Review of Kylling et al. 2021: Impact of 3D Cloud Structures on the Atmospheric Trace Gas Products from UV-VIS Sounders -- Part III: bias estimate using synthetic and observational data

General comments:

The authors did a comprehensive analysis of 3D cloud effects on synthetic and observational data of NO2 concentrations. A lot of simulations were performed regarding the NO2 and AMF bias due to 3D cloud effects, and the biases as functions of physical parameters such as illumination geometry, viewing geometry and surface albedo were explored.

The presentation of the results needs an improvement. Most figures are difficult to read: they contain too much information (Fig. 9 and 10) and/or are too small (e.g., Fig. 12) and/or a different type of graph could be more suitable (why are the dots in Fig. 5, 9 and 10 connected by lines?) and/or could better be skipped (Fig. 9.b and 10.b) (see detailed comments). Please avoid legends overlapping the graphs (e.g., in Fig 9 and 10). The description of some figures is sometimes insufficient (Fig 2. for example, does not have any explanation in the text), or is unclear (e.g., what is the added value of the skewness etc. in Figs 9.b and 10.b to your conclusions)?

The VIIRS cloud mask and cloud shadow mask are used, but are not reliable for cloud and cloud shadow identification in the analysis of cloud shadow signatures in TROPOMI data. This is because cloud shadows are small-scale features (1 or a few TROPOMI pixels). The overpass measurement time difference between TROPOMI and VIIRS, which is a couple of minutes, is enough to move clouds at least 1 TROPOMI pixel (see ESA-ATMOS symposium oral presentation by Trees et al., 2021 https://atmos2021.esa.int/iframe-agenda/files/Contribution_171_final_extabs.pdf). This is particularly true for clouds that produce cloud shadows visible from space, because those shadow-producing clouds must be located at high altitudes where the wind speeds are relatively high. Additionally, near the cloud edges cloud evolution (i.e., cloud shape change) occurs. Cloud shadows should therefore be identified using measurements taken at the TROPOMI measurement time.

In Section 4.2.1, zoomed in areas are considered (a few pixel rows) of only two cases, while the spatial natural variability of the NO2 VCD is actually very high (see e.g. your Figs. 2.c, 12.d and 13.d). In order to make sure that the observations are caused by 3D cloud effects instead of the natural variability, statistics of more observations are needed. In Section 4.2.2 shadowed pixels are compared with shadow-free neighboring pixels, but the cloud movement and evolution (see former paragraph) could consistently result in the situation where the actual shadows are located inside the neighboring pixels, while the identified shadow pixels are in fact shadow-free. Consistently confusing the shadowed pixels with non-shadowed pixels may result in false conclusions about the observed shadow induced NO2 signatures.

The results shown in Section 4.2 are highly scattered (indeed no clear relation between TROPOMI NO2 and VIIRS shadow fraction can be derived from Fig. 11.g, Fig. 14 and Fig 15). However, in the conclusions and abstract, it is written that NO2 appears low-biased in observations. Considering the scattered results, together with the questionable approach that was followed (see former paragraphs), no reliable conclusions can be drawn from this analysis about the observed NO2 bias dependence on shadows. Therefore, I suggest removing Section 4.2 and Appendix A from this paper and limit the analysis to only the modelling part (Section 4.1).

Detailed comments:

- Page 1, line 10-11: limit your conclusion to only synthetic data.
- Page 1, line 11-14: the low NO2 bias in the observations is not significant (see later comments below).

- Page 4, Fig 1.b and Fig 1.c: Please avoid using a white color for the cloud pixels, since white is also one of the colors in the color bar.
- Page 4, Fig 1.c: The cloud shadow index is 1 below the clouds, but this would be invisible from space. I suggest removing the cloud pixels in this plot (such as in Fig. 1.b), such that the cloud shadows become visible.

