing the vertical distribution of NO2 and its agreement with models and satellite retrievals. It is improving our understanding of atmospheric composition and the validation of satellite measurements. The results can contribute to improving the TROPOMI NO2 retrieval as well as model simulations. Therefore, I recommend publication in AMT after addressing the following comments.

We thank the reviewer for their constructive comments. For replies to the comments please see below. An updated version of the manuscript including the changes will be uploaded soon.

General comments:

The study is based on a relatively small dataset. Even if ten profiles are included in the analyses, these measurements have been performed on only four days, from which 3 out of 4 days are dominated by land outflow. When results are discussed, this should be mentioned, and representativeness should be included more in the discussions.

The referee makes a valid point, the number of sampled profiles and days is limited and we acknowledge that. In fact, in line 384-385 we recommend more such measurements at different locations, seasons and meteorological conditions. When presenting the profiles without land outflow (#3-6), we additionally state that these have been sampled at the same day, see line 367/268.

However, the presented profiles over a polluted sea are (to our knowledge) the very first of its kind and offer unique insights in the interplay of land and sea-based pollution and the uncertainties of satellite retrievals under these conditions. As land outflow from either the British islands or the Netherlands/Belgium is very common in this part of the North Sea, we believe our sampling to be representative of typical summer days and the two different sampling scenarios (outflow vs clean) as

beneficial to our study. Additional discussion of sample size and representativeness will be added in the discussion section.

The low NO2 concentrations in TM5 close to the surface compared to the measurements are mainly driven by a few profiles (3 & 6), which are too low; most profiles show an excellent agreement in the surface level. CAMS and LOTOS-EUROS, on the other hand, show similar low surface values for the same profiles as TM5, but since they show too high values on the other profiles at the surface level compared to the measurements and TM5, the mean looks better.

How would Fig. 1 (right panel) look like if the data are separated into land-outflow and clean cases, maybe this would be a useful additional plot.

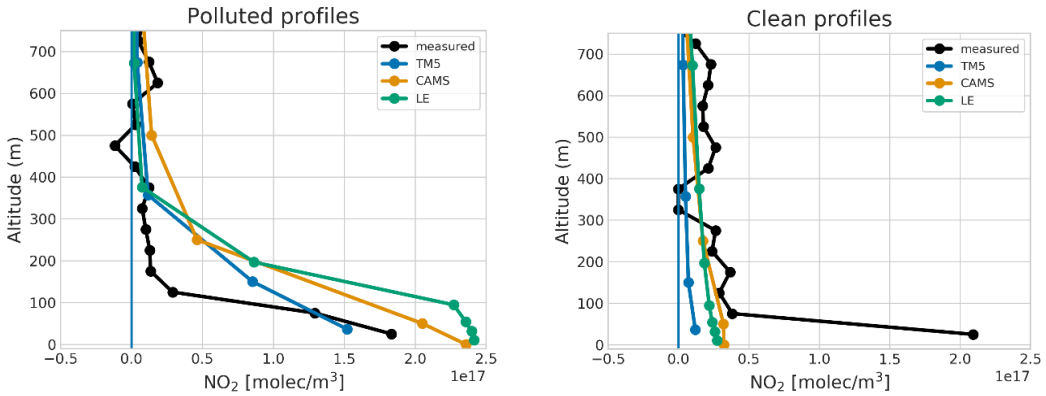
Regarding this comment see also, for example, Page 18, lines 362-368, and further comments in the specific comments section.

TM5 indeed gives a good agreement with the observed NO2 close to the surface for several outflow profiles (1,2,7,8,9), but not for others. Especially profile 10 shows very large differences, while LOTOS-EUROS and CAMS capture this situation a lot better. We will put more emphasize on the differences in agreement of surface level NO2 in different profiles in the result section, highlighting the differences in clean and outflow conditions.

As suggested by the referee, we show below the mean profiles for the clean ('Clean profiles') and outflow ('Polluted profiles') cases as classified in the main text. These two plots will also be added to the manuscript as an additional figure.

