

Atmos. Meas. Tech. Discuss., referee comment RC1
<https://doi.org/10.5194/amt-2021-253-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Review

Anonymous Referee #3

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-RC1>, 2021

General Comments

The authors introduce the kriging technique, originally from the geostatistics community, to the atmospheric science community. To properly compare point samples (such as Pandora spectrometer) and large satellite pixels (such as Ozone Monitoring Instrument (OMI) pixels), the authors take the following steps: 1) construct a semivariogram that takes account for spatial variances among the point samples, and subsequently produce a kriging estimate based on the point samples (as well as an error estimate) over a 2D grid, and 2) convolve the kriging estimate and error using a spatial response function that represents the large satellite pixel size. The authors first show detailed examples of this process using typical theoretical cases. Finally presented is an actual case comparing NO₂ columns from point measurements from Pandora instruments and OMI.

This paper is well organized and fits well in the scope of Atmospheric Measurement Techniques. As this paper introduces a new, useful technique to the atmospheric science community, adding explicit statements to help readers' better understanding will greatly benefit the community. I would recommend the paper for publication in AMT after addressing the specific comments listed below.

Specific Comments

Overall: I suggest using the term "grid box" rather than "grid" when it actually means a grid box.

Line 75: Studies to downscale satellite pixels using high-resolution model simulations worth to be mentioned, e.g., Kim et al., 2018, and Choi et al., 2010 (already referred to in the text).

Kim, H. C., Lee, S.-M., Chai, T., Ngan, F., Pan, L., and Lee, P.: A conservative downscaling of satellite-detected chemical compositions: NO₂ column densities of OMI, GOME-2, and CMAQ, *Remote Sensing*, 10, 1001, <https://doi.org/10.3390/rs10071001>, 2018.

Line 115, Eq. (1) and Eq. (2): I think h in $g(h)$ should be boldfaced, as in $f(x+\mathbf{h})$ in the later part of the equation? Although this study deal with isotropic cases only, the length of vector \mathbf{h} , i.e., $h = |\mathbf{h}|$, needs to be explicitly defined or mentioned before being used in the text.

Line 122: Parallel to the above point, I suggest explicitly mentioning that $\gamma(h)$ here is only dependent on the distance between samples, not direction/angle.

Line 130, Eq. (4) While the authors state that $a_o, b_o, c_o=1.5$, they fit paired samples into the given Gaussian function. Therefore, these coefficients cannot be fixed values. Moreover, red lines in all figures variate as well. Please check.

Line 140, Eq. (7): Does Z_o means $Z(x_o)$ (true value at x_o)? Please define x_o and Z_o .

Line 145, Eq. (9): Do $\gamma_{j_1j_2}, \gamma_{j_1o}, \gamma_{oo}$ mean $\gamma(x_{j_1}-x_{j_2}), \gamma(x_{j_1}-x_o)$, and $\gamma(x_o-x_o)$, respectively? Please define them explicitly in the text.

Overall Sect.2.1: There are too many subscripts $_o$ or o . The coefficients of the Gaussian function are $a_o, b_o,$ and c_o . A specific, random point we want to estimate Z is x_o . l_o is a constant weight. Z_o is (probably) the true value at a point x_o . Some of these subscripts are relevant but others might not. Removing unnecessary 0 subscripts may help readers understand better.

Line 163-164: The authors take 200 samples and make 100 pairs. However, the maximum number of pairs seems ${}_{200}C_2$, or ${}_nC_2$ (n =number of samples), according to the first row of Fig.2 and the second panel of Fig. 13. This is worth mentioned here.

Line 165 and Fig. 1: The authors mention that the semivariograms except C1 fit to the Gaussian function. I am not sure if the semivariogram of C2 is really Gaussian.

Line 223, Fig 2, and Fig. 3: The authors mention the relative error of C5 in line 223, but it is not shown in Fig 2 and Fig 3. Locations with large Z values will naturally exhibit larger error values as shown in the fourth columns of Fig. 2 and Fig.3. However, showing relative error in these figures might be meaningful, as reasonably illustrating the plumes is more important than the absolute value of the error in this case.

Line 248: I suggest replacing "realization" with "kriging attempts" for better understanding.

Line 337: The authors mention a two-dimensional super Gaussian spatial response function used in Sun et al. (2018) study. Although this function is a critical component in the actual OMI-Pandora case study (line 473), it is not used with the theoretical cases. Therefore, this sentence is hanging in this line and may confuse readers. It needs a better location in the paper.

Regarding the super Gaussian function: Again, although convolution with this function is a critical component in the actual OMI-Pandora case, no explanation or visual illustration has been made. Instead, only the uniform spatial response function/ideal box kernel is visually illustrated in Fig. 5. I suggest showing a figure showing the convolved C5opt with the super Gaussian function (possibly using various parameters, comparable to Fig. 5) when introducing the super Gaussian function. This figure may go into Supplement.

Line 339 and Fig.5: The authors suggest $S[m,n] = 1/(m*n) J_{m,n}$ (J is the matrix of ones) as a uniform spatial response function. Also mentioned is that the panels in Fig. 5 are convolved with an ideal box kernel in the caption of Fig. 5. In summary, the authors mention that "If the spatial response function is assumed to be an ideal box, the resulting grid will represent the average." Putting them all together, " $S[m,n] = 1/(m*n) J_{m,n}$ (J is the matrix of ones)" is the "ideal box kernel", and convolution with an ideal box kernel actually means taking the average within a grid box, which should be explicitly stated here.

Line 339-340, Eq. (14): Does $S^2[m,n]$ in Eq. (14) means the squared $[m,n]$ -th element in the matrix $S[m,n]$ in line 339? If so, please explicitly state it. Also, taking a matrix notation for the matrix $S[m,n]$ will be helpful.

Line 358: Is there a reason for not showing the synthetic satellite measurements (upscaled truth)? I am interested to see how synthetic satellite measurements compare with the kriging estimate in the 30x30 resolution (although it is before converging) and the converged kriging estimate (1x1) resolution convolved into the 30x30 resolution.

Technical corrections:

Line 130, Eq. (4): In the equation, some subscripts look like alphabet o while others look like number 0. Please make them consistent.

Line 138 & Eq. (10): Also, while the subscript of x_0 in line 138 looks like a number 0, subscript of x_o in Eq. (10) looks like an alphabet o.

Fig. 10: Dots showing uniform sampling locations are barely visible. Can they be more visible?

Fig. 12: The resolution is $0.2^\circ \times 0.2^\circ$ (approximately $20 \times 20 \text{ km}^2$), instead of $20 \times 20 \text{ km}^2$.