- Page 5, Fig 2.a: It is difficult to see the colors of the RGB. Can you try to increase the brightness and/or enhance the colors?
- Page 5: Fig 2: Please increase the size of the images. Discuss every subfigure when you introduce Fig. 2 in the text or remove the subfigure.

- Page 6, line 13: Why do you mention FRESCO here? Are you using FRESCO for your analysis? This is not yet clear for the reader at this point.
- Page 6, line 28: You do not consider ocean cases. On Page 3, line 18, you mention that your focus is on Europe and different cloud types, and that therefore the results are expected to be general and applicable elsewhere. Would your results also be representative over ocean? And over desert, or over snow/ice?

- Page 7, line 1-3: 'they compare the spectral procedure', 'and found the latter to be far superior'. Could you rephrase this? What is 'the spectral procedure'? What do you mean by 'far superior'?
- Page 7: what are the accuracies of the VIIRS cloud mask and VIIRS shadow mask (regardless of the mapping to the TROPOMI grid)?
- Page 7, line 16: From this explanation, it seems that the Cloud Shadow Index (CSI) also indicates fully cloudy pixels (as also shown in Fig. 1.c). Why is this the case? Wouldn't excluding cloudy pixels be better for the analysis of shadows which uses the CSI, such as Fig. 3? Or did you indeed apply a cloud filter? Please explain in this subsection, this is not clear for the reader at this point.
- Page 7, line 29-30: Why are the "" used here? Is this a citation? This is not a proper explanation of the AAI. Please rephrase and add a reference to de Graaf et al. (2005): https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2004JD005178 . For example: 'The AAI is a measure of the UV color of a cloud-, aerosol- and shadow-free 1-D atmosphere-surface model with respect to the measured UV color (de Graaf et al., 2005). When absorbing aerosols are present, the AAI tends to be positive, while the AAI is approximately zero or negative in the presence of clouds (see e.g. Kooreman et al., 2020; Penning de Vries et al., 2009).'

- Page 8, Fig 3: 'Please remove cloud pixels from your analysis, and if you did (already), please mention this in the caption of the figure.'
- Page 8, line 14-16: "The cloud shadow impact … respectively." Please rephrase, this sentence is difficult to read. Are the percentages you mention here average values for a CSI of 1?

- Page 8, line 16: “As the solar zenith angle increases a linear relationship appears …”. Is there really a linear relationship? The data looks scattered. Can you please quantify the linear relationship with the corresponding uncertainty?
- Page 8, line 20: “For both geometries the NO2 AMF is high … is between 1-3%.” Why? Can you physically explain these numbers here already? If not, please refer to this location in the paper explicitly when you can. For example, on page 12 line 11, you investigated the cause, and you can refer back: ‘This explains the high AMF biased that we observed in Fig. 4a for CFw < 1%’.

- Page 9, line 3: “… there are comparatively more pixels with a negative bias for LEO geometry”. Why? In lines 4 to 14 you explain that this is because the SAA en SZA are different, giving different sensitivity to cloud shadows for LEO and GEO geometries. Can you explain why this is the case?
- Page 9, first paragraph: Please discuss Figure 5 in a separate paragraph.
- Page 9 and 10, general comment: The results of the parameters such as SZA and surface albedo are discussed. The physical explanation is missing. After each finding, can make a connection here with the theory from your first paper (Emde et al. 2021)?
- Page 9 and 10, please explain better from theory of Emde et al. (2021) what the reader should be aware of when comparing LEO and GEO images (given the different SAA and SZA). What are the interesting differences between LEO and GEO results that you expect to see? And do you also observe in these simulated results what you expect from theory (Emde et al. 2021)?

- Page 10, Fig. 5: Why are the dots connected by lines, for example for (SZA=20 deg; SAA=45 deg) and (SZA = 40 deg, SAA = 270 deg)? Please reconsider the presentation of these results. Using 9 lines (for different albedo and SZA) instead of 3, or a bar chart, would suit better here.