This separation provides two messages which will be incorporated in the discussion of the profiles in the manuscript:

1. All models successfully simulate the occurrence of outflow and match the observed surface pollution, but overestimate the (vertically integrated) amount of NO2. This hints at too long NO2 lifetimes in the model, and possibly at too much vertical mixing after emission.
2. None of the models captures the clear enhancement in the lowest 50 m due to NOx emissions from ships. This is due to the smeared out (in time and space) nature of ship emissions in the models, so we should not have expected to see clear ship NO2 enhancement in the first place.



The used heights for the analysis or plots are often 100 m or 200 m, for the plume dispersion flights in 40m are used. I don't see a clear explanation of why these specific heights were used. Can you comment on this? You mentioned that the modeled potential temperature profiles show a strong gradient in the lowest 400 m, so why not higher levels than 100m/200m for the analysis?

We use the zig-zag flights (flown at around 40m) to account for the in/out-of-plume ratio in the lowest 100m. The 40m were chosen by where the aircraft would usually find the center of the plumes and by where the gradient between in-plume and outside-plume are the largest. A clarification will be added to line 140 where the zig-zag flights are introduced.

We then applied the retrieved relationship between ships present and the ratio of area covered by the plume in the lowest 100m, because that is where the plumes sit, as can be seen e.g. in Fig1 (right) and Fig4 and the new figures above. We do not observe any signatures of plumes higher up in the profile flights, so there was not reason to apply the correction in what seems to be horizontally homogeneous air layers.

The only place where 200m was used the left panel in Fig1, which mainly serves as illustration of the area and NO2 range covered in the flights.

Specific comments:

Page 2 Line 43: "as well as urban OH concentrations" I think this is not very relevant to your study, I think this part can be deleted

While OH is not of immediate importance for our study, we believe the cited study show casts the use of the high resolution TROPOMI data for interpreting pollution plumes.

Page 3 line 67: "of up to 30%" is quite general. I think it would be good to be a bit more precise "typically between 5-30% strongly dependent on location and time"

This will be adjusted in the manuscript.

Page 4 Table 1: Please change the covered periods to:

29 November 2020 – July 2021

July 2021 – 14 November 2021

April 2018 – November 2021

April 2018 – November 2021 (?)

This will be adjusted in the manuscript.

What is meant by adjustment of surface albedo? Is the albedo database changed between the version? If yes, add the database which is used in the individual product.

Starting with version 2.3.1, the surface albedo is adjusted for individual scenes where the cloud retrieval gives negative cloud fractions using the original albedo database(Van Geffen et al., 2022). KFB: refer her to the TROPOMI ATBD version where this is introduced – Jos van Geffen knows where to find/how to cite this.The database itself has not changed. A clarifying sentence will be added to the manuscript.

Page 4 line 104, page 5 line 109: What is the phrase “residual cloud fraction” and “residual cloud scattering” mean? I would delete the word residual in line 104. Regarding line 109: the improvement of the new cloud algorithm is mainly that it provides a more realistic estimate of the cloud pressure for scenes with low cloud fractions.

We now remove the word ‘residual’. Indeed, the new cloud algorithm has an improved cloud pressure estimate for partially cloudy pixels.

Page 6 Figure 1: Strange references, they should perhaps instead appear at the first time mentioning the different models, here perhaps even simply omit or provide a reference to the data product itself, for TM5 probably from the TROPOMI product...? Why is there no reference for LOTOS-EUROS?

This is the first time TM5 and CAMS are mentioned in the Material part of the manuscript and therefore seems like the appropriate point to include references. Additionally, they are provided again in Table2 for the overview of all the model products used. A reference for LOTOS-EUROS will be added in the figure caption.

The CAMS reference is the correct reference for the regional AQ ensemble which is used in the manuscript (https://confluence.ecmwf.int/display/CKB/CAMS+Regional%3A+European+air+quality+analysis+and+forecast+data+documentation). A ‘last accessed’ date will be included. Additionally, a reference to details of the multi-model forecasting system will be added (Marécal et al., 2015).

The two references for TM5 refer to the model description and the product user manual of the TM5-MP product used for TROPOMI. In the data availability (lines 412-416) section the access point for these data is provided.

Page 6 line 146: You provide more detailed information about LOTOS-EUROS here, but please also provide some more information about the used TM5 and CAMS model data you used. Also, change the section title to Model data then.

