

Reviewer 1:

I would like to thank the authors for their reply to my initial review.

The first major point of my initial review regarding molybdenum converters was reasonably addressed. The second major point was addressed, but I still have an open comment regarding the discussion of AMFs. Otherwise, I see the scientific content of the manuscript fit for publication. I have attached a fairly long list of fairly basic comments.

Thank you for your careful re-review. Below in red are our responses and modifications.

I recommend publication after their consideration:

-Line. 24: This sounds like WRF-Chem has issues with photo-chemistry under cloudy conditions, but in l. 359-364, the authors argue that the modelled jNO₂ agrees to satellite measurements. It would be good to express more clearly, what other aspect of photo-chemistry WRF-Chem could be struggling with.

Thank you for this suggestion. We have modified, "This suggests that NO₂ in WRF-Chem is more responsive to sunlight and associated photochemistry than in reality."

to

"We additionally find that modeled jNO₂ values are reasonable, suggesting this WRF-Chem NO₂ bias could be arising from overly rapid OH removal of NO₂ in sunlight, too slow NO₂ regeneration of NO₂ in sunlight, or differing boundary layer depth biases under cloudy versus clear skies."

To facilitate the additions of these words, we removed 28 words elsewhere in the Abstract. The primary change was a shortening of the second sentence of the Abstract from, "While the lack of high-quality column measurements during cloudy conditions precludes validation of the satellite data, *in situ* surface measurements and model simulations can provide insight on the quantitative understanding of NO₂ during cloudy conditions."

to

"Although clouds cause satellite data gaps, *in situ* surface measurements and model simulations can provide insight on NO₂ during cloudy conditions."

Elsewhere in the Abstract we modified:

“insight on the quantitative understanding of NO₂” → “insight on NO₂”

“are different” → “differ”

“compared to restricting to clear sky days” → “compared to clear-sky days”

“is on average much larger” → “is larger”

Finally, to maintain wording consistency with the Abstract, we also modified Lines 360 – 363 from, “Second, it is possible that OH concentrations in WRF-Chem are fluctuating too rapidly in the presence of and lack of clouds (Duncan et al., 2024) causing NO₂ to be removed too rapidly in the model. Third, there might be insufficient NO₂ recycling of organic nitrates and/or particulate nitrates in the model which could buffer photolysis-related changes”

to

“Second, it is possible that *radical concentrations (OH, HO₂, and/or RO₂)* in WRF-Chem are fluctuating *improperly* in the presence of and lack of clouds (Duncan et al., 2024) causing NO₂ to either be removed too rapidly in the model *or regenerated too slowly*. Third, there might be insufficient *photolysis* of organic nitrates and/or particulate nitrates in the model which could buffer NO₂ photolysis-related changes”

-Line. 29-37: This leaves out a few considerable sources, e.g. lightning NO_x and soil emissions.

We modified the original phrasing of, “In urban areas, the majority of ambient NO₂ originates from local NO_x emissions (=NO+NO₂; most NO_x is emitted as NO which rapidly cycles to NO₂) during high-temperature fossil fuel combustion (Crippa et al., 2021).”

to

“NO_x (=NO+NO₂; most NO_x is emitted as NO which rapidly cycles to NO₂) is released into the atmosphere by biogenic microbial activity in soils, high-temperature lightning, wildfires, and thermal fossil fuel combustion; in urban areas, the majority of ambient NO₂ originates from the latter (Crippa et al., 2021).”

-Line. 53: column amounts → column densities

Modified here as requested, and also in Lines 147 and 175.

-Line. 59: The sentence about oversampling is not clear (neither grammatically, nor semantically). To my knowledge, downscaling means to achieve truly higher resolution, not just a higher “nominal” resolution by cutting a pixel into N parts with the same value. It is also not clear what purpose downscaling/oversampling has in the context of the paper.

Good catch. This sentence is not relevant to this version of the manuscript. It has now been removed.

-Line. 63: This is incorrect: The measurements are of spectra, or perhaps, in the broader sense, of column densities. The tropospheric column densities require additional retrieval steps, that go far beyond pure measurement.