- Page 12, line 4: “east/west”. How is the solar azimuth precisely defined? Make clear which SAAs belong to west and east.
- Page 12, Fig 8: please relocate the tick labels of the color bar such that it is clear to which color they belong.
- Page 12, Fig 8: What are the tick labels 40 and 20?
- Page 12, line 7-11: “Generally … effects.” These lines are floating in the rest of the text, because they are a discussing of Fig. 4. Please move those lines to the discussion of Fig. 4, or make a connection to the former paragraphs.

Page 13-14, general comment: Please reconsider Figures 9 and 10. Consider replacing Figures 9 and 10 by figures that show the NO2 bias as functions of physical quantities such as SZA and albedo. This could make it easier to connect with the theory of Emde et al. (2021).

- Page 13, Fig 9: Fig. 9 contains too much information. Why are the lines connected? A bar chart may suit better here. What are the different case numbers? It is not clear from the figure or the caption. Please prioritize the results you want to show and possibly compute averages of the cases. Think about the message you want to convey with this figure.

- Page 13, Fig 9b: Is Figure 9.b really needed for the conclusions of your paper? Similar comment for Fig. 10b.
- Page 13, line 11: 'for similar reasons': What reasons? Additionally, are the reasons of the contamination as functions of SZA and VZA really (expected to be) identical? If yes, why?
- Page 13, line 11-12: 'under- and overestimates', under- and overestimates of what? The NO2 bias or AMF bias?
- Page 13, line 12: 'Cloud shadows are a cloud feature metric that may be used to identify affected pixels, Fig 3.' What do you mean by this sentence?

- Page 14, line 1: 'occurr.' -> 'occur.'
- Page 14, line 1: '… also occur.' Can you refer to the figure(s) where this was shown?
- Page 14, line 1-2: '…, such as cloud top altitude and cloud optical thickness, are also of importance.' How did you come to this conclusion? Can you show this or refer to the figure where this has been shown?
- Page 14, line 5: How precisely is the cloud enhancement effect visible in Figs. 9 and 10? Please explain.
- Page 14, line 7: 'theta = 20 – 30 degrees'. How can the viewing zenith angle be observed in Fig. 9a?
- Page 14, line 12 to Page 15, line 4: What is the message of this paragraph? Do you mean that 3D cloud and cloud shadow effects are smaller than the NO2 retrieval uncertainty?

- Page 15, line 7-12: Should these sentences be part of Section 4.2.2 instead?
- Page 15, line 20-22: please rephrase: : "For a cloud shadow fraction … standard deviation". Please add: "The scatter in Fig. 11.g is too large to draw conclusions about the dependence of NO2 on shadow fraction."
- Page 15, line 24-25: "Thus indicating that … cloud cases." Please rephrase this sentence: what is the subject of this sentence?

- Page 16: can you please make the figures bigger?. Also, in Fig. 11.g, the lime green squares are not visible.
- Page 16, Fig. 11: can you explain the oscillatory pattern in the geometric cloud fraction (Fig. 11.d) and the cloud shadow fraction (Fig. 11.f) in the shadow band?
- Page 16, fig 11.g: how do you precisely define "CSF pixels"? Doesn't each pixel has a certain CSF? Please clarify this in the caption.
- Page 16, Fig 11.g: the variability is much larger than the differences between NO2 (all pixels) and NO2 (CSF pixels). No significant relation between NO2 and CSF can be identified with this figure. Please clarify this explicitly in the text.

- Page 17, line 4: "For the cloud shadow band the NO2 TVCD is on average reduced by 17%". Please add a sentence here explaining that only a few pixel rows are analyzed, while the NO2 natural spatial variation is actually very large (Fig. 12.d).
- Page 17, line 22-23: "All cases show that the NO2 TVCD in the cloud shadow is lower by 8-46% (average of 25%) compared with the NO2 TVCD around the shadow." What about pixel row 396? Pixel row 396 seems to have a higher NO2 TVCD in the shadow than south of shadow.