We provide some more details on LOTOS-Euros because we did specific simulations with that model for this work. As we did not perform any dedicated model runs with TM5 or any of the CAMS models, and merely used the data products, we decided to simply refer to the primary reference for those two models in Table 2.

Page 6 line 157: Isn't wind data also needed for this? If you also use wind data, please mention which dataset is used.

Indeed, wind data is needed. This is measured by the aircraft as hinted at in line 134. A clarification of the wind data used will be added to line 157.

Page 7 Table 2: I think IFS is not a well-known abbreviation. Please add something like ECMWF Integrated Forecasting System (IFS).

This will be added to the manuscript.

Page 8 line 163: I think heterogeneity in general can be driven also by other factors like wind, lifetime, and other emission sources. The here observed heterogeneity might be mainly caused by the given factors. Please rewrite the sentence to avoid misunderstandings.

This will be rephrased to 'The *observed* heterogeneity…'

Page 8 line 178: How have you decided on the 100m height? Please give some additional information.

As also mentioned above, we did not observe plume signatures above 100m, which makes renders a correction for higher altitudes obsolete. Below 100m we deliberately approach plumes and spend over-proportionally much time in the plumes, making this correction necessary. The 100m value thus follows from the observed vertical appearance of plumes.

Page 9 line 1: How well can these horizontal scans represent the dispersion, especially for plumes not covered by these horizontal scans? How many plumes are analyzed in total, and how many have been covered by horizontal scans to account for plume dispersion? These in- and out-of-plume patterns were flown at an altitude of ~40 m (page 5, line 141). How representative is this height for the plume dispersion in 100m?

For the lowest part of each profile we typically crossed the plume of one ship 1-2 times. In the 3 scans, 12 clear plume crossings from 5 different ships have been sampled.

Ideally, a relationship between ship presence and plume covered area would have been determined at different altitudes. However, our priority was to sample all present plumes so we took the altitude where we observed the core of the plume (around 40m) in order not to miss a plume above or below the aircraft. With this, we might slightly overestimate the in-plume fraction at 50-100m but we have no reason to believe this to be significantly different from 40m and a major factor given the other uncertainties involved.

Page 10 line 226: Not all outflow cases show pollution close to the surface. Profile 9 doesn't show pronounced pollution close to the surface but is an outflow case profile. Are there any reasons why this profile looks different? Please comment on this.

In Fig S5 we can see that Profile 9 was sampled just North of where most of the outflow happens. We argue that this is the reason why not much NO2 was observed. However, profile #8 also seems to be at the edge of the outflow covered area in the model products and TROPOMI columns (Fig S5) where clear pollution is visible. It thus remains partially unclear.

Page 11 Fig. 4: The figure would benefit from mentioning the wind direction, respectively, the kind of scenario (land outflow or clean). This could be added in the title of the individual plots or at least in the caption.

We will highlight the scenario (clean or outflow) in the title and caption as well as adding wind arrows, good suggestion!

The plots in Fig. 4 are small and hard to read, especially the legend.

Legend and axis label fonts will be increased.

Page 12 line 233: "… boundary layer dynamics are a mix of sea and land characteristics for pixels at the coast"

This will added.

Page 12 line 233: It is stated that the "TM5 profiles show too little NO2 in the lowest layer" compared to measurements, but here you are discussing the land outflow cases with profiles 1,2,7,8,9,10, and for these profiles, the agreement at the surface level with the measurements is nearly perfect except for profile ten which is also stated in line 235.

Line 235 refers to the columns as shown in Figures 5, 7 and S5 and not the profiles. Indeed, the lowest level NO2 in TM5 is in very good agreement with measurements for most of the outflow cases. As mentioned above, this will be clarified in the manuscript. We want to thank the referee for pointing this out.

See also Page 13 line 262.

See also general comments.

See our replies above.

Page 12 line 234: You wrote, "Nonetheless, the coarse TM5 columns show reasonable agreement with TROPOMI retrieved columns with the exception of profile 10." I don't get the context. Where are the TM5 and TROPOMI columns compared for the respective profiles? Fig. 5 shows the columns, but only for profile 1, Fig. 4 shows profiles but no columns. Do you refer to a Figure in the supplement?