Thank you for catching this important missing nuance. We have now rephrased to add additional context, from “Satellite NO₂ measurements are of the tropospheric column.”

to

“Level 2 satellite NO₂ measurements – which are retrieved from observed NO₂ spectra using geophysical and model-based assumptions – are of the tropospheric column.”

-Line. 80: oxidation environment → oxidative capacity?

Modified as suggested.

-Line. 88: Even without dry deposition, a more shallow boundary layer will enhance the surface concentrations of trace gases, simply due to compression of the air mass into a smaller total volume.

Yes, we agree. This sentence was missing a key word that may have been leading to your confusion. This sentence originally and incompletely stated that dry deposition increases with a shallower boundary layer, but we meant that *dry deposition fluxes* increase with a shallower boundary layer since the dry deposition velocity itself does not increase but the concentration does (i.e., dry deposition flux = dry deposition velocity x near-surface concentration). The word “fluxes” has been added to this and the following sentence.

-Line. 99-102: It is not clear what the difference between point 1) and point 2) is, except that 1) refers to TROPOMI and 2) to geostationary satellites. Besides, the two points seem to express the same, namely to quantify the cloud-bias of NO₂ satellite measurements.

Good catch. This sentence has been shortened to only refer to point 1, from, "The motivation of this project is two-fold: 1) to determine what the scientific community may be missing when excluding clouds from TROPOMI-based NO₂ analyses and 2) to understand how geostationary NO₂ satellite measurements may be affected by such a bias and potentially partially remediate such a bias."

to

"The motivation of this project is to determine what the scientific community may be missing when excluding clouds from satellite-based NO₂ analyses.

-Line. 108-109: I suggest to make it more clear that the intent of the measurement is to quantify NO₂ and NO, and that the conversion of NO_y is an unwanted side-effect.

We added the word "unintendedly" to this sentence. We also modified NO_y to NO_z which is a better reference to the specific nitrogen species (PAN, HNO₃, etc.) that we are referring to.

-Line. 111: NO₂^{*} (meaning "true NO₂ + false NO₂") is not explicitly defined.

Thank you for catching this oversight. NO₂^{*} is now defined as NO₂ + unintended NO_z

-Equation (1): The formula of Lamsal et al. (2008) contains an additional term for alkyl nitrates that does not occur here. This is fine, as WRF-Chem simulations often do not include alkyl nitrates in the first place, but it should be mentioned.

We have now added an additional sentence here to clarify this, "In Equation 1, we show the Lamsal et al. (2008) correction factor with a modification to exclude alkyl nitrates which are not explicitly included in our WRF-Chem simulation."

-Line. 132-133: Does this statement refer only to TROPOMI or also to TEMPO?

This is in reference to both instruments, which measure NO₂ in this wavelength range. I think your confusion arises from the fact that we only reference a TROPOMI paper in this section (van Geffen et al. 2021). At the time of original submission, the TEMPO ATBD was unavailable, but it has since been published (Feb 2025) and is now referenced in addition to van Geffen et al., 2021.

-Line. 244: Redundant "13:30"

Thank you for flagging this. We have opted to keep this as-is to maintain clarity. A reader could be confused that the "cloudy day average" or "annual average" are 24-hour averages if any of the "13:30" qualifiers are removed.

-Line. 248: The two sentences with "17.2 %" appear to state the same.

The second sentence has been removed.

-Line. 260: Here (and in other places), change "no cloud days" to "cloud free days".

All six instances have been modified as requested: Figure 3 caption (twice), Line 260, Figure 4 caption, Figure 6 caption, Figure 7 caption.

-Fig. 3: Mention in the caption that one scatter point corresponds to one measurement station.

Thank you for requesting this additional clarification. We have added the following clarification to the caption of Figure 3, "Each scatter point corresponds to each of the 374 measurement stations."

-Line. 271: decreased photolysis is also another reason.

Added "decreased photolysis" to this sentence as suggested.

-Line. 286-287: "small (but larger)" is not clear to me.

