

Atmos. Meas. Tech. Discuss., referee comment RC3  
<https://doi.org/10.5194/amt-2021-253-RC3>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on amt-2021-253

Anonymous Referee #1

---

Referee comment on "Dealing with spatial heterogeneity in pointwise-to-gridded- data comparisons" by Amir H. Souri et al., Atmos. Meas. Tech. Discuss.,  
<https://doi.org/10.5194/amt-2021-253-RC3>, 2021

---

'Dealing with Spatial Heterogeneity in Pointwise to Gridded Data Comparisons' demonstrates a new methodology for comparing point measurements to those that are more spatially representative of an area (e.g., satellite or models) for the purpose of data product validation. This topic takes an important challenge with these types of data comparisons for the purposes of validation and attempts to lay out a solution going forward. The manuscript is very well written and organized and it fits well within the scope of AMT and should be considered for publishing after some minor changes.

General comments:

- The paper switches back and forth between the nomenclature of using the method in relation to validation vs. data comparisons. While the methodology and conclusions made related to using point measurements for validating grid-like data are completely valid and important, the language in this paper is dismissive of other potential reasons to compare point-like measurements with something like a model or satellite. Therefore, the suggestion is to clarify that this methodology is for validation and not generalize conclusions for data comparisons.
  - Examples of this type of language include, but are not necessarily limited to:
    - Line 80-81 with the question of whether a 'comparison ever logical' – in the right case the comparison could be logical when trying to learn how satellite data can be interpreted in relation to a ground-based measurements
    - Line 551: change 'comparisons' to something along the lines of validation
    - Line 572: 'point-pixel comparisons' should say pixel validation with point measurements or something along that line.
    - Figure 9: in the right orange box change 'comparison between satellites and observations' to 'validation of satellites (models) with point observations'
  - Could the authors comment on the reality of ground-based networks that could actually contribute to satellite/model validation with this methodology? Does the required observational density exist anywhere? What are some paths forward/recommendations?

- Section 4: Why is v3.0 OMI data using instead of the most up to date v4.0?

Lamsal, L. N., Krotkov, N. A., Vasilkov, A., Marchenko, S., Qin, W., Yang, E.-S., Fasnacht, Z., Joiner, J., Choi, S., Haffner, D., Swartz, W. H., Fisher, B., and Bucsela, E.: Ozone Monitoring Instrument (OMI) Aura nitrogen dioxide standard product version 4.0 with improved surface and cloud treatments, *Atmos. Meas. Tech.*, 14, 455–479, <https://doi.org/10.5194/amt-14-455-2021>, 2021.

#### Specific comments:

- The first sentence in the abstract implies that the two communities have zero realization of the point vs grid problem, which is not true. This is a known problem with the lack of an easy solution. Please consider rephrasing.
- Lines 28-29: This study demonstrates a method but it doesn't actually prove that the only available method 'must taking kriging variance...'etc. State what the paper demonstrates without implying there is no other alternative to this exact method.
- Line 50: consider adding some clarity to the hypothetical scenario by adding after 'atmospheric model' the phrase 'simulating CO2 emissions'...
- Line 191: is there a reference for the terminology of 'the sill'?
- Paragraph spanning lines 232-239: Use the word stratified somewhere to connect to the second row of the figure.
- Line 385: add that this is also a roadmap for model evaluation as well.
- Line 414: add a reference for the length scale of NO2 if connecting it to the range found in those months.
- Lines 433-440: More clearly explain the rational more clearly between 10 vs. 15 vs. 20. In the figure alone it isn't all that clear. Was the choice of 15 quantified somehow as the best option or was it subjective?
- Line 449: Should be Herman et al. 2009
- Line 454-456: what resolution is OMI oversampled to?
- Lines 456-457 and Figure 12 caption. The wording states that the total column was subtracted from the stratospheric column, when it is the stratospheric column that should be subtracted from the total column. Please fix this wording.