Yes, the columns can be seen in Fig5,7 and S5 (for profile 10 they are in S5). A reference to Fig S5 will be added here.

Page 12 line 256: “profile 2 (which was sampled right before)” right before what, the flight path which is shown in Fig. 6?

Yes, it is meant right before the flight path of Fig6 as mentioned in Line 253 ‘… after taking profile 2.’

Page 14 Fig. 7: Profiles 4,5,6 are taken on the same day. Are they taken in collocation to the same TROPOMI overpass? If yes, adding the locations of the other profile flights in the same figure would be helpful to see the location without looking into the supplement.

As shown in table 3, profile 3 & 4 are best matched with TROPOMI orbit 19551, whereas profiles 5 & 6 are sampled closer to the overpass of orbit 19552. As the locations are also shown in Fig1 (left), we don’t see benefit in adding it here as well. However we will mention in the caption, that Profile #4 is taken in collocation to the same TROPOMI orbit.

Page 14 line 273: You were just talking about profiles, but I think now you mean columns or both? Please be more precise.

It can be seen in both but we are referring to the column illustrations of Fig 7 and S5. We will make this more precise in the manuscript.

Page 14 line 281: “…are spatially diluted over the area of the model grid cell” add ”especially for the coarse TM5 model”

This will be added.

Page 15 line 301: “total tropospheric columns of NO2” is misleading since the total column describes the column to the top of the atmosphere, better write only “We compare tropospheric vertical columns of NO2”

Yes, this should indeed be vertical tropospheric columns.

Page 15 line 302: How are the “measured columns” retrieved? Please add more information. From the profile flights up to 1400m added with TM5…

This is described in lines 294-295: We used TM5 to fill the gap between 1400m and the tropopause and then integrated along the resulting profile.

Page 15 line 309: columns determined from aircraft measurements

This will be changed in the manuscript.

Page 16 Table 4: Why only a table, why not an additional scatter plot comparing the individual profiles?

A scatter plot showing 4 different TROPOMI products gets very busy. We believe the table is quickest and cleanest way to show the results, especially given the low number of samples.

Since measurements are only from 4 days, how many satellite pixels go into this comparison (correlation coefficient,…), are the profiles in different pixels when they were taken on the same day or different orbits (on days with two overpasses)?

For each measured profile we selected the single pixel with the largest overlap of the covered area. As each of the profiles is taken over and made representative for an area of the size of a single pixel, this is the logical choice.

None of the profiles cover the same pixel, and for some days the used pixels come from different orbits. In total, 10 unique TROPOMI pixels from 6 different orbits were used (see Table 3).

Page 16 line 326: Measured how at the tower or aircraft measurements at the tower?

Measured at the tower. This will be clarified in the manuscript.

Page 17 Fig. 8: Make clear how measurements left/right plot were made, aircraft/tower.

The left plot uses the aircraft data and TM5/AK profiles. The TM5/AK profiles in this panel are all sampled at the time of the aircraft profiles, for the sea case at the location of the aircraft profiles (as in Fig 1), for the land case at the Cabauw tower.

The right plot shows tower measurements and TM5 co-sampled at Cabauw during the tower measurements which were at a different day, but taken under similar meteorological conditions.

The TM5-land case on the left is included to ‘translate’ for the different sampling times of the sea and land measurements: TM5 shows deeper mixing of pollution over land not only during conditions of the TROLIX/land campaign but also during the day and time of the aircraft flights, compared to the (modeled and observed) sea profiles.

Page 17 line 336: This is a bit misleading since you also showed that land outflow dominates for 6 of the ten measured profiles, and ship emissions and land outflow are difficult to distinguish.

While land outflow is important in many of our profiles, they remain representative over the North Sea, which is often dominated by land outflow. In fact, we focused on the polluted North Sea exactly because it is not a remote and clean area and our study highlights this complication of different pollution sources. We do not claim to have evaluated TROPOMI over the North Sea for ship emissions only. As a matter of fact, this work can provide guidance under which circumstances to look for ship NO2 plumes in TROPOMI data (during northerly winds)

Page 17 line 346: Of course, it is correct that your measurements show no significant pollution above 150m with a very shallow boundary layer height. Nevertheless, these findings are the result of only four measurement days. Please mention here also possible other meteorological conditions, different seasons…

This is mentioned later on in the discussion in Lines 384-385.

