

Atmos. Meas. Tech. Discuss., referee comment RC2
<https://doi.org/10.5194/amt-2020-433-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on amt-2020-433

Christian Frankenberg (Referee)

Referee comment on "Monitoring the Tropospheric Monitoring Instrument (TROPOMI) short-wave infrared (SWIR) module instrument stability using desert sites" by Tim A. van Kempen et al., Atmos. Meas. Tech. Discuss., <https://doi.org/10.5194/amt-2020-433-RC2>, 2021

The manuscript Monitoring the TROPOMI-SWIR module instrument stability using desert sites by van Kempen et al outlines an approach for TROPOMI SWIR instrument stability monitoring using PICS site over deserts. While the manuscript is in general well written, I do have some high level questions/concerns regarding the ultimate utility of the study for TROPOMI validation apart from stating that it is in line with the onboard calibration routines. In the following, I will briefly describe some of these concerns, followed by a few detailed comments on typos, etc.

Major comments:

As far as I can see at the moment, the entire outcome of this investigation is a bulk characterization of the 2D detector response at one specific small wavelength band averaged over the entire spatial domain. However, I haven't seen any discussion on the implication of this limiting factor at all, which I consider rather large. First of all, the across-track pixels might all degrade to a different degree or have an overall different absolute calibration error. However, their response is also impacted by potential BRDF effects as each across track element has its own viewing zenith angle. Given the data density of TROPOMI, I was somewhat surprised that the authors didn't try to at least disentangle some of the across-track element variations. At the moment, I am somewhat uncertain what new information this manuscript reveals, especially as the scatter is surprisingly and the variations in slopes quite variable too. The authors would have to better explain the added value of this method (on top of the on-board calibration, which can characterize the entire FPA response). The origin of the scatter would have been an interesting feature to dig deeper into, but the authors chose not to, which is somewhat dissatisfying, as this could have been valuable to the community.

Minor issues:

Line 31: "calibrated column densities" These are retrieved products from calibrated spectra, not itself "calibrated" datasets (maybe validated and some post-hoc "calibration" like bias correction applied)

Line 48: allpart

Line 50++ Here you mention all kinds of impact factors but then chose to ignore all of them. Why not use TROPOMI and its large swath to actually see whether you can detect BRDF effects that can clearly be separated from detector effects across the spatial domain.

Line 57: "but also suffers from inaccuracies" Across the manuscript, statements like this are scattered. If you point out a weakness of an instrument, you will have to justify the statement with a citation or elaborate how you come to the conclusion. However, you can't just make a statement like this out of thin air in a peer-reviewed publication. Also, what does "most complete" mean in this context?

Line 57: "due to its very wide swath opening" --> just swath is fine, swath opening sounds awkward

Line 65 --> used for monitoring the stability of a large number of ...

Line 76: Why only every 5 days? The cloud cover should be low, so I don't understand the 5 day limit. Is it the overlap requirement? With MODIS data being available, you should also be able to determine the impact of the exact spatial overlap across variable surfaces with slightly varying albedos. As far as I can see, no attempt has been made to compare against MODIS data (e.g. to look at sub-pixel variability, etc).

Table 1 caption typos: coordinates... variability...

Line 86++ Please better explain the cloud screening, at least give a proper citation. Has any screening for desert dust events been performed? Any other filters?

Line 94: Why did you choose 50 degrees as cutoff even though this basically omits a non-significant fraction of TROPOMI's FPA? Have you checked whether adding the few additional degrees make any differences? Did you consider separating out the FPA (and thus VZA) dependence as mentioned above?

Line 105: "are of insufficient quality to reliably improve the data" Please see my comment above. Without citation or justification with analysis, this statement is misplaced at best and mean-spirited at worst. Any judgement statement like this requires corroboration.

Line 106: "A choice was made" What was the rationale of that choice? Did you consider the tradeoffs? Why not bin the analysis by viewing angles and see whether the scatter is reduced?

Line 112: affect -> affects

Line 116: "using standard mathematical rules" like what? Just gaussian error propagation? Please be specific if you can, esp. if it doesn't take up more space than "using standard mathematical rules", which is rather vague.

Figure 2: This scatter is indeed large and clear outliers exist. Are these actually single measurements? If yes, can they be color-coded by the detector position or VZA (plus and minus)? Did you try to figure out why a few were low outliers by looking at the conditions during that time (or the specific detector position?)

sine wave correction: This is rather vaguely described to be honest. It would be good to show such a fit. Does it look better if fitted against AMF or SZA? Is this something that is also seen in MODIS data? This is interesting and curious but again, the authors chose to not go the extra mile, which would have made this paper much more interesting. In general I have no problem with not diving deeper into all the issues but given that the overall relevance of this manuscript for TROPOMI validation or validation schemes using PICS in general is rather thin, I would have expected a somewhat deeper analysis into these small curiosities.

Line 131: "We attribute..." What is the basis of this, a hunch? You could actually look whether there is a VZA dependence! Why not do that, I really don't understand that choice.