Thank you for pointing this out this unclear phrasing. We meant "small - but generally still positive - difference", and have modified the phrase as such.

Line 314-326: A few (minor) comments to this section: Small AMFs are not caused by the model "filling in" the missing NO₂, but the other way around: If the AMFs are small, it implies that the satellite had reduced sensitivity, and the retrieval depends more on a priori assumptions (here: profile shapes taken from the 1° × 1° model TM5). Note, that clouds can also have an opposite effect, where the sensitivity can be enhanced above bright clouds. In its current form the section generally goes in the right direction, as it identifies the challenge of computing AMFs under cloudy conditions. However it still does not explicate the fundamental limitation I mentioned in my first review.

I try to explain my concern again: Assume, that the TM5 model works fine under cloud-free conditions, but has shortcomings under cloudy conditions. This is

theoretically possible and cannot be ruled out (the authors' own results described in section 3.4 seem to be affected by this). For example, if the model had biased estimates of radiative transfer through clouds, this would result in faulty NO₂ photolysis rates, and possibly faulty NO₂ profile shapes. Note, that there exist plenty of other physical effects that could be modelled incorrectly. Then, the comparison is made between cloudy days (that are affected by said bias) and cloud-free days (that are unaffected by said bias). The differences one sees in this comparison could (in theory) be produced by the faulty a priori assumptions instead of actual differences in the NO₂ abundance. The authors mention a very similar notion wrt. the surface NO₂ in their WRF-Chem simulation in l. 356-357. The same logic should apply to vertical column density retrievals. In other words, the presented work can answer the question: "How different are retrieved NO₂ VCDs under cloudy/cloud-free conditions?" but not "How different are the actual NO₂ VCDs under cloudy/cloud-free conditions?". The first question acknowledges that our a priori knowledge could be faulty, while the second question asks for more fundamental results that cannot be obtained in the presented methodological framework. I think the authors must distinguish clearly between pure measurement and retrieval, also wrt. to specific phrasing, e.g. in Line 63, see above.

We are in full agreement with these comments. We presented the analysis in Lines 314-326 to answer the question of "How different are *retrieved* NO₂ VCDs under cloudy/cloud-free conditions?" and never intended to answer the question of "How different are the actual NO₂ VCDs under cloudy/cloud-free conditions?"

We have now qualified this section to state our intention more clearly, "We then compare TROPOMI NO₂ measurements under varying sky conditions *to understand how the retrieved NO₂ columns differ under cloudy and clear-sky conditions, but not to answer how they actually differ.*"

Next, we apologize for an incorrect statement attributing differences to the vertical shape profiles; we meant scattering weights.

"primarily driven by the assumed vertical shape profiles in the retrieval."

to

"primarily driven by the assumed scattering weights in the retrieval."

Additionally, we have modified the sentence “During measurements when the crf >0.5 as compared to measurements when crf <0.5, the model is “filling in” the missing NO₂ and causing small air mass factors”

to

“During cloudy scenes, scattering and reflection by clouds reduce the satellite’s sensitivity to near-surface NO₂, leading to smaller air mass factors”

Finally, we have added the word “retrieved” in several instances throughout the section to more clearly articulate the TROPOMI columns as “retrieved” instead of “actual”.

Line 346: Two comments here: Firstly, as stated by the authors previously, NO₂* is also affected by PAN and alkyl nitrates.

Thank you for catching this. We have now added “PAN and alkyl nitrates” to this sentence.

Secondly, the issue with the molybdenum-based measurements is not that HNO₃ is a terminal NO₂ sink, but that it simply introduces a strong measurement bias (as in: the measurements are not correct).

We agree that mentioning the phrase “terminal sink” was unnecessary and confusing. We meant that the NO₂+OH → HNO₃ pathway is photochemically driven and therefore by measuring some amount of HNO₃ this would not capture the full photochemically driven change of NO₂. We have modified from, “the NO₂ + OH → HNO₃ reaction is a terminal sink for NO₂ during daytime”

to

“the NO₂ + OH → HNO₃ reaction is a photochemically-driven pathway for NO₂ during daytime”