Additionally, we have no reason to consider the days of our measurements as outliers for summer as temperatures and wind conditions are variable and not extreme. The fact that we sampled clean AND outflow conditions additionally gives us confidence that our findings are generalizable at least for summer days. Nonetheless, we will change line 348 to '…in the models compared to observations *on four summer days in 2021*.' for more transparency.

Page 18 line 363: "This is likely an effect of the coarse TM5 resolution of 1◦x1◦ where ship emissions are smeared out over a larger area and time." See general comments and earlier comments. This sounds like a general statement, but it can only be shown for the "clean" cases.

This will be clarified in the manuscript. 'the TM5 profiles *during clean conditions* show less…'

Page 19 conclusion: The conclusion and discussion on page 18 are very similar, especially lines 399-402 and 407-409 are very repetitive. Maybe the sections can be combined to avoid repetitions in the conclusion.

We indeed considered merging these two sections, but settled on separate sections for the following reasons:

We believe a separate conclusion section gives the reader the advantage of quickly grasping the main points of our study without burying them in a technical discussion. Naturally, some points of the discussion are repeated, but while the discussion focusses more on possible uncertainties, representativeness, shortcoming and future work, the conclusion is more concise and highlights the main findings of our study.

Technical corrections:

The suggested technical corrections will be implemented in the revised manuscript, unless otherwise commented below.

Page 1 line 4: delete "which was already", since it is not so important here

Page 1 line 10: add "the" to "and the LOTOS-EUROS model"

Page 1 line 11: improves instead of improve

Page 1 line 11: Very long sentence. Maybe split it into: "The higher horizontal resolution in the regional CAMS ensemble mean, and the LOTOS-EUROS model improves the surface level pollution estimates." and "However, the models still systematically overestimate NO2 levels at higher altitudes, indicating exaggerated vertical mixing in the models over the North Sea."

Page 2 line 29: change “intensity of ocean going ships” to “the intensity of ocean-going ships”

Page 2 line 39: For a consistent, temporally and spatially complete approach

Page 2 line 42: add “NOx” to “study emissions patterns”

Page 2 line 51: change “therefor” to “therefore”

Page 2 line 51: used for the TROPOMI NO2 column retrieval

Page 3 line 71: In this study, we investigate aircraft-based in-situ measurements of NOx (and more) over the polluted North Sea with major shipping routes and nearby industrial and densely populated centers.

Page 3 line 75: delete “profile”

Page 3 line 80: delete one of the “of” and change “satellite trace gas retrieval” to “satellite trace gas retrievals”

Page 4 line 86: delete “and” and replace with a comma

Page 4 line 89: the abbreviation S5P was already introduced earlier

Page 4 line 90: delete “scientific quality”

Page 4 line 90: retrievals of various trace gases, including NO2, since April 2018

Page 4 lines 91-93: Make two sentences and replace “the instrument” with TROPOMI or S5P. Multiple sounds like a lot, but I think it's a maximum of twice a day. “With a swath width of approximately 2600 km TROPOMI has near daily coverage at the equator. At the latitude of the North Sea (52◦N) TROPOMI (or S5P?) frequently overpasses twice a day.”

Page 4 line 96: provide reference earlier: “with the DOAS-method (Platt and Stutz, 2008) in the visible spectrum (405-465 nm)”

Page 4 line 107: delete “these”

Page 5 line 109: change improved to modified

Page 5 line 114: and vertical profiles within the lower troposphere, from the sea surface (<30 m) to 1500 m.

Page 5 line 117: delete one “of” or replace it with “the”: for the purpose of monitoring the compliance

Page 5 line 126: replace “make sure” with “ensure”

Page 6 Figure 1: for measurements in flight heights below 200 m. Blue circles indicate the location of the spiral flights.

Page 8 line 163: add a reference to Fig. 2: within the spatial extent of a TROPOMI pixel, see Fig. 2.