- Page 17, line 26-28: “If it is assumed that the clouds are the main reason for the variations in the NO2 TVCD over the cloud shadow bands, then these cases are examples of how cloud shadows give underestimates of NO2 TVCD, in agreement with the theoretical idealized box cloud results presented by Emde et al. (2021) and Yu et al. (2021).” I don’t think you can conclude this, given the high spatial NO2 variability (Figs. 2.c, 12.d, 13.d), the limited number of cases and pixel rows that were analyzed, the high scatter of the NO2 bias as functions of shadow fraction (Fig. 11.g), and the questionable approach to mask clouds and shadows using VIIRS masks on the TROPOMI grid (due to the cloud movement and cloud evolution during the TROPOMI-VIIRS overpass time difference, the undiscussed VIIRS mask accuracy, and the oscillatory features in the geometric cloud fraction and shadow fraction in the shadow band (Fig.s 11d and 11f)).
- Page 17, line 34: TROPOMI processes 25 million pixels per day. Why do you use for October 2018 and March 2019 only 1023081 pixels? What is the study region precisely?

- Page 18, line 1: 35% of what precisely? 35% is a large percentage for cloud shadows, even in months where you expect cloud shadows. Can you please verify this number? How does this number relate to the overall cloud fraction of the data set?
- Page 18 and 19, Fig 12 and 13: Fig 12 and 13 are hard to read. Please make the figures bigger. Figure 12.b and 13.b are problematic: can you ensure that the cloud movement and evolution during the TROPOMI-VIIRS overpass time difference did not consistently affect the shadow identification?
- Page 18 and page 19, Fig.12.g and 13.g: only a couple of pixel rows are analyzed, and even within this small sample, the low NO2 bias is not consistent. For example, in Fig. 12.g, the NO2 TVCD is higher in the shadow than outside the shadow for rows 262 and 265.

- Page 20, line 10: “no true NO2 TVCD is available as for the synthetic data” -> do you mean “observational data”?

- Section 4.2.2 general comment: The results in Figs. 14 and 15 are highly scattered, and no clear negative NO2 bias from cloud shadows can be determined. This should be clear in the text, conclusions and abstract.

- Section 4.2.2 general comment: neighbor pixels in a 3x3 pixel matrix where used, and the true NO2 TVCD is then taken to be the average of the cloud-free neighbors. Cloud movement and evolution during the measurements time difference of TROPMI and VIIRS could consistently result in the situation where the actual shadows are located inside the neighboring pixels, while the identified shadow pixels are in fact shadow-free.

- Page 23, line 15-19: “For clearly identified cloud shadow bands … with the theoretical findings.” Why can you assume that the clouds are the main reason for the spatial NO2 variations / assume that the NO2 background is horizontally homogeneous?
- Page 23, line 20-21: “For a solar zenith … to be impacted by cloud effects larger than 20%”. Where did you show this? Also, please mention that the data is very scattered and comment on the uncertainty of your conclusions.

- Page 24, line 1: You mention that there are “large changes between versions” of the VIIRS cloud shadow product. Could you elaborate on that? What is the accuracy of the VIIRS cloud shadow product itself (regardless of the mapping onto the TROPOMI grid)?

Appendix A

- Page 24, line 12: “As cloud shadow impact NO2 TVCD retrievals, …” Do you mean instead: “As cloud shadow impact both AAI and NO2 retrievals, …”?
- Page 24, line 15: “Indeed, over land the AAI is more negative over cloudy pixels, compare Fig. 11d and Fig. A1a”. -> This seems not really to be the case when looking at Fig. 11d and Fig. A1a: the large cloud deck between 52 deg N and 52.5 deg N does not give more negative AAI. Clouds do not always decrease the AAI, they usually just don’t *increase* the AAI (see e.g. Penning de Vries et al., 2009).
- Page 25, line 2: “…, while the NO2 TVCD shows some dependency on cloud shadow fraction, Fig. 11g.” -> Please remove this part, the dependency on cloud shadow fraction from Fig. 11g is insignificant given the high variability.