Page 8 line 165: The text is repetitive: “In general, aircraft spatial sampling characteristics are not uniform across a TROPOMI pixel as evident from Fig. 2.” Is already mentioned two sentences earlier.

Page 8 line 168: Maybe change the section title to: Representativeness of NO2 vertical profile measurements

Page 8 line 170: I am not sure if I understand the meaning of the first sentence.

We will rephrase this to ‘We first take care to ensure representativeness of the aircraft profiles at the scale of a TROPOMI pixel’

Page 8 line 176: “Fig 3 (left panel)” instead of “Fig. 4(a)”

Page 9 line 187: add (right panel)

Page 10 line 210: replace “sharp drop above the lowest 100 m” with “strongly decreasing within the lowest 100m”

Page 10 line 211: Split into two sentences, start with a new sentence after (Fig. 1): This is in agreement with…

Page 10 line 227: replace “at 200m and above” with “usually up to 200m and above”

Page 10 line 228: Split into two sentences: …and above. This leads to an overestimation in column NO2.

Page 11 line 232: replace “are originating” with “can originate”

Page 12 Fig.5: Caption is hard to read and sometimes misleading. Please rewrite, maybe something like this: NO2 columns (indicated by the bottom color bar) as seen by TROPOMI and several model products for the time of the first profile measurement. The aircraft measurements are overlayed in grey for flights above 200 m and in colors for flight heights below 200 m, indicated by the color bar on the right. Wind speed and direction from ERA 5 in 10m (?) are indicated by the arrows in the left panel.

Shown is indeed the ERA5 10m wind.

Page 12 line 248: Please add the date to highlight that it is a different day than the 2nd June on which profile 2 was taken, which is mentioned later.

Page 13 line 264: The very shallow pollution layer visible in the NO2 measurements is also…

Page 14 line 269: "…as observed by TROPOMI, see Fig. S1 and Fig. 7.

Page 14 line 272: While the profiles were taken right above the shipping lane, marked by the blue circle in Fig. 7, in CAMS and LOTOES-EUROS the shipping pollution can be seen south of the profile location, caused by the northerly winds.

Page 15 line 291: replace "more accurate" with "modified"

Page 15 line 292/294: Move the sentence from line 294 "As the measured NO2 profiles only extend to 1400 m, 295 we use TM5 profiles to fill the gap to the tropopause." to the beginning of the section line 292 after "…based vertical profiles."

Page 15 line 293/294: the layer index "l" is sometimes italics, sometimes not

Page 15 line 298: "The shallow boundary layer depth over sea in combination…"

Page 15 line 302: "re-calculation of the AMFs and tropospheric NO2 vertical columns"

Page 15 line 304: Pearson correlation coefficient

Page 15 line 313: RMSE already defined in line 305

Page 15 line 314: Unit is missing

Page 15 line 320: replace contrast with a different word, maybe "differences"

We believe the word contrast is the best fit.

Page 17 line 338: replace "retrieved NO2 columns" with "TROPOMI NO2 column retrieval"

Page 17 line 343: recalculate the AMFs and tropospheric NO2 columns

Page 17 line 348: in contrast to the model profiles on these measurement days

Page 18 line 382: missing bracket

References

Marécal, V., Peuch, V. H., Andersson, C., Andersson, S., Arteta, J., Beekmann, M., Benedictow, A., Bergström, R., Bessagnet, B., Cansado, A., Chéroux, F., Colette, A., Coman, A., Curier, R. L., Van Der Gon, H. A. C. D., Drouin, A., Elbern, H., Emili, E., Engelen, R. J., … Ung, A. (2015). A regional air quality forecasting system over Europe: The MACC-II daily ensemble production. *Geoscientific Model Development*, *8*(9), 2777–2813. https://doi.org/10.5194/GMD-8-2777-2015

Van Geffen, J. H. G. M., Eskes, H. J., Boersma, K. F., & Veefkind, J. P. (2022). *TROPOMI ATBD of the total and tropospheric NO 2 data products*. *2.4.0*. https://sentinel.esa.int/documents/247904/2476257/sentinel-5p-tropomi-atbd-no2-data-products